Discussiones Mathematicae
Probability and Statistics 37 (2017) 79–99
doi:10.7151/dmps.1196

DEFAULT PROPENSITY IMPLICIT IN PULLED TO PAR V@R FOR BONDS

Manuel L. Esquível[1,2]

Departamento de Matemática
Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa
Quinta da Torre, 2829-516, Caparica, Portugal
Centro de Matemática e Aplicações (CMA/FCT/UNL)

Raquel M. Gaspar[3]

ISEG & CEMAPRE, Universidade de Lisboa

and

João B. Sousa

ISEL & Centro de Matemática e Aplicações (CMA/FCT/UNL)

This work is dedicated to Roman Zmyslony as a token of gratitude for his joy and enthusiasm in the productive research collaborations he actively promoted between Poland and Portugal.

Whatever way we keep looking at bond prices there is no diffusion model that works.
– Pedro Corte Real, Cofounder at MAGENTAKONCEPT, Lda, Portugal.

Abstract

Using the *pulled to par* returns, proposed by [27] for computing historical V@R of bonds, we develop a way of extracting – at any reference date before maturity – *implicit default propensities* from observed bond quotes. This method is new to the literature and it has the advantage on focusing directly on loss given default.

---

[1]Corresponding author.

[2]This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações).

[3]Project CEMAPRE–UID/MULTI/00491/2013 financed by the Portuguese Science Foundation (FCT/MCE) through national funds.

To illustrate the method we present two examples of actual computation with real data – on German and Portuguese bonds. The market data seems to support the proposed method.

In the case of a very concrete simple Gaussian model, we establish the connection between our *implicit default propensity* and the more traditional notions of *default probability* and *recovery given default* of a bond.

Keywords: value-at-risk, bonds, default probability, recovery given default.

2010 Mathematics Subject Classification: 91G40, 91B28, 91B70.

1. Introduction

Whenever assessing risks, it is clear that we may distinguish, mainly, two approaches:

- the *model based approach* in which the behavior of the *profit and loss* random variable – or process, if time variation is considered – is supposed to derive of some explicit mathematical model most frequently given by some stochastic differential equations;
- a *model free approach* in which no mathematical model for the *profit and loss* random process is assumed and the risk evaluation relies only in the statistical descriptive properties of the observed realizations of the *profit and loss* random process.

The existent credit risk literature uses the *model based approach*.

Structural models rely on extensions of the classical Black & Scholes [7] and Merton [22]. In this type of models corporate liabilities are understood as contingent claims on the assets of a firm and one directly models the value process for the assets of a firm and the conditions under which default may (or may not) occur. For a recent overview on this type of models we refer to [8].

Reduced-form, or intensity models, on the other hand, choose to model the default intensity process and assume default occurs at the first jump of some counting process. These models were first proposed by [20] and have been extensively used to price credit risk derivatives. See also [18]. Nonetheless most reduced-form models focus on modeling the default probability, ignoring our roughly modeling the recovery given default, which represents a major drawback when one is interested in accessing the actual loss given default. One exception is the work of [16] where one can also find a discussion concerning the treatment of recovery in reduced-form models. An important reference concerning the importance of modeling recovery is [2].

Incomplete information models proposed by [10] became popular as they allow to establish a bridge between structural and reduced-form models. However, this occurs at a cost – models become quite intractable from a computational point of view. On incomplete information models we refer to [19]. A good survey on existent models is that of [14].

What seems to be consensual from the empirical literature on credit risk models, is that "*there is really no model* [*up to now*] *that works*". For further details and results on the performance of existing models can be found in [3, 15, 9] and [5]. This goes inline with the feedback we get from the industry.

The textbooks of [6, 26, 11] and [21] are recommended readings for those new to the credit risk literature.

In this paper we decide to use the alternative approach. That is, our method aims at being a *model free approach* to deal with credit risk. As far as we know this is the first attempt. The proposed method relies mostly on statistical descriptive properties of the observed realizations of a random process, so we find this idea a particularly good fit to celebrate the carreer of Roman Zmyslony.

The idea of *pulled to par* bond returns were first introduced in the PhD dissertation of the third author and then further developed in [27]. There and here, the motivation is the belief that financial markets are efficient (in some form[4]) and, thus, all relevant information is embeded in market quotes.

[27] uses pulled to par bond returns to propose an alternative method for computing value-at-risk (V@R) of bonds, that rely only on the observed quotes of the particular bond we are concerned with, and no other market information.

In the same spirit, here we assume that bond quotes perfectly reflects all the information concerning the financial asset under analysis – thus also the possibility of default and its severity in case of occurence.

We consider that the whole set of prices – prior to a certain reference date – may be used for pulled to par V@R calculation from which we then derive the implicit default *propensity*.

The remaining of the paper is organized as follows. Section 2 introduced the notation, the definition of *pulled to par* bond returns, some of its properties in connection with the usual notion of return. Section 3 introduces a toy bond price model via the integral of some process approximation of a white noise and studies some of its properties. Section 4 explains the method of extracting implicit default propensities from bond V@R computations and how to financially interpret them. Section 5 proposes a definition for the default probability whenever the law of the bond price process is known, computes this default probability for the model introduced in Section 3 and, uses the default propensity introduced in Section 4 for computing a recovery rate given default. In Section 6 we present the application of the concepts introduced previously to two bonds: a German

[4] The justification of this method relies on the *efficient market hypothesis*, see [12] and [13].

Government bond and a Portuguese bond; the results are coherent with our expectations given the events covered by the dates range chosen. Finally, Section 7 concludes this paper discussing its main results.

2. The *pulled to par* (*p*2*p*) returns

In this section we describe the methodology of the *pulled to par* prices and the corresponding *pulled to par* (*p*2*p*) multiplicative returns. For what follows standard references for concepts and notations are [23, p. 25] and [4].

Consider a *zero coupon bond* with maturity T, at date t and let $(p(t,T))_{0\leq t\leq T}$ be the price process of this bond (see, for instance, [4, p. 302] for a broad technical introduction to the present day mathematical models of this notion). Although this is not necessary for what follows, we may suppose that it is a stochastic process – defined on a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ – from which some realization is observed. Being so, let $p^\natural(t_\star, T)$ denote the market observed price of this *zero coupon bond* at a date $t_\star < t$. Considered as a realization of the stochastic process $(p(t,T))_{0\leq t\leq T}$, we have that for some $\omega_\natural, \in \Omega$, the probability space:

$$\forall t_\star < t, \;\; p(t_\star, T)(\omega_\natural) = p^\natural(t_\star, T).$$

To this observed zero coupon price $p^\natural(t_\star, T)$ there corresponds a – multiplicative – *yield to maturity* $y^\natural(t_\star, T)$ given by:

$$1 = p^\natural(T,T) = \exp\left(y^\natural(t_\star, T)(T - t_\star)\right) p^\natural(t_\star, T),$$

that is

$$y^\natural(t_\star, T) := \frac{1}{T - t_\star} \log\left(\frac{1}{p^\natural(t_\star, T)}\right). \tag{2.1}$$

As this *yield to maturity* is computed with market prices – which we suppose to reflect complete information – it should include in its formation the market perception of possibilities of default. This hypothesis – of completely informed market prices – is determinant for the methods we will explore next. We may define then $p^{\mathrm{p2p}}(t_\star, t, T)$ the *pulled to par* projected historical price from the past date $t_\star$ to future date t, of the do *zero coupon bond*, by:

$$p^{\mathrm{p2p}}(t_\star, t, T) = p^\natural(t_\star, T) \cdot \exp\left(y^\natural(t_\star, T)(t - t_\star)\right) = \left[p^\natural(t_\star, T)\right]^{\frac{T-t}{T-t_\star}}, \tag{2.2}$$

that is, the observed price at the past date $t_\star$ capitalized to the future date t with the implicit yield to maturity $y^\natural(t_\star, T)$ verified at $t_\star$.

The main idea in what follows is that, for any fixed date t, the family of projected prices $(p^{\mathrm{p2p}}(t_\star, t, T))_{t_\star < t}$ represents ficticious realizations of *zero coupon bond* prices at date t (based upon historically observed yields). The corresponding V@R may be computed as a quantile of an appropriate $(P\&L)$ distribution – for instance the returns' distribution. For that purpose, consider Δ some time interval which is the V@R time horizon. The usual **observed return** $R_t^{\natural,\Delta}$ between dates t and $t+\Delta$ is given in such a way that:

$$p^\natural(t+\Delta, T) = \exp\left(R_t^{\natural,\Delta} \cdot \Delta\right) p^\natural(t,T),$$

and so, we have,

$$R_t^{\natural,\Delta} = \frac{1}{\Delta} \log\left(\frac{p^\natural(t+\Delta,T)}{p^\natural(t,T)}\right). \tag{2.3}$$

Then, the *pulled to par* return between dates t and $t+\Delta$ – corresponding to the *p2p* prices – is given in such a way as to satisfy:

$$p^{\mathrm{p2p}}(t_\star+\Delta, t+\Delta, T) = \exp\left(R_{t_\star,t}^{\mathrm{p2p},\Delta} \cdot \Delta\right) p^{\mathrm{p2p}}(t_\star, t, T),$$

given, finally,

$$R_{t_\star,t}^{\mathrm{p2p},\Delta} = \frac{1}{\Delta} \log\left(\frac{p^{\mathrm{p2p}}(t_\star+\Delta, t+\Delta, T)}{p^{\mathrm{p2p}}(t_\star, t, T)}\right). \tag{2.4}$$

Returning to the definition of the *pulled to par* bond prices in formula (2.2) we have that,

$$R_{t_\star,t}^{\mathrm{p2p},\Delta} = \frac{1}{\Delta} \log\left(\frac{p^\natural(t_\star+\Delta,T)\cdot \exp\left(y^\natural(t_\star+\Delta,T)(t-t_\star)\right)}{p^\natural(t_\star,T)\cdot\exp\left(y^\natural(t_\star,T)(t-t_\star)\right)}\right),$$

from which we may derive a relation between the *pulled to par* returns and the usual returns, together with the increments of the yields to maturity,

$$R_{t_\star,t}^{\mathrm{p2p},\Delta} = \frac{1}{\Delta}\left[y^\natural(t_\star+\Delta,T) - y^\natural(t_\star,T)\right](t-t_\star) + R_{t_\star}^{\natural,\Delta}, \tag{2.5}$$

from which we see that unless the yields to maturity satisfy some vey restrictive condition – e.g. like being constant – the two types of returns defined are different. For instance in Figure 2.1 we have an illustration – by means of histograms – of the difference between the usual and the *p2p* returns (see Section 6 for the full description of the data used).

We observe that in formula (2.5), the natural returns appear at date $t_\star$ while the *p2p* returns appear as pulled to the reference date t. It is possible to have a

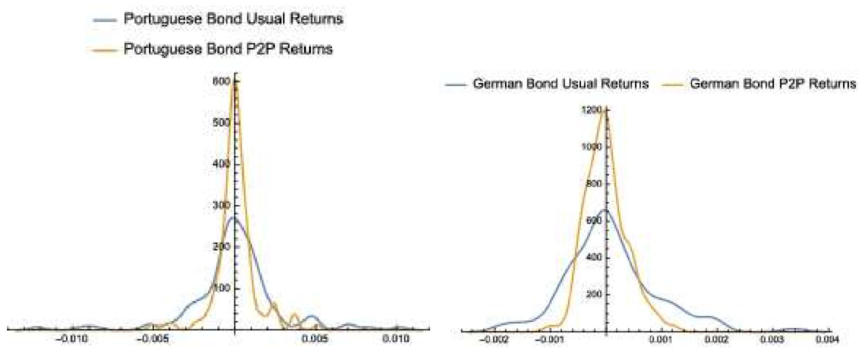

Figure 2.1. Two types of returns from some real data bonds.

relation between the two types of returns, both considered at the reference date t. We may take the expression of the increments of yields to maturity given by

$$\frac{1}{\Delta}\left[y^{\natural}(t_{\star}+\Delta,T)-y^{\natural}(t_{\star},T)\right](t-t_{\star}) \tag{2.6}$$

and consider $t_{\star}$ varying, for instance starting at zero, taking the values $0,\Delta,2\Delta$, $\dots,t-\Delta$ with $t=N\Delta$ or $\Delta=t/N$, that is, $t_{\star}^{k}=k\Delta$, with $k\in\{0,1,\dots,N-1\}$.

$$\begin{aligned}
k=0 \text{ gives } &\left[y^{\natural}(\Delta,T)-y^{\natural}(0,T)\right]N\Delta\\
k=1 \text{ gives } &\left[y^{\natural}(2\Delta,T)-y^{\natural}(\Delta,T)\right](N-1)\Delta\\
\dots\dots&\\
k=m \text{ gives } &\left[y^{\natural}((m+1)\Delta,T)-y^{\natural}(m\Delta,T)\right](N-m)\Delta\\
\dots\dots&\\
k=N-2 \text{ gives } &\left[y^{\natural}((N-1)\Delta,T)-y^{\natural}((N-2)\Delta,T)\right]2\Delta\\
k=N-1 \text{ gives } &\left[y^{\natural}(N\Delta,T)-y^{\natural}((N-1)\Delta,T)\right]\Delta.
\end{aligned}$$

Now, summing from $k=0$ to $k=N-1$ we have that:

$$\sum_{k=0}^{N-1}\frac{1}{\Delta}\left[y^{\natural}(t_{\star}^{k}+\Delta,T)-y^{\natural}(t_{\star}^{k},T)\right](t-t_{\star}^{k})=-y^{\natural}(0,T)N+\sum_{k=1}^{N}y^{\natural}(k\Delta,T),$$

and, as a consequence of formula (2.5), we have that,

$$\frac{1}{N}\sum_{k=0}^{N-1}R_{t_{\star}^{k},t}^{\mathrm{p2p},\Delta}=\frac{1}{N}\sum_{k=0}^{N-1}R_{t_{\star}^{k}}^{\natural,\Delta}+\left(\frac{1}{N}\sum_{k=1}^{N}y^{\natural}(k\Delta,T)\right)-y^{\natural}(0,T), \tag{2.7}$$

which says that a certain average of the *pulled to par* returns is equal to the same average of the usual returns plus the correction term

$$\left(\frac{1}{N}\sum_{k=1}^{N} y^{\natural}(t_{\star}^{k},T)\right) - y^{\natural}(0,T),$$

which is an average of the natural yields to maturity minus the yield to maturity $y^{\natural}(0,T)$. We observe that, following formula (2.3), we have, as $t = N\Delta$,

$$\begin{aligned}
&\sum_{k=0}^{N-1} R_{t_{\star}^{k}}^{\natural,\Delta} = R_{0}^{\natural,\Delta} + R_{\Delta}^{\natural,\Delta} + \cdots + R_{(N-1)\Delta}^{\natural,\Delta} \\
&= \frac{1}{\Delta}\left[\log\left(\frac{p^{\natural}(\Delta,T)}{p^{\natural}(0,T)}\right) + \log\left(\frac{p^{\natural}(2\Delta,T)}{p^{\natural}(\Delta,T)}\right) + \cdots + \log\left(\frac{p^{\natural}(N\Delta,T)}{p^{\natural}((N-1)\Delta,T)}\right)\right] \\
&= \frac{1}{\Delta}\left[\log\left(\frac{p^{\natural}(t,T)}{p^{\natural}(0,T)}\right)\right],
\end{aligned}$$

and so,

$$\frac{1}{N}\sum_{k=0}^{N-1} R_{t_{\star}^{k}}^{\natural,\Delta} = \frac{1}{t}\log\left(\frac{p^{\natural}(t,T)}{p^{\natural}(0,T)}\right), \tag{2.8}$$

and finally we have that formula (2.7) becomes

$$\frac{1}{N}\sum_{k=0}^{N-1} R_{t_{\star}^{k},t}^{\mathrm{p2p},\Delta} = \frac{1}{t}\log\left(\frac{p^{\natural}(t,T)}{p^{\natural}(0,T)}\right) + \left(\frac{1}{N}\sum_{k=1}^{N} y^{\natural}(k\Delta,T)\right) - y^{\natural}(0,T). \tag{2.9}$$

Formula (2.9) expresses the usual return at the reference date t as an average of the $p2p$ returns – pulled to date t – from date zero to date t minus an average of the yields to maturity taken at the successive dates from time zero to t plus the yield to maturity taken at date zero.

Remark 1. In case $(y^{\natural}(k\Delta,T))_{k\geq 1}$ is a realization of a strictly stationary ergodic stochastic process $(y(k\Delta,T))_{k\geq 1}$ we have that, almost surely and in the mean,

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{k=1}^{N} y(k\Delta,T) = \mathbb{E}\left[y(\Delta,T)\right],$$

and this will entail the convergence of the average of the $p2p$ returns, for instance when Δ tends to zero.

Remark 2. As a consequence of formula (2.9), considering, as in the previous remark, stochastic processes – from which the observed prices are realizations – a reasonable condition for the equality, in law, of the usual returns – taken at a reference date t – and the average of the $p2p$ returns – pulled to date t – from date zero to date t may be that:

$$\left(\frac{1}{N}\sum_{k=1}^{N} y(k\Delta, T)\right) - y(0,T) \stackrel{\text{Law}}{=} 0.$$

Intuitively, the above remarks, state the condictions under which (historical) $p2p$ returns are statistically good proxies for the future actual bond returns.

3. A white noise discrete time approximation based model for bond prices

In this section we consider a naive concrete model for the bond price process with the main goal of illustrating the connection between the implicit default propensity, given in definition 1 and the more standard measures of default probability and recovery given default such as the one presented in definition 2. We will also show that under some conditions, with this model, it may be – statistically – hard to distinguish the two return types, the usual returns and the $p2p$–returns. Let us suppose that,

$$p(t,T) = \exp\left(-\int_t^T w_s ds\right), \tag{3.1}$$

for some stochastic process $(w_s)_{s\geq 0}$ having properties that we will describe in the following.

Remark 3. For this, it is also enough to assume that – for almost all trajectories $\omega \in \Omega$ – the function $-\log(p(t,T)(\omega))$ is absolutely continuous in $[0,T]$ (see theorem 7.20 in [25, p. 148]); this assumption entails that $p(t,T)(\omega)$ is almost everywhere differentiable – with respect to the variable t – which implies some restrictions on what models to use with the assumption in formula (3.1). In particular, no diffusion model for the short rate – in the martingale measure – can be used for an affine term structure.

This amounts to defining a *instantaneous rate* similar to the one in the dynamics of the bank account but, instead of being derived from an usual differential equation Cauchy problem, this *instantaneous rate* comes from a backward differential equation, namely,

$$\begin{cases} \frac{d}{dt}p(t,T) = p(t,T)w_t \\ p(T,T) = 1. \end{cases} \tag{3.2}$$

With definition (3.1), having in mind formula (2.5) – for some realization of the stochastic process $(w_s)_{s\geq 0}$ – and with the definition of the yield to maturity in formula (2.1), we have that,

$$y^{\natural}(t_\star + \Delta, T) - y^{\natural}(t_\star, T) = \frac{1}{T-(t_\star+\Delta)}\int_{t_\star+\Delta}^{T} w_s ds - \frac{1}{T-t_\star}\int_{t_\star}^{T} w_s ds,$$

and so, the following condition,

$$\frac{1}{T-(t_\star+\Delta)}\int_{t_\star+\Delta}^{T} w_s ds - \frac{1}{T-t_\star}\int_{t_\star}^{T} w_s ds \overset{\text{a.s.}}{=} 0, \tag{3.3}$$

for $t_\star < t_\star + \Delta < T$, is sufficient for the equality – in law – of the returns. Obviously, this condition is verified for every constant random process; an interesting question is to find non trivial stochastic processes solving this equation.

3.1. A discrete approximation to a white noise process

For $n \geq 1$ let for any $\sigma > 0$ be

$$w_s^n = w_s = \sum_{k=0}^{+\infty} \sigma e_k \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(s), \tag{3.4}$$

with $(e_k)_{k\geq 1}$ a sequence of independent standardized Gaussian random variables. The stepwise process $(w_s)_{s\geq 0}$ is a Gaussian process with constant mean equal to zero and a covariance given by,

$$\begin{aligned}
\mathbb{E}\left[w_s^n w_t^n\right] &= \mathbb{E}\left[\left(\sum_{k=0}^{+\infty} \sigma e_k \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(s)\right)\left(\sum_{l=0}^{+\infty} \sigma e_l \mathbb{1}_{[\frac{l}{2^n}, \frac{l+1}{2^n}[}(t)\right)\right] \\
&= \mathbb{E}\left[\sum_{k=lk,l=0}^{+\infty} \sigma^2 e_k^2 \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(s)\cdot \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(t)\right. \\
&\quad \left. + \sum_{k\neq l;k,l=0}^{+\infty} \sigma^2 e_k \cdot e_l \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(s)\cdot \mathbb{1}_{[\frac{l}{2^n}, \frac{l+1}{2^n}[}(t)\right] \\
&= \sigma^2 \left(\sum_{k=0}^{+\infty} \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(s)\right)\left(\sum_{k=0}^{+\infty} \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n}[}(t)\right) \\
&= \sigma^2 \mathbb{1}_{[0,+\infty[}(s)\cdot \mathbb{1}_{[0,+\infty[}(t) = \sigma^2,
\end{aligned}$$

as $\mathbb{E}\left[e_k^2\right] = 1$ and $\mathbb{E}\left[e_k \cdot e_l\right] = \mathbb{E}\left[e_k\right]\mathbb{E}\left[e_l\right] = 0$ for $k \neq l$. By observing that,

$$\begin{aligned}&\frac{1}{T-(t_\star+\Delta)}\int_{t_\star+\Delta}^{T} w_s ds - \frac{1}{T-t_\star}\int_{t_\star}^{T} w_s ds \\ &= \frac{\Delta}{(T-(t_\star+\Delta))(T-t_\star)}\int_{t_\star+\Delta}^{T} w_s ds - \frac{1}{T-t_\star}\int_{t_\star}^{t_\star+\Delta} w_s ds,\end{aligned}$$

we may detail the joint distribution of the expression by studying each of the integrals separately, as these will give rise to independent Gaussian random variables. Let $k_\star^\Delta \geq 0$ be such that,

$$t_\star + \Delta \in \left[\frac{k_\star^\Delta}{2^n}, \frac{k_\star^\Delta + 1}{2^n}\right[.$$

We have for the distribution of the integral on the left

$$\int_{t_\star+\Delta}^{T} w_s^n ds = \sum_{k=0}^{+\infty} \sigma e_k \int_{t_\star+\Delta}^{T} \mathbb{1}_{\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right[}(s) ds = \sum_{k=k_\star^\Delta}^{k=k_T : T \in \left[\frac{k_T}{2^n}, \frac{k_T+1}{2^n}\right[} \frac{\sigma e_k}{2^n},$$

which is a Gaussian random variable with mean zero and approximate variance,

$$(T-(t_\star+\Delta))\frac{\sigma^2}{2^n},$$

as there are approximately $2^n(T-(t_\star+\Delta))$ Gaussian random variables in the interval $[t_\star+\Delta, T]$. Keeping in mind that $\Delta \geq 1$, as it is a V@R horizon, by choosing $n \geq 1$ large enough we will have, using a similar reasoning that the second integral is a Gaussian random variable with mean zero and variance approximately equal to,

$$\frac{\Delta\sigma^2}{2^n},$$

and so, finally, for $n \geq 1$ large enough and the variance being an approximate value the approximation being of the order of $\sigma^2/2^{2n-1}$,

$$\frac{1}{T-(t_\star+\Delta)}\int_{t_\star+\Delta}^{T} w_s ds - \frac{1}{T-t_\star}\int_{t_\star}^{T} w_s ds \frown \mathcal{N}\left(0, \frac{\Delta^2 + \Delta(T-(t_\star+\Delta))}{(T-(t_\star+\Delta))(T-t_\star)^2}\frac{\sigma^2}{2^n}\right).$$

Formulas (2.5) and (3.3), together with the observation just made, show that that for $n \geq 1$ large enough the usual returns – for the the case of the process $(w_s)_{s\geq 0}$ – are statistically indistinguishable of the *pulled to par* returns despite these two kind of returns being essentially distinct.

3.2. The yield process for the white noise discrete approximation

Let us consider now the yield process – defined by (2.1) – associated to the bond price process defined by formula (3.1) with $\rho > 0$ and $w'_s = \rho + w_s$ instead of w_s. Of course, w_s is as defined above in formula (3.4). Let Δ be some time period – e.g. the usual small period for V@R computations, that is, between one to fifteen days – and the discrete time stochastic process defined by:

$$\forall k \geq 1\ ,\ X_k := y(k\Delta, T) := \frac{1}{T - k\Delta} \int_{k\Delta}^{T} w'_s ds = \rho + \frac{1}{T - k\Delta} \int_{k\Delta}^{T} w_s ds.$$

for $k = 0, 1, \ldots$ and $k \leq T/\Delta$. Let us observe that $(X_k)_{1 \leq k \leq T/\Delta}$ is a Gaussian discrete time stochastic process – being defined by the the usual integral of a Gaussian process with piecewise continuous paths (see, for instance, [24, p. 45]) – with mean given by:

$$\mathbb{E}\left[X_k\right] = \rho + \mathbb{E}\left[\frac{1}{T - k\Delta} \int_{k\Delta}^{T} w_s ds\right] = \rho + \frac{1}{T - k\Delta} \int_{k\Delta}^{T} \mathbb{E}\left[w_s\right] ds = \rho,$$

and constant covariance given by $\sigma^2 - \rho^2$ as,

$$\mathbb{E}\left[(X_k - \rho)\left(X_{k+m} - \rho\right)\right] = \mathbb{E}\left[X_k \cdot X_{k+m}\right] - \rho^2,$$

and,

$$\mathbb{E}\left[X_k \cdot X_{k+m}\right] = \mathbb{E}\left[\frac{1}{(T - k\Delta)(T - (k+m)\Delta)} \int_{k\Delta}^{T} \int_{(k+m)\Delta}^{T} w_s w_r ds dr\right];$$

as we have $\mathbb{E}\left[w_s w_r\right] = \sigma^2$ we finally get $\mathbb{E}\left[(X_k - \rho)\left(X_{k+m} - \rho\right)\right] = \sigma^2 - \rho^2$.

Remark 4. The process $(X_k)_{1 \leq k \leq T/\Delta}$ being Gaussian and second order stationary is strictly stationary (see [1, p. 20]). Moreover, in the case $\sigma^2 = \rho^2$ the process $(X_k)_{1 \leq k \leq T/\Delta}$ is also ergodic and so, in this case we have – by the ergodic theorem – that, almost surely and in the mean (see [28, p. 413]), if $\Delta \ll T$,

$$\frac{1}{N} \sum_{k=1}^{N} X_k = \frac{1}{N} \sum_{k=1}^{N} y(k\Delta, T) \approx \mathbb{E}\left[y(\Delta, T)\right] = \rho.$$

This remark may be useful in the statistical estimation of this kind of model.

4. The implicit V@R default propensity for bonds

As already stated in [27] and references therein, the usual returns of a bond are not suitable for V@R computations by reason of non stationarity of the returns implied by the pulled to par effect in the constraint $p(T,T) = 1$. In previous publications on this subject we have proposed to use *p2p* returns to compute the V@R of the bond. The method may be described as follows.

1. Consider the a reference date t – which may be thought as the present day – and a pseudo-sample of *p2p* returns $R^{\mathrm{p2p},\Delta}_{t_\star,t}$ for dates $t_\star$ and $t_\star + \Delta$, for $t_\star + \Delta \leq t$, computed – always with respect to the reference date – for disjoint periods of time, that is, for instance, starting with some initial date t_0,

   $$[t_0, t_0+\Delta]\ ,\ [t_0+\Delta, t_0+2\Delta]\ ,\ [t_0+2\Delta, t_0+3\Delta], \dots, [t_0+k\Delta, t_0+(k+1)\Delta], \dots$$

2. The α level V@R at date t for the time horizon Δ, denoted by $\mathrm{V@R}^{\mathrm{p2p},\alpha}_{t,\Delta}$, is defined to be the α level quantile of the pseudo-sample of *p2p* returns $R^{\mathrm{p2p},\Delta}_{t_\star,t}$ determined according with the preceding step.

A natural question that arises is: how accurate is the V@R determination proposed above? An obvious answer is that we expect that, for generic dates t and time horizons Δ, the proportion of times that the usual return of the bond $R^{\natural,\Delta}_t$, defined in formula (2.3), verifies

$$R^{\natural,\Delta}_t \leq \mathrm{V@R}^{\mathrm{p2p},\alpha}_{t,\Delta}\ , \tag{4.1}$$

is approximately $1 - \alpha$. In order to implement this verification procedure we, chose some Δ and for successive dates t we observe whether or not formula (4.1) is verified; this corresponds to backtesting the method.

Remark 5. Let us observe that if the observed yields to maturity are constant, say, for all $0 \leq t_\star \leq t$, $y^\natural(t_\star, T) = r > 0$ – which, by formula (2.5), ensures the two types of returns are equal – we will have, by formula (2.1), that,

$$\forall t_\star \in [0,t],\ p^\natural(t_\star, T) = e^{-r(T-t_\star)},$$

and so, as there is no uncertainty, the propensity for default is null. Being so, again by formula (2.5), we postulate that there is information about the propensity for default in the observed difference of the two types of returns which, in turn, derives from the increments of the yields to maturity.

Conjecture. *The main thesis of this note is that we postulate that any appreciable difference between the proportion referred above and* $1 - \alpha$ *is due to an implicit default propensity carried by the bond prices.*

For clarity we formulate the following definition.

Definition 1. The *α-implicit default propensity* of a bond with observed returns $R^{\natural,\Delta}_t$ given by formula 2.3 and the defined above $\mathrm{V@R}^{\mathrm{p2p},\alpha}_{t,\Delta}$, the α level V@R at date t for the time horizon Δ for the pseudo-sample of $p2p$ returns $R^{\mathrm{p2p},\Delta}_{t_\star,t}$ is given by:

$$\frac{1}{\alpha}\left(\alpha - \#\left\{t : R^{\natural,\Delta}_{t} \leq \mathrm{V@R}^{\mathrm{p2p},\alpha}_{t,\Delta}\right\}\right).$$

In Section 6 we apply the method proposed to real data and we get results that correspond to our natural expectations.

Remark 6 (An example of a random model with almost null propensity for default)**.** If w_s is defined by formula (3.4), consider, for a constant $\rho > 0$, $w'_s := \rho + w_s$ for an *instantaneous rate*; then, by formula (3.1), we have – with a reasoning similar to the one used above – that for $n \geq 1$ large enough, the variance being a close approximation of order $1/2^{2n-1}$,

$$p(t,T) = e^{-\rho(T-t)} \cdot e^{-Z_t} \text{ with } Z_t \frown \mathcal{N}\left(0, (T-t)\frac{\sigma^2}{2^n}\right). \tag{4.2}$$

As a consequence, for $n \geq 1$ large enough $p(t,T) \approx e^{-\rho(T-t)}$ – an approximation to be considered under a statistical perspective – and, similarly as observed in remark 5, the propensity for default of such a bond may be considered – statistically – null for $t \leq T$.

In Section 5 we present a first attempt to formalize the definition of the bond default probability and in an example we discuss some of its properties.

5. On a definition for a bond default probability

The usual approaches for the computation of default probabilities and recovery rates can be efficiently grasped in [17, p. 624–630]. Consider the stochastic process $(p(t,T))_{0\leq t\leq T}$ giving the price – at time t – of a zero coupon bond with maturity at time T. The corresponding contract ensures one monetary unit at maturity and so there is a default if, at maturity, the lender receives $0 \leq \gamma < 1$; let us observe that the amount received at maturity, γ, can be considered as a recovery rate.

We propose also the following definition of a default probability, given a recovery rate γ, definition which is tied to the *p2p*–returns.

Definition 2. $\mathcal{P}^{\mathrm{p2p},\Delta}_{\tau}(\gamma)$, the Δ*–p2p–default probability* at time τ with *recovery rate* γ is given by:

$$\mathcal{P}^{\mathrm{p2p},\Delta}_{\tau}(\gamma) := \mathbb{P}\left[p(\tau,T)\exp\left[\left(\frac{1}{N}\sum_{k=0}^{N-1} R^{\mathrm{p2p},\Delta}_{k\Delta,\tau}\right)(T-\tau)\right] \leq \gamma\right].$$

That is, we consider the bond price at time τ *projected* to the maturity by means of the average of the Δ–$p2p$ returns, computed until date τ. The effective computation of this default probability requires the probability law of $(p(t,T))_{0\leq t\leq T}$ which – unless a model is hypothesized – is not available if we only observe one trajectory of the zero coupon bond. As an illustration, consider the example given in Remark 6. By formulas (2.1), (2.9), (3.1) and (3.4) – see the deduction presented in the appendix – we then have that:

$$(5.1)\quad \mathcal{P}_{\tau}^{\mathrm{p2p},\Delta}(\gamma) = \mathbb{P}\left[Z \leq \left(\sqrt{2}\right)^{n} \frac{\log(\gamma) + \left[\rho - \frac{1}{N}\sum_{k=1}^{N} y(k\Delta,T)\right](T-\tau)}{\sigma\sqrt{\frac{(T-\tau)^4}{T^2\tau} + \frac{(T-\tau)^3}{T^2} + 2\frac{(T-\tau)}{T} + (T-\tau)}}\right],$$

with $Z \frown \mathcal{N}(0,1)$. It is then clear that if

$$\rho \leq \frac{1}{T-\tau}\log\left(\frac{1}{\gamma}\right) + \frac{1}{N}\sum_{k=1}^{N} y(k\Delta,T)\ ,$$

then the default probability of the bond will go to zero as n grows indefinitely. In Figure 5.3 we have for the choice of parameters $\sigma = 0.2$ and ρ – corresponding to some usual volatilities and appreciation rates – the implicit default propensities of the bond in the example studied. These implicit default propensities can be quite large – comparable to ones of the Portuguese Government bond presented – even with $n = 8$.

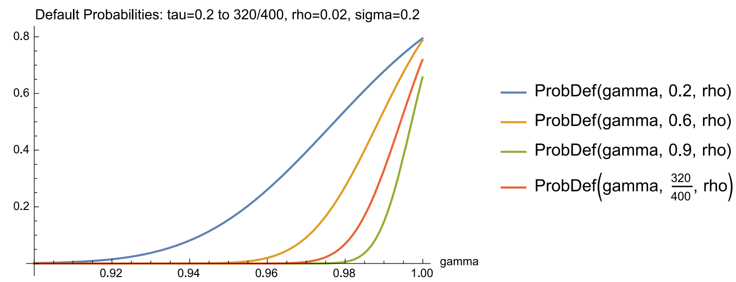

Figure 5.1. Bond default probabilities for an instantaneous rate given by w_t^8.

We observe in Figure 5.1 that there may be a significative default probability with recovery rates larger than 92% and far from maturity – for instance at $\tau = 0.2$ the maturity being at $T = 1$. We will now present the results of a simulation study that will compare numerically the implicit default propensity with the default probability introduced in this section.

We observe that for a sample of 400 observations, if the reference date is at 320, then for a $\mathrm{V@R}_{t,3}^{\mathrm{p2p},.01}$ we obtain a implicit default propensity in the $p2p$

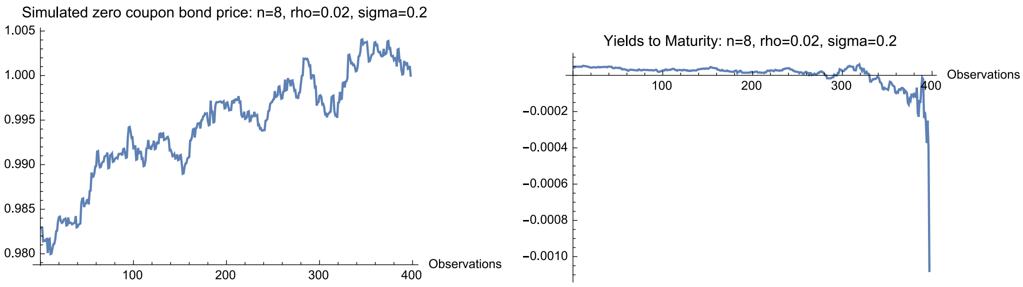

Figure 5.2. Simulated bond processes and correspondent yields to maturity for a w_t^8 trajectory.

returns equal to 0.0823057. To this value there corresponds a recovery rate of $\gamma = 0.980875$ taken at the point 8.23057% correspondent to the curve colored in red, that is, is the curve for the date 320/400. It is so clear that, in the case where there is a model we may, based on the $p2p$–returns, determine the default probability at any period ahead – in the case, three days – and, identifying probability with propensity, determine the correspondent recovery rate at any specified date, previous to maturity.

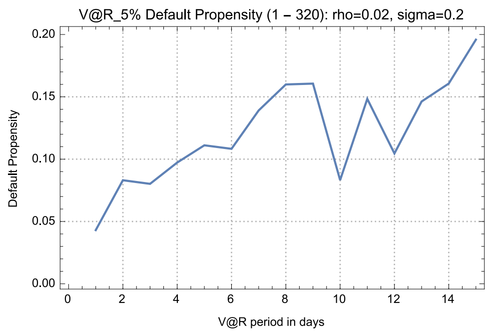

Figure 5.3. Bond $p2p$ – implicit default propensities for by w_t^8.

6. Some real data examples

6.1. A German government bond

We consider first a German government bond, EH375794Corp – Bloomberg denomination – starting May 27, 2008 and maturity date April 7, 2018. In Figure 6.1 we have the observed trajectory of an example of the pseudo-sample of $p2p$

returns and the implicit default propensities for $\mathrm{V@R}^{\Delta}_{0.0001}(t)$ with $t \in [200, 1700]$ and $t \in [600, 1700]$ for different V@R horizons $\Delta = 1, 2, \ldots, 15$.

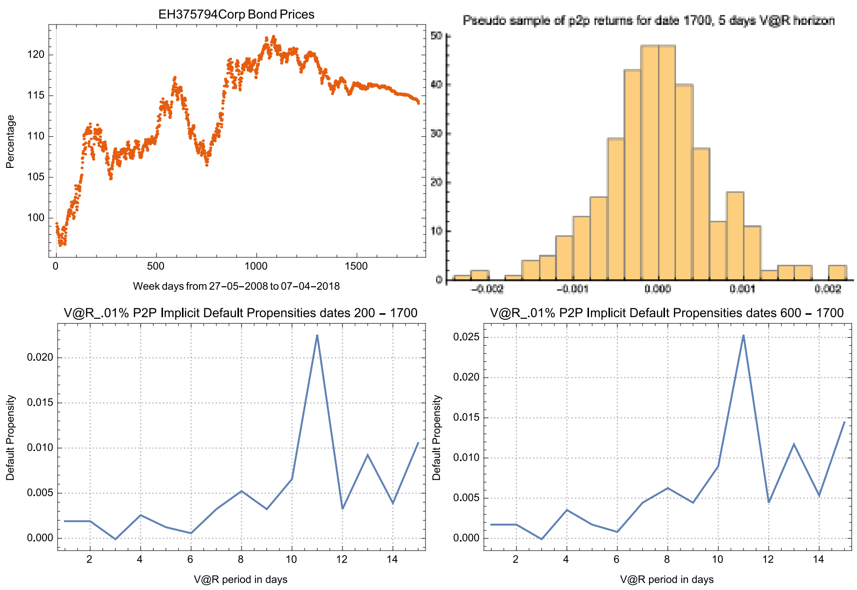

Figure 6.1. Some German government bond implicit default propensities.

We observe the following.

1. We had to choose $\alpha = 0.0001$ in order to obtain *non negative* values for the difference between the V@R of the pseudo sample and the non-violations proportion. This suggests that the α level – of the $\mathrm{V@R}^{\mathrm{p2p},\alpha}_{t,\Delta}$ – may be seen as a sort of *scale* in which we observe the implicit default propensities.
2. The implicit default propensities remain quite low attaining a maximum of approximately 2% on a V@R horizon of 11 days.
3. There is no appreciable difference between the default propensity computed starting in date 200 or in date 600; this shows some *indiference* to risk of this German bond.

6.2. A Portuguese government bond

We now consider a Portuguese bond, EG398877Corp – Bloomberg denomination – starting April 25, 2007 and with maturity date October 16, 2017. In Figure 6.2 we have the observed price trajectory, an example of the pseudo-sample of $p2p$ returns and the implicit default propensities for $\mathrm{V@R}^{\Delta}_{0.05}(t)$; with $t \in [200, 1300]$ and $t \in [600, 1300]$ – for different V@R horizons $\Delta = 1, 2, \ldots, 15$.

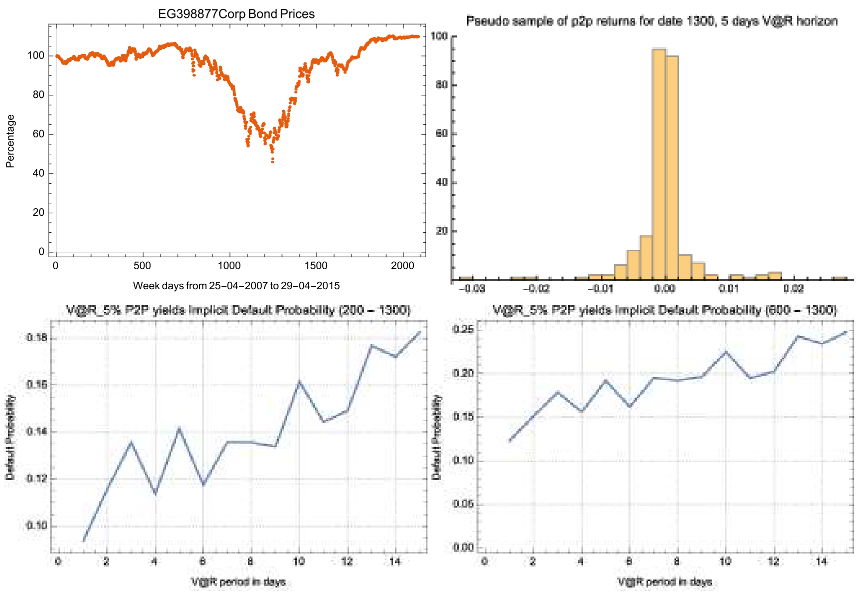

Figure 6.2. Some Portuguese government bond implicit default propensities.

We observe the following.

1. Already with a choice of $\alpha = 0.05$ we obtained very significative implicit default propensities.
2. The implicit default propensities are quite large – compared to the ones of the German bond – using $p2p$-returns from date 200 to date 1300, that is, from the initial dates until the full installment of the subprime crisis.
3. The implicit default propensities are even larger if we consider the $p2p$-returns from dates 600 to 1300, that is, from dates where the bond prices dropped abruptely.
4. As expected, the implicit default propensity clearly increases with the V@R time horizon.

Remark 7. A natural idea would be to try to fit an approximated white noise model – such as the one studied in Sections 3.1, 3.2 and 5 – to the data of the German and Portuguese bonds presented above. Unfortunately – as can be predicted by visually comparing a typical trajectory of the approximated white noise model seen in Figure 5.2 with the bond prices trajectories of the German and Portuguese bonds above – the actual computations we performed show that the results are useless[5].

[5] All Wolfram Mathematica 10 computational files used in this work are available at http://ferrari.dmat.fct.unl.pt/personal/mle/pps/pm-mle2009a.html.

7. Conclusions

In this note, using *pulled to par* prices introduced in [27], we proposed a computation of the implicit default propensity in the bond prices. A preliminary application to real data – a German and a Portuguese bond, with comparable timespans and maturities – shows convincing results varying in line with the reasonable expectations of the market; that is, the German bond has implicit default propensities attaining at most 2 % while the Portuguese bond topped a default propensity of almost 25% during the period of development of the subprime crisis. There is strong evidence that by adjusting two parameters – α, the quantile level and, Δ, the V@R time horizon – we have a reasonable quantification of the implicit default propensities implicit in the bond prices. We noticed that knowing the law of the price process allow us also to propose a definition of the default probability of the bond for a time τ before maturity and for a recovery rate $0 \leq \gamma < 1$. For the example of the approximated white noise, we matched the default propensity to the default probability curve to determine the recovery rate; this can be done for all reference dates before maturity.

Acknowledgments

We express our gratitude to Professor Pedro Corte Real for very useful discussions about the current practices – in some industry practitioners – of V@R computations for bonds. To the anonymous referee our thanks for the excellent set of comments and suggestions that led us to revise the initial text in order to improve rigor and clarity.

References

[1] R. Azencott and D. Dacunha-Castelle, Series of Irregular Observations. Forecasting and Model Building (Springer-Verlag, 1986).

[2] E. Altman, A. Resti and A. Sironi, *Default recovery rates in credit risk modelling: a review of the literature and empirical evidence*, Economic Notes **33** (2004) 183–208.

[3] R. Anderson and S. Sundaresan, *A comparative study of structural models of corporate bond yields: An exploratory investigation*, Journal of Banking & Finance **24** (2000) 255–269.

[4] T. Björk, Arbitrage Theory in Continuous Time, second edition, Springer Finance (Oxford University Press, Oxford, New York, 2004).

[5] G. Bakshi, D. Madan and F.X. Zhang, *Investigating the role of systematic and firm-specific factors in default risk: Lessons from empirically evaluating credit risk models**, The Journal of Business **79** (2006) 1955–1987.

[6] T.R. Bielecki and M. Rutkowski, Credit Risk: Modelling, Valuation and Hedging (Berlin, Springer, 2002).

[7] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, The Journal of Political Economy (1973) 637–654.

[8] C.A.F. Casimiro, Structural Models in Credit Risk (PhD thesis, 2015).

[9] M. Crouhy, D. Galai and R. Mark, *A comparative analysis of current credit risk models*, Journal of Banking & Finance **24** (2000) 59–117.

[10] D. Duffie and D. Lando, *Term structures of credit spreads with incomplete accounting information*, Econometrica **69** (2001) 633–664.

[11] D. Duffie and K. Singleton, Credit Risk: Pricing, Measurement and Management (Princeton University Press, Princeton and Oxford, 2003).

[12] E.F. Fama, Efficient capital markets: A review of theory and empirical work, The Journal of Finance **25** (1970) 383-417.

[13] E.F. Fama and K.R. French, *Size, value, and momentum in international stock returns*, Journal of Financial Economics **105** (2012) 457–472.

[14] K. Giesecke, *Credit risk modeling and valuation: An introduction*, Available at SSRN 479323, (2004).

[15] M.B. Gordy, *A comparative anatomy of credit risk models*, Journal of Banking & Finance **24** (2000) 119–149.

[16] R.M. Gaspar and I. Slinko, *On recovery and intensity's correlation-a new class of credit risk models*, Journal of Credit Risk **4** (2008) 1–33.

[17] J.C. Hull, Options, Futures and Other Derivatives, Options, Futures and Other Derivatives (Prentice Hall International Inc., fourth edition, 2000).

[18] D.S. Jones and J. Mingo, *Industry practices in credit risk modeling and internal capital allocations: Implications for a models-based regulatory capital standard: Summary of presentation*, Economic Policy Review **4** (1998).

[19] R. Jarrow and P. Protter, *Structural versus reduced form models: a new information based perspective*, Journal of Investment Management, **2** (2004) 1–10.

[20] R.A. Jarrow and S.M. Turnbull, *Pricing derivatives on financial securities subject to credit risk*, The Journal of Finance **50** (1995) 53–85.

[21] D. Lando, Credit Risk Modeling: Theory and Applications (Princeton University Press, 2004).

[22] R.C. Merton, *On the pricing of corporate debt: The risk structure of interest rates*, The Journal of Finance **29** (1974) 449–470.

[23] A. McNeil, R. Frey and P. Embrechts, Quantitative Risk Management, Princeton Series in Finance (Princeton University Press, Princeton and Oxford, 2005).

[24] J. Neveu, Processus Aléatoires Gaussiens, Publications du Séminaire de Mathématiques Supérieures, (Les Presses de L' Université de Montréal, 1968).

[25] W. Rudin, Real and Complex Analysis, 3rd ed. (New York, NY, McGraw-Hill, 1987).

[26] P.J. Schönbucher, Credit Derivatives Pricing Models: Models, Pricing and Implementation (John Wiley & Sons, 2003).

[27] J.B. Sousa, M.L. Esquível, R.M. Gaspar and P. Corte Real, *Historical VaR for bonds – a new approach*, in: Proceedings of the 8th Finance Conference of the Portuguese Finance Network, L. Coelho and R. Peixinho, eds., (PFN, Portugal, 2014), 1951–1970.

[28] A.N. Shiryaev, Probability, volume 95 of Graduate Texts in Mathematics (Springer-Verlag, New York, second edition, 1996). Translated from the first (1980), Russian edition by R.P. Boas.

Received 20 April 2017
Accepted 25 July 2017

8. APPENDIX

For completeness sake we present the deduction of the result in formula (5.1). We start with formula (2.9), written for the correspondent stochastic processes, which we represent as,

$$\frac{1}{N}\sum_{k=0}^{N-1} R^{\mathrm{p2p},\Delta}_{t^k_\star,\tau} = \frac{1}{\tau}\int_0^\tau w'_s ds + \widetilde{\rho} - \frac{1}{T}\int_0^T w'_s ds,$$

using the yield definition in formula (2.1) and the price process definition in formula (3.1) and,

$$\frac{1}{\tau}\int_0^\tau w'_s ds = \frac{1}{\tau}\log\left(\frac{p(\tau,T)}{p^\natural(0,T)}\right), \ \ \widetilde{\rho} := \left(\frac{1}{N}\sum_{k=1}^{N} y(k\Delta,T)\right)\ \frac{1}{T}\int_0^T w'_s ds = y(0,T).$$

As a consequence and as $w'_s = \rho + w_s$ we have that,

$$\begin{aligned}X := \log\left\{p(\tau,T)\exp\left[\left(\frac{1}{N}\sum_{k=0}^{N-1} R^{\mathrm{p2p},\Delta}_{k\Delta,\tau}\right)(T-\tau)\right]\right\} &= -\int_\tau^T w'_s ds \\ + (T-\tau)\left[\frac{1}{\tau}\int_0^\tau w'_s ds + \widetilde{\rho} - \frac{1}{T}\int_0^T w'_s ds\right] & \\ = (\widetilde{\rho}-\rho)(T-\tau) + \frac{(T-\tau)^2}{T\tau}\int_0^\tau w_s ds - \left[\frac{T-\tau}{T}+1\right]\int_\tau^T w_s ds. &\end{aligned}$$

Now, as $[0,\tau[\cap[\tau,T] = \emptyset$, we have that the the Gaussian random variables

$$\int_0^\tau w_s ds \frown \mathcal{N}\left(0,\tau\frac{\sigma^2}{2^n}\right) \text{ and } \int_\tau^T w_s ds \frown \mathcal{N}\left(0,(T-\tau)\frac{\sigma^2}{2^n}\right) \tag{8.1}$$

are independent and so we have that,

$$X \frown \mathcal{N}\left((\widetilde{\rho}-\rho)(T-\tau), \left(\frac{(T-\tau)^4}{T^2\tau} + (T-\tau)\left[\frac{T-\tau}{T}+1\right]^2\right)\frac{\sigma^2}{2^n}\right),$$

thus showing, as expected, that

$$(\sqrt{2})^n \frac{X + (\rho-\widetilde{\rho})(T-\tau)}{\sigma\sqrt{\frac{(T-\tau)^4}{T^2\tau} + \frac{(T-\tau)^3}{T^2} + 2\frac{(T-\tau)}{T} + (T-\tau)}} \frown \mathcal{N}(0,1).$$

For simplicity we have supposed that both T and τ are dyadic numbers, that is, of the form $m/2^n$ for some integer m. Otherwise the variances in formula (8.1) are approximations of order $\sigma^2/2^{2n-1}$, which is negligible even for moderate values of n.