American Option Pricing using Simulation: An Introduction with an Application to the GARCH Option Pricing Model

Lars Stentoft*

HEC Montréal, CIRANO, CIRPEÉ, and CREATE

March 13, 2012

Abstract

This paper has been prepared for the Handbook of Research Methods and Applications in Empirical Finance to be published by Edward Elgar Publishing with Adrian Bell, Chris Brooks, and Marcel Prokopczuk as editors. It contains an introduction to how simulation methods can be used to price American options and a discussion of various existing methods. An application using one of these methods, the regression based method, to the GARCH option pricing model is also provided.

JEL Classification: C22, C53, G13

Keywords: American options, GARCH models, simulation.

*Address correspondence to Lars Stentoft, HEC Montreal, 3000 chemin de la Cote-Sainte-Catherine, Montreal (Quebec) Canada H3T 2A7, or e-mail: lars.stentoft@hec.ca. Phone: (+1) 514 340 6671. Fax: (+1) 514 340 5632.
1 Introduction

To this date, pricing options in general and options with early exercise in particular remains a challenge and an area of interest in empirical finance. There are several reasons for this. For example, in reality many exchange traded options do have the possibility of early exercise and hence this feature needs to be considered when pricing these claims. In particular, this is the case for most if not all options written on individual stocks and if neglected their values could be severely underestimated. However, potentially much more importantly is the fact that many decisions made in real life, not only of the financial kind, can be regarded as a problem of deciding on the optimal stopping time and hence corresponds to the American option pricing problem. A classical example is that of real options, for example the option to develop an area for mining. The decision of when to develop the mine is essentially one of deciding on the optimal stopping time, and the value of the mine should be calculated taking this into consideration.

As a result of the wide applicability of the American option pricing problem it is of immediate interest to develop a flexible pricing framework. For example, theoretically it is well known that the price of an option on a financial asset depends on multiple factors like the strike price and the maturity of the option, the interest rate, and the value of the underlying asset, the volatility of this asset, and the amount of dividends paid on it. While some of these factors, e.g. the strike price and the time to maturity, are fixed or varies only deterministically, the other factors are potentially stochastic and if so should be modelled accordingly. More generally, any real option will likely depend on several stochastic factors. Therefore a valuation framework should be able to accommodate multiple factors and this in a realistic way.

Historically, the American option pricing problem has been solved using numerical procedures such as the binomial model, see e.g. Cox, Ross & Rubinstein (1979), or the finite difference method, see e.g. Brennan & Schwartz (1977). However, though it is possible to incorporate the early exercise feature in these models, it is not computationally feasible to include more than a
couple of stochastic factors. For example, the finite difference method simply cannot be extended to more than two or three stochastic factors and though it is possible to extend the binomial model, see e.g. Boyle (1988), Boyle, Evnine & Gibbs (1989), and Amin (1991), the problem is that the computational complexity increases exponentially with the number of stochastic factors. For example, whereas in one dimension a recombining binomial model with \( k \) steps has \( k + 1 \) final nodes in two dimensions the number of final nodes is \( (k + 1)^2 \). In the literature this problem is known as the “curse of dimensionality”, and in principle it is thus computationally difficult to accommodate all of the potential stochastic factors and instead alternative solutions are called for. One alternative suggestion is to use simulation techniques.

Simulation methods are very flexible and can easily accommodate multiple stochastic factors.\(^1\) Indeed, simulation methods can be used immediately to price European options and the method was in fact introduced to finance at least as early as in Boyle (1977) and even before the binomial model was suggested. Simulation methods do generally not suffer from the curse of dimensionality as the computational complexity only grows linearly with the number of stochastic factors. However, for a long period it remained a challenge to solve the optimal stopping time problem using this technique. This paper contains an introduction to how simulation methods can be used to price American options with an application to the GARCH option pricing model.

Section 2 presents the option pricing framework and explains how the price of an American option can be calculated using dynamic programming. Section 3 introduces how simulation techniques can be used to solve this problem and discusses various existing methods for how to accommodate the early exercise feature of the American option. Section 4 explains how the flexibility of simulation techniques can be used to model more realistically the dynamics of financial assets using the GARCH framework and in Section 5 this framework is used to examine the effect on estimated option prices of the specification of the price of risk. Finally, Section 6 provides some concluding remarks.

\(^1\)For an excellent introduction to the use of Monte Carlo methods, not just for option pricing but for many other problems in finance, see Glasserman (2004).
2 Option pricing framework

The theoretical framework we use follows closely that of Stentoft (2008b). In particular, we first of all assume that the time to expiration can be divided into $K$ periods, $t_0 = 0 < t_1 \leq t_2 \leq \ldots \leq t_K = T$. Thus, we are essentially approximating the American option price with the so-called Bermudan option price.\footnote{The reason for this name is that this type of option is somewhere in between a European and American option, just as Bermuda geographically is located between Europe and North America.} However, as the number of periods increase this approximation becomes better and better. Next, we assume a complete probability space $(\Omega, \mathcal{F}, P)$ equipped with a discrete filtration $(\mathcal{F}(t_k))_{k=0}^K$ and we assume that the underlying model is Markovian with state variables $(X(t_k))_{k=0}^K$ adapted to $(\mathcal{F}(t_k))_{k=0}^K$. We further denote by $(Z(t_k))_{k=0}^K$ an adapted payoff process for the derivative, satisfying $Z(t_k) = h(X(t_k), t_k)$ for a suitable function $h(\cdot, \cdot)$. As an example, consider the American put option for which the only state variable of interest is the stock price, $X(t_k) = S(t_k)$. We have that $Z(t_k) = \max (S - S(t_k), 0)$, where $S$ denotes the strike price. Finally, we assume that $X(0) = x$ is known and hence $Z(0)$ is deterministic and we let $T(t_k)$ denote the set of all stopping times with values in $\{t_k, \ldots, t_K\}$.

Following e.g. Karatzas (1988) and Duffie (1996), in the absence of arbitrage we can specify the object of interest, the American option price, as the solution to the following problem

$$P(0) = \max_{\tau(t_1) \in T(t_1)} E[Z(\tau)]. \quad (1)$$

The problem of determining the American option price in (1) is referred to as a discrete time optimal stopping time problem and one of the preferred ways to solve this problem is to use the dynamic programming principle. Intuitively this procedure can be motivated by considering the choice facing the option holder at time $t_k$: that is to exercise the option immediately or to continue to hold the option until the next period. The optimal choice is to exercise immediately if the value of this is positive and larger than the expected payoff from continuing to hold the option. In this section we explain in detail how this is done.
2.1 Exact algorithms

Let $V(t_k)$ denote the value function of the option at a time $t_k$ prior to expiration. If the option holder keeps the option until the next period and acts optimally from this time onwards the expected payoff will be $E[Z(\tau(t_{k+1})]|X(t_k)]$. On the other hand, if the option is exercised immediately the payoff is $Z(t_k)$. Thus, the value of the option at time $t_k$ may be written as

$$V(t_k) = \max(Z(t_k), E[Z(\tau(t_{k+1})]|X(t_k)]) .$$

Furthermore, since it is always optimal to exercise at expiration, the optimal stopping time $\tau(t_k)$ can be generated iteratively according to the following algorithm:

$$\begin{align*}
\tau(t_K) &= T \\
\tau(t_k) &= t_k 1\{Z(t_k) \geq E[Z(\tau(t_{k+1}))|X(t_k)]\} + \tau(t_{k+1}) 1\{Z(t_k) < E[Z(\tau(t_{k+1}))|X(t_k)]\}, 
\end{align*}$$

1 < k \leq K - 1, \quad (3)

From the algorithm in (3), the value of the option in (1) can be calculated as

$$P(0) = E[Z(\tau(t_1))|X(0)].$$

By definition this price corresponds to the true option price in (1).

Alternatively one can formulate the dynamic programming problem directly in terms of the value function $V(t_k)$. To see this note that if the option holder keeps the option until the next period and acts optimally from this time onwards, the expected payoff will be $E[V(t_{k+1})|X(t_k)]$. Thus, the value of the option at time $t_k$ may equivalently be written as

$$V(t_k) = \max(Z(t_k), E[V(t_{k+1})|X(t_k)]) .$$

Furthermore, since the value at expiration equals the intrinsic value, the value functions can be

3 Though $V(0) = P(0)$, at all other times $V(t_k)$ is a function of the option parameters and the state variables.
generated iteratively according to the following algorithm:

\[
\begin{aligned}
V(t_K) &= Z(t_K) \\
V(t_k) &= \max \{Z(t_k), E[V(t_{k+1}) | X(t_k)]\}, \quad 1 < k \leq K - 1.
\end{aligned}
\]  

(6)

From the algorithm in (6), the value of the option in (1) can be calculated as

\[
P(0) = E[V(t_1) | X(0)].
\]  

(7)

The backward induction theorem of Chow, Robbins & Siegmund (1971) provides the theoretical foundation for the algorithm in (6) and establishes the optimality of the price in (7).

Although it might not be clear that these two algorithms lead to the same optimal stopping time and hence the same price this is in fact the case. To show this, note that it can be optimal to exercise the option at time \(t_k\) if and only if the value of the option equals the immediate exercise value, i.e. if and only if \(V(t_k) = Z(t_k)\). Hence, the optimal stopping time identified at any time \(t_k\), \(\tau(t_k)\), can be written as

\[
\begin{aligned}
\tau(t_k) &= \sum_{n=k}^{K} t_n 1\{Z(t_k) < V(t_k), k \leq h < n \text{ and } V(t_n) = Z(t_n)\} \\
&= t_k 1\{Z(t_k) = V(t_k)\} + \sum_{n=k+1}^{K} t_n 1\{Z(t_h) < V(t_h), k+1 \leq h < n \text{ and } V(t_n) = Z(t_n)\} 1\{Z(t_k) < V(t_k)\} \\
&= t_k 1\{Z(t_k) = V(t_k)\} + \tau(t_{k+1}) 1\{Z(t_k) < V(t_k)\} \\
&= t_k 1\{Z(t_k) \geq E[V(t_{k+1}) | X(t_k)]\} + \tau(t_{k+1}) 1\{Z(t_k) < E[V(t_{k+1}) | X(t_k)]\}.
\end{aligned}
\]  

(8)

where the last equality follows from (5). Finally, by definition it follows that

\[
E[Z(\tau(t_{k+1})) | X(t_k)] = E[V(t_{k+1}) | X(t_k)],
\]  

(9)

and substituting this into (8) one obtains the algorithm in (3).
2.2 Approximate algorithms

We first consider how to implement an approximate method in the case of the stopping time algorithm in (3). We define the function $F(\omega, t_k) \equiv E[Z(\omega, \tau(\omega, t_{k+1})) | X(\omega, t_k)]$, where $\omega$ represents a sample path, and we let $F_M(\omega, t_k)$ denote an approximation to $F(\omega, t_k)$ based on $M$ “parameters”, where parameters is used in a broad sense. We assume that these parameters can be estimated based on $N$ observations and we denote the estimated approximation by $\hat{F}_M^N(\omega, t_k)$. The estimated approximate stopping time can then be derived from the following algorithm:

\[
\begin{align*}
\hat{\tau}_M^N(\omega, t_K) &= T \\
\hat{\tau}_M^N(\omega, t_k) &= t_k 1\{Z(\omega, t_k) \geq \hat{F}_M^N(\omega, t_k)\} + \hat{\tau}_M^N(\omega, t_{k+1}) 1\{Z(\omega, t_k) < \hat{F}_M^N(\omega, t_k)\}, \\
&1 < k \leq K - 1. 
\end{align*}
\]  

(10)

From the algorithm in (10) a natural estimate of the option value in (4) can be calculated as

\[
\hat{P}_M^N(0) = \frac{1}{N} \sum_{n=1}^{N} Z(\omega_n, \hat{\tau}_M^N(\omega_n, 1)).
\]  

(11)

For the value function algorithm in (6) a similar approach can be taken. Now we define the function $H(\omega, t_k) \equiv E[V(\omega, t_{k+1}) | X(\omega, t_k)]$, we let $H_M(\omega, t_k)$ denote an approximation to $H(\omega, t_k)$ based on $M$ parameters, and we assume that these parameters can be estimated based on $N$ observations and we denote the estimated approximation $\hat{H}_M^N(\omega, t_k)$. The estimated approximate value functions can then be derived from the following algorithm:

\[
\begin{align*}
\hat{V}_M^N(\omega, t_K) &= Z(\omega, t_K) \\
\hat{V}_M^N(\omega, t_k) &= \max \left(Z(\omega, t_k), \hat{H}_M^N(\omega, t_k)\right), \\
&1 < k \leq K - 1.
\end{align*}
\]  

(12)

From the algorithm in (12) a natural estimate of the option value in (7) can be calculated as

\[
\hat{P}_M^N(0) = \frac{1}{N} \sum_{n=1}^{N} \hat{V}_M^N(\omega_n, t_1). 
\]  

(13)
In the case of the exact algorithms it was straightforward to show that the price in (4) and (7) coincide, and one might think that this is also the case for the approximate price obtained with (11) and (13). Unfortunately this is not the case and instead it can be shown that the two methods lead to approximations with bias of opposite sign. In particular, the approximate stopping time algorithm leads to low biased estimates since by definition any approximate stopping time is suboptimal. The approximate value function algorithm, on the other hand, leads to high biased estimates because of the “max” operator in (12). Since the biases are of different sign a direct comparison of the two methods is somewhat difficult. However, it is in fact possible to quantify the magnitude of the bias and based on this argue that the absolute size of the bias will be negligible for the approximate stopping time algorithm. The approximate value function algorithm, on the other hand, will generally have a non-negligible positive bias, see Stentoft (2008a).

Finally, it may be argued that there should be less accumulation of errors for the approximate stopping time algorithm than for the approximate value function algorithm. Consider, as an example, the case above with an approximate value function which is biased high compared to the true value function. In this situation the resulting price estimate from the approximate value function method is affected by two sources of errors. First of all, the bias means that exercise is chosen relatively late compared to the optimal time due to the higher value attributed to continuation. Secondly, the future value attributed to a path, for which the choice is to continue to hold the option, is too high because of the biased estimation. For the approximate stopping time algorithm, on the other hand, only the error due to choosing to exercise relatively late occurs. Hence we can expect there to be less accumulation of errors with this method, see Stentoft (2008a).

---

\[ E \left[ \tilde{V}_N (t_{K-1}) | X (t_{K-1}) \right] = E \left[ \max \left( Z (t_{K-1}), \tilde{H}_N (t_{K-1}) \right) | X (t_{K-1}) \right] \]
\[ \geq \max \left( Z (t_{K-1}), E \left[ \tilde{H}_N (t_{K-1}) | X (t_{K-1}) \right] \right) \]
\[ = \max (Z (t_{K-1}), H (t_{K-1})) \]
\[ = V (t_{K-1}) . \]
3 Option pricing using simulation

It should be clear from the previous section that pricing American style options is a complicated task. Indeed, it remains a challenge to develop flexible methods that can be used in e.g. a multidimensional settings. As discussed in the introduction many of the methods that have been used, like the binomial model and the finite difference method, cannot be extended to accommodate more than a couple of stochastic factors. Therefore alternative methods are needed and this is where simulation techniques are interesting.

In this section we explain how simulation can be used to price both European and American style options. We first of all consider the case of the European option which is a “special case” of the framework above where $\tau (t_1) = T$. Though this type of options is not the main focus of this paper, considering this case first allow us to discuss some important properties of the simulation approach. Next we discuss how American style option can be priced using simulation. We review several existing methods and discuss the advantages and potential problems.

3.1 European options

Using the fact that for the European option it is optimal to exercise at time $T$, i.e. $\tau (t_1) = T$ by definition, and substituting this into (1) we obtain the following formula

$$p (0) = E \left[ Z (T) \right],$$

where we use lower case to denote that this is the European price and where $Z (t_k) = h (X (t_k), t_k)$ is the payoff from exercising the option at time $t_k$. From (14) it is clear, that all that is needed to price the option are the values of the state variables, $X (t_k)$, on the day the option expires. Thus, an obvious estimate of the true price in (14) can be calculated using $N$ simulated paths as

$$\hat{p}^N (0) = \frac{1}{N} \sum_{n=1}^{N} h (X (\omega_n, T), T),$$

where
where $X(\omega_n, T)$ is the value of the state variables at the time of expiration $T$ along path number $n$.

As it is (15) does not necessarily rely on the discretized framework since all that is needed is the values of the state variables at maturity. For example, if we consider options in the constant volatility log-normal model of Black & Scholes (1973) and Merton (1973), or BSM, these values can be generated directly. However, there are several situations were this is not the case. First of all, the framework may be discrete by definition as it is the case in the GARCH framework we use in Sections 4 and 5. Next, when simulating from many models with stochastic volatility and jumps discretization methods are needed. Finally, as we discuss below in some situations the payoff of the option is in fact path dependent which also requires a discrete time framework as the one used here. Clearly, as the number of periods $K$ tends to infinity we would expect that our estimates converge to those obtained with at truly continuous time formulation in the two last cases.

### 3.1.1 Extensions

One of the most important advantages of the simulation method is its flexibility, and one might argue that “if we can simulate it, we can price it”. For example, simulation methods can be used to price European options whose payoff depends on a function of the path of the underlying asset and not just the terminal value. A classical example is the so called Asian option or average value option, where the option payoff is determined by the average underlying price over some pre-set period of time. This type of option is of particular interest in thinly traded markets since price manipulation is difficult. One of the first methods proposed to price this type of product was simulation, see Kemna & Vorst (1990). Other types of path dependent options are look-back options and barrier options. Note that some path dependent options depend on the entire path of the underlying asset. In this case, simulation methods only provide an approximation to the true price, the precision of which depends on the precision of the discretization used for simulation.

---

5 In the movie “The Cowboy Way” Woody Harrelson’s character Pepper says: “If it’s got hair, I can ride it. If it’s got a beat, I can dance to it”. It’s a bit the same thing here!
Simulation methods are also well suited to price European style options when the simple BSM model is generalized. For example, one could consider treating the volatility as an additional stochastic factor. This allows for a large degree of flexibility and in the literature many such models have been suggested. Some early examples, which use simulation methods for pricing, are Hull & White (1987) and Scott (1987). Simulation methods have also been used to price options in discrete time models with time varying volatility such as the GARCH option pricing model of Duan (1995). Though the binomial model can be extended to accommodate stochastic volatility and jumps, see e.g. Amin (1993), and GARCH features, see e.g. Ritchken & Trevor (1999), these extensions lack the flexibility available with the simulation method.

Finally, the simulation approach is easily extended to more complex types of European options like options on multiple stocks, see e.g. Margrabe (1978), Stulz (1982), and Johnson (1987) for multivariate extensions of the constant volatility Gaussian model, Da Fonseca, Grasselli & Tebaldi (2007) and Gourieroux & Sufana (2010) for multivariate stochastic volatility models, and Rombouts & Stentoft (2011) for a multivariate discrete time model with time varying volatility and correlations. Classical examples of options on multiple stocks are basket options, for which the payoff is based on the average of a basket of securities, spread options, where the payoff is based on the difference between the prices of two or more assets, and options which pay the best or worst of N assets. Note also that the most traded type of options, index options, are in principle options on an average of multiple assets. Simulation methods can easily accommodate all these types extensions as the computational complexity increases only linearly in the number of underlying assets.

3.1.2 Properties

Simulation methods are not only very flexible they also have very nice asymptotic properties. For example, another advantage of the simulation method is that it is well known that the price estimate in (15) will be unbiased and asymptotically normal under very mild assumptions. In fact, the only restriction is that the payoff function should be square integrable, which is a relatively
mild assumption. The reason is that after all the European price is just an expected value and estimating this by an average of independent simulation values is a (very) good idea so to say.

The fact that a central limit holds for the option price estimate in (15) immediately tells us that the standard error of this estimate tends to zero as $1/\sqrt{N}$. Therefore, in order to obtain an estimate with half the variance one needs to double the number of simulated paths. Though this is very “reasonable” in statistics, compared to many other numerical methods this would be considered slow. For this reason, a lot of research has been conducted to reduce the variance of the estimated price. Among the simplest and most widely used methods is antithetic simulation, see e.g. Boyle (1977), control variables, see e.g. Boyle (1977) and Kemna & Vorst (1990), and moment matching methods, see e.g. Barraquand (1995), who termed this quadratic re-sampling, and Duan & Simonato (1998), who call it empirical martingale simulation. Other more specialized methods are importance sampling and methods that use stratified sampling or samples from low-discrepancy sequences. For an early discussion of how these methods may be applied to security pricing in general and option pricing in particular see the review of Boyle, Broadie & Glasserman (1997), and for a detailed introduction to these method see Glasserman (2004).

Finally, it should be noted that one potential additional challenge in using simulation for pricing is that sometimes future values of the state variables cannot be simulated directly or only simulated at great costs. In particular, whereas simulation is straightforward in the simple constant volatility log-normal BSM model this is not the case with many of the stochastic volatility and jump models which generalizes this framework. Whenever this is the case discretization methods are needed and these will affect the properties of the obtained estimates. For a general introduction to this see Kloeden, Platen & Schurz (1994) and for applications of particular relevance in finance see Glasserman (2004). This issue is also of particular importance if the derivative is path dependent as discussed above. However, note that for other applications, e.g. to the GARCH framework which is discrete in nature and therefore simple to simulate from, this is not a problem.
3.2 American options

Though simulation techniques are simple to use for pricing European options, even those with path
dependence, this is by no means the case with American style options. Essentially the problem
at hand is that pricing American options is backward looking, i.e. we start at maturity and then
solve the problem in (1) by going backwards, and simulation is forward looking, i.e. tomorrow’s
stock price is generated from today’s value. Thus, for a long time it was believed that it would
be difficult if not impossible to price this type of options using simulation. For example, earlier
editions of standard textbooks such as Hull (2000) remarks that simulation methods “cannot easily
handle situations where there are early exercise opportunities”.⁶

In order to illustrate the problem, consider the simplest possible two period framework and
assume that the simulated paths are as illustrated in Figure 1. Further assume that we wish to
price a put option with a strike price of \( \bar{S} \). While the payoff is known at time \( T \), at time \( t_1 \) it has to

⁶In later versions of Hull’s classic some of the methods we review in the following are included.
be decided if early exercise would be optimal along each path. Using next period’s payoff along the path would lead to biased results as this implies perfect foresight on behalf of the option holder. To be specific, the problem is that in order to determine the optimal stopping time one needs to have the conditional expectation of the future payoff.

### 3.2.1 Overview of existing methods

In the literature several methods have been proposed to approximate this conditional expectations using simulation techniques. Examples include the bundling or state space partitioning methods of Tilley (1993) and Barraquand & Martineau (1995), the regression based methods of e.g. Carriere (1996), Tsitsiklis & Van Roy (2001), and Longstaff & Schwartz (2001), and the stochastic mesh method of Broadie & Glasserman (2004). In the following we provide an overview of these methods.\(^7\) We also consider the parameterized stopping time method of Garcia (2003) and Ibanez & Zapatero (2004) although this method does not approximate the continuation value directly.

**Parameterized stopping strategy** A first idea of how to solve the American option pricing problem could be to parameterize the early exercise decision. For example, if this can be determined by a low dimensional set of parameters \(\theta\) then solving the optimal stopping time problem amounts to optimizing over this set. To be specific, consider the put option with one early exercise time in Figure 1. Here \(\theta\) could simply be the threshold price of the underlying asset at which it becomes optimal to exercise the option. This type of methods was suggested by Garcia (2003) and Ibanez & Zapatero (2004). More generally one could estimate a separate threshold price for each early exercise date or alternatively one could approximate the threshold value across the early exercise dates by a low order function.

Though at first the method appears straightforward, in real applications the problem is that the optimization is generally not simple and often the problem fails to be well behaved. For example,\(^7\) The random tree method of Broadie & Glasserman (1997) also uses simulation to price American options. However, this method is closer in spirit to the binomial model and requires additional subsampling at each time step. Hence it is not a true simulation method and it suffers from the curse of dimensionality.
even in the simplest possible case the simulated option price is not in general continuous in the
parameter over which the optimization is conducted. As a result of this it may be impossible for
standard optimization algorithms to find the optimum. This is illustrated in Figure 1 where the
function is flat between the 3 paths that are in the money and jumps discretely at each point.
Though this becomes less of an issue as the number of paths increases, with a large number of
paths the numerical maximization becomes computationally very demanding.

Another potential problem with the parametric method is the curse of dimensionality of the
optimization problem both in terms of the number of exercise steps and in the number of stochastic
factors. For example, extending the simple two period problem to e.g. one with 50 early exercise
points would results in a 50 dimensional optimization problems. This is very difficult to solve and
typically more than 50 times as difficult as solving the one dimensional problem. Though this type
of issues could be circumvented by approximating the 50 points by a low order function this would
introduce an approximation bias unless the functional form is sufficiently flexible. Moreover, though
this type of approximation could decrease the computational burden in the one dimensional case
this may not by the case when the number of stochastic factors increase. To illustrate this consider
again the case with only one possible early exercise time but let the option be multidimensional
option on the average of several underlying assets. In one dimension the threshold is a point, in
two dimensions it is a line, and in three dimensions it is a plane.\footnote{For other option payoffs, say the maximum of several stocks, there are early exercise regions which cannot be
easily approximated by simple functions, see e.g. Ibanez & Zapatero (2004). In this case the method of Garcia (2003),
which parameterizes directly the early exercise region, also requires the use of approximations.}

**Bundling methods and state space partitioning** The real problem in using simulation to
price American options is the fact as we go backward through time we simply cannot use the future
path to determine the optimal stopping time as this would imply perfect foresight on behalf of the
option holder. One way to avoid this is to bundle several paths together and use the average future
values within a bundle as an estimate of the expected future value. This approach was one of the
first suggestions made for the American option pricing problem and dates back to Tilley (1993).
To be specific, let $M$ be the number of equally sized bundles and let $b(\omega_n, t)$ denote which bundle path number $\omega_n$ belongs to at time $t$. Then the algorithm of Tilley (1993) calculates the value along path $\omega_n$ at time $t$ as

$$\hat{V}(\omega_n, t) = \max \left[ Z(\omega_n, t), \frac{1}{N/M} \sum_{j: b(\omega_j, t) = b(\omega_n, t)} \hat{V}(\omega_j, t + 1) \right].$$ \hspace{1cm} (16)$$

That is, the continuation value is a simple average of future payoffs.\footnote{In Tilley (1993) a refined version of the algorithm called the “sharp boundary” algorithm is also considered.} Because of this averaging the issue of perfect foresight is mitigated.

More generally, the underlying idea of this type of algorithm is that when simulated paths are bundled together one can calculate the probability of transitioning to a future bundle conditional on today’s bundle. Moreover, average future values within a bundle can be used as estimates of actual future values without assuming perfect foresight. With these probabilities and the future values the conditional expected future values can be approximated. Let $M$ be the number of, now not necessarily equally sized, bundles and let $b(\omega_n, t)$ denote which bundle path number $\omega_n$ belongs to at time $t$. Furthermore, let $p(i, j, t)$ be the probability of moving from bundle $i$ to bundle $j$ from time $t$ to $t + 1$, calculated by simply counting the number of transitions. Then the bundling algorithm estimates the value along path $\omega_n$ at time $t$ as

$$\hat{V}(\omega_n, t) = \max \left[ Z(\omega_n, t), \sum_{j=1}^{M} p(b(\omega_n, t), j, t) \times \hat{V}(j, t + 1) \right],$$ \hspace{1cm} (17)$$

where $\hat{V}(j, t + 1)$ denotes the average option value of all paths in bundle $j$ at time $t + 1$. Note that in the limit, when the number of bundles equals the number of paths, we obtain the perfect foresight solution.

The main problem with the bundling algorithm is that the partition has to be chosen before the simulation and choosing the partition is particularly tricky in high dimensional problems. For
example, for a given number of partitions along each underlying asset value the total number of partitions grows exponentially in the dimension of the problem. To circumvent this problem, Barraquand & Martineau (1995) suggest that instead of partitioning the state space along the simulated paths one should use the payoff for stratification. In particular, instead of partitioning the (high dimensional) state space of the underlying asset partitions are chosen based on the (one dimensional) payoff. The algorithm is denoted stratified state aggregation along the payoff, or SSAP, and it ensures that the partition is feasible even in high dimensions.\footnote{Barraquand & Martineau (1995) also suggest an alternative formulation of the payoff calculations, or the backward integration, in which only the average values are used. The backward induction is then performed as

\[
\tilde{V}(i,t) = \max \left[ \tilde{Z}(i,t), \sum_{j=1}^{M} p(i,j,t) \times \tilde{V}(j,t+1) \right],
\]

where \(\tilde{V}(i,t)\) and \(\tilde{Z}(i,t)\) denotes the averages across bundle \(i\) at time \(t\).}

**Regression based methods** In the bundling algorithms the conditional expectation term, the second term in (17), is constant for all paths in a given bundle. Hence this method essentially approximates the conditional expectation function by a step function. While this may be a reasonable method for approximating some functions it most likely is not so for functions such as the conditional expectation. For example, in the two period case in a BSM world we know that at time \(t_1\) this expectation is given by the one period European option value. Thus, unless the number of bundles is very large the approximation is poor and so is the estimate of the option price. In Figure 2 we illustrate this using 4 partitions with a dotted line. Thus, in actual applications of the bundling method a large number of bundles, say 100, is required.

The step functions needed in the bundling algorithm can actually be estimated in one go by using simple dummy variables for the bins of the state variables as regressors in a simple linear regression using the cross section of simulated paths. This is in fact how we generated the approximation in Figure 2 in which the regression used 4 partitions, that is a constant term and 3 “steps”. However, as an obvious alternative to the step functions one could then use other more refined regression functions. For example, one could use (linear) spline functions as suggested by Carriere (1996) or
a given polynomial family as suggested by Longstaff & Schwartz (2001), and these methods are now collectively referred to as regression based methods. They have been analyzed in quite some detail, see e.g. Moreno & Navas (2003), Stentoft (2004a), and Areal, Rodrigues & Armada (2008).

To be specific, let \( \{\phi_m(\cdot)\}_{m=1}^M \) denote a “family” of functions, also called a basis or approximation architecture, to be used for the approximation. Then a regression based method would estimate the value along path \( \omega_n \) at time \( t \) by

\[
\hat{V}(\omega_n, t) = \max \left[ Z(\omega_n, t), \sum_{m=1}^{M} \phi_m(X(\omega, t_k)) \hat{b}_m(t_k) \right],
\]  

(18)

where \( \hat{b}_m(t_k) \) is found from a regression of future values on the regression functions \( \{\phi_m(\cdot)\}_{m=1}^M \).
Though the regression based methods do not involve choosing a partition as in the bundling algorithm they do involve a choice of regression functions. However, in the existing literature there is a significant amount of evidence that the actual choice of regressors is in fact of minor importance in actual applications, see e.g. Moreno & Navas (2003) and Stentoft (2004).

The regression based methods differ first of all in terms of how the approximation is carried out. In Figure 2 we also plot the approximation obtained using 3 regressors plus a constant term with linear splines, the dashed line, and polynomials, the thin solid line, respectively. From this figure it is seen that both approximate the true value quite well in the center whereas the approximation in the tails is much worse. When comparing to the step function approximation though, it is seen that the spline and polynomial approximation methods have clear advantages. Secondly, the methods differ in terms of the interval over which the approximation is done. For example, Longstaff & Schwartz (2001) suggest that only the in the money, or ITM, paths should be used. In Figure 3 we compare the approximation when using only the ITM paths and when using all the paths. This figure clearly shows how much better the fit is for a given number of parameters, \( M \), when using only the ITM paths and lends support to the suggestion of Longstaff & Schwartz (2001).

**Stochastic mesh methods** In the bundling and state space partitioning method the transition densities are approximated for the bundles. However, if these where known we could simply estimate the value along path \( \omega_n \) at time \( t \) by

\[
\hat{V} (\omega_n, t) = \max \left[ Z (\omega_n, t), \sum_{j=1}^{N} p (\omega_n, j, t) \times \hat{V} (\omega_j, t + 1) \right].
\]  

(19)

Unlike the bundling algorithm in (17) here all the \( N \) paths are used. This method is known as the stochastic mesh method, see e.g. Broadie & Glasserman (2004). The reason for the name is that at any time step all current knots are “connected” to all future knots, which creates a mesh like structure.

\[11\]However, note that with this method only the stopping time method can be used.
Figure 3: This figure shows the approximations obtained with a polynomial of order \( M = 4 \). The left hand plot shows the results when ITM paths only are used and the right hand plot when all paths are used.

A major drawback of the stochastic mesh method is that often computationally complex operations are required to obtain the transition densities. Moreover, in many cases the transition densities are not known analytically and will have to be approximated. This further adds to the computational complexity of the method. Broadie, Glasserman & Ha (2000) suggest two methods which can be used to determine the transition densities when these are either unknown or fail to exist. Both methods rely on constraining the weights to correctly replicate certain basic theoretical quantities, such as moments of the state variables. Since there will generally be many more weights than constraints this problem is underdetermined, requires imposing an optimization criterion, and it is non-trivial and complicated to solve.

Nevertheless, the stochastic mesh method is not only closely connected to the stratified state
partitioning method it is in fact also related to the regression based methods. To see this note that we can rewrite the predicted value given by \( \sum_{m=1}^{M} \phi_m(X(\omega, t_k)) \hat{b}_m(t_k) \) in (18) as

\[
\frac{1}{N} \sum_{j=1}^{N} \Phi(X(\omega, t_k)) (\Phi(X(\omega, t_k))^T \Phi(X(\omega, t_k)))^{-1} \Phi(X(\omega, t_k))^T \hat{V}(j, t + 1),
\]

where \( \Phi(X(\omega, t_k)) \) is the \( N \times M + 1 \) matrix of regressors used and \( \Phi(X(\omega_n, t_k)) \) is the \( M + 1 \) row vector of the regressors evaluated for path \( \omega_n \). Thus, the regression based methods correspond to a particular choice of weights in the stochastic mesh method.

### 3.2.2 Discussion

Of the methods discussed above, the regression based methods in general and the least squares Monte Carlo, or LSM, method of Longstaff & Schwartz (2001) in particular are some of the most popular methods for pricing American options using simulation. For example, a web search performed early 2012 using http://scholar.google.com resulted in 41 citations of Garcia (2003) and 116 citations of Ibanez & Zapatero (2004) for the parameterized stopping strategy method, 305 citations of Tilley (1993) and 306 citations of Barraquand & Martineau (1995) for the state space partitioning method, 239 citations of Carriere (1996), 272 citations of Tsitsiklis & Van Roy (2001), and a massive 1519 citations of Longstaff & Schwartz (2001) for the regression based methods, and 126 citations of Broadie & Glasserman (2004) for the stochastic mesh method.\(^\text{12}\)

However, the regression based methods are not only the most cited methods for pricing American options using simulation they are also very likely the most flexible methods. After all, the regression based methods use nothing but the simulated paths together with linear regression to determining the option price. The flexibility is illustrated by the many applications of the method to price, among other things, life insurance contracts (e.g. in Bacinello, Biffis & Millossovich (2010)), real estate derivatives (e.g. in Longstaff (2005)), real options (e.g. in Gamba (2002)), which has several

\(^{12}\) Even when correcting for the year of publication the LSM method is with 138 citations per year cited almost 10 times as often as the average of the other methods.
applications such as gas storage, mine expansion decisions, and timber harvest contracts, and executive stock options (e.g. in León & Vaello-Sebastià (2009)). On top of being very flexible it can be argued that the regression based methods have advantages both in terms of computational efficiency and in terms of the asymptotic properties.

**Computational efficiency**  One of the main advantages of the regression based methods is that they are computational efficiency. In fact, as shown in Stentoft (2004a) in multiple dimensions the LSM method is actually more efficient than the binomial model which is notoriously difficult to outperform. When comparing the methods it should be clear that computationally the challenge is to determine the parameters, i.e. the transition probabilities in the bundling algorithm, the coefficients in the regression based algorithm, and the transition densities in the stochastic mesh method. The reason that the regression based methods are computationally efficient is that they rely on nothing but simple linear regression to calculate the coefficients needed for the approximation of the conditional expectation function. We now examine the computational time of the different methods using the example of an American put option in a BSM model. The option matures in one year and has a strike price of 40, the interest rate is 6%, and the underlying asset’s price is 40, it pays no dividend, and it has a volatility of 40%.

In Figure 4 we report the computational time for the regression based method. Since this method relies on linear regression to determine the $M$ parameters it is expected to be of order $K \times N \times M^2$.\(^{13}\) The left hand plot shows computational time against the number of time steps, $K$, for different number of paths and confirms that computational time is linear in the number of time steps. The right hand plot shows computational time against the number of regressors, $M$, and shows that this indeed grows faster than linearly. However, this is to be expected as the regression is of order $N \times M^2$. Finally, when comparing computational time for a given number of time periods or regressors while increasing the number of paths, $N$, Figure 4 shows that there is

\(^{13}\)In fact, it may be slightly faster depending on the regressors and the regression method chosen.
Figure 4: This figure shows the computational time for the regression algorithm for different values of $N$, the number of simulated paths. The left hand plot shows the computational time against the number of time steps, $K$, and the right hand plot against the number of regressors, $M$.

a linear relationship. Put differently, computing a price with 50,000 paths takes roughly twice as much time as with 25,000 paths and half the time as with 100,000 paths.

In Figure 5 we report the corresponding results for the bundling algorithm. The left hand plot shows that for this method computational time is linear in the number of time steps, $K$, as it was the case with the regression based method. The right hand plot shows that computational time is also linear in the number of partitions, $M$. This is in contrast to the regression based method for which computational time grew faster. Finally, note that Figure 5 confirms that computational time increases linearly in the number of paths, $N$, with the bundling method as it was the case with the regression based method. Essentially these results show that the bundling algorithm is of order $K \times N \times M$. 

23
Figure 5: This figure shows the computational time for the bundling algorithm for different values of $N$, the number of simulated paths. The left hand plot shows the computational time against the number of time steps, $K$, and the right hand plot against the number of partitions, $M$.

Finally, consider the stochastic mesh method. Compared to the state space partitioning method, each valuation of (19) in the stochastic mesh is of order $N \times N$, and therefore the method is at least of order $K \times N^2$ even if the transition densities are known. The computational complexity is further increased when the transition densities are either unknown or fail to exist and have to be estimated using e.g. the method of Broadie et al. (2000). In particular, using the simplest method the weights are determined using linear regression and therefore is of order $N \times B^2$, where $B$ is the number of constraints imposed to obtain the weights. However, this problem has to be solved for all $N$ paths and at each time step. Therefore the overall complexity of the method is in fact of order $K \times N^2 \times B^2$. Thus, even if the number of constraints is low and the method is implemented with few paths the computational cost is large with this method due to the factor $N^2$. 

24
Based on the results above it may appear that the bundling algorithm is in fact the most efficient since the computational time only grows linearly in the number of partitions. However, this is true only if the same number of regressors and partitions are needed. In real applications, this is rarely the case since it is possible to obtain a much better approximation using e.g. a polynomial function than a step function as Figure 2 shows. Moreover, when comparing the actual performance of the methods the price estimates and not only computational time needs to be compared. In Figure 6 we plot the root mean square error, RMSE, of a particular estimate, $\hat{P}_M^N(0)$, which we define as

$$RMSE\left(\hat{P}_M^N(0)\right) = \sqrt{E\left[\left(\hat{P}_M^N(0) - P(0)\right)^2\right]},$$

against computational time. We estimate the expected value by the sample mean from 1,000 independent simulations, and we compare the estimate from the parametric method, the bundling algorithm, the SSAP method, and the regression based method.

The left hand plot of Figure 6 shows the results for an option with two time periods. From this plot it is seen that the best method is the regression based method as this is closest to the origin and offers the best trade-off between computational time and precision. The next best method is the bundling method, though the parametric method has similar performance when computational speed is important. The SSAP method is always dominated by the other methods, even the parametric method, in this example. The results for the option with five time periods in the right hand plot of Figure 6 are very similar and the ranking of the methods is the same. However, when comparing the two plots it is seen that the relative performance of the regression based method compared to e.g. the bundling method improves with the number of time steps. For example, estimates with a $RMSE \approx 0.02$ can be obtained roughly 22% faster with the regression based methods in the first case and 57% faster in the second case.\(^{14}\)

\(^{14}\)The computational times for the regression method are 11.1 and 43 and for the bundling method 14.2 and 100.2, respectively.
Figure 6: This figure shows the RMSE of the different estimators against computational time. The left hand plot shows results for an option with 2 time steps and the right hand plot for an option with 5 time steps. Results are reported for the parametric stopping time, the bundling method, the SSAP method, and the regression based method. Best models are close to the origin.

Asymptotic properties  Another advantage of the regression based methods is that they have nice asymptotic properties. In particular, it can be shown that the estimated price converges to the true price as the number of paths and regressors tend to infinity under some regularity assumptions. Stentoft (2004b) proves this for the LSM algorithm in a general multi-period and multidimensional setting and shows that to achieve convergence the number of regressors, $M$, should not be increased too fast and has to satisfy $M < C \times N^{1/3}$. Glasserman & Yu (2004) studies the case where all the paths are used in the regression and prove convergence in the normal and log-normal case provided that $M < C \times \ln(N)$ and $M < C \times \sqrt{\ln(N)}$, respectively. Thus, this shows that the speed with which $M$ can be increased is much slower when all the paths are used than when only the in the
money paths are used. See Gerhold (2011) and references therein for generalizations of these results to other processes. Though it is difficult if not impossible to derive the actual convergence rates for the price estimate, the numerical evidence provided in Stentoft (2012) shows that in many cases this happens as quickly as with the European price estimate.

For the bundling method, on the other hand, significantly less research has been conducted in terms of convergence results. In particular, though it is argued in Tilley (1993) that for a fixed bundling parameter $\alpha$, where $0 < \alpha < 1$, if the number of bundles, $M$, is picked as $M = N^\alpha$ the algorithm converges as $N$ tends to infinity, no formal proof is provided. This type of result can likely be generalized to the multidimensional situation though here the standard bundling method suffers from the curse of dimensionality. Moreover, as shown in Boyle et al. (1997) an important disadvantage of the SSAP method is its potential lack of convergence. The reason for this is that the payoff is not a sufficient statistic for determining the optimal exercise strategy. This limits the potential use of these methods in particular for multidimensional problems.\textsuperscript{15}

The asymptotic properties of the parameterized stopping strategy method are examined in Garcia (2003) who shows that if the stopping time can actually be parameterized by $\theta$ then the simulation method converges. Likewise, in Broadie & Glasserman (2004) the stochastic mesh estimator is shown to converge to the true price when the mesh weights $p(\omega_n, j, t)$ are known. Thus, the existing results are for the “idealized” algorithms and therefore of little use for practical purposes. In fact, this is similar in spirit to assuming that the exact conditional expectation in e.g. $F(\omega, t_k)$ is given by some finite element function $F_M(\omega, t_k)$ with known form. In this case, it is simple to show that $\hat{F}_M^N(\omega, t_k)$ converges to $F_M(\omega, t_k)$ as $N$ tends to infinity and therefore the regression based method converges. To obtain convergence of the two methods in the general case one would need results when the stopping time is approximated or when the weights of the mesh are estimated, respectively, and these are to our knowledge not available.

\textsuperscript{15}In Raymar & Zwecher (1997) a generalization of the SSAP method to stratification along two dimensions is used. Though this method is shown to perform better in finite samples, asymptotically it has the same drawback as the original SSAP method.
4 The GARCH option pricing model

In the previous section we illustrate how American options can be priced using simulation. In particular, we argued that a framework that relies on cross sectional regressions to determine the conditional expectations is computational efficient and very flexible. It should therefore be an obvious candidate to be used to price options in very realistic models.

In this section we illustrate this using the generalized autoregressive conditional heteroskedastic, or GARCH, model originally suggested by Engle (1982) and Bollerslev (1986). We first introduce the GARCH framework and we explain how the risk neutral dynamics needed for option pricing can be obtained. Next, we provide a detailed algorithm for American option pricing in this model. We also discuss some alternative methods and review some of the existing results.

4.1 The GARCH framework

As elegant as it is, the BSM model with constant volatility and log-normal returns is unsatisfying as a model for financial data. For example, there exist an extensive empirical literature which has documented that the return variance of many assets is time varying. For illustrative purposes we plot in Figure 7 the return through time for four assets. The assets considered are General Motors (GM), International Business Machines (IBM), Merck and Company Inc. (MRK), and the Standard and Poor’s 100 index (OEX). The sample period is January 2, 1976, to December 29, 1995, for a total of 5,055 observations. The figure clearly shows that periods of high volatility are clustered, i.e. periods of high return variance are followed by periods of low return variance, and hence the variance is time varying.

In the empirical literature the GARCH framework has been used extensively to model financial time series with similar properties as the ones considered in Figure 7, see e.g. the surveys by Bollerslev, Chou & Kroner (1992) and Poon & Granger (2003). To illustrate this framework consider a discrete time economy with the price of an asset denoted $S_t$ and the dividends of that asset denoted
Figure 7: This figure shows time series plots of the annualized continuously compounded returns for the four assets considered. The sample period is January 2, 1976, to December 29, 1995, for a total of 5,055 observations.

\[ \delta_t. \] We then assume that the continuously compounded return process, \( R_t = \ln \left( \frac{S_t + \delta_t}{S_{t-1}} \right) \), can be modelled using the GARCH framework. In the most general form we specify the dynamics as

\[
R_t = m_t \left( \cdot; \theta_m \right) + \sqrt{h_t} \varepsilon_t
\]

(21)

\[
h_t = g \left( h_s, \varepsilon_s; -\infty < s \leq t - 1, \theta_h \right)
\]

(22)

\[
\varepsilon_t = Z_t,
\]

(23)

where \( Z_t \), conditional on the information set \( \mathcal{F}_{t-1} \) containing all information up to and including time \( t - 1 \), is a standard normal variable.
In (21) the term \( m_t(\cdot; \theta_m) \) denotes the conditional mean which is allowed to be governed by a set of parameters \( \theta_m \) as long as the process is measurable with respect to the information set \( \mathcal{F}_{t-1} \). This means that besides the various exogenous variables we might consider using in the mean equation, the mean term may include lagged values of e.g. the variance as is the case for the GARCH-in-Mean specifications. Likewise, in (22) the parameter \( \theta_h \) governs the variance process which is allowed to depend on lagged values of innovations to the return process and lagged values of the volatility itself. Two classical specifications are the GARCH process of Bollerslev (1986) given by

\[
    h_t = \omega + \alpha h_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1},
\]

and the non-linear asymmetric GARCH process, or NGARCH, of Engle & Ng (1993) given by

\[
    h_t = \omega + \alpha h_{t-1} \left( \varepsilon_{t-1} + \gamma \sqrt{h_{t-1}} \right)^2 + \beta h_{t-1}.
\]

The NGARCH specification can potentially accommodate asymmetric responses to negative and positive return innovations and thus allows for the so-called leverage effect. Moreover, note that the GARCH specification is obtained when the asymmetry parameter, \( \gamma \), is zero.\(^{16}\)

A major advantage with the discrete time GARCH model in (21) – (23) is that data is readily available for estimation. Moreover, estimation is computationally simple and can be performed using maximum likelihood techniques for example. In particular, conditional on starting values the log likelihood function for a sample of \( T_{\text{obs}} \) observations is, up to a constant, given by

\[
    L_{T_{\text{obs}}} (\theta) = \frac{1}{T_{\text{obs}}} \sum_{t=1}^{T_{\text{obs}}} -\frac{1}{2} \left( \log (h_t) + \varepsilon_t^2 / h_t \right),
\]

where \( \theta \) denotes the set of parameters in \( \theta_m \) in \( \theta_h \). This is very different from the continuous time framework in which volatility is modelled as a separate stochastic process. Though this framework

\(^{16}\)The specification, however, is much more general and can accommodate most, if not all, conceivable specifications used in the extant GARCH literature.
allows for a large degree of flexibility in the specification and sometimes leads to elegant solutions for the price of European options, in real applications a problem with the continuous time models is that volatility is unobservable. Hence estimation of these models is rather complicated. Moreover, when pricing claims for which numerical procedures are required, as it is the case with e.g. American options, future values of the unobservable volatility are needed. This variable is latent and hence potentially complicated to predict.\footnote{Nevertheless it is somewhat surprising that very few papers compare the option pricing performance of these two strands of the literature, though see Stentoft (2011a) for an exception.}

4.1.1 The risk neutral dynamics

Though there are clear advantages of the discrete time GARCH framework in terms of estimation, when deriving a theoretical option pricing model discreteness poses a potential problem as the asset market models easily become incomplete. This is indeed the case for the GARCH type models. Heuristically, the reason is that, unlike in the binomial model where the stock can take only two different values next period conditional on the price today, in a GARCH model the number of future possible values is infinite. Thus, in order to obtain the appropriate risk neutral dynamics further assumptions about preferences, apart from non-satiation, or about the risk premium have to be made in order to get to a risk neutral valuation relationship.

The approach used in Duan (1995), where an extension of the risk-neutralization principle in Brennan (1979) referred to as the locally risk neutral valuation relationship, or LRNVR, is used, falls in the first category. The LRNVR can be shown to hold under some familiar assumptions on preferences and under the assumed conditional normality in (23). Using the LRNVR it can be shown that the appropriate risk neutralized dynamics are given by

\[
R_t = m_t (\cdot; \theta_m) + \sqrt{h_t} \varepsilon_t \quad (27)
\]

\[
h_t = g (h_s, \varepsilon_s; -\infty < s \leq t - 1, \theta_h) \quad (28)
\]

\[
\varepsilon_t = Z_t - \lambda_t, \quad (29)
\]
where $Z_t$, conditional on $\mathcal{F}_{t-1}$, is a standard normal variable under the risk neutral measure $Q$\textsuperscript{18}. Furthermore, in (29) above $\lambda_t$ is the solution to

$$
\lambda_t = \frac{m_t(\theta_m) - r_t + \frac{1}{2}h_t}{\sqrt{h_t}}.
$$

(30)

Thus, it is immediately observed that the risk neutral dynamics depend alone on parameters that can be estimated using historical returns. Note also that in this setup risk neutralization happens through a change of the mean of the risk neutral innovations.

Though the risk neutral system depends only on parameters that can be estimated using historical return data accommodating the time varying volatility via the GARCH framework has one drawback: in general no closed form solutions can be found for the future distribution of stock prices and hence pricing of even European options is difficult. However, it is immediately clear that using the system in (27) – (30) a large number of paths of the risk neutralized asset prices can be generated and hence the simulation based methods discussed previously can be used. In particular, this is the method used originally by Duan (1995) who recognized that the asset price $S_t$ and the conditional volatility $h_{t+1}$ together serve as sufficient statistics to simulate from the risk neutral dynamics.

4.1.2 Extensions and alternatives

There are several interesting extensions to the GARCH framework above. For example, an interesting generalization is to consider situations where the conditional distribution is non-Gaussian. For example, this situation was considered theoretically by Duan (1999) who shows how to derive the risk neutral dynamics in this case. However, though non-Gaussian distributions have been shown to be important the problem remains that this complicates pricing. For example, it is generally not possible to find a closed form expression for $\lambda_t$ like that in (30) for the Gaussian case and therefore

\textsuperscript{18}In fact, the system above is somewhat more general than that used in Duan (1995). Nevertheless, this more general specification can be shown to obtain for example by considering the restriction to the log-normal model of the generalized LRNVR of Duan (1999).
it is difficult if not impossible to examine directly the effect of different mean specifications on option prices. Moreover, when using the method of Duan (1999) with non-Gaussian innovations the risk neutral innovations in (29) are generally not available explicitly and instead one needs to generate these using a numerical procedure, see e.g. Stentoft (2008a). This complicates pricing and adds to the computational complexity.

The Gaussian GARCH option pricing model of Duan (1995) and the generalization in Duan (1999) are not the only methods for obtaining the risk neutral dynamics to be used for option pricing. More recently, Christoffersen, El Kamhi, Feunou & Jacobs (2010) have demonstrated that option pricing is possible in a non-Gaussian framework with time varying volatility using only the no-arbitrage principle. In particular, they show that under certain assumptions the resulting risk neutral return dynamics are from the same family of distributions as the physical return dynamics also in the non-Gaussian case. Thus, the paper shows that similar results can be provided under a more general pricing framework which does not rely on particular assumptions on the preferences of the representative agent. Note that in the log-normal case this method and that of Duan (1999) coincide and lead to equivalent risk neutral dynamics. In the more general non-Gaussian case however the methods differ since in the approach of Christoffersen, El Kamhi, Feunou & Jacobs (2010) risk neutralization happens directly through a change of the parameters of the risk neutral distribution instead of by transforming the risk neutral innovations.19

Finally, it should be noted that multivariate extensions of the GARCH framework exist. For example, these models have been used extensively in recent years to model the behavior of financial data. The resulting dynamics are generally more realistic as they take into account the interaction and correlation between assets. In Rombouts & Stentoft (2011) the dynamics needed for option pricing are derived for a general multivariate discrete time model with time varying volatility and correlations. This framework generalizes the method of Christoffersen, El Kamhi, Feunou & Jacobs (2010), which is obtained as a special case when only one asset is considered.

19 Yet another approach, which would provide a similar set of conditions, is to specify a candidate stochastic discount factor directly as is done in e.g. Gouriou & Monfort (2007).
4.2 American option pricing with the GARCH framework

In the following we will use the American put option as an example, and for simplicity we will assume that the option can be exercised only at the end of every trading day. Furthermore, we assume that the one period continuously compounded dividend yield, \( \delta_t \), and the one period interest rate, \( r_t \), are both constant and we denote them simply by \( \delta \) and \( r \).

4.2.1 The algorithm

To fix notation, let the present time be denoted by \( t = 0 \) and assume that the option expires in \( K \) trading days at time \( T \). For notational convenience we denote the \( K \) trading days simply as \( t_k = k \). The pricing algorithm consists of the following four steps: a parameter identification step, a simulation step, and two pricing steps.

1. Identification and initialization of the relevant parameters:

   The relevant parameters in the GARCH framework are obtained. This can be done by estimating the parameters in (21)–(23) using maximum likelihood techniques. The parameter values are stored together with the necessary historical values for \( S_s, h_s, \varepsilon_s, \) and \( m_s (\cdot; \theta_m) \), for \( s \leq 0 \) as specified in the particular choice of variance equation in (22).

2. Simulation of the risk neutral asset price:

   Given the historical values for \( S_s, h_s, \varepsilon_s, \) and \( m_s (\cdot; \theta_m) \), for \( s \leq 0 \), we can immediately calculate the conditional volatility through period one, \( h_1 \). From (28) this is simply \( h_1 = g (h_s, \varepsilon_s, S_s; s \leq 0) \), and with this a realization of the next period stock price \( S_1 \) can be calculated from the starting level \( S_0 \) as

   \[
   S_1 = S_0 \exp \left\{ m_1 (\cdot; \theta_m) - \delta + \sqrt{h_1} \varepsilon_1 \right\},
   \]

   where \( \varepsilon_1 = Z_1 - \lambda_1 \) with \( Z_1 \sim N (0, 1) \). Likewise, the conditional volatility through the next
period should now be calculated as

$$ h_2 = g (h_s, \varepsilon_s, S_s; s \leq 1). \quad (32) $$

Iterating on (31) and (32) a further $K - 1$ times until the time of expiration is reached yields a single realization of a stock path. The simulation consists of $N$ such stock paths, each of which we label $S(n), n = 1, \ldots, N$, generated using different random Gaussian variates or using different seeds in the random number generator.\(^{20}\)

3. Calculation of the payoff along each path:

Let $Z$ be an $N \times K$ payoff matrix, with typical element $z_t(n)$. At time $T$ the payoff along each path is the maximum value of zero and the value of exercising the option. Hence, the elements of the last column in the matrix $Z$ may be calculated as

$$ z_K(n) = \max (0, X - S_K(n)), \quad 1 \leq n \leq N. \quad (33) $$

Next, at any time $t, 0 < t < K$, ordinary least squares, OLS, is used to estimate the conditional expectation of the payoff if the option is kept alive, along the paths for which the option is in the money and thus immediate exercise should be considered. As the dependent variable we use the discounted pathwise future payoff

$$ y_t(n) = \sum_{t=1}^{K} e^{-(t-t)r} z_t(n), \quad (34) $$

where $r$ is the risk free interest rate. As regressors we use powers and cross products of the two state variables which in this case are the stock price $S_t$ and the level of the volatility $h_{t+1}$. Denote the state variables by $x_t(n)$ and let $\Xi(x_t(n))$ be the transformation with powers and cross products. Then the conditional expectation $F(x_t(n)) = E[y_t(n)|x_t(n)]$ from Section

\(^{20}\)Most modern computer programs allow for generating all paths at once using simple matrix operations.
2.2 is approximated by
\[ \hat{F}(x) = \mathbb{E}(x_t(n)) \hat{\beta}_t, \]  

where \( \hat{\beta}_t \) is a vector of coefficients determined by OLS. The fitted values correspond to the estimated conditional expectation of the payoff from the option assuming that the option is kept alive. We now use the following rule to determine whether to exercise or not at time \( t \) for all the in the money paths, \( n, \)

\[ z_t(n) = \begin{cases} 
S_t(n) - X \text{ and } z_t(n) = 0, & t < \tilde{t} \leq K \text{ if } S_t(n) - X > \hat{F}(x) \\
0 & \text{otherwise}
\end{cases}. \]  

(36)

4. Calculating the value of the option:

At \( t = 0 \) the value of the option is calculated from the payoff matrix by discounting the payoffs to period zero using the risk free rate of interest and averaging across the simulated paths. Since there is at most one nonzero element along each path in \( Z \) this yields

\[ \hat{p}^N(0) = \frac{1}{N} \sum_{n=1}^{N} \sum_{t=1}^{K} e^{-tr} z_t(n). \]  

(37)

This ends the pricing algorithm. \( \blacksquare \)

In relation to the algorithm a couple of remarks are in place. First of all, the European price estimate is obtained by discounting the values in (33) back to time zero and averaging as

\[ \hat{p}^N(0) = \frac{1}{N} \sum_{n=1}^{N} e^{-Kr} z_K(n). \]  

(38)

Again, it is well established that (38) is an unbiased and asymptotically normally distributed estimate of (14) if the paths are independent. Secondly, in terms of the conditional expectation approximation a choice has to be made as to which transformation to use. We choose to use
monomials and cross products of the state variables as regressors. From Stentoft (2004b) we know that as the number of regressors, \( M \), and the number of paths, \( N \), tend to infinity the price estimate in (37) converges to the American option price in (1) where exercise is allowed once at the end of every trading day.

### 4.2.2 Discussion and review of some existing results

In addition to the simulation based algorithm outlined above alternative methods have been developed for pricing American options in the GARCH framework. For example, Ritchken & Trevor (1999) develop a lattice algorithm which approximates the NGARCH processes with conditional normality. Their algorithm can be seen as a generalization of the well known binomial tree which is used as an approximation to, e.g., the constant volatility BSM model. Duan, Gauthier, Sasseville & Simonato (2003) provides an alternative way to price options using a binomial tree approach that allows reducing the dimensionality of the three from two to one. Another alternative is proposed in the paper by Duan & Simonato (2001) in which a Markov Chain approximation to the underlying asset price process for an NGARCH specification is developed and used for option pricing purposes. See also Ben-Ameur, Breton & Martinez (2009) who develops a dynamic programming approach combined with interpolation that has the method of Duan & Simonato (2001) as a special case.

The methods proposed by Ritchken & Trevor (1999) and Duan & Simonato (2001) are compared to the simulation method above in Stentoft (2005) who finds that the LSM algorithm provides estimates which are very close to what is reported for American option prices in these two papers. However, compared to these other numerical methods the simulation method has clear advantages. In particular, the simulation based method is more flexible than the other methods and therefore it is indeed a useful alternative. For example, it is not immediately obvious how to extend the algorithms developed in Ritchken & Trevor (1999) and Duan & Simonato (2001) to other GARCH specifications. Moreover, though the dynamic programming approach of Ben-Ameur et al. (2009) can be used with a large class of GARCH models it is developed under the assumption of Gaussian
innovations. The simulation approach, on the other hand, does not suffer from this shortcoming and can be applied to virtually any GARCH specification as well as to models with non-Gaussian innovations.

The widespread use of simulation methods in the empirical applications of the GARCH model is a testament to its flexibility. For example, simulation methods have been used to price options on the Standard and Poor’s 500 Index using various GARCH specifications together with Gaussian innovations in among others Bollerslev & Mikkelsen (1996), Bollerslev & Mikkelsen (1999), Christoffersen & Jacobs (2004), and Hsieh & Ritchken (2005), and with non-Gaussian innovations in Christoffersen, Dorion, Jacobs & Wang (2010) and Rombouts & Stentoft (2010). Moreover, simulation techniques have been used together with the regression based method to price American style options with success in Stentoft (2005), Stentoft (2008a), and Stentoft (2011b). In fact, the only empirical paper that prices American style options in a GARCH framework and does not use simulation is Weber & Prokopczuk (2011). In this paper a binomial method is used and the GARCH models are calibrated directly to the option prices.

The empirical literature mentioned above generally finds support for not only GARCH effects, particular with asymmetries, but also for non-Gaussian innovations and this so for both European and American style options. For example, when considering options on the Standard and Poor’s 500 Index the above papers all find that the GARCH models diminish the mispricings found when using the constant volatility BSM model. Moreover, Stentoft (2011b) provides what is likely the largest analysis ever conducted of individual stock options using 30 stocks from the Dow Jones Industrial Average, or DJIA, and prices a total of 139,879 option contracts over the 11 year period from 1996 to 2006. The results of the paper provide clear support for using an asymmetric volatility specification together with non-Gaussian distribution, particularly of the Normal Inverse Gaussian type, and statistical tests show that this model is most frequently among the set of best performing models.
5 The price of risk and the effect on option prices

As the discussion above shows, several GARCH specifications and underlying distributions have been used for empirical option pricing in the GARCH framework. However, comparatively little work has examined the importance of the price of risk. This may appear somewhat puzzling as this parameter is what drives a wedge between the historical and the risk neutral dynamics.

In this section we consider this question in detail as an illustration of the flexibility of the GARCH framework and of the simulation based approach. To keep the analysis tractable we consider only models with Gaussian innovations since this allows for a closed form solution for $\lambda_t$ given $m_t(\cdot; \theta_m)$. We first explain how alternative specifications of the risk premium can be accommodated in the above framework and we then provide pricing results for the different specifications.

5.1 Specifications of the price of risk

The original GARCH option pricing paper by Duan (1995) specifies the mean as

$$R_t = r + \lambda \sqrt{h_t} - \frac{1}{2} h_t + \sqrt{h_t} \varepsilon_t,$$  \hspace{1cm} (39)

where $\varepsilon_t$ is standard normally distributed and $r$ is the risk-free interest rate. The particular specification corresponds to one where the mean part, $m_t(\cdot; \theta_m)$, of (21) is given by

$$m_t(\cdot; \theta_m) = r + \lambda \sqrt{h_t} - \frac{1}{2} h_t.$$

Substituting this into (30) yields a very simple expression for $\lambda_t$ which is given by

$$\lambda_t = \frac{r + \lambda \sqrt{h_t} - \frac{1}{2} h_t - r_t + \frac{1}{2} h_t}{\sqrt{h_t}} = \lambda.$$ \hspace{1cm} (41)

Hence, for this model the unit risk premium is constant and $\lambda$ is often interpreted as the price of risk.
Although the specification in Duan (1995) may be considered quite realistic in terms of the way the volatility influences the mean other functional forms could equally well be considered. In particular, an equally reasonable relation is to specify the mean as

\[ m_t (\cdot; \theta_m) = r + \lambda h_t - \frac{1}{2} h_t. \]  

(42)

This corresponds to the specification used in Heston & Nandi (2000), which in terms of the risk neutral dynamics this would lead to the following specification for \( \lambda_t \)

\[ \lambda_t = \frac{r + \lambda h_t - \frac{1}{2} h_t - r_t + \frac{1}{2} h_t}{\sqrt{h_t}} = \lambda \sqrt{h_t}. \]  

(43)

Thus, with this specification the unit risk premium is increasing in the level of the volatility. Likewise, we might consider specifying the mean as

\[ m_t (\cdot; \theta_m) = r + \lambda - \frac{1}{2} h_t, \]  

(44)

which would lead to the following specification for \( \lambda_t \)

\[ \lambda_t = \frac{r + \lambda - \frac{1}{2} h_t - r_t + \frac{1}{2} h_t}{\sqrt{h_t}} = \lambda / \sqrt{h_t}. \]  

(45)

In this situation the unit risk premium is decreasing in the level of the volatility.

The mean specifications in (40), (42), and (44) are closely related. For example, if we specify the functional form of the risk premium as \( \lambda h_t^\beta \) we are essentially considering the cases of \( \beta = 1 \) and \( \beta = 0 \) as alternatives to setting \( \beta = \frac{1}{2} \) as was done in Duan (1995). Whether either one of these is a more plausible specification from a theoretical point of view will not be discussed here. Instead we examine the empirical performance of the GARCH option pricing model with these alternative specifications. In Table 1 we report estimation results for \( \lambda \) with the three specifications considered...
Table 1: Estimation results for the risk parameter $\lambda$

<table>
<thead>
<tr>
<th></th>
<th>GM</th>
<th>IBM</th>
<th>MRK</th>
<th>OEX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda$</td>
<td>s.e. ($\lambda$)</td>
<td>$\lambda$</td>
<td>s.e. ($\lambda$)</td>
</tr>
<tr>
<td>Panel A: GARCH model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.0412</td>
<td>(0.0205)</td>
<td>0.0352</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.0322</td>
<td>(0.0140)</td>
<td>0.0316</td>
<td>(0.0171)</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.0206</td>
<td>(0.0077)</td>
<td>0.0216</td>
<td>(0.0102)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>GM</th>
<th>IBM</th>
<th>MRK</th>
<th>OEX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda$</td>
<td>s.e. ($\lambda$)</td>
<td>$\lambda$</td>
<td>s.e. ($\lambda$)</td>
</tr>
<tr>
<td>Panel B: NGARCH model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.0187</td>
<td>(0.0101)</td>
<td>0.0095</td>
<td>(0.0185)</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.0191</td>
<td>(0.0102)</td>
<td>0.0100</td>
<td>(0.0140)</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.0198</td>
<td>(0.0103)</td>
<td>0.0039</td>
<td>(0.0087)</td>
</tr>
</tbody>
</table>

Notes: This table reports the estimation results for the risk parameter. Panel A reports results for GARCH models and Panel B results for NGARCH models. In each panel we report the estimate as well as the standard error for each of the four assets, GM, IBM, MRK, and OEX. Each row reports the results for the three models specified in the text where the risk premium is specified as $\lambda h_t r_t$ for $\beta$ equal to 0, 1/2, and 1, respectively model 2, model 3, model 4.

For the four assets plotted in Figure 7. Each row reports the results for the three models with $\beta$ equal to 0, 1/2, and 1, respectively model 2, model 3, and model 4.\textsuperscript{21} Model 1 is used to denote a specification without any risk premium, that is when $\lambda = 0$, and therefore no results are reported for this model.

Panel A in Table 1 reports the results for the GARCH specification and shows that in general the risk parameter is estimated to be significantly positive. In fact, the parameter is insignificant only for IBM and this so only for models 2 and 3. The panel also shows that the estimates are similar in size. The specification with the largest likelihood value, not reported here, is model 4 for three out of the four stocks. Panel B in Table 1 reports the results for the NGARCH specification and shows that once asymmetries are accommodated in the variance specification the size of the estimated risk

\textsuperscript{21}We only report estimation results for the parameter of interest, $\lambda$. The rest of the parameters are for all specifications close to those reported in Stentoft (2005).
parameter decreases. In fact, once asymmetries are considered the estimate is insignificant for half of the specifications, that is for all the models for GM and IBM. When comparing the likelihood value model 3 now has the largest value for three out of the four stocks. Thus, overall the estimation results show that $\lambda$ is statistically significant for most of the considered specifications.

### 5.2 Pricing results

We now examine the effect of the alternative specifications of the price of risk in terms of the estimated option prices. For more details on the sample of options see Stentoft (2005). As a natural benchmark we consider a model without any price of risk, i.e. a model with $\lambda = 0$. In Tables 2 and 3 pricing errors are reported for the individual stock options and the index options, respectively.\(^{22}\) We report results across models as well as for put and call options individually. We consider two metrics for option pricing comparison using the dollar errors. Specifically, letting $P_k$ and $\tilde{P}_k$ denote the $k$th observed price respectively the $k$th estimated price we use the bias, $BIAS = K^{-1} \sum_{k=1}^{K} (P_k - \tilde{P}_k)$, and the root mean square error, $RMSE = \sqrt{K^{-1} \sum_{k=1}^{K} (P_k - \tilde{P}_k)^2}$.

The first thing to note from the tables is that the overall pricing errors are generally of similar order of magnitude and close to those reported previously in the literature. In particular, for model 3, which corresponds to the specification proposed in Duan (1995) and used empirically in Stentoft (2005) for the same assets considered here, the reported results are very close to those found in Stentoft (2005, Tables 8 and 9) for the $BIAS$. Moreover, when comparing to what is obtained with the constant volatility model the results show that allowing for GARCH type volatility specifications can explain a large fraction of the mispricing often found for models with constant volatility, and this so irrespective of the specification of the price of risk.

However, though the results are similar there are differences. For example, for the individual stock options Table 2 shows that allowing for a risk premium generally decreases the overall $BIAS$ of the GARCH models. In particular, the pricing errors of model 3 and model 4 are 1.9% and 2.3%\(^{22}\) We have chosen to report only results for the combined sample of individual stock options. The results for each stock are available on request.
Table 2: Pricing errors in dollar terms for individual stock options

Panel A: BIAS

<table>
<thead>
<tr>
<th></th>
<th>All (obs=8424)</th>
<th>Put (obs=3009)</th>
<th>Call (obs=5415)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GARCH</td>
<td>NGARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>Model 1</td>
<td>-0.1212</td>
<td>-0.1417</td>
<td>-0.0839</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.1212</td>
<td>-0.1429</td>
<td>-0.0819</td>
</tr>
<tr>
<td>Model 3</td>
<td>-0.1189</td>
<td>-0.1423</td>
<td>-0.0789</td>
</tr>
<tr>
<td>Model 4</td>
<td>-0.1184</td>
<td>-0.1424</td>
<td>-0.0778</td>
</tr>
</tbody>
</table>

Panel B: RMSE

<table>
<thead>
<tr>
<th></th>
<th>All (obs=8424)</th>
<th>Put (obs=3009)</th>
<th>Call (obs=5415)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GARCH</td>
<td>NGARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>Model 1</td>
<td>0.4438</td>
<td>0.4511</td>
<td>0.4275</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.4430</td>
<td>0.4509</td>
<td>0.4267</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.4439</td>
<td>0.4513</td>
<td>0.4282</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.4449</td>
<td>0.4530</td>
<td>0.4298</td>
</tr>
</tbody>
</table>

Notes: This table reports the pricing errors in dollar terms for individual stock options. Panel A reports the BIAS and Panel B the RMSE. In each panel we report the errors for the four models specified in the text, a model with no price of risk, model 1, and 3 models where the risk premium is specified as $\lambda h_t^\beta$ for $\beta$ equal to 0, $1/2$, and 1, respectively model 2, model 3, model 4. We report results for both the GARCH and the NGARCH specification and for all options as well as put and call options separately.

lower than for model 1, respectively. For the NGARCH model the differences are, on the other hand, quite small and this is also the case when considering the RMSE metric. When considering the put and call options separately the table shows that the differences in performance are largest for the put options where the BIAS is 2.3%, 5.9%, and 7.3% smaller with models 2, 3, and 4, respectively, than with model 1 when using a GARCH specification.

For the index options the differences are more pronounced than for the individual stock options and occur both for the GARCH and NGARCH specifications. For example, Table 3 shows that compared to the model with no risk premium the pricing errors are 11.9% smaller with model 2 for the GARCH specification and 22.9% smaller with model 4 for the NGARCH specification. When considering the GARCH model the table also shows that the reason for model 2’s superior
Table 3: Pricing errors in dollar terms for index options

Panel A: BIAS

<table>
<thead>
<tr>
<th></th>
<th>All (obs=8291)</th>
<th>Put (obs=4804)</th>
<th>Call (obs=3487)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GARCH</td>
<td>NGARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>Model 1</td>
<td>-0.0893</td>
<td>-0.0453</td>
<td>-0.3909</td>
</tr>
<tr>
<td>Model 2</td>
<td>-0.0787</td>
<td>-0.0427</td>
<td>-0.3732</td>
</tr>
<tr>
<td>Model 3</td>
<td>-0.0858</td>
<td>-0.0465</td>
<td>-0.3792</td>
</tr>
<tr>
<td>Model 4</td>
<td>-0.0848</td>
<td>-0.0349</td>
<td>-0.3795</td>
</tr>
</tbody>
</table>

Panel B: RMSE

<table>
<thead>
<tr>
<th></th>
<th>All (obs=8291)</th>
<th>Put (obs=4804)</th>
<th>Call (obs=3487)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GARCH</td>
<td>NGARCH</td>
<td>GARCH</td>
</tr>
<tr>
<td>Model 1</td>
<td>0.9299</td>
<td>0.8699</td>
<td>0.9200</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.9242</td>
<td>0.8719</td>
<td>0.9112</td>
</tr>
<tr>
<td>Model 3</td>
<td>0.9210</td>
<td>0.8702</td>
<td>0.9105</td>
</tr>
<tr>
<td>Model 4</td>
<td>0.9228</td>
<td>0.8756</td>
<td>0.9119</td>
</tr>
</tbody>
</table>

Notes: This table reports the pricing errors in dollar terms for index options. Panel A reports the BIAS and Panel B the RMSE. In each panel we report the errors for the four models specified in the text, a model with no price of risk, model 1, and 3 models where the risk premium is specified as $\lambda h_t^\beta$ for $\beta$ equal to 0, 1/2, and 1, respectively model 2, model 3, model 4. We report results for both the GARCH and the NGARCH specification and for all options as well as put and call options separately.

performance is that it performs well for the put options. For the call options the difference between models 1 and 2 is on the other hand only minor. Note though that models 3 and 4 outperform model 1 for both put and call options by between 1.5% and 3.0%. For the NGARCH model things are less clear. For example, whereas model 4 outperforms model 1 for put options the opposite is the case for the call options. In fact, once asymmetries in the volatility specification are allowed for neither of the specifications with risk premium outperforms the model with no premium for both puts and calls. Finally, the table shows that when considering the RMSE metric the differences in pricing errors for the index options are generally very small and in most cases less than 1%.
6 Conclusion

This paper contains an introduction to how simulation methods can be used to price American options and a discussion of various existing methods. In particular, we provide a review of four methods: the parameterized stopping time method of Garcia (2003) and Ibanez & Zapatero (2004), the bundling or state space partitioning methods of Tilley (1993) and Barraquand & Martineau (1995), the regression based methods of e.g. Carriere (1996), Tsitsiklis & Van Roy (2001), and Longstaff & Schwartz (2001), and the stochastic mesh method of Broadie & Glasserman (2004). We compare the methods in terms of their computational efficiency and asymptotic properties and argue that the regression based methods have clear advantages.

To illustrate the flexibility of the regression based method an application to the GARCH option pricing model is provided. In particular, we present a detailed algorithm for how to price options in the GARCH framework using simulation and the regression based method. A discussion of the benefits of the proposed algorithm compared to other existing methods is provided. In an empirical application we examine the effect on estimated option prices of using different specifications for the price of risk. We find that the specification of the price of risk has a non-negligible effect on the pricing performance. In particular, the best model has pricing errors that are up to 2.3% lower than the benchmark for the individual stock options and up to 22.9% lower than the benchmark for the index options.

The empirical example provided here is only one a several interesting applications of the regression based method. For example, more general GARCH specifications as well as applications to other financial derivatives e.g. within the fixed income market could be considered. Moreover, though the regression based method has been analyzed in quite some detail it is still of interest to examine improvements to it using variance reduction techniques or ways to reduce the potential bias of the estimated price.
References


