BOUNDS ON THE AMERICAN OPTION

Anthony Neuberger

Warwick Business School
University of Warwick
Coventry
United Kingdom

January 2007

The essential feature of American style claims lies in the holder's right to time the exercise decision. The value of the claim depends on the information about future prices that the holder will acquire over time. Much of the literature makes restrictive assumptions about information revelation – for example that the underlying price process is Markov. This paper explores the upper bound on the price of an American option, placing no assumptions on the information structure. The analysis provides insight into the processes that make the American feature valuable, and points the way to hedging strategies for American options that are robust to model error.

I am grateful to comments from Jerome Detemple, Gordon Gemmill, Vicky Henderson, David Hobson, Dick Stapleton, and seminar participants at Boston University, Princeton, Manchester Business School and the University of Warwick.

1. INTRODUCTION

The essence of an American option lies in the holder's right to time the exercise decision. The value of being American depends on the information that the holder will acquire about future asset returns. Much of the literature on American options assumes that the underlying asset price follows a Markov process, and that the only information the holder acquires about future returns is contained in past asset prices. With a richer information structure, where the holder learns about the distribution of future returns, the value of being American will in general be higher. We derive the arbitrage-free upper bound on the price of an American option given only the contemporaneous prices of all European options, and explore the price processes that support these values.

American options are the most common and also the most subtle of exotic options. While many exotic options have pay-offs that are functions of the price path alone, the pay-off to an American option also depends on the exercise strategy of the holder. The decision to exercise will depend on all the information that the holder has at the time. In the standard Brownian diffusion model, the only relevant information for the holder of the option is the current asset price. The optimal exercise strategy is to exercise the option the first time the asset price breaches a barrier – the *immediate exercise boundary* (see Broadie and Detemple (2004) for a more formal treatment). With a richer information structure – for example with stochastic volatility – the exercise strategy is more complex, and depends on all state variables rather than just the underlying asset price.

In general, holding the price of European options fixed, the richer the information structure, the greater the value of the American option. As Longstaff, Santa-Clara and Schwartz (2001) argue in the context of American and Bermudan style swaptions, the costs of using a single-factor model to make exercise decisions when the term structure is driven by multiple factors are substantial. They argue that these costs remain substantial even if the model is recalibrated to fit the current set of market prices of related derivative securities (in this case European swaptions) at each possible exercise date. While Andersen and Andreasen (2001) and Svenstrup (2005) contest this latter conclusion, the need to recalibrate a single factor model to take account of new information revealed in market prices when deciding whether to exercise the option is common ground.

Taking account of new information though repeated recalibration of a model is internally inconsistent. It violates the model's own assumptions that the parameters are constant. Internal consistency could be obtained by using a multi-factor model with at least as many factors as the number of parameters required to fit the surface of call option prices. But the number of factors needed, and the associated problems of parameter estimation and stability, make this an unattractive way of accounting for new information.

We therefore follow a different route. To get a bound on how much new information could matter, we ask the question: given the prices of European options, how much could an American option be worth without allowing arbitrage opportunities? In answering the question, the only assumption (apart from the standard assumptions of frictionless markets and absence of arbitrage) is that interest rates and dividends are deterministic.

The search for arbitrage-based option pricing bounds goes as far back as Merton (1973) who shows that the price of an American put option is bounded below by the equivalent European put option; using obvious notation, $P^a\left(X,T\right) \geq P^e\left(X,T\right)$. It can readily be shown that the European put with the same maturity as the American, but a strike whose present value is equal to the nominal strike of the American option provides an upper bound; $P^a\left(X,T\right) \leq P^e\left(Xe^{rT},T\right)$.

Merton (1973) concerns the dominance relationship between two claims. Brown, Hobson, and Rogers (2001) generalize the question by searching for the tightest bounds on a claim (in their case barrier and lookback options) in the presence of a whole set of other claims (European options with the same maturity). The bounds have a number of uses. The bound is enforced by a hedging strategy which is of interest because it is robust. Robust strategies put a floor on potential hedging losses. Being model free, they avoid the problems highlighted by Green and Figlewski (1999) when conventional hedging techniques are used with an incorrect model, or poor parameter estimates. The extreme processes that maximize the value of the exotic option cast light on the features of a price process that make the exotic option particularly valuable.

The bound on the value of the American feature is not in general very close to the price from standard pricing models. As will be shown in the case of the American put option, the maximum possible value of the American premium (defined below) can readily be three or more times larger than that implied by a standard diffusion process before creating an arbitrage opportunity. This suggests that there is scope for models with richer information structures to generate American option values that differ substantially from traditional models. Examples given in this paper confirm that insight.

The model is set in a discrete time, discrete space world. This is for convenience only. It makes it possible to formulate the problem of finding the rational bounds as a finite dimensional linear program. The continuous time, continuous space limits are easily identified. The strategy for finding the arbitrage bound is as follows: in the first stage, we identify a class of strategies using European options that bound the American option. The search for the cheapest member of the family is formulated as a linear program (LP). In the second stage, a family of processes for the underlying asset is characterized. The search for the member of the family that places the greatest value on the American option is formulated as a second LP. It is shown that the second LP is the dual of the first, so they have the same solution. The solution is both a feasible price for the American option, and an upper bound on its price. So it is the highest possible price consistent with the absence of arbitrage.

Others authors (Andersen and Broadie (2004), Rogers (2002) and Haugh and Kogan (2004)) also exploit the relationship between the primal and dual problems to bound the value of an American option. Their work differs from ours in that their bounds are designed to bracket the true value of the American option under a known price process by using a near to optimal exercise strategy. We, on the other hand, are seeking price bounds that are independent of the process.

Having identified the bounds on the general American option in the next section, section 3 explores numerically the bounds on the price of the American put option. The nature of the dominating strategies and of the processes that support these extreme values are identified. These processes involve jumps and state dependent volatility. Section 4 explores the extent to which processes with jumps and state dependent volatility may lead to particularly high values for the American feature. The fifth section tightens the bounds by ruling out implausible processes. The final section concludes.

2. IDENTIFICATION OF THE BOUNDS

2.1 The Set Up

The model is set in a discrete time, discrete space framework. The time index t goes from 0 to T. There is a single risky security. It pays no dividends[1] and is traded in a frictionless market at price S_t. S_t can take one of J possible values:

(2.1) $$S_t \in \left\{ x_j \,\middle|\, j = 1, ..., J \right\} \text{ for all } t, \text{ with } x_1 < ... < x_J.$$

The spacing of the time and price nodes is not critical, though it is natural to think of the discrete times being equally spaced, and the discrete prices being placed in a geometrical progression. The interest rate is non-stochastic. We work with discounted prices, discounting the nominal price to $t = 0$. Price is used as shorthand for discounted price.

An American option is characterized by a $J \times T$ *pay-off matrix* $\mathbf{A} = \{a_{jt}\}$. If $S_t = x_j$ and the option has not previously been exercised, the holder can choose to exercise it and receive a pay-off whose discounted value is a_{jt}.[2] Once the option has been exercised, it is dead; it cannot be exercised again. The holder can chose to let the option expire without exercising it.

For each node (j, t) there exists an elementary security that pays 1 if $S_t = x_j$ and 0 otherwise[3]. Its time zero price is p_{jt}. The $J \times T$ matrix $\{p_{jt}\}$ is denoted by $\mathbf{P}$. p_{jt} can be

---

[1] The extension to assets with a constant yield, or a known dividend, is described in 1.4.

[2] It is implicitly assumed that the American option cannot be exercised at time 0; this is for notational convenience and has no effect on the results.

[3] We could instead have assumed the existence of a complete set of conventional European call options (that is an option for each strike x_j and each maturity t). The Arrow-Debreu claims we are assuming can be

interpreted as the risk-adjusted probability of state *j* occurring at time *t*, given information at time 0. The prices of the elementary claims are assumed to preclude arbitrage. This implies the existence of a process under which *S* is a martingale and $\Pr\{S_t = x_j\} = p_{j,t}$ for all (j, t) (Davis and Hobson 2007).

2.2 Dominating Portfolios

A portfolio of elementary securities is represented by a $J \times T$ matrix $\mathbf{E} = \{e_{jt}\}$. The portfolio pays the holder an amount e_{jt} at time *t* if $S_t = x_j$. The cost of the portfolio, $c(\mathbf{E})$, must be the same as the cost of the elementary securities that compose it:

$$c(\mathbf{E}) = \sum_{j,t} e_{j,t} p_{j,t}. \tag{2.2}$$

The following theorem sets out sufficient conditions under which the European portfolio **E** dominates the American option **A**.

THEOREM 2.1: *if there exist three $J \times T$ matrices* **E, D** *and* **V** *that satisfy the following conditions:*

$$\begin{aligned}
&1)\quad e_{j,t} \ge 0 \text{ for all } j,t; \\
&2)\quad v_{j,t} \le e_{j,t} + d_{j,t}\left(x_m - x_j\right) + v_{m,t+1} \text{ for all } j,m,t, \\
&\qquad \text{where } v_{m,T+1} \equiv 0; \\
&3)\quad v_{j,t} \ge a_{j,t} \text{ for all } j,t,
\end{aligned} \tag{2.3}$$

then the cost of **E** *is an upper bound on the price of the American claim* **A**.

*Proof***:** suppose the theorem is false. An agent writes the American option at time 0, and uses the proceeds to buy the European option **E**. The agent then follows the following strategy: until the time the American option is exercised, the agent does nothing. If the

---

created from the call options using a butterfly spread – a long position in calls with strikes x_{j-1} and x_{j+1} and a short position in calls with strike x_j.

option is exercised at time t^*, then at any time $t \geq t^*$, the agent hedges by holding d_{jt} [4] of the underlying asset.

By assumption, **A** costs more than **E**, so the cash flow at inception is strictly positive. At every period up to time t^*, the agent receives an amount e_{jt} which is positive by condition (1). (2) ensures that the proceeds from the delta hedged portfolio **E** from time t^* to the end is at least v_{j^*,t^*}. Condition (3) ensures that this exceeds a_{j^*,t^*}, the exercise cost of the American option. The aggregate cash flow is also strictly positive if the American option is never exercised. This is an arbitrage strategy. To avoid arbitrage, the price of the American option must be less than or equal to the price of the European portfolio. ■

To understand the theorem, interpret $v_{j,t}$ as the *intrinsic value* of the European portfolio **E** at node (j, t). That is to say, by delta hedging suitably (represented by the matrix **D**), the holder of the portfolio can ensure that the hedged proceeds from the portfolio from time t onwards will equal at least $v_{j,t}$ whatever path the asset subsequently follows. The theorem states that if **E** has positive cash flows only, and if the intrinsic value of **E** exceeds the immediate exercise value of the American option at all nodes, then **E** dominates the American option.

The theorem identifies a set of strategies that bound the price of an American option. To find the tightest such bound, it is natural to express the problem as a linear program:

LP1: Find the triplet of $J \times T$ matrices $\{\mathbf{D}, \mathbf{E}, \mathbf{V}\}$

that minimizes $\sum_{j,t} e_{j,t} p_{j,t}$

subject to the constraints:

1) $e_{j,t} \geq 0$ for all j, t;
2) $v_{j,t} \geq a_{j,t}$ for all j, t;
3) $e_{j,t} + \left(x_m - x_j\right) d_{j,t} - v_{j,t} + v_{m,t+1} \geq 0$ for all j, m, t

   where $v_{m,t} \equiv 0$ when $t = T + 1$.

---

[4] Strictly, we should write $d_{j_t,t}$, but we drop the second subscript for simplicity.

The feasible set is not empty[5]. The solution is bounded below by zero; otherwise there would exist a portfolio **E** that has positive cash flows in all states of the world, and has a negative cost, and this would be an arbitrage. So LP1 has a solution; the solution is denoted by an asterisk. Theorem 2.1 implies that the price of the American option is bounded above by c(**E***). The next step is to show that it is the least upper bound.

2.3 Feasible Prices

c(**E***) is proved to be the least upper bound by identifying a process for the underlying asset and an exercise strategy for the American option ("a feasible process"), under which the price of every traded security is a martingale, and under which the expected pay-off to the American option is c(**E***). The set of feasible processes is of very high dimensionality. We rely on the intuition that the search for the cheapest dominating strategy, and the search for the process that places highest value on a claim, are in some sense dual problems. If we can show that the dual of LP1 can be interpreted as the search within a non-empty set of feasible processes then we have attained our goal.[6]

The dual of LP1 can be written as:

---

[5] For example, define $e_{jt} = v_{jt} = \max[a_{jt}, 0]$ and $d_{jt} = 0$ for all j and t.

[6] If the set of dominating portfolios we identified in Proposition 1 is too small, so the upper bound is not the supremum, then our strategy will fail. The dual problem will still identify a process, but it will not be a feasible process.

LP2: Find the $J \times T$ matrix **F** and the $J \times J \times T$ matrix **H** that maximizes $\sum_{j,t} f_{j,t} a_{j,t}$ subject to the constraints:

1) $\sum_{m=1}^{J} \left(x_m - x_j \right) h_{j,m,t} = 0$ for all j,t;
2) $\sum_{m=1}^{J} h_{j,m,t} \le p_{j,t}$ for all j,t;
3) $f_{j,t} + \sum_{m=1}^{J} \left(h_{m,j,t-1} - h_{j,m,t} \right) = 0$ for all j,t where $h_{m,j,t-1} \equiv 0$ when $t = 1$;
4) $f_{j,t} \ge 0$ for all j,t;
5) $h_{j,m,t} \ge 0$ for all j,t.

This program has a natural interpretation. Let I_t be an indicator variable that is 0 when the option is live, and 1 from the time the option is exercised. Interpret *f* and *h* as follows:

$$f_{j,t} \equiv \Pr\left\{ I_{t-1} = 0 \wedge I_t = 1 \wedge S_t = x_j \right\},$$
$$h_{j,m.t} \equiv \Pr\left\{ I_t = 1 \wedge S_t = x_j \wedge S_{t+1} = x_m \right\}.$$

The last two constraints specify that the probabilities must be positive; the first three constraints specify that:

1) $E\left[S_{t-1} - S_t \middle| S_t, I_t = 1 \right] = 0$ for all S_t, t;
2) $\Pr\left\{ S_t = x_j \wedge I_t = 1 \right\} \le \Pr\left\{ S_t = x_j \right\}$ for all j,t;
3) $\Pr\left\{ I_t = 1 \middle| S_t \right\} = \Pr\left\{ I_{t-1} = 1 \middle| S_t \right\} + \Pr\left\{ I_{t-1} = 0 \wedge I_t = 1 \middle| S_t \right\}$ for all S_t, t.

While the conditions in LP2 are necessary for (**F, H**) to characterize a martingale process for *S* (and exercise strategy) that is consistent with the risk neutral probabilities **P**, it is less obvious that they are sufficient. If they are insufficient, the solution to LP2 (and to LP1) may not be a feasible price for the American option. Rather than prove sufficiency, we use LP2 to motivate the characterization of a family of martingale processes and exercise strategy, and then search for the process and associated exercise strategy that maximizes the American option's value.

Consider the following process: at time 0 the world is in state 0, so the state indicator function $I_0 = 0$. If at some time τ it switches to state 1, $I_t = 1$ for all $t \geq \tau$. Consider the filtration jointly generated by the state and the asset price $\{S_t\}$. The process for the underlying can be fully captured by two $J \times J \times T$ matrices **G** and **H** where:

$$g_{j,m,t} \equiv \Pr\left\{S_t = x_j \wedge S_{t+1} = x_m \wedge I_t = 0\right\}; \quad (2.4)$$
$$h_{j,m,t} \equiv \Pr\left\{S_t = x_j \wedge S_{t+1} = x_m \wedge I_t = 1\right\}.$$

Also define $f_{j,t}$, the probability that the state changes to state 1 at time t and $S_t = x_j$, so $f_{j,t} \equiv \Pr\left\{S_t = x_j \wedge I_{t-1} = 0 \wedge I_t = 1\right\}$.

For {**F**, **G**, **H**} to represent a feasible process, it is sufficient that they are positive, that S is a martingale with respect to past prices and the indicator variable, that they are consistent with **P**, and **F** is consistent with **G** and **H**. This requires that:

$$\begin{aligned}
&1)\ \ f_{j,t} \geq 0; \qquad 2)\ \ g_{j,t} \geq 0; \qquad 3)\ \ h_{j,t} \geq 0;\\
&4)\ \ \sum_m \left(x_m - x_j\right) g_{j,m,t} = 0; \qquad 5)\ \ \sum_m \left(x_m - x_j\right) h_{j,m,t} = 0;\\
&6)\ \ \sum_m \left(g_{j,m,t} + h_{j,m,t}\right) = p_{j,t}; \quad 7)\ \ \sum_m \left(g_{m,j,t} + h_{m,j,t}\right) = p_{j,t+1}\ \ (t < T);\\
&8)\ \ f_{j,t} = \sum_m \left(h_{j,m,t} - h_{m,j,t-1}\right) \qquad \text{(with } h_{m,j,t} \equiv 0 \text{ when } t = 0\text{)};\\
&\text{for } j = 1,\ldots,J \text{ and } t = 1,\ldots,T.
\end{aligned} \quad (2.5)$$

Consider the strategy of exercising the option when the state changes to state 1. Under the process {**F**, **G**, **H**} and with this exercise strategy, the value of the American option is:

$$v(\mathbf{F}) = \sum_{j,t} a_{j,t} f_{j,t}. \quad (2.6)$$

Now consider the linear program:

LP3: Find the $J \times T$ matrix **F** and the two $J \times J \times T$ matrices **G** and **H** that maximizes $\sum_{j,t} f_{j,t} a_{j,t}$ subject to the constraints in (2.5).

The feasible set is not empty; consider for example the strategy of never exercising the American claim, so **F** = **H** = **0**. **G** is then just a martingale process for the underlying asset that is consistent with the prices of the European options; such a process must exist if the market is free of arbitrage. The problem is also bounded since **F** is bounded ($p_{j,t} \geq f_{j,t} \geq 0$). So there is a solution $v(\mathbf{F^*})$.

THEOREM 2.2: *the search for the cheapest dominating portfolio (LP1) is equivalent to the dual of the search for the process that maximizes the value of the American claim (LP3). The cost of the cheapest dominating portfolio* $c(\mathbf{E^*})$ *is equal to the maximum feasible price* $v(\mathbf{F^*})$*, and they are therefore both equal to the least upper bound on the value of the American claim.*

*Proof***:** in the Appendix. The proof proceeds by identifying the dual to LP3, and showing that it is equivalent to LP1 in the sense that the solutions to the two programs are the same. The second part of the theorem immediately follows because the solution to a primal and its dual are the same. ■

2.4 The Bounding Process

In this section we explore the nature of the price process that maximizes the value of the American option. In solving LP1 to find the maximum value, the binding constraints can be identified. The Complementary Slackness Theorem then identifies the constraints in the dual problem, LP2, that are binding. These binding constraints help characterize the price processes that maximize the value of the American option and the corresponding exercise strategies.

The table sets out the primal and dual LPs with complementary constraints matched:

	Primal (LP1)	***Dual (LP2)***
Objective	Minimize $\sum_{j,t} p_{j,t} e_{j,t}$	Maximize $\sum_{j,t} a_{j,t} f_{j,t}$
Subject to:		
(1)		$\sum_m (x_m - x_j) h_{j,m,t} = 0$
(2)	$e_{j,t} \geq 0$	$\sum_m h_{j,m,t} \leq p_{j,t}$

(3)		$f_{j,t} + \sum_m \left(h_{m,j,t-1} - h_{j,m,t} \right) = 0$
(4)	$v_{j,t} \ge a_{j,t}$	$f_{j,t} \ge 0$
(5)	$e_{j,t} + \left(x_m - x_j \right) d_{j,t} - v_{j,t} + v_{m,t+1} \ge 0$	$h_{j,m,t} \ge 0$

By the Complementary Slackness Theorem:

(2.7)

1) $e_{j,t} > 0 \Rightarrow \sum_m h_{j,m,t} = p_{j,t}$ (from 2);

2) $v_{j,t} > a_{j,t} \Rightarrow f_{j,t} = 0$ (from 4);

3) $e_{j,t} + \left(x_m - x_j \right) d_{j,t} - v_{j,t} + v_{m,t+1} > 0$
$\Rightarrow h_{j,m,t} = 0$ (from 5).

The characteristics of the processes that maximize the value of the put and the corresponding optimal exercise strategies can now be identified:

THEOREM 2.3: *The extreme process that drives the option value to its maximum is given by the following:*

(a) *if node (j,t) is attained and the option is dead, then the path continues to node (m, t+1) only if $v_{.,t+1}$ is a linear function of the asset price in the range $[x_j, x_m]$.*

(b) *nodes can be divided into four complementary sets:*

(2.8)

$$W = \left\{ (j,t) \middle| e_{j,t} = 0 \wedge v_{j,t} = a_{j,t} \right\}; \quad X = \left\{ (j,t) \middle| e_{j,t} > 0 \wedge v_{j,t} = a_{j,t} \right\};$$
$$Y = \left\{ (j,t) \middle| e_{j,t} = 0 \wedge v_{j,t} > a_{j,t} \right\}; \quad Z = \left\{ (j,t) \middle| e_{j,t} > 0 \wedge v_{j,t} > a_{j,t} \right\}.$$

The optimal strategy is to exercise in the X zone, and not to exercise in the Y zone; the Z zone is attained only by paths where the option has already been exercised. The exercise strategy is not constrained in the W zone.

*Proof***:** the first part of the theorem follows from part (3) of (2.7), while the second part follows directly from parts (1) and (2).

The interpretation of Theorem 2.3 will become clearer when it is applied to specific types of American option. In the next section, the model is applied to the case of the American put option on an asset that pays no dividends, with nominal strike price K and with a constant risk-free return per period of R. In this case $a_{jt} = KR^{-t} - x_j$.

2.5 Extending the Model

The theorems apply to an American option on an asset that pays no dividends. They can readily be extended to assets that do pay dividends, provided that the present value of the dividend is a known function of the asset price. In particular, take the case where the asset has a constant yield of d per period. If the option is exercisable into the cum-dividend asset, and the yield is expressed as a proportion of the ex-dividend price, the underlying asset for the purposes of the model is a portfolio that starts with one unit of the asset, with dividends reinvested. This notional underlying asset pays no dividends by construction. The pay-off to an American option if exercised at time t when at node x_j, which has been denoted by $a_{j,t}$, is a function of the node j and the time t.

For example, take the case of an American call option on this asset with nominal strike K. The present value of the exercise value at node (j, t) is $a_{jt} = x_j(1 + d)^{-t} - KR^{-t}$ where R is the one period return on the risk-free asset. The extension to discrete proportional dividends and to known cash dividends is straightforward, as is the extension to time varying but non-stochastic interest rates and dividend yields is straightforward.

The model can also be used to bound the price of Bermudan-style options where the holder's right to exercise the option is restricted to particular times or periods. If the holder cannot exercise the option at time t then a_{jt} can be set equal to zero (or a negative number) for all j. Since the holder of a live option can chose to allow the option to expire unexercised, it is rational to defer exercise when $a \leq 0$ and so keep the option alive.

3. EXPLORING THE BOUNDS

3.1 Bounds on the American Put

In this section the model is implemented for American put options with maturity τ years. The option can be exercised at any of T+1 equally spaced time points that start at 0 and end at τ. The length of each period is $\delta = \tau/T$. Price nodes are distributed geometrically with $x_{j+1} = ux_j$ for some $u > 1$.

We assume initially that all European options trade at the same implied volatility, and the risk free interest rate is constant. In particular, since the model is set in a discrete price framework, the European option prices and the state price densities are generated as if the discounted underlying asset follows the following (continuous time) Poisson jump process with volatility σ:

(3.1)
$$S_{t+dt} = \begin{cases} S_t u & \text{with probability } \pi dt \\ S_t / u & \text{with probability } u\pi dt \\ S_t & \text{otherwise.} \end{cases}$$
$$\text{where } (1+u)\pi\left(\log u\right)^2 = \sigma^2.$$

The process followed by the asset price is independent of the coarseness of the time grid δ, but does depend on u. As u goes to 1, the jumps become smaller and more frequent, and the process converges to a diffusion, with a volatility that remains equal to σ. We will call the process with finite u a diffusion since it is the closest approximation to a pure diffusion that exists in a discrete space world. The American option can only be exercised at a node so δ does affect the value; as δ goes to zero, the permissible exercise dates converge to the continuum [0, τ].

The support of the price at any time horizon is unbounded on the positive real line. For computational convenience, it is desirable to keep the number of nodes finite, so we set absorbing barriers at $S_0 e^{\pm 4\sigma\sqrt{\tau}}$. These barriers have no significant impact on the valuation.

For the base case, the initial asset price S_0 is 100, the time to maturity τ is 1 year, and the annualized volatility σ is 10%. Table I sets out the upper bound on the price of the American put option in the base case with a number of different values for the strike price K and the interest rate r. u =1.01, so price nodes are spaced at 1% intervals, and there are 50 periods per year. J, the number of price levels is 130.

The bound is compared with the conventional valuation of an American option – the expected pay-off to the option under the process in (3.1) under the optimal exercise strategy. Two European put option valuations are also reported. One is the value of a European put option $P^E(K, T)$ with nominal strike K and that matures after T periods. The other is the value of the European put option with nominal strike K and maturity of T or less that has maximum value, $\text{Max}\{P^E(K, t)\mid t \le T\}$, which we denote by $P^E(K, \le T)$. The American premium is defined as the difference between the value of the American put and $P^E(K, \le T)$ with the same nominal strike and the same or shorter maturity. It represents the premium the holder pays for being able to specify the exercise time later, rather than having to choose the exercise time at inception.

The value of an option depends on the level of the strike relative to the spot, so two measures of the moneyness of the put option are reported, one at inception, and the other at expiry:

(3.2)
$$d_0 = \ln\left(S/K\right)/\sigma\sqrt{\tau}\,;$$
$$d_T = \ln\left(S/Ke^{-r\tau}\right)/\sigma\sqrt{\tau}\,.$$

Table I: Upper Bound on the Value of an American Put

Strike (K)	Upper Bound (1)	American, Diffusion (2)	$P^E(K, \le T)$ (3)	$P^E(K, T)$ (4)	Ratio (5)	d_0 (6)	d_T (7)
r = 2%							
95	1.64	1.40	1.35	1.35	18%	0.51	0.71
100	3.71	3.20	3.03	3.03	24%	0.00	0.20
105	6.83	6.09	5.67	5.67	36%	-0.49	-0.29
110	10.81	10.06	10.00	9.18	8%	-0.95	-0.75
r = 5%							

95	1.28	0.90	0.77	0.77	25%	0.51	1.01
100	3.31	2.38	1.92	1.92	33%	0.00	0.50
105	6.51	5.23	5.00	3.92	15%	-0.49	0.01
110	10.62	10.00	10.00	6.81	0%	-0.95	-0.45
r = 10%							
95	0.75	0.43	0.27	0.26	33%	0.51	1.51
100	2.65	1.54	0.99	0.79	34%	0.00	1.00
105	6.04	5.00	5.00	1.89	0%	-0.49	0.51
110	10.38	10.00	10.00	3.75	0%	-0.95	0.05

The table shows the upper bound on the value of an American put with maturity of 1 year for various levels of strike and interest rates, when the spot price of the underlying is 100, and European options are trading on an implied annual volatility of 10%. The second column shows the price of the American put assuming the asset follows the diffusion process; the next two columns show the prices of European options with the same strike and maturities of either up to 1 year, or of 1 year exactly. "Ratio" is the ratio of the American option premium under the diffusion process to the maximum premium, and is equal to {(2)-(3)}/{(1)-(3)}. d_0 and d_T are measures of moneyness, defined in the body of the paper. Values are calculated using 50 steps/year, and a stock price mesh of 1%.

One striking feature of table I is how small is the ratio (shown in column 5) between the value of being American assuming that the process is a diffusion, and the upper bound on the value of being American. Taking a crude average across strikes and interest rates, the standard valuation of model value is only 19% of its rational bound, and for no strike or interest rate does it rise above 37%, suggesting that the American feature may be seriously undervalued by standard models.

With five parameters in the model (S, K, r, τ, σ), table I is apparently sampling only a small part of the parameter space. But relative prices are largely determined by the moneyness vector (d_0,, d_T); for a given moneyness vector, all the option prices are approximately proportional to $S\sigma\sqrt{\tau}$[7]. The ratio column then depends essentially only on

[7] If the discounted exercise price K_t were linear in time, and if the process generating European option prices had constant variance of price *changes,* dS_t then, holding moneyness constant, prices of all options (European, American, upper bound) would be exactly proportional to the standard deviation of the terminal price. In our model, it is $\ln K_t$ that is linear in time, and $\ln(dS_t)$ that has constant variance, but for low interest rates and volatility (small values of $r\tau$ and $\sigma\sqrt{\tau}$) the homogeneity property holds approximately.

the (two-dimensional) moneyness vector. The range of parameters in Table I is chosen to give a suitably wide range of values for the moneyness vector.

To exhibit the bounding European portfolios, the intrinsic value $v_{j,t}$ of the portfolio is plotted as a function of x_j for different values of t. The lines are convex; it follows that the intrinsic value $v_{j,t}$ is identical to the sum $\sum_{s=t}^{T} e_{j,s}$. The values of e, the pay-out from the portfolio in each period, can be read off by looking at the difference between successive curves. Figure 1 shows these portfolios for an out of the money option (K = 98) and an in the money option (K = 102). In both cases it is easy to verify that the portfolios do dominate the American put: the fact that they are all positive and decreasing in t shows that e is positive. The values for $v_{j,t}$ either equal or strictly dominate the immediate exercise value of the American put, $K_t - x_j$ where K_t denotes the present value of the strike price discounted back from time t to time 0.

For the out of the money American put, the dominating portfolio takes a particularly simple form:

$$
\begin{aligned}
& v_{j,t} = \left[K_t - x_j \right]^+ ; \\
(3.3) \qquad & e_{j,t} = \begin{cases} K_t - K_{t+1} & \text{if } e_j \leq K_{t+1}; \\ K_t - x_j & \text{if } K_{t+1} < e_j < K_t; \\ 0 & \text{if } e_j \geq K_t. \end{cases}
\end{aligned}
$$

While this portfolio always dominates the American put, the lower panel of Figure 1 shows that it is not always the cheapest dominating portfolio. To understand how and why it dominates the American put option it is best to consider its counterpart in a continuous time world, and work with nominal rather than discounted prices.

Consider a portfolio that consists of a European put with maturity τ and strike K, and also pays out at the rate rK per unit time so long as the option is in the money ($S_t < K$). This portfolio is the continuous time limit of (3.3). An agent who is long this portfolio and short the American option can ensure positive future cash flow in the following manner (assume for convenience that exercise is physical): the agent does nothing until the option is exercised. In the unlikely event that the option is exercised when it is out of the money, the agent gets positive cash flow from the portfolio, and also positive cash flow from the exercise itself because he can sell the share immediately for more than he has paid for it. If the option is exercised when it is in the money, the agent borrows K and holds the share. The cash flow from the portfolio offsets the interest paid on the debt, whose value remains K so long as the option remains in the money.

If the option stays in the money until maturity, the agent can use the European put option to exchange the share for K and use the cash to repay the debt. If the option goes out of the money prior to maturity, the agent can sell the share and repay the debt immediately, knowing that future cash flows from the portfolio will be positive.

3.2 Extreme Processes

To characterize the processes that lead to the high value of the American option, look first at the simple portfolio (3.3). From Theorem 2.3, the optimal strategy for the holder of the American put is to exercise it when the option is first in the money (all nodes that are in the money are members of X, while all nodes that are out of the money are members of Y. No nodes are in zones W [8]or Z). Another implication of the theorem concerns the extreme process itself. Define the node (j, t) as *deep in the money* if $x_j < K_{t+1}$ define it as *just in the money* if $K_{t+1} \leq x_j \leq K_t$, and as *out of the money* if $x_j > K_t$. Define a path as being *dead* if the option has been exercised. Then Theorem 2.3 says that if the option is deep in the money this period, it will be in the money next period; if it is just in the money this period, it will be out of the money next period; and that if it is out of the money this period, and the path is dead, then it will remain out of the money.

The extreme process is illustrated in Figure 2, where the zone of each node is shown, as is the prohibited transitions for dead paths. It is always optimal to exercise the option as soon as it is in the money. The logic is as follows: if the option is just in the money this

[8] We implicitly assume that none of the nodes is exactly at the money; if this is not the case the discussion becomes rather more complex but the conclusions are substantially unaltered.

period, it will be out of the money and remain out thereafter, so it is worth exercising now while it still has value. But if the option is deep in the money now, it will still be in the money next period, so it is better to exercise the option now and get the strike price immediately rather than wait till next period.

Turning now to the higher strike option, the picture as shown in Figure 3 becomes more complex. There is a large and growing wedge of *Z* zone between the *X* zone on the left where the option is exercised, and the *Y* zone on the right where the option is not exercised. Taking a live path (in the *Y* zone) it can stay in the *Y* zone or jump across the *Z* zone into the *X* zone. It is then exercised. It stays in the *X* zone, unless it moves close to the border with the *Z* zone when it can drift across the border. Once in the *Z* zone, it stays there.

4. PROCESSES AND AMERICAN OPTION VALUE

Table 1 shows that the upper bound on the American premium is a substantial multiple of the actual value computed using conventional valuation methods. The high bounds are actual prices under specific price processes. These findings admit of two (not necessarily incompatible) interpretations: the American premium is substantially underestimated by conventional methods, and the set of theoretically admissible processes is so wide that they generate implausibly high prices.

In this and the following section we attempt to cast some light on the validity of the two interpretations. In this section we specify processes that do generate high American option premia and ask how plausible they are. In the following section the no arbitrage restrictions are tightened by excluding some implausible processes. Plausibility is largely subjective, so the evidence presented is suggestive rather than conclusive.

We start with a postulated process for the asset price and use it to generate European option prices. Three estimates of the value of the American option premium are computed: the value under the given process (Pr^T – the "true" value), the upper bound using the linear program (Pr^B – the bound) and the value if the asset price follows a

Markov diffusion process with a deterministic volatility $\sigma(S_t, t)$ that is consistent with the prices of the European options (Pr^D – the diffusion value).

Table 1 assumes that the true process is a diffusion, so $Pr^D = Pr^T$, and Pr^D is small compared with Pr^B, ($Pr^D < 0.4Pr^B$ in all the cases examined). In general, Pr^D and Pr^T differ. If for a plausible process, the true value Pr^T approaches the upper bound Pr^B and is much higher than the diffusion value Pr^D , this suggests that the size of the American premium does depend heavily on the process. The standard assumption that the asset follows a Markov diffusion would need to be treated with caution. Conversely, if the true value and the diffusion value remain close, this would suggest that the high upper bound simply reflects the weakness of the no arbitrage condition.

A useful starting point for the choice of processes is the extreme processes that support the bounds. Two features of the processes described in section II.2 above are striking: price jumps, and stochastic volatility. The asset jumps down from being deep out of the money (zone *Y*) to deep into the money (zone *X*). It would be interesting to see the extent to which large downward jumps increase the value of American puts relative to European puts.

An alternative interpretation of the extreme process does not necessarily have jumps in it at all. Rather the two state process discussed in Section I.3 can be regarded as a regime switching model, with a high and low volatility state. It is plausible that the American feature is particularly valuable when the holder acquires new information about future volatility. For both jumps and stochastic volatility, we develop simple, if rather extreme models, to explore whether either phenomenon generates a substantially higher American premium than does a Markov diffusion .

4.1 Testing Procedure

Starting with the jump hypothesis, take a jump diffusion process of the form:

$$S_{t+dt} = \begin{cases} Su^n & \text{with probability } \lambda dt \\ S_t u & \text{with probability } \pi_u dt \\ S_t / u & \text{with probability } \pi_d dt \\ S_t & \text{otherwise} \end{cases}$$

(4.1) where $(\pi_u + \pi_d)(\log u)^2 = \sigma^2, \quad u^n = e^J$,

and $\lambda\left(u^n - 1\right) + \pi_u\left(u - 1\right) + \pi_d\left(1/u - 1\right) = 0.$

This process has fixed jump size J with frequency λ, and diffusion with volatility σ. The American put, and also all European call options are valued by taking the expected pay-off to the option under this process. The upper bound on the American put is computed using the linear program LP1.

To compute the value of the American put under a Markov diffusion, first calculate the local implied volatility from the European call prices (using the method of Dupire (1994) and Derman and Kani (1994)). Specifically, a local transition rate $\pi_{j,t}$ is computed:

(4.2)
$$\pi_{j,t} = \frac{C_{j,t+1} - C_{j,t}}{\left(C_{j+1,t} - (1+u)C_{j,t} + uC_{j-1,t}\right)\left(x_{j+1} - x_j\right)}.$$

$C_{j,t}$ is the price of a European call with strike x_j and expiry t under the jump diffusion process. The American option is valued using a diffusion as in equation (3.1), but with the transition rate varying according to the node. This process generates European option values that approximate those from the jump diffusion model $\{C_{j,t}\}$ as the grid becomes finer, but of course the American option valuation is different.

A similar procedure is used to test whether stochastic volatility is important. The stochastic volatility model is designed to be simple and to represent an extreme example of substantial new information about future price processes becoming available in good time t be exploited by the holder of the American option. The world is in one of two states: a high volatility state where the asset price follows a diffusion with constant volatility σ_h and a low state with volatility σ_l. The state of the world is revealed immediately after time 0, and it does not change thereafter. The diffusion is modeled in the discrete space world by equation (3.1) with the appropriate volatility. The value of the

American option at time zero is the value in each of the two states weighted together by the state probabilities. We then follow exactly the same procedure as for the jump process.

4.2 Results - The Jump Process

For the jump diffusion process, we take J = -0.1, λ = 0.5, and σ = 7.07%. The standard deviation of annual continuously compounded returns is therefore 10%, as with the diffusion process used previously. On average jumps occur 0.5 times a year; when they occur the asset price declines by about 9.5%; in the absence of jumps the asset price diffuses with an annualized volatility of 7.07%. The jump and the diffusion contribute equally to the overall volatility of returns.

The parameters are chosen to represent a process that is rather jumpier than would be observed in many markets. Bates (1996) estimates a deterministic volatility jump diffusion process from the prices of $/DM currency options with 38% of the volatility coming from jumps, jump size roughly symmetrical around zero with a standard deviation of 7.7%, and jump frequency of 0.76/year.

Table II presents the value of the American option under the jump diffusion process, the bounds on the value of the American option, and the value of the asset price followed a pure diffusion supporting exactly the same European option values.

Table II: Value of an American Put when European Option Prices support a Jump Diffusion

Strike	**American Prices**			**European Prices**	**Relative prices**	
(K)	**Bound (1)**	**Jump (2)**	**Diffusion (3)**	**$P^E(K, \leq T)$ (4)**	**Jump (5)**	**Diffusion (6)**
r = 2%						
95	1.74	1.55	1.52	1.46	32%	21%
100	3.63	3.24	3.16	3.01	37%	25%
105	6.63	5.97	5.86	5.48	42%	33%
110	10.60	10.00	10.00	10.00	0%	0%
r = 5%						
95	1.43	1.13	1.06	0.92	42%	29%

100	3.24	2.56	2.42	1.99	46%	35%
105	6.35	5.15	5.00	5.00	11%	0%
110	10.47	10.00	10.00	10.00	0%	0%
r = 10%						
95	1.01	0.71	0.62	0.40	50%	36%
100	2.60	1.81	1.64	1.09	48%	36%
105	5.97	5.00	5.00	5.00	0%	0%
110	10.32	10.00	10.00	10.00	0%	0%

The table shows the upper bound on the value of an American put with maturity of 1 year for various levels of strike and interest rates, when the spot price of the underlying is 100, and European options are trading as if the asset price process is a jump diffusion with jump size of -0.1, jump frequency of 0.5/year, and diffusion volatility of 7.07%. The next column shows the price of the American put assuming the asset follows the jump-diffusion process. Column (3) shows the price assuming the asset price follows a pure diffusion process which is consistent with the same European put prices. Column (4) shows the price of the highest valued European option with the same strike and maturities of up to 1 year. The last two columns show the American premia (columns (2) and (3) respectively less column (4)) as a proportion of the maximum premium (column (1) less column (4)). Values are calculated using 200 steps/year, and a stock price mesh of 1%, with exercise being possible every 20 steps.

Comparing the cells of Tables I and II, the effect of the jumps on option prices is quite complex, and varies according to maturity and strike. But it is possible to draw several conclusions: first, although the jump process has some impact on the prices of all options, European or American, the level of the American premium under a diffusion process as shown by column (5) is not greatly affected by the change in the shape of the European implied volatility surface in going from Table I to Table II. Second, given the European option prices, American puts are up to 50% more valuable under the jump process than they are under a pure diffusion process. Third, despite the size and frequency of the jumps, jumps go only a part of the way to capturing the types of process that yield very high American option values. Taking a crude arithmetic average across the strikes and interest rates, the value of the American premium under the Markov diffusion process averages 18% of its maximum possible value, about the same as in Table I. Under the jump diffusion process, this ratio is 26%.

The assumption that jumps are all downwards is crucial. Jumps that are symmetrically distributed around zero have only a small effect on American option prices, while jumps that are predominantly positive make the American premium lower than it would be under pure diffusion. This example should be seen as providing an upper bound on the impact of jumps on the American premium.

4.3 Results - The Stochastic Volatility Process

For the stochastic volatility process, we take σ_l = 5%, σ_h = 15%, and $\Pr\{\sigma = \sigma_h\}$ = 37.5%. This means that the asset price follows a pure diffusion process. Immediately after time 0, it is revealed that the volatility, which thereafter remains constant, is either 5% and 15%. The mixing probability is chosen to ensure that unconditional standard deviation of annual continuously compounded returns is 10%, as with the diffusion process used previously.

The model is chosen to represent an extreme form of stochastic volatility where a lot of new information about volatility is revealed very early in the life of the option. By comparison, Bates (1995) in estimating a stochastic volatility model from the prices of $/DM currency options, has an unconditional distribution for the squared volatility with a coefficient of variation of 1.14 and a half-life of innovations in volatility of just over 6 months; the process used in this paper has a coefficient of variation of 1.50 and volatility shocks are permanent.

Table III presents the value of the American option under this process, the bounds on the value of the American option, and the value of the asset price followed a pure diffusion supporting exactly the same European option values.

Table III: Value of an American Put when European Option Prices support a Stochastic Volatility Process

Strike	**American Prices**			**European Prices**	**Relative prices**	
(*K*)	**Bound (1)**	**S Vol (2)**	**Diffusion (3)**	**$P^E(K, \leq T)$ (4)**	**S Vol (5)**	**Diffusion (6)**
r = 2%						
95	1.43	1.26	1.26	1.22	17%	19%
100	3.23	2.75	2.72	2.57	26%	22%
105	6.60	6.10	5.74	5.24	63%	37%

110	10.88	10.52	10.10	10.00	59%	11%
r = 5%						
95	1.20	0.91	0.92	0.82	23%	26%
100	2.87	2.05	2.00	1.64	33%	29%
105	6.38	5.69	5.00	5.00	50%	0%
110	10.75	10.14	10.00	10.00	18%	0%
r = 10%						
95	0.90	0.57	0.58	0.43	30%	32%
100	2.27	1.38	1.35	0.88	36%	34%
105	6.07	5.24	5.00	5.00	22%	0%
110	10.57	10.00	10.00	10.00	0%	0%

The table shows the upper bound on the value of an American put with maturity of 1 year for various levels of strike and interest rates, when the spot price of the underlying is 100, and European options are trading as if the asset price process is a pure diffusion with a volatility of 5% or 15%. The next column shows the price of the American put assuming the asset follows the mixed diffusion process. Column (3) shows the price assuming the asset price follows a pure diffusion process which is consistent with the same European put prices. Column (4) shows the price of the highest valued European option with the same strike and maturities of up to 1 year. The last two columns show the American premia (columns (2) and (3) respectively less column (4)) as a proportion of the maximum premium (column (1) less column (4)). Values are calculated using 200 steps/year, and a stock price mesh of 1%, with exercise being possible every 20 steps.

Table III shows that American puts are generally more valuable when the true volatility is revealed immediately, with the major gain occurring for options that are deeper in the money. The revelation of the true volatility does go some way to capturing the types of process that yield very high American option values. The value of being American under the pure diffusion is 17% of its maximum possible value, about the same as in Table I. Under the jump diffusion process, this ratio is 31%.

The analysis in this section suggests that asymmetric jumps and stochastic volatility can have a significant impact on the American option premium, and to that extent traditional approaches to valuing American options may be subject to significant error. However, even when the chosen processes that are quite extreme – large, very asymmetric jumps, or high degrees of volatility uncertainty that are largely resolved shortly after the American option is bought – the prices do not approach the rational bounds.

5. TIGHTENING THE BOUNDS

The analysis in the previous section suggests that the high bounds may be due in part to allowing extreme and implausible processes. In this section we explore the possibility of generating tighter but still robust bounds by excluding some implausible processes.

The "good-deal" literature starting with Cochrane and Saá-Requejo (2000) and Bernardo and Ledoit (2000) is motivated by a similar desire to find restrictions on pricing in incomplete markets that are tighter than no arbitrage. They assume that the objective process for the underlying is known, and impose restrictions on the behavior of the stochastic discount factor. Given our set-up, it is more natural to focus on the robust hedging strategy, and impose restrictions on the evolution of the prices of the European options that are used for constructing a dominating portfolio.

The strategy that dominates the American option, as described in section I, involves buying a specific (convex, positive) European portfolio. When the American option is exercised, the European portfolio is liquidated. The strategy is conservative in that it allows for the possibility that the liquidation value of the European portfolio is no more than its intrinsic value. In other words, the strategy allows for the possibility that at some future time the long option position held as a hedge will be trading on an implied volatility of zero.

While implied volatilities do vary widely over time, they never seem to approach zero. For example, the VIX index, which is traded on the Chicago Board Options Exchange, measures the implied volatility of an at-the-money option on the S&P 500 Index with 30 days to expiry. Since its introduction in 1993 to end 2006, it has averaged 19% and reached a minimum of 8.6%. Inspection of the time series of the VIX index in Figure 4 suggests strongly that it should be possible to nominate some non-trivial floor level for the implied volatility of a European portfolio over its lifetime without a large sacrifice of robustness.

Avellaneda, Levy, and Paras (1995) follow a related approach when they set bounds on the volatility of the underlying price process to bound option prices, but there is a subtle distinction. Since we do not directly restrict the path of the underlying asset we do not exclude the possibility that the realized volatility of the price path will be very low. What we do stipulate however that there is a floor on future implied volatility.

To implement this idea, we stipulate that the probability of the price moving from its current level must exceed some critical level $\underline{\lambda}$ per unit time. With the minimum total return on a price change in the discrete space world being of size u this corresponds to a floor volatility of $\sqrt{\underline{\lambda}}\ln(u)$. This restriction can readily by incorporated within the linear program. Recall the matrix $H = \{h_{j,m,t}\}$, which represents the probability of being at node j at time t, at node m at time t+1 and the option having been exercised by time t. The minimum volatility constraint, to be imposed in LP2, can be written as:

(5.1) $$h_{j,j,t} \leq \left(1-\underline{\lambda}\delta\right)\sum_{m} h_{j,m,t} \quad \text{for all } j,m,t,$$

where δ is the length of the time period.

In terms of the dual problem (LP1) this introduces a new positive variable $Q = \{q_{j,t}\}$ into the third constraint:

(5.2) $$3)' \quad e_{j,t} + \left(x_m - x_j\right)d_{j,t} + \left(\underline{\lambda}\delta - I_{m\neq j}\right)q_{j,t} - v_{j,t} + v_{m,t+1} \geq 0,$$
where $I_{m\neq j}$ is an indicator function.

The interpretation of the constraint is that in addition to delta hedging by going long $d_{j,t}$ units of the underlying, the agent can write $q_{j,t}$ (> 0) instantaneous variance contracts. These contracts pay 1 if $S_{t+1} \neq S_t$ and zero otherwise. The constraint in (5.2) specifies that the price of an instantaneous variance contract is greater than or equal to $\underline{\lambda}\delta$. Implied variance (at least on one period contracts, but in fact on all convex claims) never falls below $\underline{\sigma}^2 = \underline{\lambda}\left(\ln u\right)^2$.

The effects of imposing a floor on implied volatility are shown in Table IV. The parameters are the same as in Table I, so in particular all European options trade on an implied volatility of 10%, and the floor on the future implied volatility of options, $\underline{\sigma}$, is 5%.

Table IV: Upper Bound on the Value of an American Put with a Floor Level of Implied Volatility

Strike	***Upper Bound***		***American***	***European***	***Ratio***	***Ratio***
	Not restricted	***Min implied vol of 5%***	***value with diffusion***	**P^E(*K*,≤*T*)**	$\frac{(2)-(4)}{(1)-(4)}$	$\frac{(3)-(4)}{(1)-(4)}$
(*K*)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
r = 2%						
95	1.64	1.50	1.40	1.35	50%	18%
100	3.71	3.48	3.20	3.03	66%	24%
105	6.79	6.66	6.09	5.67	88%	37%
110	10.69	10.65	10.06	10.00	93%	9%
r = 5%						
95	1.28	1.09	0.90	0.77	62%	25%
100	3.31	3.01	2.38	1.92	78%	33%
105	6.39	6.26	5.23	5.00	91%	17%
110	10.31	10.28	10.00	10.00	90%	0%
r = 10%						
95	0.75	0.62	0.43	0.27	73%	33%
100	2.65	2.35	1.54	0.99	82%	34%
105	5.74	5.61	5.00	5.00	83%	0%
110	10.00	10.00	10.00	10.00	-	-

The table shows the upper bound on the value of an American put with maturity of 1 year for various levels of strike and interest rates, when the spot price of the underlying is 100, and European options are trading on an implied annual volatility of 10%. The first column is the unrestricted upper bound, while the second assumes that European options will never trade in future on an implied volatility of less than 5%. The third column shows the price of the American put assuming the asset follows the diffusion process; and the fourth shows the price of the most valuable European option with the same strike and maturity of up to 1 year. The "Ratio" columns standardize columns (2) and (3). Values are calculated using 10 exercise dates and a stock price mesh of 1%.

We have put a lower bound on the implied volatility of European options equal to half the current level of implied volatility. This does tighten the bounds substantially, reducing the maximum American premium by around 20% on average, with the effect being most pronounced for low strike options. But even so, the value of the American premium assuming a Markov diffusion process is less than half the upper bound in all cases, suggesting that the width of the robust bounds can only be attributed in part to allowing implausible price processes.

6. CONCLUSIONS

The value of being American depends on the amount of information about the process that is revealed over the life of the option. In the standard Markov process, the only information revealed is the asset price itself. In a richer model, much more information may be revealed about the probabilities of future price paths, and this information may greatly increase the value of the right to determine the date of exercise.

In order to understand better how much information about future returns could affect the American option premium, we have shown how to determine the upper bound on the value of an American option, and demonstrated how to construct a hedging strategy that enforces those bounds. The model applies to a variety of Bermudan and American structures. The key restrictions or assumptions is that there is just one underlying asset with a known yield or cash dividend stream, and that interest rates are non-stochastic.

By applying the model to the pricing of an American put, we have demonstrated that the potential value of the American feature may be several times larger than is obtained from standard models. The price processes that yield these extreme values tend to assign as much volatility as possible to states of the world where the option is live, where the American option has not yet been exercised, and low volatility to states where the option will have been exercised. While we have been unable to identify plausible processes that generate American option values that are close to the rational bounds, we have seen that downward price jumps and, rather more significantly, information about future volatility do go some way to justify substantial increases in the value of being American.

We have also shown that the very high American premia that are possible without allowing arbitrage are unlikely to materialize in practice since they would allow near arbitrages that would fail only if options were to trade at some future date at very low implied volatilities.

The strategies that enforce the bounds on the American option price have the useful property that they strictly limit the potential downside. The loss is limited to the difference between the rational bound on the American option and the price at which it is written. Whether there are better hedging strategies for American options using European options that are robust to changes in the surface of implied volatilities is a matter for future research.

REFERENCES

Andersen, L., and J. Andreasen, 2001, "Factor Dependence of Bermudan Swaptions: fact or fiction?", *Journal of Financial Economics*, 62(1), 3–37.

Andersen, L., and M. Broadie, 2004, "Primal-Dual Simulation Algorithm for Pricing Multidimensional American Options", *Management Science*, 50(9), 1222–1234.

Avellaneda, M., A. Levy, and A. Paras, 1995, "Pricing and hedging derivative securities in markets with uncertain volatilities", *Applied Mathematical Finance*, 2, 73–88.

Bates, D., 1996, "Jumps and Stochastic Volatility: Exchange Rate Processes implicit in Deutsche-Mark Options", *Review of Financial Studies*, 9(1), 69–107.

Bernardo, A.E., and O. Ledoit, 2000, "Gain, Loss, and Asset Pricing", *Journal of Political Economy*, 108(1), 144–172.

Britten-Jones, M., and A. Neuberger, 2000, "Option Prices, Implied Price Processes and Stochastic Volatility", *Journal of Finance*, 55(2), 839-866.

Broadie, M., and J. B. Detemple, 2004, "Option Pricing: Valuation Models and Applications" , *Management Science*, 50(9), 1145–1177.

Brown, H., Hobson, D. and L.C.G. Rogers, 2001, "Robust Hedging of Barrier Options", *Mathematical Finance*, 11, 285–314.

Cochrane, J. H., and J. Saá-Requejo, 2000, "Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets", *Journal of Political Economy*, 108(1), 79-119.

Davis, M. H. A., and D. G. Hobson, 2007, "The Range of Traded Option Prices", *Mathematical Finance*, 17(1), 1-14.

Dupire, B., 1994, "Pricing with a smile", *Risk*, 7, 18-20.

Derman, E., and I. Kani, 1994, "Riding on a Smile", *Risk*, 7, 32-39.

XEl Karoui, N., and M. C. Quenez, 1995, "Dynamic Programming and the Pricing of Contingent Claims in an Incomplete Market", *SIAM Journal of Control and Optimization*, 33, 29-66.

Green T, C., and S. Figlewski, 1999, "Market Risk and Model Risk for a Financial Institution Writing Options", *Journal of Finance*, 54, 1465-1499.

Haugh, M., and L. Kogan, 2004, "Pricing American Options: A Duality Approach", *Operations Research,* 52(2), 258-270.

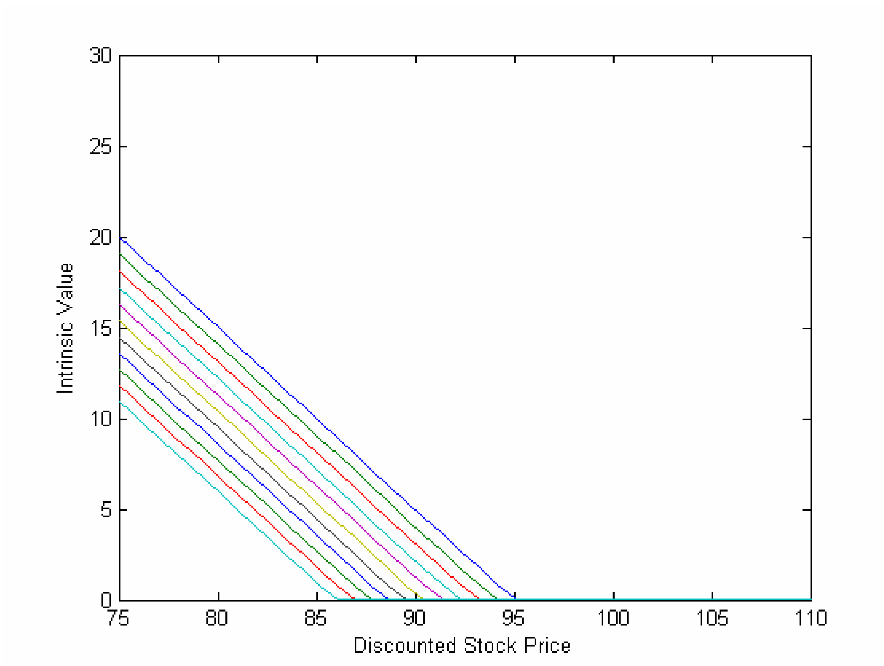
Longstaff, F., P. Santa-Clara and E. Schwartz, 2001, "Throwing away a billion dollars: the cost of suboptimal exercise strategies in the swaption market", *Journal of Financial Economics*, 62(1), 39-66.

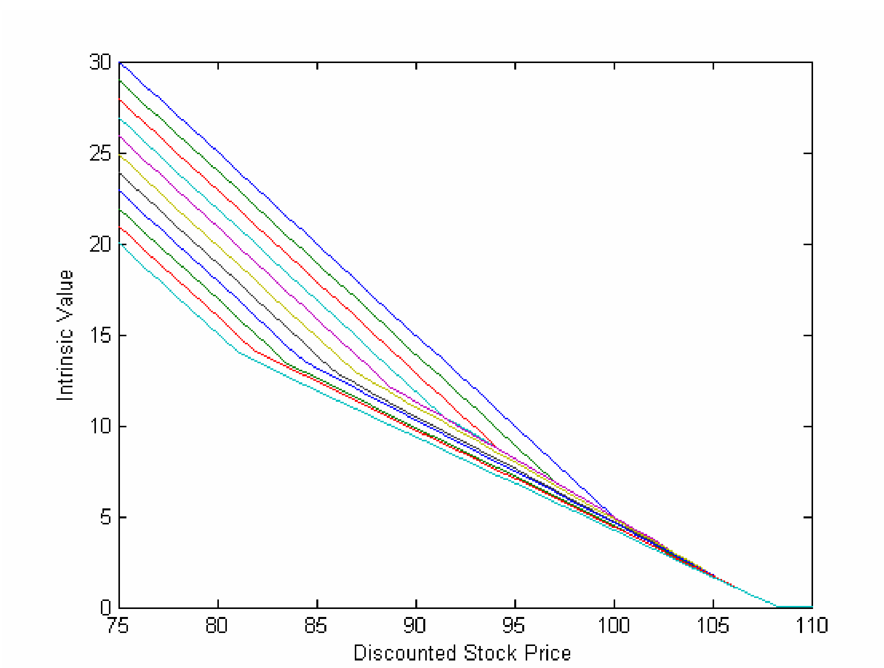
Merton, R. C., 1973, "The Theory of Rational Option Pricing", *Bell Journal of Economics*, 4, 141-183.

Rogers, L. C. G., 2002, "Monte Carlo Valuation of American options", *Mathematical Finance,* 12 (3), 271-286.

Svenstrup, M., 2005, "On the suboptimality of single-factor exercise strategies for Bermudan swaptions", *Journal of Financial Economics*, 78(3), 651-684.

Figure 1: The Cheapest Dominating Portfolios for an American Put

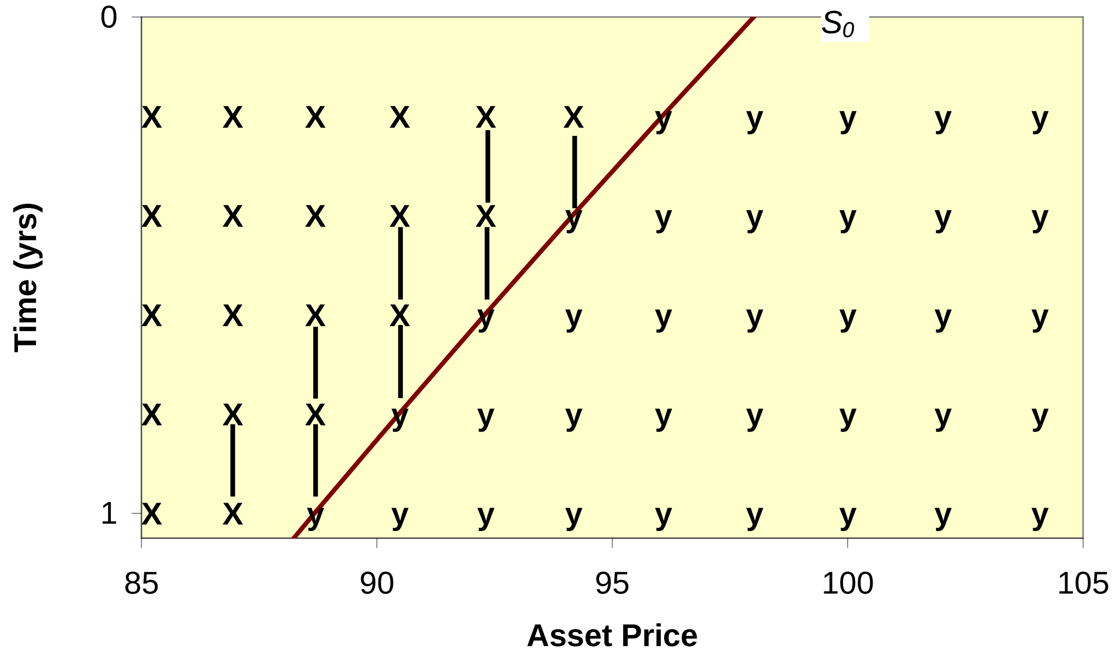




The two charts show the European options that dominate an American put option. In the top chart the strike price of the American option is 95, while in the bottom it

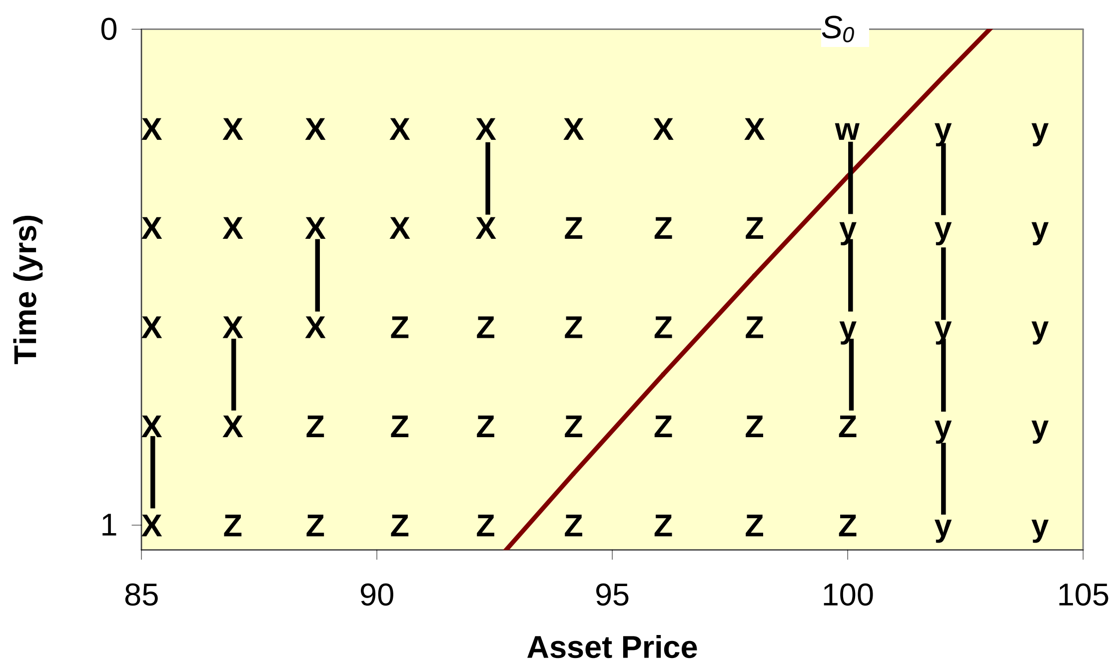
is 105. The difference between successive lines shows the cash flow that period as a function of the asset price. The other parameters are: initial asset price 100, risk free rate 10%, option maturity 1 year, European options have an implied annual volatility (using a binomial model with a mesh of 1%) of 10%.

Figure 2: Bounding Process for Out-of-the-money Put



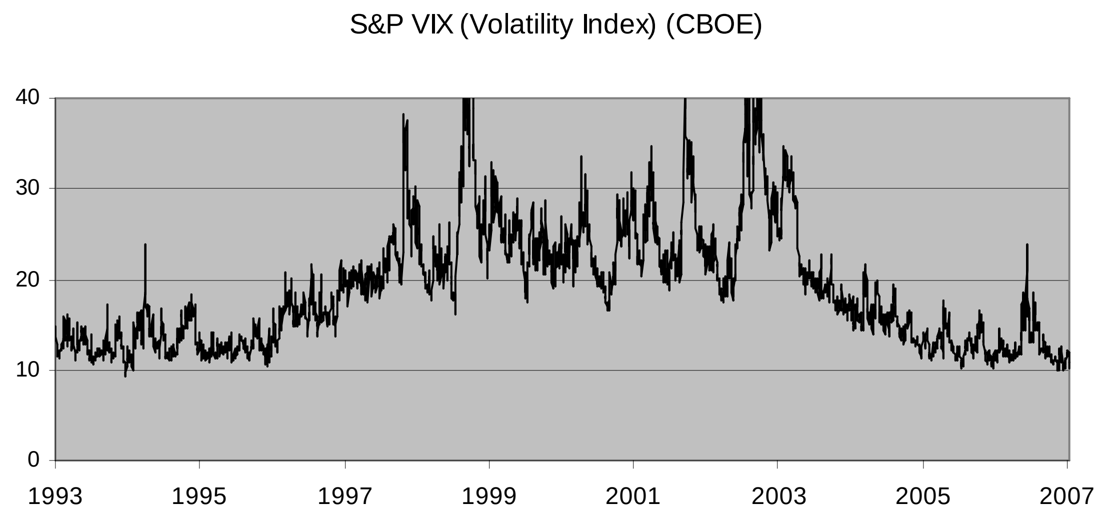
Each node is marked by a letter. The *X* shows that the put option must be exercised at that node, the *y* that the put option must not be exercised at that node. The oblique line is the present value of the exercise price. The short bold lines block transitions between nodes that cannot occur once the option has been optimally exercised. The parameters are as in the top panel of Figure 1: S = 100, K = 98, σ = 10%, r = 10%, T = 1 year.

Figure 3: Bounding Process for In-the-money Put



Each node is marked by a letter. The *X* shows that the put option must be exercised at that node, the *y* that the put option must not be exercised at that node, the *Z* that all paths going through that node have already been exercised. The oblique line is the present value of the exercise price. The short bold lines block transitions between nodes that cannot occur once the option has been optimally exercised. The parameters are as in the bottom panel of Figure 1: *S* = 100, *K* = 103, σ = 10%, *r* = 10%, *T* = 1 year.

Figure 4: The dynamics of implied volatility



The VIX index is the implied volatility of a synthetic at-the-money 30 day option on the S&P500 index, and it is traded on the Chicago Board Options Exchange. It is measured in annualized percentage points. Data collected from Yahoo Finance.

APPENDIX

Proof of Theorem 2.2:

Rewrite the problem LP2 as:

Find $\{\mathbf{F},\mathbf{G},\mathbf{H}\} \geq \mathbf{0}$ that maximizes $\sum_{j,t} a_{j,t} f_{j,t}$ subject to:

1) $\sum_m \left(g_{j,m,t} + h_{j,m,t}\right) \geq p_{j,t}$; 2) $-\sum_m \left(g_{j,m,t} + h_{j,m,t}\right) \geq -p_{j,t}$;

3) $\sum_m \left(g_{m,j,t} + h_{m,j,t}\right) \geq p_{j,t+1}\ (t < T)$; 4) $-\sum_m \left(g_{m,j,t} + h_{m,j,t}\right) \geq -p_{j,t+1}\ (t < T)$;

5) $\sum_m \left(x_m - x_j\right) g_{j,m,t} \geq 0$; 6) $-\sum_m \left(x_m - x_j\right) g_{j,m,t} \geq 0$;

7) $\sum_m \left(x_m - x_j\right) h_{j,m,t} \geq 0$; 8) $-\sum_m \left(x_m - x_j\right) h_{j,m,t} \geq 0$;

9) $f_{j,t} - \sum_m \left(h_{j,m,t} - h_{m,j,t-1}\right) \geq 0$; 10) $-f_{j,t} + \sum_m \left(h_{j,m,t} - h_{m,j,t-1}\right) \geq 0$;

for $j = 1,\ldots,J$ and $t = 1,\ldots,T$, with $h_{m,j,t} \equiv 0$ when $t = 0$.

Compare this with the linear program in standard form:

LP3: Find $\mathbf{x} \geq \mathbf{0}$ that maximizes $z = \mathbf{c}^{\mathbf{T}}.\mathbf{x}$ subject to $\mathbf{A}.\mathbf{x} \leq \mathbf{b}$, where $\mathbf{x}$, $\mathbf{c}$ are n x 1 vectors, $\mathbf{b}$ is m x 1, and $\mathbf{A}$ is m x n.

It can be seen that n, the number of variables, is $JT + 2J^2T$, while m, the number of constraints, is $10JT - 2J$. The vector of variables **x** consists of three blocks: the elements of **F**, then the elements of **G**, and finally the elements of **H**. The vector **c** consists of the elements of **A**, followed by $2J^2T$ zeros. The vector **b** consists of the elements of **P**, followed by –**P**, then **P** again (minus its first row) and the negative of this, followed by $6JT$ zeros. **A** is not readily describable in detail, but its general structure is illustrated below where the blocks of zeros are identified by a minus sign, and the blocks with some non-zero elements are denoted by a plus sign:

	JT	J^2T	J^2T
JT	+	-	+
JT	+	-	+
JT	-	-	+
JT	-	-	+
JT	-	+	-
JT	-	+	-
$JT-J$	-	+	+
$JT-J$	-	+	+
JT	-	+	+
JT	-	+	+

The dual of LP3 is:

LP4: Find $\mathbf{y} \geq \mathbf{0}$ that minimizes $w = \mathbf{b}^\mathbf{T}.\mathbf{y}$ subject to $\mathbf{A}^\mathbf{T}.\mathbf{y} \geq \mathbf{c}.$

If we break **y** into 10 blocks so $\mathbf{y} = (\mathbf{y^1, y^2, \ldots, y^{10}})$ then LP4 can be written as:

LP5:

Find $\left\{\mathbf{y^1,...,y^{10}}\right\} \geq \mathbf{0}$ that minimizes $\sum_{j,t}\left(y^1_{j,t} - y^2_{j,t} + y^3_{j,t} - y^4_{j,t}\right) p_{j,t}$ subject to:

1) $\left(y^9_{j,t} - y^{10}_{j,t}\right) \geq a_{j,t};$

2) $\left(y^1_{j,t} - y^2_{j,t} + y^3_{j,t} - y^4_{j,t}\right) + \left(x_m - x_j\right)\left(y^5_{j,t} - y^6_{j,t}\right) \geq 0;$

3) $\left(y^1_{j,t} - y^2_{j,t} + y^3_{j,t} - y^4_{j,t}\right) + \left(x_m - x_j\right)\left(y^7_{j,t} - y^8_{j,t}\right) + \left(y^{10}_{j,t} - y^9_{j,t}\right) - \left(y^{10}_{m,t+1} - y^9_{m,t+1}\right) \geq 0;$

where $y^3_{j,t} = y^4_{j,t} = 0$ when $t = 1$ and $y^9_{m,t} = y^{10}_{m,t} = 0$ when $t = T+1$,

for $j, m = 1, \ldots, J$ and $t = 1, \ldots, T$.

The bracketing shows where terms are always associated. Define the four $J \times T$ matrices **C, D, V** and **E** as:

$$\mathbf{C} \equiv \mathbf{y}^5 - \mathbf{y}^6;$$
$$\mathbf{D} \equiv \mathbf{y}^7 - \mathbf{y}^8;$$
$$\mathbf{V} \equiv \mathbf{y}^9 - \mathbf{y}^{10};$$
$$\mathbf{E} \equiv \mathbf{y}^1 - \mathbf{y}^2 + \mathbf{y}^3 - \mathbf{y}^4.$$

Then LP5 becomes:

LP6:

Find $\{\mathbf{C},\mathbf{D},\mathbf{V},\mathbf{E}\}$ that minimizes $\sum_{j,t} e_{j,t} p_{j,t}$ subject to:

1) $v_{j,t} \geq a_{j,t}$;

2) $e_{j,t} + \left(x_m - x_j\right) c_{j,t} \geq 0$;

3) $e_{j,t} + \left(x_m - x_j\right) d_{j,t} - v_{j,t} + v_{m,t+1} \geq 0$;

where $v_{m,t} \equiv 0$ when $t = T+1$,

for $j, m = 1,\ldots,J$ and $t = 1,\ldots,T$.

Consider constraint (2). By setting $m = j$ it is seen to imply that $\mathbf{E} \geq \mathbf{0}$. But that is all that (2) implies, because this is the only place in the LP6 where **C** is mentioned, so provided $\mathbf{E} \geq \mathbf{0}$, we can ensure that (2) is satisfied by setting $\mathbf{C} = \mathbf{0}$. The solution to LP6 remains unchanged if constraint (2) is replaced by $e_{j,t} \geq 0$, and **C** is dropped from the list of variables. But then the problem is identical to LP1.