Highlights for ‘The Implication of Missing the Optimal-Exercise Time of an American Option’

1. Upper bounds on the loss from delaying exercise of American options are found.
2. No knowledge of the optimal policy or price is required for computing bounds.
3. The bound is valid for both put and call options.
4. The bound is valid for a variety of underlying asset price dynamics.
5. We examine the impact of investor risk preferences on value lost using our bounds.
The Implication of Missing the Optimal-Exercise Time of an American Option

Arun Chockalingam\textsuperscript{a,}\textsuperscript{∗}, Haolin Feng\textsuperscript{b}

\textsuperscript{a}Department of Industrial Engineering, Eindhoven University of Technology, Eindhoven 5600MB, The Netherlands
\textsuperscript{b}Lingnan College, Sun Yat-sen University, Guangzhou, 510275, China

Abstract

The optimal-exercise policy of an American option dictates when the option should be exercised. In this paper, we consider the implications of missing the optimal exercise time of an American option. For the put option, this means holding the option until it is deeper in-the-money when the optimal decision would have been to exercise instead. We derive an upper bound on the maximum possible loss incurred by such an option holder. This upper bound requires no knowledge of the optimal-exercise policy or true price function. This upper bound is a function of only the option-holder’s exercise strategy and the intrinsic value of the option. We show that this result holds true for both put and call options under a variety of market models ranging from the simple Black-Scholes model to complex stochastic-volatility jump-diffusion models. Numerical illustrations of this result are provided. We then use this result to study numerically how the cost of delaying exercise varies across market models and call and put options. We also use this result as a tool to numerically investigate the relation between an option-holder’s risk-preference levels and the maximum possible loss he may incur when adopting a target-payoff policy that is a function of his risk-preference level.

Keywords: Finance, American options, Delaying exercise, Exercise policy, Suboptimal exercise policy, Free boundary problem

1. Introduction

The American put option specifies a strike price $K$, an expiry time $T$, and affords the holder of the option the right to sell the underlying asset the option is written on for $K$ at any time until and including $T$. Its European counterpart allows the holder to exercise the option only at $T$. At the time of exercise, the payoff of the option is the positive part of the difference between the strike price $K$ and the current asset price (i.e., $K$ less the asset price). The price of the option at any time is the maximum expected discounted payoff of the option over the remainder of its life.

To receive the maximum discounted payoff possible, the option holder needs to exercise the option optimally. Knowledge of the option price is essential to decide if the option should be exercised. If the immediate payoff exceeds the current price of the option, the holder should exercise the option, holding on to it otherwise. The optimal-exercise time is the first time that the immediate payoff, also known as the intrinsic value of the option, is equal to the price of the option. Essentially, there exists an optimal-exercise policy that dictates when the option holder should exercise the option.

The problem of pricing the American option is closely related to the determination of this optimal-exercise policy. In fact, the option pricing problem gives rise to a free-boundary problem in partial (integro) differential equations (PIDEs). Free-boundary problems are a class of problems in which the domain over which the PIDE is to be solved is not known \textit{a priori} and needs to be solved for simultaneously with the
solution to the PIDE. The free boundary in the option pricing problem is the optimal-exercise policy. The solution to the PIDE is the option price function which describes the option price for a range of asset prices and times to expiry. Thus, solving the free-boundary problem yields the option price function and the optimal-exercise policy as this policy is the boundary of the domain over which the PIDE is solved.

The problem of pricing these options has received significant attention in literature. Under the assumption of constant volatility, Black and Scholes (1973) derives the celebrated Black-Scholes equation to compute the price of a European call option. The put-call parity can then be used to compute the price of the corresponding European put. Closed-form solutions for the price of European options have also been derived for market models which overcome the shortcomings of the Black-Scholes model, see for example, Heston (1993) and Madan et al. (1998). Closed-form solutions for European option prices may be computed in these cases because the exercise time is known with certainty.

Similar solutions do not exist for American options however, even under the Black-Scholes model (except for an American call option written on a stock that does not pay any dividends). The difficulty in computing such a solution arises as a result of the early-exercise feature of the American option. The lack of a closed-form solution has led to substantial research in developing approximations and computational schemes to price American options under a variety of market models. These schemes utilize simulation or numerical schemes such as finite difference and finite-element methods. Simulation-based methods typically compute the price of an option for a single time to expiry and asset price pair. The option price is then compared with the immediate payoff to decide if the option should be exercised (see for example Longstaff and Schwartz (2001) and Jin et al. (2013). Schemes based on finite difference and finite-element methods typically compute the entire price function. The optimal exercise policy can then be obtained either directly or as a post-processing step if the free-boundary problem was not directly solved. See, for example, Pressacco et al. (2008), for an overview and comparison of finite-difference and lattice-based methods.

Empirical evidence suggests that suboptimal exercise of American options occurs frequently in the markets. Diz and Finucane (1993) studies exercise decisions in the S&P 100 market and concludes that many exercise decisions are inefficient (suboptimal). Bauer et al. (2009) examines the impact of options trading on individual investor performance and find that investors incur significant losses on their options investment. These losses are indicative of suboptimal exercise of the options. Pool et al. (2008) also finds that investors have incurred significant losses over a ten year period by exercising call options suboptimally (including allowing the option to expire when it should have been exercised during its lifetime). Barraclough and Whaley (2012) studies exercise decisions on put options and also finds that a large number of options remain unexercised when they should have been, and that these suboptimal-exercise decisions have yielded significant losses to put option holders over a twelve-year period.

Suboptimal exercise of options is attributed to two main causes in literature, the first being model misspecification, which, as the name implies, refers to the incorrect specification of the underlying market model (for example assuming constant volatility when volatility is in fact stochastic). Model misspecification results in investors incorrectly pricing options, obtaining an incorrect exercise policy and consequently making a suboptimal-exercise decision. Longstaff et al. (2001) studies the effect of using single-factor models to make exercise decisions on swaptions when the underlying term structure is actually driven by several factors. The authors conclude that the total possible present value of the costs of following single factor strategies could be several billion dollars, even when the single factor is re-calibrated frequently. Chockalingam and Muthuraman (2011) demonstrates that pricing American put options assuming a constant volatility when the volatility of the underlying asset price is stochastic leads to mispricing of the American option, even when the constant volatility is set equal to the mean level of the stochastic process describing the evolution of volatility. As mentioned before, this would lead to misidentification of the optimal-exercise policy value and therefore suboptimal-exercise decisions. Irrational investor behavior is cited as the second cause for suboptimal-exercise decisions of options. Classical option pricing theory computes option prices and exercise policies assuming that investors are rational agents. Empirical studies, however, have demonstrated that investors do indeed exercise options irrationally. Overdahl and Martin (1994) and Poteshman and Serbin (2003) find evidence of this in exchange traded call options. Engstrom (2002) studies exercise decisions on call options in the Swedish stock market and finds evidence of the same. Finucane (1997) also studies empirically exercise decisions on call options, comparing these results to results obtained under the assumption
that there is no friction present in the market. The author finds that 20% of exercises are not optimal, and that even if a large number of these irrational trades can be explained by the presence of transaction costs, a significant number of exercises still appear to be irrational.

The literature on characterizing the cost of suboptimal exercise is relatively scarce. In Ibáñez and Paraskevopoulos (2010), the authors study the sensitivity of the American option to suboptimal-exercise policies by considering policies that advance exercise, and those that delay exercise. In the put option case, the advancing (delaying) exercise corresponds to the exercise boundary lying above (below) the optimal-exercise boundary. They find that the cost of suboptimal exercise is a function of the gamma of the option at the optimal-exercise boundary and the bias in the suboptimal-exercise policy (difference between the optimal and sub-optimal boundaries).

In this paper, we consider the cost of delaying exercise of an American option. As mentioned above, in the put option case this corresponds to exercise policies that dictate the holder should hold on to the option till it is deeper in the money when the optimal action would have been to exercise the option. Naturally, the optimal action, along with the optimal-exercise policy, is unknown to us. Such suboptimal policies delay exercise of the option. This situation is illustrated below in Figure 1 for the classic Black-Scholes model. In the figure, the stopping region (where the option should be optimally exercised is shaded. The optimal-exercise policy dictates that the option should be exercised at time $t_1$ when the asset price strikes the policy from the above). The suboptimal-exercise policy prescribes waiting till time $t_2$ to exercise the option. While the structure of the exercise policies will change when the underlying asset price is assumed to be modeled by different processes, this concept of delaying exercise will remain unchanged throughout the paper.

By decomposing the price of the American put option into the price of the corresponding European put and an early exercise premium, Carr et al. (1992) derive an expression for the delayed-exercise premium, i.e., the additional benefit gained by delaying exercise till the optimal-exercise time. This perspective, and consequently, the delayed-exercise premium, differ significantly from our approach and the upper bound we compute. In our case, we consider delaying the exercise of the put option past the optimal-exercise time. As such, the premium from Carr et al. (1992) will not coincide with the upper bound on the cost of delaying exercise that we compute later in the paper.

Ibáñez and Paraskevopoulos (2010) notes that computing the cost of delaying exercise is more difficult than computing its advancing exercise counterpart as the cost of delaying exercise depends on the entire suboptimal-exercise policy whilst the cost of advancing exercise depends only on the value of the suboptimal-exercise policy at the time of exercise. The central contribution of this paper is a theoretical result which provides an upper bound for the cost of delaying exercise. Even though this paper and Ibáñez and Paraskevopoulos (2010) consider the suboptimal exercise of American options, the works differ significantly. For one, Ibáñez and Paraskevopoulos (2010) consider the sensitivity of American option prices to suboptimal exercise and shows that the amount an investor may lose by suboptimally exercising an American option is a function of the gamma of the option. In this paper, we derive tight upper bounds on the value an investor holding an American option could lose. More importantly, the cost estimate presented in Ibáñez and Paraskevopoulos (2010) is a function of the unknown optimal-exercise policy. In this paper, the upper bound on the cost is solely a function of the suboptimal-exercise policy used by an investor and does not require knowledge of the optimal-exercise policy.

The paper is organized as follows. Section 2 describes the dynamics of the price of the underlying asset and the free-boundary problem that the option pricing problem gives rise to. The main theoretical result is proved in Section 3. We illustrate the upper bounds in relation to the cost of delaying exercise for a variety of scenarios and make a number of observations about these bounds in Section 4. We conclude in Section 5. Appendix A presents the proof of the maximum principle presented in Section 3. Appendix B discusses the application of the main result to call options and compares the maximum possible loss associated with put and call options.
2. Free-boundary Problem Formulation

As mentioned above, the problem of pricing an American option gives rise to a free-boundary problem in PIDEs, a class of problems where the domain over which the PIDE is solved is not known a priori and has to be solved for as part of the solution. In the context of American put options, the lower boundary (free boundary) of this domain describes the optimal-exercise policy. Any point on or below the free boundary belongs to the optimal stopping region. The main result of this paper is derived from the free-boundary problem formulation. As such, in this section, we describe in detail the free-boundary problem.

Consider an American put option with strike price $K \in \mathbb{R}^+$ and expiry time $T \in \mathbb{R}^+$. Let $X(t)$ denote the price of the underlying asset at time $t \in [0, T]$. Assuming that the evolution of the volatility process is described by the Heston (1993) model, under the risk-neutral measure, asset prices evolve according to the dynamics

$$\frac{dX(t)}{X(t)} = (r - \delta - \lambda \xi)dt + \sqrt{Y(t)}dW(t) + (\eta - 1)dQ(t),$$

where

- $r > 0$ is the risk-free rate of return,
- $\delta \geq 0$ is the continuous dividend yield,
- $Q(t)$ is an independent Poisson process with intensity $\lambda$, and can be approximated by
  $$dQ(t) = \begin{cases} 
0 & \text{with probability } 1 - \lambda dt \\
1 & \text{with probability } \lambda dt,
\end{cases}$$

$\eta - 1$ is an impulse function producing a jump from $X(t) = x$ to $x \cdot \eta$, where $\eta$ is a positive random variable with density function $\nu$. 

![Figure 1: Illustration of delaying exercise](image)
\[ \xi = \mathbb{E}^\nu[\eta - 1], \text{ with } \mathbb{E}^\nu[\cdot] \text{ being the expectation operator with respect to } \nu, \]

the instantaneous volatility at time \( t \), \( \sqrt{Y(t)} \) is derived from the stochastic process \( Y \) and \( W(t) \) is a standard Brownian motion.

Under the Heston (1993) stochastic-volatility model, the stochastic process \( Y \) is modeled by a square-root process of the form
\[
dY(t) = \kappa(m' - Y(t))dt + \nu \sqrt{Y(t)}dZ(t),
\]
where \( \kappa \) represents the speed of mean reversion of \( Y(t) \) to its long-term mean \( m' \), \( \nu \) represents the volatility of volatility and \( Z(t) \) is another standard Brownian motion correlated with \( W(t) \). We assume a constant correlation \( \rho \in [-1, 1] \), i.e., \( dW(t) \cdot dZ(t) = \rho dt \).

A variety of models, besides the Heston (1993) model have been proposed in the literature to capture the evolution of volatility in the market, including those presented in (Stein and Stein 1991, Scott 1987, Hull and White 1987). Similar to Heston (1993), these papers model the evolution of the volatility term with a second diffusion process and a mapping from the realization of the diffusion process to the instantaneous volatility term. Papers such as Cheang et al. (2013) and Hanson and Yan (2007) consider American put option pricing when the evolution of the price of the underlying asset is governed by a stochastic-volatility jump-diffusion model, with the stochastic volatility component being modeled after the Heston (1993) model. We have chosen to focus on the Heston (1993) model in this paper for ease of exposition. The obtained theoretical results are, however, readily extendable and applicable to these other models as well. Other models of stochastic volatility exist in the literature, such as the Generalized AutoRegressive conditional Heteroskedasticity (GARCH) model proposed in Bollerslev (1986). Our focus in this paper is on models that utilize a second diffusion process to model the evolution of volatility and as such, other stochastic volatility models are beyond the scope of this paper.

To state the free-boundary problem concisely, we first define notation. Let \( b \) represent the optimal-exercise policy, \( C = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 | x > b(\tau, y)\} \) the optimal continuation region, and \( S = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 | x \leq b(\tau, y)\} \) the optimal stopping region where \( \tau = T - t \) represents the time to expiry, \( X(\tau) = x \) and \( Y(\tau) = y \). In the Black-Scholes setting \( b \) would only be a function of \( \tau \), making it a curve (as shown above in Figure 1) as opposed to a surface, as is the case here (illustrated in Figure 2). Delaying exercise in this situation would refer to an exercise surface that lies in the stopping region prescribed by \( b \).

We denote by \( p(\tau, x, y) \) the price of the American put option when the time till expiry of the option is \( \tau \), the underlying asset price is \( x \), and the volatility of volatility is \( y \). The function \( p \) then represents the price function of the American put and solves the following PIDE
\[
\frac{1}{2} y x^2 p_{xx} + \rho \nu x y p_{xy} + \frac{1}{2} \nu^2 y^2 p_{yy} + (r - \delta - \lambda \xi) x p_x + (\kappa(m' - y) - \Lambda(\tau, x, y)) p_y - (r + \lambda) p + p_\tau + \lambda \int_0^\infty p(\tau, x, y, \eta) \nu(\eta)d\eta = 0
\]
(3)

(with subscripts referring to partial derivatives) in the continuation region as shown in Cheang et al. (2013). Note that even though Cheang et al. (2013) consider American call options, the PIDE that the price function solves depends only on the dynamics governing the evolution of the underlying asset price and not on the nature of the option. As such, the put and call price functions solve the same PIDE in their respective continuation regions. The function \( \Lambda(\tau, x, y) \) in Equation (3) represents the compensation an investor requires in order to bear the additional risk that arises from fluctuations in volatility, known as the market price of volatility risk. We make the natural assumption that \( \Lambda(\tau, x, y) = 0 \) at \( \tau = 0 \), i.e., investors require no compensation when the volatility of volatility \( y = 0 \). The choice of using a function to represent the market price of volatility risk allows us to capture various choices for the form of the market price of volatility risk. For example, \( \Lambda(\tau, x, y) \) directly corresponds to \( \lambda \nu \) in Equation (20) of Cheang et al. (2013) and the proportional risk premium \( \lambda v \) considered in Heston (1993). We point out here that our results do not depend on the specification of the function \( \Lambda \). Bakshi (2003) show empirically that the market price of volatility
risk is negative, implying that investors are willing to pay a premium to bear this additional risk so as to receive protection from downside risk during times of increased volatility in the market.

For notational simplicity, for any given function $f$ that is appropriate for the operations, we define the following:

$$E_{\nu}[f(\tau, x \cdot \eta, y)] \equiv \int_{0}^{\infty} f(\tau, x \cdot \eta, y) \nu(\eta) d\eta,$$

$$A f \equiv \frac{1}{2} \rho y x f_{xx} + \rho y x f_{xy} + \frac{1}{2} \nu^{2} y f_{yy} + (r - \delta - \lambda \xi) x f_{x} + (\kappa(m' - y) - \Lambda(x, y, \tau)) f_{y} - (r + \lambda) f - f_{\tau}.$$

With this notation, we now define the free-boundary problem. The option price function $p$, together with the optimal-exercise policy $b$, solves

$$A p + \lambda E_{\nu}[p(\tau, x \cdot \eta, y)] = 0 \quad \text{in} \ C,$$

$$p(\tau, x, y) = (K - x)^{+} \quad \text{in} \ S,$$

$$p(0, x, y) = (K - x)^{+},$$

$$\lim_{x \to \infty} p(\tau, x, y) = 0,$$

$$\lim_{y \to \infty} p_{y}(\tau, x, y) = 0,$$

$$(r - \delta - \lambda \xi) x p_{x} + \kappa m' p_{y} - (r + \lambda) p - \tau + \lambda E_{\nu}[p(\tau, x \cdot \eta, y)] = 0 \quad \text{at} \ y = 0,$$

$$p_{x}(\tau, x, y) = -1 \quad \text{as} \ x \downarrow b, \quad \text{and}$$

$$p(\tau, x, y) \geq (K - x)^{+} \quad \forall (\tau, x, y) \in C \cup S,$$

The optimal-exercise policy $b$ and price function $p$ together uniquely solve Equations (4) - (11). Explanations of the various boundary conditions can be found in standard texts and references on American option pricing.
To find the optimal-exercise policy, one needs to solve the free-boundary problem defined by Equations (4) - (11).

**Fixed-boundary problem.** If the exercise policy is specified, one is left with a fixed-boundary problem. Suppose we are given an exercise policy (not necessarily optimal) represented by a continuous $c(\tau, y)$ such that $c(\tau, y) > 0 \ \forall (\tau, y) \in (0, T] \times \mathbb{R}_+$. The exercise rule for this policy would be similar to that of the optimal-exercise policy. Namely, at any time $\tau$, hold the option if $x > c(\tau, y)$ and exercise otherwise. Then, the solution to the following fixed-boundary problem is the value function $v$ associated with the exercise policy defined by $c$:

$$Av + \lambda E^v [v(\tau, x \cdot \eta, y)] = 0 \quad \text{in } \hat{C} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 | x > c(\tau, y)\},$$  

$$(r - \delta - \lambda \xi) v(x, y) + \kappa m v_y - (r + \lambda) v - v_\tau + \lambda E^v [v(\tau, x \cdot \eta, y)] = 0 \quad \text{at } y = 0,$$  

With the the option pricing problem defined and the fixed-boundary problem established, we present the main result of the paper in the next section.

3. Cost of Delaying Exercise

We first state Theorem 1, which establishes the maximum principle that is needed for subsequent proofs. Theorem 1 shows that the solution to the fixed-boundary problem will achieve its global maximum in the stopping region and its global minimum at $\tau = 0$ provided it is non-negative in the stopping region.

**Theorem 1.** For a given $T \in (0, \infty)$ and a continuous $c(\tau, y) > 0$ for all $(\tau, y) \in (0, T] \times \mathbb{R}_+$, let $\hat{C} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 | x > c(\tau, y)\}$ and $\hat{S} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 | x \leq c(\tau, y)\}$. Also, let $g$ be the solution to

$$Ag + \lambda E^v [g(x \cdot \eta, y)] = 0 \quad \text{in } \hat{C},$$

(A defined as above), with boundary conditions given by,

$$g(0, x, y) = 0,$$  

$$g(\tau, x, y) = F(\tau, x, y)^1 \quad \text{in } \hat{S}$$  

$$\lim_{\tau \to \infty} g(\tau, x, y) = 0,$$  

$$\lim_{\tau \to \infty} g_\tau(\tau, x, y) = 0,$$  

$$(r - \delta - \lambda \xi) x v_x + \kappa m v_y - (r + \lambda) v - v_\tau + \lambda E^v [v(\tau, x \cdot \eta, y)] = 0 \quad \text{at } y = 0,$$  

$$G^v_g(y) \overset{df}{=} \max \{g(\tau, x, y) : (\tau, x) \in (0, T] \times \mathbb{R}_+, \ x > c(\tau, y)\} \to 0 \quad \text{as } y \to \infty.$$  

If $r > 0$ and $F(\tau, x, y) \geq 0$ in $\hat{S}$, then

$$\max_{(\tau, x, y) \in \hat{C} \cup \hat{S}} g(\tau, x, y) = \max_{(\tau, x, y) \in \hat{S}} g(\tau, x, y) \equiv \max_{(\tau, x, y) \in \hat{S}} F(\tau, x, y),$$  

and

$$\min_{(\tau, x, y) \in \hat{C} \cup \hat{S}} g(\tau, x, y) = g(0, x, y) \equiv 0.$$

$^1F$ is a dummy function.
Proof. See Appendix A.

**Theorem 2.** For a given \( T \in (0, \infty) \) and a continuous \( c(\tau, y) > 0 \) for all \((\tau, y) \in (0, T] \times \mathbb{R}_+\), let \( \mathcal{C} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 \mid x > c(\tau, y)\} \). Also, let \( g \) be the solution to the problem posed in Equations (18) - (24).

If \( r > 0 \) and \( F(\tau, x, y) = 0 \) in \( \hat{S} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 \mid x \leq c(\tau, y)\} \), then \( g(\tau, x, y) = 0 \) in \( \hat{C} \).

**Proof.** Applying Theorem 1 with \( F(\tau, x, y) = 0 \) in \( \hat{S} \) yields the conclusion.

As explained in section 2, if an exercise policy is specified by a continuous boundary \( c \) such that \( c(\tau, y) > 0 \) for all \((\tau, y) \in (0, T] \times \mathbb{R}_+\), then the value function \( v \) associated with such a policy could be found by solving the fixed-boundary problem posed in Equations (12) - (17). As before, let \( b \) denote the boundary corresponding to the optimal-exercise policy and \( p \) denote the true price function associated with \( b \). A delaying exercise policy could then be determined by a boundary \( c \) such that \( c(\tau, y) \leq b(\tau, y) \) for all \((\tau, y) \in (0, T] \times \mathbb{R}_+\). The main goal of this paper is to study the cost of delaying exercise. Recall that this cost is defined as the difference between the value function associated with the optimal exercise policy \( b \) and that of the delaying exercise policy \( c \).

Mathematically, we denote this cost as \( g \equiv p - v \) by definition. Then, it is expected that \( \mathcal{G}_i^T \) satisfies Equation (24), which implies that within the optimal continuation region, the difference between \( p \) and \( v \) is negligible (since the probability of exercising the option diminishes at high levels of volatility). Theorem 3 then gives a tight upper bound for \( g \), the cost arising from delaying exercise.

**Theorem 3.** Suppose \( c \) is a continuous, non-increasing boundary such that \( c(\tau, y) \leq b(\tau, y) \) for all \( \tau \in [0, T] \) and \( y \in \mathbb{R}_+ \), where \( b \) is the optimal exercise policy with price function \( p \). Let \( v \) be the value function associated with \( c \). If \( (K - x)^+ - v \leq \epsilon \) in the region \( \hat{C} \cup \hat{B} \), then we have \( 0 \leq p - v \leq \epsilon \), where \( \hat{C} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 \mid x > c(\tau, y)\} \) and \( \hat{B} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 \mid x = c(\tau, y)\} \).

**Proof.** We define \( \bar{g}(\tau, x, y) = p(\tau, x, y) - v(\tau, x, y) \). Owing to the definition of an optimal-exercise policy, clearly \( \bar{g} \geq 0 \). The rest of the proof serves to establish that \( \bar{g} \leq \epsilon \).

Since the price function \( p \) solves (4) - (9) while \( v \) solves (12) - (17), \( \bar{g} \) will solve the following

\[
A\bar{g} + \lambda E^\nu[\bar{g}(x \cdot \eta, y)] = 0 \quad \text{in} \quad \hat{C} \cap \bar{C} = \mathcal{C},
\]

\[
\bar{g}(\tau, x, y) = (K - x)^+ - v(\tau, x, y) \quad \text{def} \quad F(\tau, x, y) \quad \text{in} \quad \mathcal{S} = \{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2 \mid x \leq b(\tau, y)\},
\]

\[
\bar{g}(0, x, y) = 0,
\]

\[
\lim_{x \to -\infty} \bar{g}(\tau, x, y) = 0,
\]

\[
\lim_{y \to -\infty} \bar{g}(\tau, x, y) = 0,
\]

\[
(K - \lambda)\bar{g}_x + k^p \bar{g}_y = (r + \lambda)\bar{g}_\tau - \bar{g}_\tau + \lambda E^\nu[\bar{g}(x \cdot \eta, y)] = 0 \quad \text{at} \quad y = 0.
\]

In the stopping region \( \mathcal{S} \), \( p(\tau, x, y) = (K - x)^+ \) by definition. Thus, \( \bar{g}(\tau, x, y) = p(\tau, x, y) - v(\tau, x, y) = (K - x)^+ - v(\tau, x, y) \leq \epsilon \) in this region. Applying Theorem 1 to \( \bar{g} \), we have

\[
\max_{(\tau, x, y) \in \mathcal{C} \cup \mathcal{S}} \bar{g}(\tau, x, y) = \max_{(\tau, x, y) \in \mathcal{S}} F(\tau, x, y) \leq \epsilon.
\]

This ends the proof of the theorem.

Given that the optimal-exercise policy is seldom known, Theorem 3 allows us to compute the cost of delaying exercise with knowledge of only the option-holder’s exercise policy, and not the optimal-exercise policy. With the option-holder’s exercise policy, we may compute the the value function, \( v \), associated with the policy, and as a consequence of Theorem 3, compute the upper bound on the cost of delaying exercise as \( \sup_{\tau, x, y} (K - x)^+ - v(\tau, x, y) \).
Remark 1. The maximum principle holds true for call options as well (when modified to accommodate the appropriate boundary conditions). As such, with some minor adjustments to Theorem 3, we arrive at a similar result for call options. A numerical illustration of the application of Theorem 3 to call options is presented in Appendix B.

With Theorem 3, the option holder may compute the cost of delaying exercise over the life of the option. Thus at any time during the life of the option, the holder can use this theorem to determine the maximum possible loss of delaying exercise of the option. After some time has elapsed from the inception of the option, this maximum still holds true but will overstate the loss that the holder may factually face. Naturally, as the time to expiry of the option decreases, the cost of delaying exercise will decrease as well. The following theorem, which gives us a tighter upper bound on the cost of delaying exercise, can be used to determine the maximum loss possible from delaying exercise of the option over the remainder of its life.

Theorem 4. Suppose \( c \) is a continuous, non-increasing boundary such that \( c(\tau, y) \leq v(\tau, y) \) for all \( \tau \in (0, T] \) and \( y \in \mathbb{R}_+ \), where \( b \) is the optimal-exercise boundary with price function \( p \). Let \( v \) be the value function associated with \( c \). If \( (K - x)^+ - v \leq c_{\tau'} \) in the region \( \mathcal{C}_{\tau'} \cup \mathcal{B}_{\tau'} \), then we have \( 0 \leq c_{\tau'} - v \leq c_{\tau'} \) for all \( (\tau, x, y) \in (0, \tau'] \times \mathbb{R}_+^2 \), where \( \mathcal{C}_{\tau'} = \{ (\tau, x, y) \in (0, \tau'] \times \mathbb{R}_+^2 \mid x > c(\tau, y) \} \) and \( \mathcal{B}_{\tau'} = \{ (\tau, x, y) \in [0, \tau'] \times \mathbb{R}_+^2 \mid x = c(\tau, y) \} \).

Proof. Replacing \( T, \epsilon, \hat{C} \) and \( \hat{B} \) with \( \tau', \epsilon_{\tau'}, \hat{C}_{\tau'} \) and \( \hat{B}_{\tau'} \), respectively and proceeding as we did in Theorem 3 will yield the result.

Remark 2. As a consequence of Theorem 4, when the time to expiry is equal to \( \tau' \), the cost of delaying exercise \( p(\tau', x, y) - v(\tau', x, y) \) is bounded by \( \sup_{(\tau, x, y) \in (0, \tau'] \times \mathbb{R}_+^2} (K - x)^+ - v(\tau, x, y) \) which is no greater than \( \sup_{(\tau, x, y) \in (0, T] \times \mathbb{R}_+^2} (K - x)^+ - v(\tau, x, y) \). The implication of this observation is that as the life of an option increases (i.e., as \( T \) increases), so too does the cost of delaying exercise. This implication can be explained by considering two options with time to expiry \( T \) and \( \tau' \), with \( \tau' < T \) and identical otherwise. If the holder of the option expiring at \( T \) suboptimally delays the exercise of the option until \( \tau' \) and continues to follow a suboptimal policy thereafter, the holder incurs a cost associated with the time value of the option, as well as the cost of following a suboptimal policy after \( \tau' \). The cost of following the same suboptimal policy for the option expiring at \( \tau' \) can therefore be no greater than the cost associated with the option expiring at \( T \). This result is also illustrated numerically in Figure 3.

4. Numerical Illustrations and Investigations

In this section, we briefly illustrate the applications of Theorem 3 and 4 to various scenarios. Recall that the results apply to both put and call options for the SVJD model. The results naturally apply to the Black-Scholes model, jump-diffusion model (geometric Brownian motion with jumps) and stochastic volatility models as well as these are specific cases of the SVJD model. We provide illustrations to highlight these bounds as follows. We first present illustrations of the upper bound for American put options under the Black-Scholes and SVJD models (the simplest and the most complex of the models we have covered, respectively).

4.1. Upper bounds for the Black-Scholes model

Figure 3 depicts upper bounds obtained under 4 scenarios for the American put option when assuming that the market under the Black-Scholes model. Each row in the figure depicts a scenario, with the parameters for the scenario listed on the right. The common parameters for all the scenarios are the strike price, \( K = 10 \), risk-free rate \( r = 5\% \) and volatility \( \sigma = 20\% \). The parameters that vary across the scenarios are the time to expiry \( T \) (either 6 months, 1 year or 3 years) and the continuous dividend yield \( \delta \) (either 0 or 8\%).

The left figure in each row shows the optimal-exercise policy for that particular scenario (dashed line) and an exercise policy that delays exercise of the option. The optimal policies \( b \) are computed using the
Figure 3: Exercise policies and upper bounds on the cost of delaying exercise under the Black-Scholes model for an American put option
moving-boundary approach described in Muthuraman (2008). The delaying exercise policies \( c \) (solid lines) are taken to be \( c(\tau, x) = b(\tau, x) - 0.5\tau \). The figure on the right of each row shows the difference between the true option price function and the value function associated with the suboptimal-exercise policy at two different times during the life of the option (at \( T \) and 0.57). The dotted horizontal lines represent the upper bound obtained as a consequence of Theorem 3. As the figure shows, this upper bound corresponds to the maximum difference between the true price function and the suboptimal value function. Note that these bounds were obtained without knowledge of the true price function but only with the intrinsic value and the value function associated with the suboptimal-exercise policy.

Looking at Figure 3 we also observe that, as highlighted by Remark 2, the upper bound increases as the life of the option increases (i.e. as \( T \) increases).

Table 1: Comparison of upper bound from Theorem 3 and true cost of suboptimally delaying exercise

<table>
<thead>
<tr>
<th>( T ) = 3 years, ( \delta = 0 )</th>
<th>( T ) = 1 year, ( \delta = 0 )</th>
<th>( T ) = 0.5 years, ( \delta = 0 )</th>
<th>( T ) = 3 years, ( \delta = 8% )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi )</td>
<td>True*</td>
<td>U. bound*</td>
<td>( \psi )</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0166</td>
<td>0.0152</td>
<td>0.1</td>
</tr>
<tr>
<td>0.02</td>
<td>0.1102</td>
<td>0.1102</td>
<td>0.0867</td>
</tr>
<tr>
<td>0.03</td>
<td>0.2311</td>
<td>0.2311</td>
<td>1.6333</td>
</tr>
<tr>
<td>0.04</td>
<td>0.3511</td>
<td>0.3511</td>
<td>2.4</td>
</tr>
<tr>
<td>0.05</td>
<td>0.4591</td>
<td>0.4591</td>
<td>3.1667</td>
</tr>
<tr>
<td>0.06</td>
<td>0.5517</td>
<td>0.5517</td>
<td>3.9333</td>
</tr>
<tr>
<td>0.07</td>
<td>0.6296</td>
<td>0.6296</td>
<td>4.7</td>
</tr>
<tr>
<td>0.08</td>
<td>0.695</td>
<td>0.695</td>
<td>5.4667</td>
</tr>
<tr>
<td>0.09</td>
<td>0.7499</td>
<td>0.7499</td>
<td>6.2333</td>
</tr>
</tbody>
</table>

*Maximum difference between true option price function and value function associated with suboptimal-exercise policy

b Upper bound obtained using Theorem 3

Figure 3 considers delaying exercise policies of the form \( c(\tau, x) = b(\tau, x) - 0.5\tau \) for illustrative purposes. In Table 1, we consider generic suboptimal-exercise policies taking the form \( c(\tau, x) = b(\tau, x) - \psi \tau \) for various values of \( \psi \) and compare the upper bound obtained as a consequence of Theorem 3 and the maximum difference between the true option price function and the value function associated with each of these suboptimal-exercise policies. Looking at the table, we find that Theorem 3 always yields upper bounds in alignment with the true cost of suboptimally delaying exercise. Also note that the parameter \( \psi \) essentially controls the difference or gap between the optimal exercise-policy and suboptimal exercise-policy. Table 1 shows that as \( \psi \) increases, so too does the upper bound on the cost of delayed exercise, verifying the insights obtained from the result in Nuñez and Paraskevopoulos (2010), that the cost of delaying exercise depends on the difference between the two policies.

4.2. Upper bounds for the SVJD model

We now illustrate the upper bounds on the cost of delaying exercise under SVJD models. In particular, we assume that the stochastic-volatility component of the model is represented by the Heston (1993) model and that jumps in the asset price are lognormally distributed, as in Merton (1976) with parameters \( \mu_\eta \) and \( \sigma_\eta \), i.e.,

\[
\nu(\eta) = \frac{1}{\eta \sigma_\eta \sqrt{2\pi}} e^{-(\ln(\eta) - \mu_\eta)^2/2\sigma_\eta^2}.
\]

The parameters of the SVJD model for this illustration are as follows: \( K = 10, \ T = 3 \) years, \( r = 5\% \), \( \delta = 0 \) for the dividend yield, \( \kappa = 3 \), \( m_\eta = 0.04 \), \( \psi = 0.9 \), \( \rho = 0.1 \), \( \Lambda = 0 \), \( \lambda = 0.1 \), \( \mu_\eta = 0.6 \) and \( \sigma_\eta = 0.45 \). The optimal-exercise policies and true price functions were computed by modifying the procedure presented in Chockalingam and Muthuraman (2011) to accommodate for jumps. Under the Black-Scholes model, the delaying exercise policy was chosen by modifying the optimal-exercise policy along the time direction. Under stochastic volatility, the optimal policy may now be modified along two directions, the time direction and the volatility direction. Accordingly, in our illustration, we consider two scenarios. Figure 4 depicts the first scenario.
Figure 4: Exercise policies and upper bounds under the SVJD model when changing the optimal policy along $\tau$. 

$y = 0$

$y = 1$
The leftmost figure shows both the optimal (top) and delaying exercise policies. As in the Black-Scholes case, the delaying-exercise policy $c$ is taken as $c(\tau, y) = b(\tau, y) - 0.5y$. The difference between the price function and the value function for different levels of $y = \sigma^2$ for two different times to expiry are also shown, with the upper bound obtained using Theorem 3 shown as a horizontal dotted line. From the figure, we observe that the maximum difference occurs at $y = 0$ and $\tau = T$. Consistent with the observation under the Black-Scholes model, the maximum cost is attained at $\tau = T$. The explanation for the maximum being attained at $y = 0$ is as follows; the difference between the two policies is constant across the $(\tau, y)$ plane. We see that the optimal policy decreases as $y$ increases implying that waiting for the option to go deeper into the money is optimal for large values of $y$. Therefore, for large values of $y$, the effect of delaying exercise is marginal. On the other hand, the effect of delaying exercise for smaller values of $y$ becomes more significant, resulting in a higher cost of delaying exercise. One may also view the distance between the two policies as the instantaneous drop that must occur in the asset price to induce the holder to exercise the option. The probability of such a drop occurring is lower for smaller values of $y$ and therefore results in a larger cost of delaying exercise.

Figure 5 depicts the second scenario, in which the delaying-exercise policy is obtained by modifying the optimal-exercise policy along the volatility direction.

As above, the leftmost figure shows both policies (with the optimal policy being the higher policy). Now, $c$ is taken to be $c(\tau, y) = b(\tau, y) - 0.5y$. Contrary to what was observed in Figure 4, the maximum difference between the price and value functions is attained at $y = 1$ and $\tau = T$. Again, the maximum occurring at $\tau = T$ is consistent with our observations from the Black-Scholes model. This time, however, the maximum occurs at the other end of the $y$-axis. The explanation for this observation is that in this situation the difference between the two policies depends only on $y$. At $y = 0$, the two policies coincide resulting in no delay in the exercise of the option. As $y$ increases, the difference between the policies also increases. And as the cost of delaying exercise is a function of the difference between the two policies, it also increases. Similar to above, if we view the difference between the two policies as the instantaneous drop that must occur in the asset price for the holder to exercise the option, the larger this drop becomes, the lower the probability of it occurring. This results in a larger cost of delaying exercise.

Comparing Figures 3 and 5, we find that the magnitude of the upper bound decreases as more factors are added into the model governing asset price evolution. When looking at the Black-Scholes case for $T = 3$ and $\delta = 0$ for the dividend yield, we find that while the upper bound for delaying exercise exceeds 0.2, under the SVJD model, the upper bound is about 0.16. The probability of an instantaneous drop occurring in the asset price under the Black-Scholes model is zero as the model assumes that the asset price process is continuous. As more factors are added to the model, particularly, the inclusion of a jump component, the probability of such a drop occurring increases, thereby increasing the probability that the option will be exercised and reducing the cost of delaying exercise. This observation is in line with empirical observations showing that model mis-specification results in suboptimal exercise of the option. The observation that the inclusion of a jump component in particular can reduce the cost of delaying exercise suggests that at the very least, a model used to represent the dynamics of the price of the underlying asset should incorporate jumps.

4.3. Tighter upper bounds for the cost of delaying exercise

As mentioned before, the upper bound on the cost of delaying exercise for the life of the option will overstate the possible loss faced by the option holder some time after the inception of the option. This statement is supported by observing Figures 3 - 5. We find that the upper bound on the cost of delaying exercise can be much higher than the maximum difference between $p$ and $v$ at times $\tau < T$. If one is interested in determining the maximum loss from following the policy $c$ at some time $\tau' < T$, the upper bound obtained as a consequence of Theorem 3 overstates the maximum possible loss. A tighter upper bound may be determined using Remark 2. Rather than determining the difference between $(K - x)^+ - v$ over all $\tau \in [0, T]$, one needs to determine the difference between $(K - x)^+ - v(\tau, x)$ only until $\tau'$ in the case of the Black-Scholes model. The upper bound at any time $\tau'$ may thus be found as $\sup_{x, \tau \in [0, \tau']} (K - x)^+ - v(\tau, x)$. Figure 6 shows these tighter bounds for each of the scenarios depicted in Figure 3 for a variety of times $\tau'$. 
Figure 5: Exercise policies and upper bounds under the SVJD model when changing the optimal policy along $y$.
As the figure shows, these tighter upper bounds are equal to the maximum of \( p(\tau', x) - v(\tau', x) \), with the tighter bounds being monotonic in \( \tau \).

![Graph 1](image1.png)

![Graph 2](image2.png)

![Graph 3](image3.png)

**Figure 6:** Tighter upper bounds on the cost of delaying exercise in the Black-Scholes model

In the case of SVJD models, by Remark 2, one needs to determine the difference \((K - x)^+ - v\) over \( \tau \in [0, \tau'] \) and all \((x, y) \in \mathbb{R}_+^2\) to obtain a tighter upper bound. These tighter bounds are illustrated in Figures 7 and 8 when modifying the exercise policy along the \( \tau \) and \( y \) directions respectively. We find that the monotonic nature of these bounds is preserved under the SVJD case. We also find that the tighter bounds are monotonic in \( y \), i.e., in the instantaneous volatility \( \sigma \) of the SVJD model.

**Remark 3.** In each of the scenarios discussed above, an exercise policy that prescribes delaying exercise of the option was chosen by modifying the optimal exercise policy. As pointed out earlier, one may compute a tight error bound without knowledge of the optimal exercise policy and needs only to check that the policy does in fact prescribe delaying exercise. To conduct this check for American put options, one needs to only compute the value function associated with this exercise policy. If the first derivative of this value...
Figure 7: Tighter upper bounds under the SVJD model when changing the optimal policy along $\tau$

Figure 8: Tighter bounds under the SVJD model when changing the optimal policy along $y$
function with respect to the stock price $x$ is less than $-1$ at a stock price just above the exercise policy, i.e., if $v_x(\tau, c(\gamma)) < -1$ for the SV and SVJD models ($v_x(\tau, c(\gamma))$ for the Black-Scholes and Jump-Diffusion models), then the policy prescribes delaying exercise (lies below the optimal-exercise policy). The first derivative being less than $-1$ implies that the associated exercise policy lies below the optimal exercise policy, because at a stock price just above the exercise policy, the value function falls below the option’s intrinsic value $(K - x)^+$. Such a behavior could not occur if the associated exercise policy prescribes premature exercise of the option, i.e., if it lies above the optimal exercise policy. A similar check can be used for American call options.

4.4. Costs related to ‘risk-preference’

In this section, we numerically investigate the cost of adopting a type of simple suboptimal-exercise policies, which we call ‘target payoff’ policies. We suppose that an option holder using a ‘target payoff’ policy will exercise before expiry only when the stock price is no higher than $\gamma \cdot K$, where $\gamma$ represents the holder’s risk-aversion level. Thus, the investor will only exercise the option when the (nominal) payoff is at least equal to a target level $(1 - \gamma) \cdot K$ when exercising the option. Clearly, with everything else fixed, as $\gamma$ decreases, the target payoff increases while the probability of exercising the option to gain the payoff is diminishing. Therefore, the choice of $\gamma$ for an option holder using the ‘target payoff’ policy can be thought of as reflecting the risk preference of the option holder. As such, we call $\gamma$ the risk aversion factor in this target payoff context.

Mathematically, given the strike price $K$, the $\gamma$-target-payoff strategy is determined by the following exercise:

$$c(\tau) = \begin{cases} K & \text{if } \tau = 0 \\ \gamma K & \text{otherwise.} \end{cases}$$  \hspace{1cm} (34)

The results presented thus far have been established for continuous exercise policies. The policy represented by Equation (34) is discontinuous at $T$. In order to apply the established results, we approximate the policy given in Equation (34) by

$$c(\tau) = \begin{cases} K e^{\gamma \frac{\tau}{T}} & \text{for } \tau \leq \hat{\tau} \\ \gamma K & \text{for } \tau > \hat{\tau}, \end{cases}$$  \hspace{1cm} (35)

for some small $\hat{\tau} > 0$, and where $\alpha = -\frac{1}{2} \log \gamma$. Using Equation (35) to represent Equation (34) ensures continuity of the suboptimal exercise policy in the interval $[0, T]$ and allows the use of the results presented in the paper. When the dividend yield is positive, one needs to only replace $K$ in both equations with $\min(K, \frac{y}{\tau})$.

We first explore the interaction between $\gamma$ and the maximum cost incurred when adopting a suboptimal exercise policy that prescribes delaying exercise, with all parameters held constant. We then allow each parameter $(K, T, \sigma, \delta)$ to vary as $\gamma$ varies to study the joint interaction between $\gamma$, the parameters and the maximum cost.

For the first study (with constant parameters), we use the following parameters: $K = 10$, $T = 1$ year, $r = 5\%$, $\sigma = 20\%$ and $\delta = 0$. For a given $\gamma$, we obtain a suboptimal exercise policy given by Equation (35). We can check that such a policy lies below the optimal-exercise policy by using the check discussed in Remark 3. Solving the fixed-boundary problem in Equations (12) - (17) with this exercise policy yields the associated value function. We can then use Theorem 3 to compute the maximum cost incurred by the option holder when adopting the delaying strategy dictated by the policy.

Figure 9 plots the maximum cost obtained using Theorem 3 for a range of risk-aversion levels. Looking at the figure, we find that the maximum cost monotonically decreases as $\gamma$ increases. This is to be expected, since an increasing $\gamma$ results in an exercise policy that prescribes earlier exercise, therefore minimising the penalty incurred due to delaying exercise.

Next, we allow the strike price $K$ to vary and take values in the range $[5, 125]$. Figure 10 plots the maximum cost incurred as $\gamma$ and $K$ vary. We find that the maximum cost is monotonically increasing in $K$ for all levels of risk aversion. The greatest cost is incurred for low levels of $\gamma$ and high levels of $K$. That a
larger cost is incurred for low levels of $\gamma$ was seen above in Figure 9. This observation can be explained by considering the price of the option in conjunction with the option’s strike price. The price of an option is monotonically increasing in the strike price of the option. With lower risk aversion (hence a lower $\gamma$), the option holder will wait significantly longer to exercise the option, thereby increasing the probability of never exercising the option and foregoing the amount paid for the option.

![Figure 9: Maximum cost when all parameters are constant](image)

Figure 9: Maximum cost when all parameters are constant

Figure 10 shows how the maximum cost varies as a function of $\gamma$ and $T$. For this study, $T$ varied from 3 months to 15 years. As above, we find that the maximum cost increases monotonically as $T$ increases, consistent with results from studies earlier in the section. In particular, the maximum cost is higher for lower values of $\gamma$. The reason for this stems from the relation between the price of the option and the time to expiry of the option. Since the price of the option is monotonically increasing in $T$, a lower $\gamma$ leads to an increased probability of the option not being exercised, even though $T$ is larger, therefore increasing the loss faced by the option holder.
An important factor affecting the price and exercise of options is the volatility of the price process of the underlying asset. As such, we allow $\sigma$ to take values between 0 and 100%. Figure 12 shows how the maximum cost varies as $\gamma$ and $\sigma$ vary. Looking at the figure, we find that the maximum cost is decreasing in $\sigma$. As above, the maximum cost is higher for lower values of $\gamma$. The figure shows that delaying exercise is less costly when the volatility of the price process of the underlying asset is higher. A higher volatility implies that there is a higher probability for large downward movements in the price of the underlying asset while the option has not been exercised. Delaying exercise will allow the option holder to take advantage of such a movement in the asset price and lock in a large payoff.

Finally, we allow the dividend yield $\delta$ to vary from 0 to 10%. Figure 13 shows how the maximum cost varies as $\gamma$ and $\delta$ vary. Looking at the figure, we find that while the cost is monotonically increasing in $\delta$, it is barely noticeable in comparison with the decrease in cost as $\gamma$ increases. Contrary to intuition, however, the change in the maximum cost does not significantly vary as $\delta$ increases. Furthermore, the magnitude of
the cost is much lower than in the other cases.

![Figure 13: Maximum cost for various dividend yields $\delta$](image)

5. Conclusion

In this paper, we consider the cost of delaying exercise of the option. Specifically, we consider exercise policies that dictate the option should be held until it is deeper in the money when the optimal action would have been to exercise the option and derive upper bounds on the cost of delaying exercise by following these suboptimal policies. These upper bounds are computed using only the intrinsic value of the option and the value functions associated with the suboptimal-exercise policies and therefore require no knowledge of the optimal-exercise policy to determine the maximum possible loss the holder of the option could face by following this policy. These bounds apply to a variety of market models ranging from the simple Black-Scholes model to the complex stochastic volatility with jump-diffusion models. Under stochastic-volatility with jump-diffusion models, these bounds apply to existing stochastic volatility models described in literature and to jump-diffusion models for which the distribution of the jump amplitude can be described by a probability density function. After showing that the upper bound may be calculated after determining the entire value function for a given suboptimal-exercise policy, we illustrated that tighter upper bounds may in fact be computed by only evaluating the value function until the time of interest for parameter values of interest.

Avenues for future research include extending the result to exercise policies that advance exercise of the American option. Knowledge of an upper bound for scenarios such as these could be used to extend policy iteration algorithms that converge to the optimal-exercise policy for put options from above above and not from below. Naturally, such an extension can be applied to call options where the algorithms will converge to the optimal-exercise policy from below. With knowledge of such bounds, one can also determine which is more costly, advancing or delaying exercise. Another interesting study would be analyzing the sensitivity of these upper bounds to changes in the parameters of the underlying model so as to gain insight about which factors significantly affect optimal exercise of the option. Deriving these results for larger classes of processes such as Lévy and GARCH processes as well as other derivatives which permit early exercise is also another avenue for future direction. In the case of other early exercise derivatives, we conjecture that similar, if not the same, results will hold. The verification of this conjecture is also a possible direction for future research. A comparison of the losses arising in derivatives markets from delaying exercise against these upper bounds would be an illuminating empirical study.
Acknowledgement

We’d like to thank the three anonymous referees and the editors for their feedback. Additionally, Haolin Feng would like to acknowledge the support from the National Natural Science Foundation of China (Grant No. 71401187) and the Fundamental Research Funds for the Central Universities.

Bibliography


Appendix A. Proof of Theorem 1

Proof. Re-arranging terms, Equation (18) becomes

\[ rg + \lambda g - \lambda \mathbb{E}^\nu [g(\tau, x \cdot \eta, y)] = \cdots \]

the payoff function for the put option needs to be replaced with the payoff function for the call option. As

Subsequently, Theorem 3 can also be applied to American call options. In the case of the call option, the Hessian needs to be a negative semidefinite matrix due to the necessary conditions. Hence at \((\tau', x', y')\), the market price of volatility risk, as \(\Lambda\) and suppress its dependence on the variables to simplify notation.

We show that the maxima of \(g\) is attained in the stopping region \(\tilde{\mathcal{S}}\) by showing that there does not exist any \((\tau', x', y') \in \hat{\mathcal{C}}\) such that

\[ g(\tau', x', y') > \max_{(\tau, x, y) \in \tilde{\mathcal{S}}} F(\tau, x, y) \equiv M \geq 0. \]  

(A.2)

We show the above by assuming the contrary. First, say an internal global maxima exists and is attained at some \((\tau', x', y') \in \{ (\tau, x, y) \in (0, T) \times \mathbb{R}_+^2 \mid x > c(\tau, y) \}\) such that \(g(\tau', x', y') > M \geq 0\). By the necessary conditions for an internal maxima, we have that \(\frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} = 0\) and \(\frac{\partial g}{\partial \tau} \geq 0\). Substitution into Equation (A.1) yields

\[ rg + \lambda \{ g - \mathbb{E}^\nu [g(\tau', x' \cdot \eta, y')] \} = \frac{1}{2} \frac{\partial^2 g}{\partial x^2} y \left[ x + \left( \frac{\partial^2 g}{\partial \tau^2} + \frac{\partial^2 g}{\partial x^2} \rho^2 \right) \right] + \frac{1}{2} \left( \frac{\partial^2 g}{\partial \tau^2} + \frac{\partial^2 g}{\partial x^2} \rho^2 \right) \left( \frac{\partial^2 g}{\partial x^2} \rho^2 \right) - \frac{\partial g}{\partial \tau} \]  

(A.3)

The Hessian needs to be a negative semidefinite matrix due to the necessary conditions. Hence at \((\tau', x', y')\), \(\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} - \left( \frac{\partial^2 g}{\partial x \partial y} \right)^2 \geq 0\) and \(\frac{\partial^2 g}{\partial x^2} \leq 0\). Moreover, since \(\rho^2 \leq 1\), \(\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} - \left( \frac{\partial^2 g}{\partial x \partial y} \right)^2 \geq \frac{\partial^2 g}{\partial x^2} \geq 0\) and \(\frac{\partial^2 g}{\partial y^2} \leq 0\). Now, since the maxima is attained at \(g(\tau', x', y')\), we have \(g(\tau', x', y') \geq g(\tau', x' \cdot \eta, y')\) for all possible \(\eta\), resulting in \(g \geq \mathbb{E}^\nu [g(\tau', x' \cdot \eta, y')]\). Hence, from Equation (A.3), we have a positive LHS but a non-positive RHS, implying a contradiction.

Next, assume that the maximum is attained at some \((\tau', x', 0)\) on the boundary \(y = 0\) such that \(g(\tau', x', 0) > M \geq 0\). Again at \((\tau', x', 0)\), by conditions of maxima, we have \(\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y} = 0\) and \(\frac{\partial g}{\partial \tau} \leq 0\). Substitution into Equation (23) gives

\[ rg + \lambda \{ g - \mathbb{E}^\nu [h(\tau', x' \cdot \eta, 0)] \} = \kappa m \frac{\partial g}{\partial y} \]  

(A.4)

which has a positive LHS but a non-positive RHS, leading to a contradiction.

Now, due to Equations (19) and (21), a global maximum \(M \geq 0\) clearly cannot be attained on the boundary \(\tau = 0\) or as \(x \to \infty\). Nor can it be achieved as \(y \to +\infty\), due to Equation (24). This leads to the fact that there does not exist any \((\tau', x', y') \in \hat{\mathcal{C}}\) such that \(g(\tau', x', y') > \max_{(\tau, x, y) \in \mathcal{C} \cup \mathcal{S}} F(\tau, x, y)\), resulting in Equation (25).

Arguments similar to those made above yield Equation (26).

Appendix B. Upper bounds for the American call option

As mentioned in Remark 1, Theorem 1 holds for call options as well (with some minor adjustments). Subsequently, Theorem 3 can also be applied to American call options. In the case of the call option, the payoff function for the put option needs to be replaced with the payoff function for the call option. As
mentioned above, one may compute the upper bound on the cost of delaying exercise for a call option by numerically evaluating the follow expression:

\[ \sup_{\tau, x, y} (x - K^+) - v(\tau, x, y). \]

We briefly illustrate these bounds for the call option under the Black-Scholes model. The optimal-exercise policies for the call option may be computed with a straightforward modification to the moving boundary approaches presented in Muthuraman (2008). Figure B.14 depicts the exercise policies and cost of delaying exercise for two scenarios. The leftmost figure shows the optimal-exercise policy (dashed line) and the policy prescribing delaying exercise (solid line), taken to be \( c(\tau, x) = b(\tau, x) + 0.5\tau \). As we are now dealing with a call option, a policy that prescribes delaying exercise will now lie above the optimal-exercise policy. The figure on the right of each row shows the difference between the true option price function and the value function associated with the sub-optimal exercise policy at two different times during the life of the option (at \( T \) and \( 0.5T \)). The dotted horizontal lines represent the upper bound obtained as a consequence of Theorem 3. As the figure shows, similar to the put option scenarios, this upper bound corresponds to the maximum difference between the true price function and the suboptimal value function.

---

**Appendix B.1. Comparison of the costs of delaying call and put option exercise**

In this subsection, we explore numerically the difference in delaying the exercise of call options against put options. The obtained results provide us with a measure to compare the two types of options and study if delaying the exercise of one type is costlier than delaying the exercise of the other. As Theorem 3 provides us with an upper bound on the cost of delaying exercise, given policies that prescribe delaying exercise of the respective options, we can compute an upper bound on the cost of delaying exercise and compare these. We limit our comparison to the Black-Scholes model as the insights drawn remain the same for all the other models.

Figure B.15 plots the upper bounds on the cost of delaying exercise of both put and call options as a function of the spread between the risk-free rate \( r \) and the continuous dividend yield \( \delta \). Following the previous illustrations, the values of the other relevant parameters are \( T = 3 \) years, \( K = 10 \) and \( \sigma = 20\% \).
The risk-free rate $r$ is fixed at 5%. The dividend yield is allowed to vary from 1% to 15%. We do not consider the case of $\delta = 0$ as in such a situation, exercising an American call option prior to $T$ is never optimal. The optimal-exercise policy would then be infinite for $\tau < T$ and $K$ for $\tau = T$. For each of the scenarios considered, the suboptimal-exercise policies were selected as $c(\tau, x) = b(\tau, x) + 0.5\tau$ for call options and $c(\tau, x) = b(\tau, x) - 0.5\tau$ for put options.

Looking at the figure, we first observe that for negative spreads, the upper bound on the cost of delaying the exercise of put options is close to zero, implying that there is little cost in delaying the exercise of put options on stocks which have dividend yields greater than the risk-free rate. In this situation, the dividends earned by holding on to the stock longer (by not exercising the option) help to almost offset the loss resulting from suboptimal exercise of the option. As the spread increases, we observe that the upper bound increases. This is due to the fact that as the spread increases, delaying exercise of the put option would result in the holder forfeiting the interest he could have earned on the strike price received from exercise of the option. The dividend earned from holding the stock does not offset the loss substantially, causing the upper bound to increase with the spread, somewhat linearly.

On the other hand, we find that as the spread increases, the upper bound on the cost of delaying exercise of a call option decreases towards zero. Due to the increase in the spread, delaying exercise earns interest on the strike price that would have been spent if the option was exercised. At lower spreads, the interest earned helps to offset some of the loss incurred from sub-optimal exercise of the option, while at higher spreads, it helps to offset almost all of the incurred loss.

We also find that the cost of delaying the exercise of a call option does not consistently dominate the cost of delaying the exercise of a put option. For negative spreads, the maximum loss incurred from delaying the exercise of a call option exceeds that of delaying the exercise of a put option. For negative spreads, exercise of the call option would grant the option holder the underlying stock and its dividend payments which exceed the risk-free rate. As such, delaying exercise of the option results in the possible dividend payments being forfeited. For positive spreads, the option holder forfeits the interest that could be earned on the strike price by exercising the option, making the delay of exercising a put option more costly than the delay of exercising a call option for positive spreads.