Dept. of Math/CMA. Univ. of Oslo
Statistical Research Report No 4
ISSN 0806–3842 October 2009

VALUING TARGET REDEMPTION NOTES BY A STRATIFIED LONGSTAFF SCHWARTZ ALGORITHM

FRED ESPEN BENTH AND PÅL NICOLAI HENRIKSEN

Abstract. A great challenge using the traditional regression based Bermuda option valuation based on Longstaff and Schwartz (LS) (see Longstaff and Schwartz [10]) is the stability of solutions for different basis functions. In this paper we develop an alternative method in the spirit of LS which is less challenging with respect to proper choice of basis functions. The method also makes it possible to quantify the probability of exercise at future nodes in a Bermuda option when moving backward in time. We will apply the method to valuation of target redemption notes with early exercise features under stochastic interest rates based on a LIBOR market model.

1. Introduction

A Target Redemption Note (TARN), in the sense of Brigo and Mercurio [5], is a note where the sum of all coupons must equal an agreed amount within the note's maturity. The size of the coupons are typically driven by an underlying process (an exchange rate, say) and the structure of the coupons may be arbitrary but is typically given by a derivative structure, *e.g.* a call option. The coupons are stochastic and sometimes equal to zero, but in the end the sum of coupons equals the agreed amount. This means that the actual maturity of the note may be much shorter than the nominal maturity. In this paper we will generalize the TARN such that it also includes an early exercise feature. Specifically, this means that if the buyer of the note finds receiving the remaining target amount as beneficial at some time point she may exercise and receive this amount. We will only let the early exercise feature be possible at certain discrete time points, implying that we are in a Bermuda option setting.

We will use a mean reverting Ornstein Uhlenbeck process as a model for the underlying process and a LIBOR market model for the stochastic interest rates. In general, Bermuda options in low dimensional settings may be evaluated by solving moving boundary partial differential equations. Our setting is not low dimensional and we will resort to simulations and regression based methods to value the Bermudan option. There are, however, several ways of valuing Bermuda options numerically, and we refer to Glasserman [8] for an extensive presentation of these. In general, regression based methods underestimate

Date: October, 2009.
Key words and phrases. Target redemption note, LIBOR market model, Bermuda option, Regression, Longstaff Schwartz algorithm.

American or Bermudan option prices because they suboptimally estimate stopping times. Several authors, see e.g. Broadie and Cao [6], Andersen and Broadie [2] or Kolodko and Schoenmakers [9], have worked with upper and lower bound algorithms to analyse the gap between the upper and lower estimates. Other authors, see *e.g.* Bender, Kolodko and Schoenmakers [4], study the impact of applying different stopping rules in a early exercise option setting. In this paper use a stratification idea to propose an improvement on the regression estimates in the LS algorithm for valuing Bermuda options.

The regression based Bermuda option valuation developed by Longstaff and Schwartz [10] is very popular due to its intuitive build-up and easy implementation. However, one major drawback of the method is that one needs to make a choice of basis function for the regression. Making the right choice is far from trivial, and often solutions are very sensitive to the selected basis functions used. A lot of studies have been done in this area, see e.g. Beveridge and Joshi [7], Piterbarg [12] or Bauer, Bergmann and Reuss [3], where different methods have been suggested to determine the regression basis. In this paper we will point out beneficial aspects of the stratified LS method regarding this issue. We will also see in numerical examples that the new method seems to be more robust to different choices of basis functions. The stratified LS method is based on the traditional LS where the continuation value of the Bermuda option is approximated by regression. But instead of approximating the continuation value by one regression, the continuation value is split into different maturities conditioned on the early exercise on these maturities. Probabilities of exercise at different maturities is estimated through logistic regression, and continuation value at each maturity is approximated by its own regression. When regressing on one specific maturity the response data will be much less noisy. Thus, these regressions are easier to control and therefore lead to less sensitivity and higher accuracy. Another feature of our stratification approach is that the stochastic discount rate quit conveniently can be moved outside a conditional expectation operator giving the continuation value.

The main drawback of the new alternative method is that it is harder to implement than the traditional LS as one needs to have under control all the different regressions. The numerical calculations are also slightly more time consuming. But, as we will see in a numerical example, the new method is easier, more robust, more stable and more accurate. The choice of basis functions is much less sensitive to different settings and setups.

Our numerical example is motivated by new life insurance legislation as of 1 January 2008 in Norway. This legislation makes a clear separation between client assets and company assets in an insurance company. In many ways, this is good, but it is less good with regards to surplus allocation. If e.g. the return on client assets is negative, this will hit the company equity. This was the case for many insurance companies in 2008. However, if the return on client assets is extremely good, as might be the case in 2009, it would be natural to be able to rebuild company equity by allocating parts of the return to this purpose. This is not possible. But, the financial product we introduce in our numerical example could be of help to take back some control of risk management in insurance company.

This paper is organized as follows: in Section 2 we will discuss the underlying of the TARN, while in Section 3 we present the LIBOR Market Model used for discounting the coupons of the note. We continue in Section 4 with introducing the TARN contract in a general setting, and present in Section 5 our Stratified LS method. In Section 6 we will see how the TARN is valued based on the LS methods and in Section 7 we will look at numerical examples for different bond life times and different volatilities of the underlying. Finally, in Section 8, we conclude and comment on possible further work.

2. The underlying process

The underlying process of the TARN may be just about anything; a foreign exchange rate, an interest rate, energy spot rates, asset values, etc.. In the example we use for illustration in Section 7, the underlying will be net periodic gains or losses from statement of account of the fiscal year of an insurance company.

Since our methods are based on simulation, any Markovian stochastic process will work. In many situations, *e.g.* for rates or profit and loss in financial statements of account, a mean reverting process is natural. Thus, in this paper we will apply an Ornstein-Uhlenbeck process $X(t)$ for the underlying given by:

$$dX(t) = \lambda(\mu - X(t))\,dt + \sigma\,dW(t). \tag{2.1}$$

Here λ is the mean reversion rate, μ is the mean level and σ is the volatility of the underlying. We suppose that the Brownian motion W is defined on a filtered complete probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. By applying Ito's formula to (2.1), it is easily seen that $X(t)$ is normally distributed with mean and variance given by:

$$\mathrm{E}(X(t)) = X(0) \cdot e^{-\lambda t} + \mu(1 - e^{-\lambda t}),$$
$$\mathrm{Var}(X(t)) = \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda t}).$$

This means that we can write the stochastic process $X(t_i)$ as:

$$X(t_i) = X(t_{i-1}) \cdot e^{-\lambda\Delta t} + \mu(1 - e^{-\lambda\Delta t}) + \sigma\sqrt{\frac{1}{2\lambda}(1 - e^{-2\lambda\Delta t})}\epsilon, \tag{2.2}$$

where $\epsilon \sim \mathrm{N}(0, 1)$ and $\Delta t = t_i - t_{i-1}$. This makes simulation of $X(t_i)$ for any time t_i, $t_i = t_1, t_2, \cdots$, easy.

In this paper, coupons of the TARN will be given as a function of the underlying, $f(X(t_i))$, at the different nodes on coupon dates $t_1, \cdots, t_M$. Typically, $f(\cdot)$ is the payoff function from some derivate, *e.g.* an option. Here t_M is the final coupon date of the TARN, and the remaining amount must be paid at this date if the target has not been met earlier.

We will denote the first time t_i where $\sum_{j=1}^{i-1} f(X(t_j)) < N$ and $\sum_{j=1}^{i} f(X(t_j)) \geq N$ as the knock out time t_{KO}. There are closed form expressions for first hitting times of Ornstein Uhlenbeck processes, see *e.g.* Alili, Patie and Pedersen [1]. However, since the knock out time will interfere with possible early exercise times in the Bermuda option, we will not apply such methods here. We will come back to this in Section 4.

In many situations, there will be dependency between interest rates and the underlying $X(t)$. One option is that this dependency is introduced by correlating ϵ with corresponding variables in the interest rate model. In this paper, we will not explicitly introduce such dependencies.

3. The LIBOR market model (LMM)

In this section we will briefly illustrate the LIBOR market model and motivate the application of this in the TARN setting. For a general introduction to the LMM, or the lognormal forward LIBOR model as it also is known as, see e.g. Brigo and Mercurio [5].

Let the contract value at time t of a zero-coupon bond with maturity at time T be denoted by $P(t,T)$. Further, let the times $T_0, T_1, \cdots, T_M$ represent expiry-maturity pairs of LIBOR rates. We will let the coupon dates of TARN, $t_1, \cdots, t_M$, be a subset of the expiry-maturity dates. The LIBOR rate at time t with expiry T_{i-1} and maturity T_i, $F(t, T_{i-1}, T_i)$, is defined as

$$F(t, T_{i-1}, T_i)P(t, T_i) := \frac{1}{\tau_i}(P(t, T_{i-1}) - P(t, T_i)), \tag{3.1}$$

where $\tau_i = T_i - T_{i-1}$ is the time interval from expiry to maturity. One usually says that the rate is "alive" until time T_{i-1}, at which it equals the LIBOR spot rate $F(t, t, T_i) = L(t, T_i) := -(P(t, T_i) - 1)/((T_i - t)P(t, T_i))$ for $t \in [T_{i-1}, T_i]$. In shorthand notation we will write the LIBOR rate as $F_i(t) = F(t, T_{i-1}, T_i)$.

Notice in Equation (3.1), that $F_i(t)P(t, T_i)$ is the price of the tradable asset $P(t, T_{i-1}) - P(t, T_i)$ with notional $1/\tau_i$. Let Q^i be the probability measure associated with the numeraire $P(t, T_i)$. Thus, from the change of numeraire theory, under the measure Q^i we know that the price divided by the numeraire is a martingale. But the price divided by the numeraire is simply $F_i(t)$. Hence, $F_i(t)$ is a martingale under Q^i. Letting $F_i(t)$ be modelled by a diffusion process, the dynamics of $F_i(t)$ under Q^i must be driftless. Thus,

$$dF_i(t) = \sigma_i(t)F_i(t)\, dZ_i^i(t), \tag{3.2}$$

for $t \leq T_{i-1}$. Here, $\sigma_i(t)$ represents the volatility of the LIBOR rate with expiry maturity pair (T_{i-1}, T_i) at time t, and Z_i^i represents the ith element of the M dimensional Brownian motion vector $Z^i(t)$ under Q^i. The correlation matrix between the M elements in the Brownian motion vector will in the following be denoted ρ.

We are now modelling $F_i(t)$ as a lognormal process under the forward measure Q^i. However, we will need to model $F_k(t)$ under the measure Q^i for $k, k+1, \cdots, M$ and k such that $t \leq T_{k-1}$. To obtain these LIBOR rates, we need the results from the following Proposition (Proposition 6.3.1. in Brigo and Mercurio [5]).

Proposition 3.1. *Under the lognormal assumption, we obtain that the dynamics of $F_k(t)$ under the forward-adjusted measure Q^i in the three cases $i < k$, $i = k$ and $i > k$ are, respectively,*

$$\begin{aligned} dF_k(t) &= \sigma_k(t)F_k(t) \sum_{j=i+1}^{k} \frac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1 + \tau_j F_j(t)}\, dt \\ &\qquad + \sigma_k(t)F_k(t)\, dZ_k^i(t)\,, \qquad i < k, t \leq T_i\,, \end{aligned} \tag{3.3}$$

$$dF_k(t) = \sigma_k(t)F_k(t)\, dZ_k^i(t), i = k\,, t \leq T_{k-1}\,, \tag{3.4}$$

	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\cdots$	$(T_{M-1}, T_M]$
$F_1(t)$	$\sigma_{1,1}$	-	-	$\cdots$	-
$F_2(t)$	$\sigma_{2,1}$	$\sigma_{2,2}$	-	$\cdots$	-
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$F_M(t)$	$\sigma_{M,1}$	$\sigma_{M,2}$	$\sigma_{M,3}$	$\cdots$	$\sigma_{M,M}$

TABLE 1. Piecewise constant volatility for different expiry maturity pairs at different times t.

$$dF_k(t) = -\sigma_k(t)F_k(t)\sum_{j=k+1}^{i}\frac{\rho_{k,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)}\,dt + \sigma_k(t)F_k(t)\,dZ_k^i(t)\,, i > k, t \leq T_{k-1}\,, \tag{3.5}$$

where Z^i is a M dimensional Brownian motion under Q^i. All of the equations above admit a unique strong solution if the volatilities σ are bounded.

Proof. See Brigo and Mercurio [5]. □

Using Proposition 3.1, it is easy to model lognormal forward LIBOR rates under any forward measure Q^i.

One great challenge using LMM, is the calibration of the coefficients ρ and σ to market data. We will not go into detail on this subject here, but refer the reader to Brigo and Mercurio [5] for details. However, we do have to assume certain parametric forms on ρ and σ which will be applied later in the paper. The following parametric form of ρ was first introduced in Schoenmaker and Coffey [13],

$$\rho_{i,j} = \exp\Big(-\frac{|i-j|}{M-1}\big(-\ln\rho_\infty + \eta\frac{i^2+j^2+ij-3Mi-3Mj+3i+3j+2M^2-M-4}{(M-2)(M-3)}\big)\Big). \tag{3.6}$$

Important advantages with this form is that it by construction gives a positive semidefinite correlation matrix and correlations decrease as distance between rates increase.

Further, we will use piecewise constant σ as displayed in Table 1 motivated by Morini and Brigo [11].

In Section 5 we will argue for using zero coupon bond prices as discount factors in option valuation. These zero coupon bond prices will be derived from the LIBOR rates. In this section we have presented a model for LIBOR rates for different expiries and maturities under a given forward measure Q^i. However, each of these rates only apply in the expiry maturity period, *e.g.* (T_{i-1}, T_i). Often we will need to discount over longer periods and also discount to other times than the present. Applying simple algebra to Equation (3.1), we see that discounting at time t from arbitrary T_k to T_i, $T_k > T_i$, may be expressed as:

$$\frac{P(t,T_k)}{P(t,T_i)} = \prod_{j=i+1}^{k} \frac{1}{1+\tau_j F_j(t)}$$

Since in our setting the coupon dates of the TARN is a subset of the expiry maturity dates, we introduce the discount factors

$$D(t,T_k) = \prod_{j=i+1}^{k} \frac{1}{1+\tau_j F_j(t)} \tag{3.7}$$

for the discrete time points $t \in \{T_{-1} = 0, \cdots, T_{k-1}\}$. Notice that at all time points $t = T_i$ the LIBOR rates $F_j(t)$, $j > i$ are $\mathcal{F}_t$ adapted processes. This also means that the discount rates (3.7) are $\mathcal{F}_t$ adapted processes for all expiry maturity times.

Simulation of the LMM rates is easily done by a Milstein scheme as described in Brigo and Mercurio [5].

4. Target Redemption note with early exercise feature

The structured bond we will consider is called a Target Redemption note (TARN). We will let it have a notional of N on a n year life time. The name Target Redemption origins in the fact that the note expires as soon as the sum of coupons equals the notional N. If this has not happened at time n, the remaining amount will be paid in the final coupon. The coupon at time t_i, $t_i \in \{t_0 = 0, \cdots, t_M = n\}$, is arbitrary in the sense that it depends on some underlying process $X(t_i)$ and the stochastic interest rates. In this paper, we will add two features to the TARN; (i) the coupons will be general in the sense that they consist of one TARN part, one constant part and one interest rate related part, (ii) the bond has an early exercise feature in the sense that if it is beneficiary for the buyer to terminate the bond at any coupon time t_i, $i = 1, \cdots, M$, by receiving the remaining TARN amount she may require this. This time will be denoted as the early exercise time, t_{EE}. The last part means that our TARN must be valued as a Bermuda option.

Coupons, C_i, are paid at every coupon date t_i, $i = 1, \cdots, M$. In practise, this usually means every month, quarter or year. In mathematical terms, the coupons will be given by:

$$C^{(i)} = C^{(i)}_{\text{Constant}} + C^{(i)}_{\text{Interest Rate}} + C^{(i)}_{\text{TARN}} \quad i = 1, \cdots, \min(KO, EE, M). \tag{4.1}$$

$C^{(i)}_{\text{Constant}}$ denote the constant coupons and $C^{(i)}_{\text{Interest Rate}}$ are coupons dependent on the interest rate given by the LMM. The $C^{(i)}_{\text{TARN}}$'s are the TARN coupons which are dependent on the underlying $X(t_i)$ and a TARN coupon payoff function $f(\cdot)$. One possibility of TARN coupon payoff function is a call option type of payoff. That is $f(X(t_i)) = \alpha \max(X(t_i) - K, 0)$. This is what we will use in our numerical example in Section 7 and it may be interpreted as the α fraction of surplus exceeding K in the statement of account of the insurance company. This means that $C^{(i)}_{\text{TARN}}$ is given by:

$$C^{(i)}_{\text{TARN}} = \begin{cases} f(X(t_i)) & t_i = t_1, \cdots, \min(t_{KO}, t_{EE}, t_M) \\ N - \sum_{j=1}^{i} f(X(t_j)) & t_i = \min(t_{KO}, t_{EE}, t_M) \end{cases}$$

Multiplying the discount rates (3.7) with the coupons, the value of the TARN with early exercise feature may be expressed as:

$$\pi_0 = \sup_{\tau \in \{1, \cdots, M\}} \mathrm{E}^Q \Big(\sum_{i=1}^{\tau} D(0, t_i) C^{(i)} \Big). \tag{4.2}$$

The Q in this expression represents a risk neutral probability measure. We have seen that the zero coupon bond prices given by the LMM easily may be given a martingale structure through Proposition 3.1. Thus the discount rate may be regarded as risk neutral, *e.g.*, defined under Q. On the other hand, if the underlying $X(t)$ is a tradable asset, the dynamics (2.1) must be changed under Q to incorporate the possibility to hedge in $X(t)$. The change of dynamics of $X(t)$ would be based on a Girsanov transform such that the discounted price is a Q-martingale. In our numerical example in Section 7 the underlying is supposed to be not tradable, and from an arbitrage viewpoint there exists many pricing measures for this process. To simplify the exposition, we will not introduce any particular risk premium in the dynamics for X but suppose that it is already modelled under a risk neutral measure. From now on we write $\mathrm{E}(\cdot) = \mathrm{E}^Q(\cdot)$ without any further reference to the pricing measure Q which would incorporate the measure under which the LMM model is stated and the pricing measure of X.

A TARN with early exercise feature as described above is sometimes called a cancelable TARN. An alternative to this is a callable TARN, where one enters a contract, at time $t = 0$, where one party pays an amount $N - \sum_{j=1}^{i} f(X(t_j))$ at an arbitrary time t_i to another party to obtain future coupons $C^{(j)}$ for $j = i+1, \cdots, \min(KO, EE, M)$. The connection between the cancelable and callable TARN can be seen through a non-cancelable/non-callable bond. Notice that the cash flow of a non-cancelable product, Equation (4.3), a cancelable product, Equation (4.4), and a callable product, Equation (4.5), respectively may be written as:

$$C^{(1)} + C^{(2)} + \cdots + C^{(i)} + \cdots + C^{(M \wedge KO)}, \tag{4.3}$$

$$C^{(1)} + C^{(2)} + \cdots + C^{(i)} + N - \sum_{j=1}^{i} f(X(t_j)), \tag{4.4}$$

$$-(N - \sum_{j=1}^{i} f(X(t_j))) + C^{(i+1)} + \cdots + C^{(M \wedge KO)}, \tag{4.5}$$

if we assume exercise of both the cancelable and the callable product at time t_i. Since the stochastic discount rate applied to the three products are the same, there should also be a connection between the prices of these three products. The cash flows above indicate that this connection should be given by

$$\pi_0^{\text{Callable}} = \pi_0^{\text{Non-Cancelable}} - \pi_0^{\text{Cancelable}}.$$

However, this is only true if the time of canceling and the time of calling is the same. It is not a priori known that the exercise strategies correspond to each other for the cancelable and the callable product.

The connection between the cancelable, callable and the non-cancelable, and the hedgeability in the underlying and the market consistent interest rate, illustrate the importance of understanding the early exercise structure of the TARN from a risk management perspective - both for the issuer and the buyer of the note. We will have a strong focus on the early exercise structure in the rest of this paper.

5. Bermudan option valuation

Before we move on to present the pricing algorithms applied to value Bermuda option, we will briefly discuss the LMM probability measure used in the valuation. The aim of this discussion is that we choose the simplest method with regards to implementation of the pricing algorithms, *i.e.* we want to use the same forward measure throughout the whole calculation.

We will use Q^{t_M}, the notional maturity forward measure as described in Section 3, as pricing measure when evaluating the Bermudan option. The lognormal forward LIBOR rate is by construction a martingale under the Q^{t_M} measure. It is well known that since C_i is a $\mathcal{F}_{t_i}$ measurable function, the following equality holds:

$$\mathrm{E}(D(0,t_i)C_i) = \mathrm{E}\Big(\frac{D(0,t_n)C_i}{P(t_i,t_n)}\Big).$$

Hence, we may write (4.2) as:

$$\begin{aligned}\pi_0 &= \sup_{\tau\in\{t_1,\cdots,t_M\}} \sum_{i=1}^{\tau} \mathrm{E}\Big(\frac{D(0,t_n)C_i}{P(t_i,t_n)}\Big)\\ &= \sup_{\tau\in\{t_1,\cdots,t_M\}} \sum_{i=1}^{\tau} P(0,t_M)\mathrm{E}^{Q^{t_M}}\Big(\frac{C_i}{P(t_i,t_n)}\Big), \end{aligned} \tag{5.1}$$

where we in the last equality change the numeraire to the forward measure Q^{t_M}. This means that all forward rates, and therefore all discount rates through (3.7) and zero coupon prices, are modelled under the same measure Q^{t_M} in our valuation. This simplifies implementation.

We will value the TARN using two different algorithms; the traditional Longstaff and Schwartz (LS) algorithm [10] and the stratified LS. Both algorithms are based on regression to estimate continuation value of the option. Before we describe the algorithms, we will see what we mean by continuation value.

In Bermuda option valuation we need to calculate for each market scenario the stopping times, that is, the optimal times to terminate the note. At each coupon date, t_i, one strikes if the value of immediate exercise is greater than the value of continuing. This rationale simply says that the buyer prefers to receive more than less, in the sense that it is of greater value to get paid now than to receive future coupons and wait with the final payback. To find the first early exercise time, one needs to move backwards from terminal time, $\min(t_{KO}, t_M)$, to the initial time $t_0 = 0$. In mathematical terms, at time t_i, we compare the current payment if exercising

$$C_i | t_{EE} = t_i$$

with the value of continuing,

$$C_i + P(t_i, t_M) \mathrm{E}^{Q^{t_M}} \Big(\sum_{j=i+1}^{\min(EE,KO,M)} \frac{C_j}{P(t_j, t_n)} | \mathcal{F}_{t_i} \Big), \tag{5.2}$$

where t_{EE} in (5.2) is the first stopping time later than t_i, i.e. $t_i < t_{EE}$. Since we only have market information up to time t_i, i.e. $\mathcal{F}_{t_i}$, at time t_i we need to approximate the expectation in (5.2). The two algorithms are different ways of doing this.

In the algorithms, we start by simulating m paths of the underlying X and the LMM interest rates. Each of these paths are simulated up to the minimum of the bond life time t_M and the time of hitting the target t_{KO}. Make a $m \times M$ cash flow matrix. Into the last column, write the cashflow, C_M, for the realisations of the accumulated underlying that do not exceed the target before the bond life time. If all realisations exceeds the target earlier, move to column $M-1$ and repeat the procedure with corresponding coupons. Consider now column $j = M-1, \cdots, 1$. For case j, fill in column j in the cash flow matrix the payments if we exercise the note for all scenarios which are alive up to at least time t_j. For those scenarios with cash flow occurring also later than t_j, we compare the value of exercising and approximate of the value of continuing (5.2). If the former is larger than the latter we exercise, else we continue. In case of exercise, we let all entries in columns $j+1, \cdots, M$ be set to zero, since the note is terminated. Continue this recursion until the first possible exercise time. This leaves us with a cash flow matrix from which we can read off early exercise time t_{EE} for each of the m simulations.

For the traditional LS we approximate the expectation (5.2) as follows. At time t_i, for those scenarios with cash flow occurring also after time t_i we calculate the current and future discounted coupon payment stream up to time $\min(t_{KO}, t_{EE}, t_M)$. This gives us a sequence of numbers, y, for a subsample of scenarios. We regress this y against a basis of orthogonal function of the stochastic processes composing the entiry of market information at time t_i, $\mathcal{F}_{t_i}$. Since both the underlying process $X(t)$ and the LMM forward rates $F_j(t)$ are Markovian, we only need a basis of orthogonal functions of $X(t_i)$ and $F_j(t_i)$, $j = i+1, \cdots, \min(KO, EE, M)$. In our numerical examples in Section 7 we will let this basis be of polynomial form, *e.g.*

$$\begin{gathered}\{1, X(t_i), X(t_i)^2, \Xi(t_i), \Xi(t_i)^2, D(t_i, t_{i+1}), D(t_i, t_{i+1})^2,\\ X(t_i)D(t_i, t_{i+1}), \Xi(t_i)D(t_i, t_{i+1}), X(t_i)\Xi(t_i), X(t_i)\Xi(t_i)D(t_i, t_{i+1})\},\end{gathered}$$

where $\Xi(t_i) = \sum_{j=1}^{i} C_{\mathrm{TARN}}^{(j)}$. Many other basises are possible, and as mentioned earlier a lot of studies have been dedicated to finding the optimal basis. Since a regression needs to be performed for each time step, there is also a possibility to make the basis time dependent. Having regressed y against the basis of function, we may plug the underlying and the LMM rates back into the regression to predict values for the expectation (5.2) for each simulated path. This prediction is then used to determine if we exercise at time t_i or not for each simulated path. Repeat this procedure for all times $t_{M-1}, \cdots, t_1$ to find the structure of early exercise at time $t_0 = 0$. If early exercise only is possible from time t_j, we only repeat the procedure for all times $t_{M-1}, \cdots, t_j$.

From the algorithm above, we see that one single regression needs to take into account a lot of information. Through the basis of orthogonal function we need to be able to describe future early exercise, future hitting of the target, the stochastics in future interest rates, the stochastics in future underlying, the non-linearity of the coupons, the potential remaining target that needs to be paid at exercise and potential covariations between all these factors. Of course, this makes the choice of basis functions very important and highly complex. The stratified LS which we will present now is an alternative to this.

Looking at the expectation (5.2), we see that this may be written as

$$
\begin{aligned}
\mathrm{E}^{Q^{t_M}}\Big(\sum_{j=i+1}^{\min(EE,KO,M)} \frac{C_j}{P(t_j,t_n)}|\mathcal{F}_{t_i}\Big) = \\
\sum_{k=i+1}^{M} \mathrm{E}^{Q^{t_M}}\Big(\sum_{j=i+1}^{k} \frac{C_j}{P(t_j,t_n)}|\mathcal{F}_{t_i} \cap \{\min(t_{EE}, t_{KO}, t_M) = t_k\}\Big) \\
\times \Pr(\min(t_{EE}, t_{KO}, t_M) = t_k)\,. \qquad (5.3)
\end{aligned}
$$

At time t_i, we use our simulations to estimate all the conditional expectations given that $\min(t_{EE}, t_{KO}, t_M) = t_k$ for each $k = i+1, \cdots, M$. I.e. for all k, from the m simulations register those that have $\min(t_{EE}, t_{KO}, t_M) = t_k$. Use the sample paths from these simulations to estimate the conditional expectations given that $\min(t_{EE}, t_{KO}, t_M) = t_k$ by regression. This way one will only have data with one given maturity in the regression, which means that the data will be less noisy. The regression for each k is based on an orthogonal function basis just like in the traditional LS case. But since the data is much less noisy, the process of choosing the basis turns out to be much easier. In our numerical examples in Section 7 we will see that the simple function basis choice

$$\{1, X(t_i), \Xi(t_i), D(t_i, t_{i+1})\}\,,$$

works well.

Notice that using LMM rates for discounting, the discount rate $D(t_i, t_j)$ is known for all future times $t_j = t_{i+1}, \cdots, t_M$. *I.e.*, the discount rates are $\mathcal{F}_{t_i}$-adaptable. However, for the traditional LS we do not know the maturity of the note at time t_i. Hence, we can not exploit the $\mathcal{F}_{t_i}$-adaptiveness. But, in the stratified LS we condition on the maturity, *i.e.*, we condition on $\min(t_{EE}, t_K, t_M) = t_k$. This way we may use the $\mathcal{F}_{t_i}$-adaptiveness and write

$$\mathrm{E}[D(t_i, t_j)C_j|\mathcal{F}_{t_i} \cap \{\min(t_{EE}, t_{KO}, t_M)] = t_k\}]$$

$$= D(t_i, t_j)\mathrm{E}[C_j|\mathcal{F}_{t_i} \cap \{\min(t_{EE}, t_{KO}, t_M) = t_k\}],$$

for all $j \le k$. This further reduces the noise in the data of our regression to estimate the expectation, leading to more accurate and robust results.

To estimate the probabilities $\Pr(\min(t_{EE}, t_{KO}, t_M) = t_k)$ at time t_i we will again use regression. In this paper we use multivariate logistic regression with a logit link. That is, at time t_i register where each simulated path terminates in the set $(t_{i+1}, \cdots, t_M)$. Give the terminal time an index 1 and all other times index 0. As for the conditional expectations, we use an orthogonal function basis as explanatory variables. In our numerical examples in Section 7, we use

$$\{1, X(t_i), \Xi(t_i), D(t_i, t_{i+1})\}$$

which seems to give reasonable results in most cases.

We may ask what happens if we have very few simulations that terminate at time t_k? If very few or no simulations give $\min(t_{EE}, t_{KO}, t_M) = t_k$, this is reflected in the probability $\Pr(\min(t_{EE}, t_{KO}, t_M) = t_k)$. Hence these amounts will not contribute in any significant way to the total.

The probabilities $\Pr(\min(t_{EE}, t_{KO}, t_M) = t_k)$ in itself raise an interesting discussion. In the traditional LS, these probabilities are not explicitly given. But knowing the probability distribution for future early exercise or knock out as a function of different underlying and stochastic interest rates at present time is very interesting both for the issuer and the buyer of the note. This information may be applied both in hedging and risk management. The probabilities also make the stratified LS method attractive in situations where the exercise probabilities are known. This is the case in life insurance, say, where the probability of future payoffs depend on conditional mortality rates at times $t_k \ge t_i$. We will not go further into details about this here.

In the stratified LS we use alot more regressions, both normal and logistic, than the single regression in the traditional LS. The computational time is a bit longer than for the traditional LS, but we will see in our numerical examples in Section 7 that the robustness of the stratified LS more than makes up for the extra computational time.

6. Valuation of the TARN

In the previous sections we have described methods to find the early exercise time t_{EE} structures from Monte Carlo simulations using traditional and stratified LS. In this Section we will look at how we can use this to value the Target Redemption note.

At time $t = 0$, the value of the TARN is given by

$$V^{SB} = \mathrm{E}\Big(\sum_{i=1}^{M} D(0, t_i)\big(C_i \times \mathbf{1}(t_i \le \min(t_{EE}, t_{KO}, t_M))\big)\Big)$$

where $D(0, t)$ is the discount rate.The value can than be expressed as

$$V^{SB} = \sum_{i=1}^{M} P(0, t_M)\mathrm{E}^{Q^{t_M}}\Big(\frac{C_i \times \mathbf{1}(t_i \le \min(t_{EE}, t_{KO}, t_M))}{P(t_i, t_M)}|\mathcal{F}_{t_i}\Big). \tag{6.1}$$

Here $P(0,\cdot)$ comes from the initial values of the LIBOR market model and $P(t_i,t_M)$ is $\mathcal{F}_{t_i}$-adaptable. The expectation $\mathrm{E}^{Q^{t_M}}(\cdot)$ is estimated as the mean of the simulations.

To simplify notation, let $t_K^{(j)} = \min(t_{EE}^{(j)}, t_{KO}^{(j)}, t_M)$ for each simulation $j = 1,\cdots,m$. Hence, we end up with the estimated value of the structured bond at time $t=0$ as given by

$$\hat{V}^{SB} = \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{M}\frac{P(0,t_M)}{P^{(j)}(t_i,t_M)}\cdot C_i^{(j)}\cdot \mathbf{1}(t_i \le t_K^{(j)}). \tag{6.2}$$

7. Numerical example

One possible situation for the structured bond as described earlier is if a company, e.g. an insurance company, wants to issue a bond where the coupons are dependent on the yearly profit and loss of the company, and the buyer of the bond has the option of early exercise, i.e. getting paid the remaining amount of the target. Why would this be interesting? In Norway we have seen new legislation for life insurance companies where the company has no way of disposing any surplus generated from insurance assets or insurance risk. All surplus from insurance assets and insurance risk have to be allocated back to the clients of the company. However, a deficit would have to be covered by the insurance company equity as long as buffer funds are not adequate. I.e. the insurance company has no upside but a huge downside of actively managing the asset portfolio. This is a severe reduction of risk management tools of the company, and thus the structured bond could be one instrument to regain some control of the surplus. The bond is also a nice tool to build up the company equity. In 2008, a lot of pension funds and insurance companies experienced a reduction in equity due to bad results. So issuing the bond could help rebuild equity. The early exercise feature of the buyer is incorporated in the bond to make it more attractive on the market.

In our numerical example we will let the underlying have the following parameters:

$$\lambda = 1, \qquad \mu = 10, \qquad X(0) = 0$$

In addition we will display numerical results for both $\sigma = 10$ and $\sigma = 50$.

The LMM in our examples have parameters for the correlation (3.6) given by $\rho_\infty = 0.24545$ and $\eta = 1.04617$ based on Schoenmaker and Coffey [13]. Further we let the correlation matrix corresponding to Table 1 be given by

$$\sigma = \begin{bmatrix}
0.2685 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.2280 & 0.2340 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.1950 & 0.1950 & 0.2490 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.1785 & 0.1965 & 0.1830 & 0.2505 & 0 & 0 & 0 & 0 & 0 & 0\\
0.1680 & 0.1725 & 0.1800 & 0.1890 & 0.2460 & 0 & 0 & 0 & 0 & 0\\
0.1680 & 0.1725 & 0.1500 & 0.1695 & 0.1890 & 0.2565 & 0 & 0 & 0 & 0\\
0.1695 & 0.1545 & 0.1785 & 0.1470 & 0.1800 & 0.1785 & 0.2445 & 0 & 0 & 0\\
0.1830 & 0.1860 & 0.1620 & 0.1230 & 0.1365 & 0.1815 & 0.1785 & 0.2400 & 0 & 0\\
0.2070 & 0.1395 & 0.1950 & 0.1935 & 0.1095 & 0.0705 & 0.1845 & 0.1695 & 0.2235 & 0\\
0.1815 & 0.1935 & 0.1590 & 0.1470 & 0.1380 & 0.1350 & 0.0345 & 0.2160 & 0.1770 & 0.2205
\end{bmatrix},$$

1. order	$\{1,\, X(t_i),\, \Xi(t_i),\, D(t_i,t_{i+1})\}$
2. order	$\{1,\, X(t_i),\, X(t_i)^2,\, \Xi(t_i),\, \Xi(t_i)^2,\, D(t_i,t_{i+1}),\, D(t_i,t_{i+1})^2,$ $X(t_i)\cdot\Xi(t_i),\, X(t_i)\cdot D(t_i,t_{i+1}),\, \Xi(t_i)\cdot D(t_i,t_{i+1})\}$
3. order	$\{1,\, X(t_i),\, X(t_i)^2,\, X(t_i)^3,\, \Xi(t_i),\, \Xi(t_i)^2,\, \Xi(t_i)^3,$ $D(t_i,t_{i+1}),\, D(t_i,t_{i+1})^2,\, D(t_i,t_{i+1})^3,\, X(t_i)\cdot\Xi(t_i),$ $X(t_i)\cdot D(t_i,t_{i+1}),\, \Xi(t_i)\cdot D(t_i,t_{i+1}),\, X(t_i)\cdot D(t_i,t_{i+1})\cdot\Xi(t_i)\}$

TABLE 2. Different sets of basis functions applied in the linear regression both in the traditional and the stratified LS.

which is based on a calibration in Morini and Brigo [11]. We have increased their estimates heuristically to mimic interest rate volatilities observed during the recent credit crunch.

Further to this, the structured bond is valued using $m = 500,000$ simulations and a notional amount of $N = 100$. We will also vary the life time of the bond and display results for $t_M = 3$ years and $t_M = 10$ years. The coupons of the bond are given by:

$$C^{(i)} = 0.01 \cdot N + 0.6 \cdot F_i(t_i) \cdot N + \max(X(t_i), 0) \quad \text{for all } i \geq 1.$$

We value the structured bond based on three different sets of basis functions both for the traditional and the stratified LS; 1. order polynomials, 2. order polynomials and 3. order polynomials as given in Table 2. In all cases we let the basis of the logistic regression be given by:

$$\{1, X(t_i), \Xi(t_i), D(t_i, t_{i+1})\}.$$

These basises could probably be further refined, but as we will see there is little or no use for this in our numerical examples - especially for the stratified LS.

We find early exercise times from simulations based on the traditional LS and stratified LS algorithms as described in Section 5. From these we estimate the value of the TARN based on Equation (6.2). The bond values for four different scenarios with three different sets of basis functions are displayed in Table 3. As we can see from the table, the biggest differences in prices are in Scenario 2. This is the scenario which is most challenging for the algorithms because it has a long life time of 10 years and high volatility of the underlying. The high volatility is also reflected in the t_{KO} structure, which can be seen in Figure 1. Thus, there is lots of uncertainty that the regressions need to pick up. The prices of the TARN are much larger for the stratified LS algorithm than for the traditional LS algorithms. Thus, we can conclude that the stratified LS algorithm works better in the most challenging case. This conclusion is strengthened when looking at Figure 2, where we see that the early exercise structure seems to converge for the low order set of basis functions for the stratified LS algorithm, but does not seem to have converged at all for the traditional LS algorithm. Our experience has been that further increasing the uncertainty of the underlying as well as the volatility of the LMM increase the difference in bond prices and convergence of early exercise structure. The stratified LS seems to be more robust and more accurate than the traditional LS in highly complex situations.

Also in Scenario 4, where the life time of the note is 3 years and the volatility of the underlying is 50, the stratified LS seems to work better than the traditional LS. Also here, Scenario 1: $t_M = 10$, $\sigma = 10$

Basis fun. order	1. order	2. order	3. order
Stratified LS	108.35	108.32	108.32
Traditional LS	108.38	108.48	108.52

Scenario 2: $t_M = 10$, $\sigma = 50$

Basis fun. order	1. order	2. order	3. order
Stratified LS	105.07	105.58	105.77
Traditional LS	102.48	103.46	104.70

Scenario 3: $t_M = 3$, $\sigma = 10$

Basis fun. order	1. order	2. order	3. order
Stratified LS	99.70	99.70	99.70
Traditional LS	99.76	99.77	99.78

Scenario 4: $t_M = 3$, $\sigma = 50$

Basis fun. order	1. order	2. order	3. order
Stratified LS	100.69	100.72	100.72
Traditional LS	100.53	100.61	100.62

TABLE 3. Values of the TARN derived from the stratified LS and the traditional LS for two different life times of the note and two different volatilities of the underlying. The linear regressions in the algorithms are based on 1. order, 2. order and 3. order basis functions as described in Table 2.

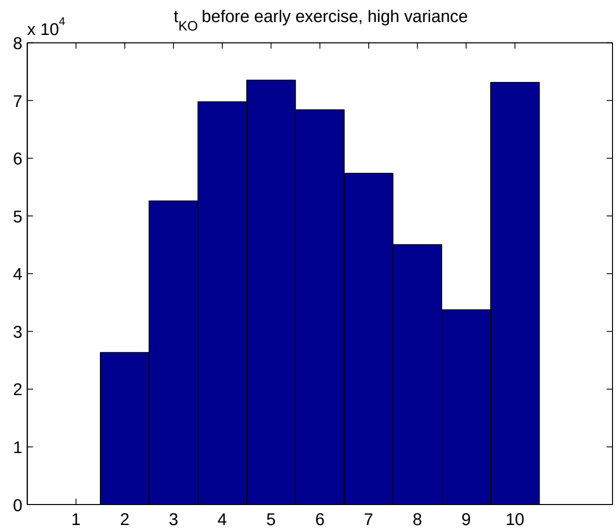

FIGURE 1. The structure of t_{KO} before early exercise has been taken into account.

we sow that the early exercise structure for the stratified LS method seemed to converge

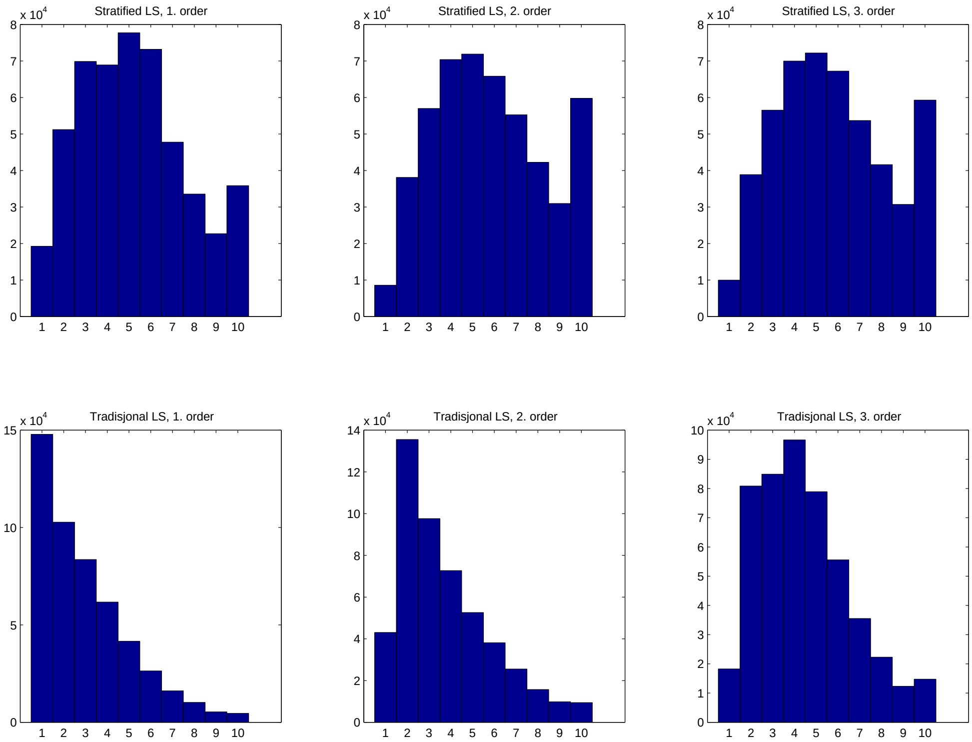

FIGURE 2. The early exercise structure for the stratified LS and the traditional LS in Scenario 2.

for 2. order basis functions, whilst the traditional LS converged slower. Notice that there is less difference in the price in Scenario 4 than in Scenario 2. This is due to the decrease in bond maturity which makes the valuation problem less sensitive.

In Scenario 1 and Scenario 3, the traditional LS works slightly better than the stratified LS. The reason for this is found in Figure 3 and Figure 4. Here we see that the algorithms have hardly changed the early exercise structure at all from the initial knock out, t_{KO}, structure. This also means that several of the nodes have close to zero probability of early exercise. Our experience in general has been that the stratified LS algorithm performs slightly worse than the traditional LS for such early exercise structures. The reason for this is that the linear regression for nodes with very little early exercise probability is based on very few data points. This makes prediction inaccurate. However, there is always a question whether or not these cases are interesting. When almost all early exercise or knock out happens either at the first few nodes or the last couple of nodes, the complexity of the early exercise algorithm should be low to make transparency and analysis easier. The stratified LS algorithm might be a too heavy tool in these situations.

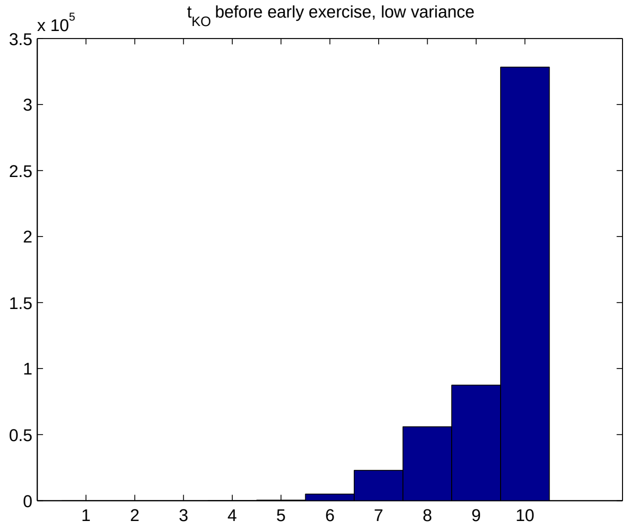

FIGURE 3. The structure of t_{KO} before early exercise has been taken into account.

One further experience has been that including higher order polynomials, *i.e.* higher order than 3, in both algorithms is a non-trivial exercise. The polynomials are easy to implement in our computer program, but higher orders does not give us much more accuracy. We often run into problems with overfitting in our regressions. Also, the scale of the data which is fed into the regression may differ alot for higher orders. This is a robustness problem when changing the parameters of the underlying process or the LMM.

We have also experienced that both the price and the early exercise structure of the note described here is highly sensitive to changes in coupon structures, notional amounts and parameters of the underlying and the LMM. In *e.g.*, Bauer, Bergmann and Reuss [3] one develops an algorithms which automizes the choice of basis functions in the traditional LS regression. Because of the sensitivity in prices and early exercise structures one would think that such algorithms had to be executed for each set of model parameters. We have seen that the stratified LS converges, in early exercise structure and bond prices, for low order of basis functions. This makes the stratified LS algorithm more robust and algorithms for choice of basis functions are less relevant.

8. Conclusions and further work

In this paper we have proposed a refinement of the traditional LS regression method for pricing of derivatives with early exercise features. The new method, called stratified LS, splits the estimated continuation value of the Bermuda option into a sum of conditional expectations and probabilities of exercise at different future time nodes. The main advantages of this split are that the data used in the regressions are less noisy and that the LMM discount rate may be excluded in the regression. The probability of early exercise, as a

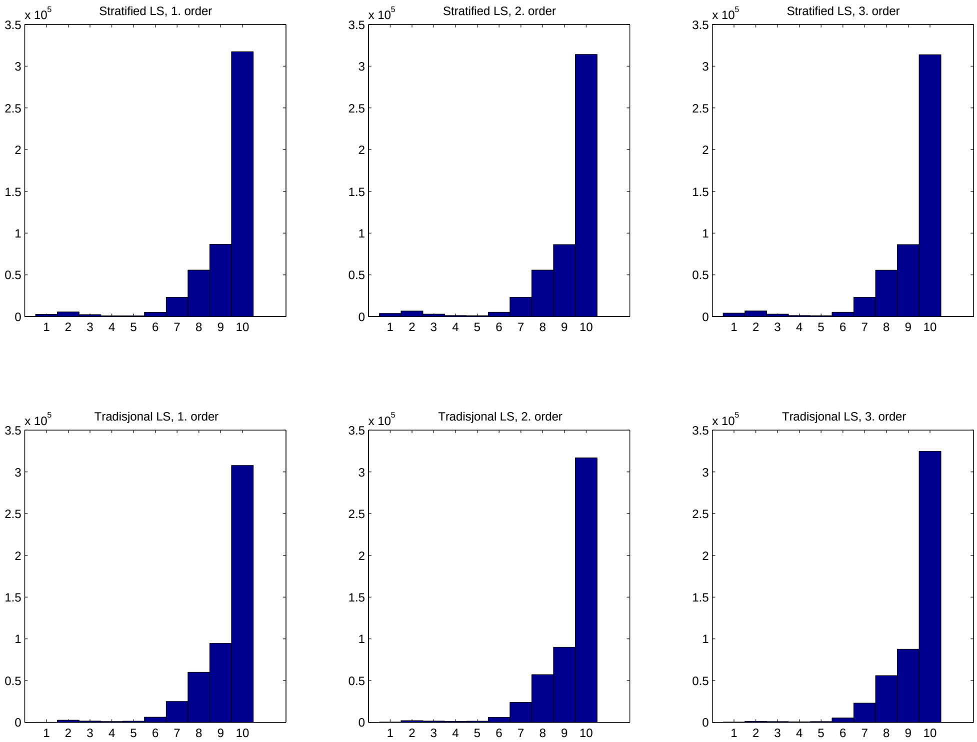

FIGURE 4. The early exercise structure for the stratified LS and the traditional LS in Scenario 1.

function of the stochastic processes generating the entiry of market information, at future time nodes may be useful in itself in e.g. risk management setting.

In numerical examples we have seen that the stratified LS converges faster than the traditional LS with respect to the polynomial order of the basis functions in the regressions. This makes stratified LS alot more robust and choice of basis functions easy. The stratified LS outperforms the traditional LS in scenarios where the underlying processes and stochastic interest rates experience high uncertainty. Also, the outcomes of the algorithms differ more as the life time of the note increases. These are the most challenging cases in derivative pricing. We have also seen that accuracy of the stratified LS is more or less the same as the traditional LS in less challenging scenarios.

We see a high potential for further work to this paper. For instance, by using the forward measure Q^{t_M}, we use all the parameters calibrated in the LIBOR market model at each time step in the LS algorithm. If one accepts model and parameter calibration uncertainty in the LMM model, this introduces more uncertainty in the early exercise times than necessary. In future work we want to address such problems in the context of TARN contracts.

Further, the logistic regression of $\Pr(\cdot)$ in (5.3) has not been analysed properly. In this paper we have used a logit link and this gave reasonable results, but it should be analysed whether or not this link gives the best fit.

Also, in the stratified LS and the traditional LS we use a linear combination of basis functions in the regression estimation of the conditional expectation in Equation (5.3) and Equation (5.2). However, we know from the payoff structure of the coupons , i.e. the call option, that we are in a highly nonlinear situation. Thus, the residuals of the regression are not symmetric. This indicates that another form of basis functions should be used in the regressions. Other forms have not been analysed in this paper since the linear form gave reasonable results, but this would be natural in a future study.

References

[1] Alili, L., Patie, P. and Pedersen J.L. (2005). Representations of First Hitting Time Density of an Ornstein-Uhlenbeck Process. *Stochastic Models* No. 21

[2] Andersen, L. and Broadie, M. (2004). Primal-Dual Simulation Algorithm for Pricing Multidimensional American Options. *Management Science*, Vol.50, No. 9, pp. 1222–1234.

[3] Bauer, D., Bergmann, D. and Reuss, A. (2009). Solvency II and Nested Simulations - a Least-Squares Monte Carlo Approach. *Ulm University, Preprint series*, No. 2009-05.

[4] Bender, C., Kolodko, A. and Schoenmakers, J. (2006). Iterating Cancelable Snowballs and Related Exotics. *Risk* September, pp. 126–130.

[5] Brigo, M. and Mercurio, F. (2001). Interest Rate Models: Theory and Practice. *Springer Finance*.

[6] Broadie, M. and Cao, M. (2008). Improved lower and upper bound algorithms for pricing American options by simulation. *Quantitative Finance*, Vol. 8, No. 8, pp. 845–861.

[7] Beveridge, C. and Joshi, M. (2008). Juggling Snowballs. *Risk Magazine*, No. 21.

[8] Glasserman, P. (2004). Monte Carlo Methods in Financial Engineering, *Springer*.

[9] Kolodko, A. and Schoenmakers, J. (2006). Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics*, No. 10, pp. 27–49.

[10] Longstaff, F. and Schwartz, E. (2001). Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies*, Vol. 14, pp. 113-147.

[11] Morini, M. and Brigo, D. (2003). New developments on the analytical cascade swaption calibration of the LIBOR Market Model. Working paper available at www.damianobrigo.it

[12] Peterbarg, V. V. (2003). A Practitioner's Guide to Pricing and Hedging Callable Libor Exotics in Forward Libor Models. Available at SSRN: http://ssrn.com/abstract=427084 or DOI: 10.2139/ssrn.427084

[13] Schoenmaker, J and Coffey, C. (2000). Stable implied calibration of a multi-factor LIBOR model via semi-parametric correlation structure. *Weierstrass Institute Preprint* No. 611.

(Fred Espen Benth), Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway

E-mail address: `fredb@math.uio.no`

URL: `http://folk.uio.no/fredb/`

(Pål Nicolai Henriksen), Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway

E-mail address: `paalnhe@math.uio.no`