Portfolios of American Options Under General Preferences: Results and Counterexamples†

Vicky Henderson†‡
Oxford-Man Institute
University of Oxford

Jia Sun §
Warwick Business School
University of Warwick

Elizabeth Whalley ¶
Warwick Business School
University of Warwick

February 1, 2011

Abstract

We consider the optimal exercise of a portfolio of American call options in an incomplete market. Options are written on a single underlying asset but may have different characteristics of strikes, maturities and vesting dates. Our motivation is to model the decision faced by an employee who is granted options periodically on the stock of her company, and who is not permitted to trade this stock.

The first part of our study considers the optimal exercise region of single options. We prove results under minimal assumptions and give several counterexamples where these assumptions fail - describing the shape and nesting properties of the exercise regions.

The second part of the study considers portfolios of options with differing characteristics. The main result is that options with co-monotonic strike, maturity and vesting date should be exercised in order of increasing strike. It is true under weak assumptions on preferences and requires no assumptions on prices. Potentially the exercise ordering result can significantly reduce the complexity of computations in a particular example. This is illustrated by solving the resulting dynamic programming problem in a CARA utility indifference model.

1 Introduction

Our aim in this paper is to provide results concerning the optimal exercise of American options under a minimal set of assumptions on prices, valuation methodology, and the agent’s preferences. The goal is to characterize optimal behavior in terms of exercise ordering for an agent with a portfolio of American calls of differing strikes, maturities and vest dates.

The first part of the paper considers single American options and revisits a classic Cox and Rubinstein [8] result concerning the dependence of the option’s continuation region upon the

†The authors thank participants at seminars at Imperial College London and the A.Ma.Me.F Workshop (Berlin, September 2010) for comments.
‡Oxford-Man Institute, University of Oxford, Eagle House, Walton Well Road, Oxford. OX2 6ED. UK. Email: vicky.henderson@oxford-man.ox.ac.uk
§Finance Group, Warwick Business School, University of Warwick, Coventry. CV4 7AL. UK. Email: Jia.Sun07@phd.wbs.ac.uk
¶Finance Group, Warwick Business School, University of Warwick, Coventry. CV4 7AL. UK. Email: elizabeth.whalley@wbs.ac.uk
strike and option maturity. In particular, a basic result is that if two identical agents each held an American call, the agent with the option with the lower strike/earlier maturity/shorter vest period would exercise first. Put another way, the continuation region for the option with lower strike/earlier maturity/shorter vest period is contained within that of the high strike/later maturity/longer vest period. We call this the “nesting” property. We show this result only requires that agents prefer more to less and prefer cash sooner, and that the valuation mechanism has a monetary property (which we define later). We give several examples where nesting holds, (in particular, in a utility indifference model with CARA) and a counterexample (for which we can perform explicit computations) where nesting fails as the valuation is not monetary. On a similar theme, under the assumption that the price is a diffusion, we give conditions under which a higher dividend rate leads to a smaller continuation region. This generalizes results given for exponential Brownian prices in a utility indifference model by Carpenter, Stanton and Wallace [7].

We give stochastic proofs for each of our results, in some cases using couplings, and we avoid calculations. Indeed, for each of the results described already, we do not need to know anything about the actual shape of the exercise region(s) or boundary (ies) themselves, and as our examples show, these regions may not be simply described by a single threshold separating the continuation and exercise region.

To illustrate how complex exercise boundaries can be, we also study a far more specialized model with risk neutral pricing and diffusion prices. We include these results because some are new and original in their own right, and also they provide a useful comparison to the utility indifference model. We give conditions on the dividend rate such that the exercise region is characterized by a single threshold. A counterexample with state-dependent dividends results in disconnected exercise regions. To obtain a single threshold which is also monotonic in time requires the additional restriction of a time-homogeneous diffusion. This is well known but not always correctly stated (see Kim [26]). A simple counterexample with time dependent volatility violates the result.

The second part of the paper considers to what extent the results for single American options can be extended to a portfolio of options with differing characteristics. In a complete market, or under pricing by risk-neutral expectation, the consideration of a portfolio of different options is no more difficult than that of a single option. In such a situation, options can be treated independently as the presence of other options does not alter the agent’s strategy. However, our interest lies in incomplete markets, for an agent with a non-linear preference structure. Our canonical example of such a market will be the situation of an employee who is restricted from hedging her portfolio of employee stock options and for whom, the exercise strategy for a particular option depends upon the rest of her portfolio.

Our main result is that given a portfolio of American call options with a co-monotonic set of strikes, maturities and vest dates, it is optimal to exercise the options in order of increasing strike. We only require our minimal assumptions on preferences - our agent must prefer more to less and prefer cash sooner - and we do not make any assumptions on the price or on the valuation methodology. The proof relies on a coupling style argument where we consider two agents and show that regardless of the strategy the “sub-optimal” agent follows, the “optimal” agent can do better by exercising in order of increasing strike.

Our primary application of the exercise ordering result is to consider a utility indifference
pricing model for a portfolio of American options. An introduction to the large literature on utility indifference pricing can be found in Carmona (ed) [5], in particular the survey of Henderson and Hobson [20]. The portfolio exercise problem results in a mixed optimal control and multiple stopping problem reflecting the choices over investment and exercise of multiple options.\footnote{Several papers consider simulation approaches to multiple exercise problems, see Bender [3], Ibáñez [23] and Meinshausen and Hambly [30] and a Malliavin approach of Carmona and Touzi [6]. These are all concerned with pricing under risk neutral expectation, under specific price models and thus there is no mixed control problem. Additionally, since the main application is swing options they are concerned with exercise constraints which take the form of intervals of time.} In general, if there is no co-monotonicity, we need to consider all possible exercise orderings. Conditional on an ordering, we optimize over hedging strategies and exercise times. Finally, we optimize over exercise order. In contrast, if there is co-monotonicity, our results show that the problem is reduced to a single dynamic programming exercise. With \( n \) distinct options, this is a saving of a factor of \( n! \). As an illustration, we solve the dynamic programming problem for CARA and give examples in both the co-monotonic and non-co-monotonic cases. This example extends the model of Leung and Sircar [28] who treat single (and multiple identical) American options under the CARA utility indifference pricing model. This will be discussed further in the next section.

2 Employee Stock Options

The aim of this section is to give a description of our canonical example and the related literature and briefly discuss the relevance of our results for employee stock options. We consider an employee who has been granted American call options on the stock of her company on a periodic basis. According to the US National Center for Employee Ownership (2008), there exist about 3,000 broad-based employee stock option plans in the US with about 9 million participants. Commonly these options are granted at-the-money, with maturities of ten years, and vesting periods of three years, although many variations exist. Prior to the vest date, the option is not permitted to be exercised by the employee. At the vest date, the option becomes American and the employee can exercise it at any time up to and including the expiry date. After a few years, the employee will have accumulated a portfolio of American calls with various strikes, maturities and vest dates. We consider the optimal exercise of such an option portfolio - an important step in understanding the cost of granting options from the company perspective - and it is pertinent to consider exercise behavior since it can be observed in practice (unlike say, the subjective value of options to the employee which is unobservable). Note that since both subjective (here, utility indifference) value and the cost to the company are obtained easily once optimal exercise is known, it is sufficient to concentrate on the exercise decision.

Since employees are not permitted to trade in the company stock, and cannot directly sell or transfer their options, they can only unwind their risk exposure by exercising options or perhaps by trading other assets such as a market index. As such, the employee faces an incomplete market. Models based on utility-indifference pricing of American options capture many of the important aspects of the employee’s situation. Indeed, there are a number of papers in this vein, notably Carpenter, Stanton and Wallace [7], Grasselli and Henderson [16], Henderson [18], Leung and Sircar [28] and Rogers and Scheinkman [33] all of whom study American option pricing.
under utility indifference for single or multiple identical options. (There is also a long literature in finance focussing on one-period or binomial models typically without partial hedging, for a discussion and references, see the survey of Henderson and Sun [21]. The first treatment of employee stock options in a utility indifference framework with partial hedging was Henderson [17] who considered European style payoffs.) The above papers make particular choices of utility function (CARA or CRRA) and all assume exponential Brownian motions for the stock price and the correlated traded asset, and solve the resulting free boundary problem. This is done explicitly in some special cases (Grasselli and Henderson [16], Henderson [18]) or via numerical solutions.

Several of our results for single American options are relevant to this literature. We highlight two here. First, we give an example (extended from Carpenter, Stanton and Wallace [7]) of preferences where there are several disconnected exercise and continuation regions. This is an important observation from the perspective of employee stock option modeling, as there is a strand of literature (see for example Cvitanic, Wiener and Zapatero [9]) which assumes an exogenous form for a single boundary of threshold form, rather than deriving the threshold(s) endogenously via a utility (or other) model. It is therefore important to recognize the limitations of such an exogenous specification. Second, we give conditions under which a higher dividend rate leads to a smaller continuation region, under the assumption that the price is a diffusion. This generalizes results given for exponential Brownian prices in a utility indifference model by Carpenter, Stanton and Wallace [7]. In fact, our result is not limited to the utility indifference setting but only requires weak assumptions on preferences.

The main interest of this paper for the employee stock option literature is the treatment of portfolios of American options with different strikes, maturities and vest dates. Carpenter, Stanton and Wallace [7] only consider a single American call option in their paper, but acknowledge that in a more realistic portfolio setting “It would be useful to understand which options are most attractive to exercise first...”. Although some of the aforementioned papers recognize the need to study multiple options, only options with identical characteristics have been studied. In this case, the employee’s risk aversion causes her to unwind risk gradually and thus exercise options intertemporally. For instance, the assumption of infinitely divisible claims (and perpetual options) in Henderson and Hobson [19] leads to singular control where options are exercised when the price reaches a new maximum, to keep the position below a smooth function.

We consider the exercise of portfolios of American options with differing characteristics. Our main exercise ordering result requires the strikes, maturities and vest dates are co-montonic. Firms often grant employee stock options which are at-the-money and with a fixed vest date of three years. If such options are granted in a bull market, then their characteristics will indeed be co-montonic. Since the theorem also holds for ordered but random maturities, its conclusion is still true if the employee faces employment termination risk or an exogenous income shock.

We solve the dynamic programming problem for CARA utility and illustrate the resulting exercise boundaries for a portfolio of options. Our examples highlight the influence of other options on the exercise of a particular option in the portfolio. For example, the existence of a longer dated, higher strike option will cause the agent to exercise a shorter dated, lower strike option earlier. This is in contrast to a risk neutral pricing model under which options can be treated independently and thus do not interact. This extends the model of Leung and Sircar [28] who solve numerically the dynamic programming problem for single American options under the
3 Exercising Single American Call Options

We work on a filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s \geq t}, \mathbb{P})\) and let \(t\) denote current time and \((X_s)_{s \geq t}\) the price process of an asset. The market consists of the risky asset \(X\) and riskless bonds. We consider an agent who holds a portfolio of American call options on \(X\). A call option on \(X\) has strike \(K\), maturity \(T\) and vesting date \(V \leq T\). During the vesting period \(t \leq s \leq V\) the agent is not permitted to exercise the option. The agent’s portfolio will contain American calls of potentially differing strikes, maturities and vesting dates. We will consider \(n\) options with characteristics \((K^{(i)}, T^{(i)}, V^{(i)})_{1 \leq i \leq n}\) with vest dates such that \(t \vee V^{(i)} \leq T^{(i)}\). In this section we develop notation and assumptions to state results for a single American option. We return to the portfolio problem in Section 6.

Let \((I_s^{(i)})_{s \geq t} = (I_s^{(i)}(\omega))_{s \geq t}\) be an increasing process representing the income or cashflow generated by the exercise of options from time \(t\) onwards. For a single American call option exercised at stopping time \(\tau = \tau(\omega)\) with \(t \vee V \leq \tau \leq T\) we have

\[
I_s^{(i)}(\tau) := I_s^{(i)}(\tau(\omega)) = (X_{\tau(\omega)} - K)^{+}1_{\{s \geq \tau(\omega)\}}.
\] (1)

Implicit in our definition is that exercise involves either receiving the net proceeds \((X_s - K) > 0\) (“cash” exercise) or discarding the option at expiry as it is worthless (if \(X_T < K\)). As is standard, we do not allow the agent to exercise before expiry if \(X_s < K\). Since \((I_s^{(i)})_{s \geq t}\) is increasing we can write \(I_s^{(i)} = \int_t^s dI_u^{(i)}\) with the convention that \(I_t^{(i)} = 0\) so that \(dI_t^{(i)} = I_t^{(i)}\).

This covers the case where an option is exercised at time \(t\), for then \(I_t^{(i)} > 0 = I_{t-}^{(i)}\). We also define the constant income process \(i^{(i)}\) by \(i_s^{(i)}(\omega) = i \quad \forall \; s \geq t, \forall \omega\). In particular \(0^{(i)}\) is the zero cashflow.

Let \(\mathcal{V}(I^{(i)}(t))\) be the valuation (at time \(t\)) by the agent of the random future cashflow \((I^{(i)}_s)_{s \geq t}\). Note \(\mathcal{V}(I^{(i)}(t))\) is a known quantity at time \(t\). A simple example is the expected value of income received during \([t, T]\), denoted \(\mathcal{V}(I^{(i)}(t)) = \mathbb{E}[I^{(i)}_T | \mathcal{F}_t]\). Our assumption is that the agent prefers more to less.

**Hypothesis 3.1** (Prefer more to less) Suppose \((I_s^{(i)})_{s \geq t}\) and \((J_s^{(i)})_{s \geq t}\) are income processes. If \(dI_s^{(i)} \geq dJ_s^{(i)}\) for all \(\omega \in \Omega\) and for all \(s \geq t\) then \(\mathcal{V}(I^{(i)}(t)) \geq \mathcal{V}(J^{(i)}(t))\).

Hypothesis 3.1 is a minimal requirement on preferences which should be valid in all non-pathological circumstances. It is much weaker than pricing by expectation. Later we will impose stronger hypotheses on preferences.

The cashflow resulting from the exercise of a single American call is given in (1). We assume that the agent is able to choose the best exercise time and hence the value of the call (with no vesting, or which has already vested) is given by

\[
\mathcal{A}(x, t; K, T) = \mathcal{A} = \sup_{\tau: \tau \leq t \leq T} \mathcal{V}(I^{(i)}(\tau))
\] (2)

where the supremum is taken over stopping times. The value \(\mathcal{A}\) depends on the current data of price \(x\), current time \(t\), and also on the characteristics of the option \(K, T\). Sometimes we will
write $A = A(K, T)$ or $A = A(x, t)$ depending on whether we want to describe how $A$ depends on the option characteristics or on the state of the market.

For the American call with vesting, we have

$$A(x, t; K, T, V) = A = \sup_{\tau: t \vee V \leq \tau \leq T} V(I^{(t)}(\tau)).$$

(3)

Note if the option has already vested, ie. $V \leq t$ then $A(x, t; K, T, V) = A(x, t, K, T)$.

We stress that at this point, we have not made any assumptions beyond Hypothesis 3.1 on the agent’s preferences, or any assumptions on the price process itself.

For the situation with no vesting period, the following results are given in Merton [31] and Cox and Rubinstein [8]. These are classic no-arbitrage results and the (now textbook) arguments to enforce these relations involve the purchase and sale of options, see Cox and Rubinstein [8] (Chapter 4, Propositions 2 and 3) for details. However since we are interested in incomplete market situations where these trades cannot be implemented (such as employee stock options) we need to give alternative proofs which do not rely on trading, but instead rely on the weak preferences in Hypothesis 3.1. These are close in spirit to the arguments of Merton [31] which rely on the notion of portfolio dominance.

**Proposition 3.2** Fix $x, t$. Then

(i) Fix $K, V$. Then $A(K, T, V)$ is increasing in maturity $T$.

(ii) Fix $K, T$. Then $A(K, T, V)$ is decreasing in vest date $V$.

Suppose $V$ satisfies Hypothesis 3.1. Then

(iii) Fix $T, V$. $A(K, T, V)$ is decreasing in $K$.

**Proof:**

(i) and (ii). These are immediate from the fact that as $T$ increases or $V$ decreases, the set of admissible strategies increases, so that if $T \leq T'$, $\{\tau: t \vee V \leq \tau \leq T\} \subseteq \{\tau: t \vee V \leq \tau \leq T'\}$ and if $t \leq V' \leq V$, $\{\tau: t \vee V \leq \tau \leq T\} \subseteq \{\tau: t \vee V' \leq \tau \leq T\}$.

(iii) Let $K' > K$. For any $t \vee V \leq \tau \leq T$ let

$I^{(t)}(\tau, K) = (X_{\tau} - K)^+ 1_{\{\tau < T\}}$ and $dI^{(t)}(\tau, K) = (X_{\tau} - K) 1_{\{t \vee V \leq \tau \leq T\}} 1_{\{\tau = s\}}$. It follows from Hypothesis 3.1 that for any such $\tau$, $V(I^{(t)}(\tau, K')) \leq V(I^{(t)}(\tau, K))$ and optimizing over $\tau$ we conclude $A(x, t; K', T, V) \leq A(x, t; K, T, V)$.

Now we want to fix option characteristics $(K, T, V)$ and vary $(x, t)$ to consider the optimal exercise decision. Define the exercise region $E = E_{K, T, V}$ by

$$E = \{(x, s) : (s \geq V), (x > K), A(x, s; K, T, V) = A(x, s; K, s)\} \cup \{(x, s) : s = T\}$$

(4)

and the continuation region $C = C_{K, T, V}$ by

$$C = \{(x, s) : (s < V)\} \cup \{(x, s) : V \leq s < T, (x \leq K)\} \cup \{(x, s) : V \leq s < T, A(x, s; K, T, V) > A(x, s, K, s)\}.$$

\(^2\)Note by convention we require the option to be in-the-money $(x > K)$ for $(x, s); s < T$ to be in the exercise region. Under the Black Scholes model this does not need comment, but if the price process is more general, and if the price $X$ can hit zero in finite time (and then zero is absorbing for $X$) then the option is worthless. Although $A(0, s; K, T, V) = A(0, s, K, s)$ we do not include $(0, s)$ in $E$. The reason for this will become clear when we discuss the shapes of $E$. Again by convention we set $\{(x, s) : s = T, x \leq K\}$ to be contained in $E$. On this set the option is discarded rather than exercised.
Note that in a Black Scholes complete market framework, the immediate exercise value is just the intrinsic value, i.e. $A(x, s; K, s) = (x - K)^+$. We define also the exercise and continuation regions for an American call which has either already vested, or had no vesting period.

Define $\mathcal{E} = \mathcal{E}_{K,T}$ by

$$\mathcal{E}_{K,T} = \{(x, s) : (x > K), A(x, s; K, T) = A(x, s; K, s)\} \cup \{(x, s) : s = T\}$$

and the continuation region $\mathcal{C} = \mathcal{C}_{K,T}$ by

$$\mathcal{C}_{K,T} = \{(x, s) : s < T, (x - K)^+\} \cup \{(x, s) : s < T, A(x, s; K, T) > A(x, s; K, s)\}.$$ Given that on $(V \leq s)$ we have $A(x, s; K, T, V) = A(x, s; K, T)$, it follows immediately that

$$\mathcal{C}_{K,T,V} = \mathcal{C}_{K,T} \cup \{(x, s) : s < V\}. \tag{5}$$

We now define some further useful concepts: an optimal stopping time, a notion of monetary valuation and a stronger hypothesis on preferences. Define (an) optimal stopping time by

$$\tau^* = \inf\{u \geq t; X_u \in \mathcal{E}\}.$$ \tag{6}

**Definition 3.3** We say $V$ is monetary if $V(I(t)) = 0$ and $I_s^t = i + J_s^t$ implies $V(I(t)) = i + V(J(t))$, i.e. a cashflow of $i \geq 0$ at time $t$ is given value $i$.

Risk neutral pricing and utility indifference pricing with negative exponential utility are both examples of monetary pricing rules. However in general, other utility functions lead to non-monetary pricing rules - the value to an agent of an income stream can depend on her wealth. Note our use of the terminology monetary is identical to its use in the axiomatic treatment of risk measures, see Föllmer and Schied [13].

**Definition 3.4** $I(t) \succeq J(t)$ if $I_s^t(\omega) \geq J_s^t(\omega)$ for all $s \geq t$ and for all $\omega \in \Omega$.

**Hypothesis 3.5** (Prefer more to less and prefer cash earlier)

Suppose $I(t) \succeq J(t)$. Then $V(I(t)) \geq V(J(t))$.

Hypothesis 3.5 says agents prefer more to less and prefer cash at earlier times. This is consistent with non-negative interest rates. Clearly this is still a very weak requirement on preferences, but if $V$ satisfies Hypothesis 3.5 then it automatically satisfies Hypothesis 3.1.

Prior to the main result of this section, we give a simple lemma which is needed later.

**Lemma 3.6** Suppose $V$ is monetary. Suppose $s \geq V$. Then $A(x, s; K, T, V) = A(x, s; K, T) \geq (x - K)^+.$

**Proof:** Fix $(x, s)$. The strategy of immediate exercise yields $V(I(s)(s, K)) = V((x - K)^+)(s)) = (x - K)^+$, where we use the monetary property of $V$. Hence $A(x, s; K, T) \geq V(I(s)(s, K)) = (x - K)^+$. \[\Box\]

Here is our first major result which says that the continuation regions are nested for different strikes/maturities. Note that this does not prevent them from being complicated in shape, see Figures 2 and 3.
Theorem 3.7 Suppose \( V \) is monetary and satisfies Hypothesis 3.5. If \( K \leq K' \), \( T \leq T' \) and \( V \leq V' \) then \( C_{K,T,V} \subseteq C_{K',T',V'} \) so that \( \tau^* \leq \tau'^* \).

Proof:
Fix \((x, s)\) and for now, ignore vesting. We prove if \( T \leq T' \) then \( C_{K,T} \subseteq C_{K',T} \). Suppose \((x, s) \in C_{K,T} \). Then \( s < T \leq T' \). Suppose \( x > K \). (Otherwise \((x, s) \in C_{K,T'} \) by definition).

Then
\[
A(x, s, K, s) < A(x, s; K, T) \leq A(x, s, K, T')
\]
by Theorem 3.2 (i). Hence \((x, s) \in C_{K,T'} \).

Now we prove if \( K \leq K' \), then \( C_{K,T} \subseteq C_{K',T} \). Suppose \((x, s) \in C_{K,T} \). Then \( s < T \). If \( x \leq K' \) then \((x, s) \in C_{K',T} \) as the option is out-of-the-money and there is nothing to prove. So suppose \( s < T \) and \( x > K' \). There exists \( \tau \) such that \( V(I^{(s)}(\tau, K)) > V(I^{(s)}(s, K)) = V(((x - K)^+)^{(s)}) = (x - K)^+ = x - K \), where we use the monetary property of \( V \). For this \( \tau \),
\[
J_u^{(s)}(\tau) := I_u^{(s)}(\tau, K') + (K' - K) \geq I_u^{(s)}(\tau, K)
\]
and hence \( J^{(s)}(\tau) \geq I^{(s)}(\tau, K) \). By Hypothesis 3.5,
\[
V(I^{(s)}(\tau, K') + (K' - K)) \geq V(I^{(s)}(\tau, K))
\]
But since \( V \) is monetary, \( V(I^{(s)}(\tau, K') + (K' - K)) = V(I^{(s)}(\tau, K')) + (K' - K) \) and it follows that
\[
A(K', T) + (K' - K) \geq V(I^{(s)}(\tau, K)) > (x - K).
\]
Hence \( A(K', T) > (x - K) \) so \((x, s) \in C_{K',T} \).

Now consider extending the proof to include vesting dates. Set \( V \leq V' \). If \((x, s) \in C_{K,T,V} \) then by (5) either \((x, s) \in C_{K,T} \subseteq C_{K',T'} \) or \((x, s) \in \{(x, s) : s \leq V = V \} \subseteq \{(x, s) : s \leq V' \} \). Hence \((x, s) \in C_{K',T',V'} \).

\( \Box \)

Theorem 3.7 tells us that under valuation methodologies which are monetary and satisfy Hypothesis 3.5, if two identical agents each hold a different American call option, the agent with the lower strike/earlier maturity/shorter vest date will exercise first. That is, the continuation regions for options with co-montonic strike, expiry and vest date are nested. Note this result does not apply to a portfolio of options held by a single agent, at least not at this level of generality on preferences. Of course if one were to make the strong assumption of risk neutrality, options can be treated independently, and the result can be applied. Since our main interest is in non-risk neutral pricing, our aim is to develop a portfolio version of Theorem 3.7 in Section 6.

We will give a number of examples which satisfy the assumptions of Theorem 3.7 and thus have nested continuation regions.

Example - Black Scholes model
A first simple example is the Black Scholes framework where \( V \) is risk neutral pricing and \( X \) follows exponential Brownian motion with constant interest rate \( r \), constant proportional dividends \( q \), and constant volatility \( \sigma \). Figure 1 gives exercise thresholds for pairs of calls. The exercise region takes the form
\[
\mathcal{E} = \{(x, s) : (s \geq V) \text{ and } x \geq x^*(s)\} \cup \{(x, s) : s = T\}
\]
Figure 1: The exercise threshold of a single American call under risk neutral pricing and exponential Brownian motion. In each panel, the solid line gives the threshold for call with \((K = 10, T = 5, V = 1)\). The dashed line in each panel is the threshold for the call with \((T = 6, V = 2)\) and different strike in each panel as indicated. Other parameter values are: \(r = 0.05, q = 0.02, \sigma = 0.4\)

for some function of time \(x^*(s)\) (see the later result in Theorem 4.2). In each panel, the solid line is the exercise threshold \(x^*(s)\) for a call with \((K = 10, T = 5, V = 1)\). Note if \(X\) is above the threshold level at the vest date of 1 year, the call is exercised at the vest date. In panel (a), the dashed line is the threshold for a call with \((K = 11, T = 6, V = 2)\). The continuation regions are nested, as expected, since the strikes, maturities and vest dates are co-monotonic.

If, for instance, we consider the pair of options with \((K = 10, T = 5, V = 1)\) and \((K = 8.5, T = 6, V = 2)\) then we violate the conditions of the theorem. As we see in panel (ii) of the figure, this results in continuation regions which are not nested.

The proof of Theorem 3.7 relied on the monetary property and Hypothesis 3.5. (An alternative proof could use convexity of \(A\) in strike \(K\), see Hobson [22], however this relies on expectation pricing, which is a stronger requirement than the monetary property). The well known textbook arguments of Cox and Rubinstein (1984) (Chapter 4, Proposition 5 (c)) use no-arbitrage arguments which involve trading in order to enforce no-arbitrage. The monetary property is, of course, implicit but unstated in such arguments as the potential arbitrage requires the agent to assign a monetary value to a cash profit. We emphasize the importance of the monetary property by the following counterexample.

Counterexample - non-monetary and non-nested continuation regions

Suppose \(t = 0\) and \(T = 1\). Let \(I_s^{(0)}(\tau) = (X_\tau - K)^+ 1_{s \geq \tau}\) be the income from option exercise at \(0 \leq \tau \leq T\) and suppose \(\mathcal{V}(I^{(0)}) = \mathbb{E}U[(X_T - K)^+ + w]\) where \(w\) is initial wealth. We make a particular choice of \(U\) to enable calculations to be carried out. (Note, here for simplicity we define \(\mathcal{V}\) as the value function rather than utility indifference price as used later in the paper). Define \([x]\) the integer part of \(x\), and \(\text{frac}(x) = x - [x]\) the fractional part of \(x\). Let \(U(y) = \lfloor y \rfloor\) and suppose \(X_s = x\) for \(0 \leq s \leq 1\) and \(X_1 = x + Z\) where \(Z\) takes values \(\{1/2, -1/2\}\) with equal probability.

Take \(w = 0\) for simplicity and consider \(K = 0\). Then it is worth the agent waiting until time
1 to exercise if \( \text{frac}(x) \geq 1/2 \), otherwise waiting can only reduce the utility. Hence

\[
V(I^{(0)}) = \begin{cases} 
\lfloor x \rfloor + 1/2 & \text{frac}(x) \geq 1/2 \\
\lfloor x \rfloor & \text{frac}(x) < 1/2 
\end{cases}
\]

and \( C_{K=0} = \{ x : \text{frac}(x) \geq 1/2 \} \). Now consider \( K = 1/2 \). Then

\[
V(I^{(t)}) = \begin{cases} 
\lfloor x - 1/2 \rfloor & \text{frac}(x) \geq 1/2 \\
\lfloor x \rfloor - 1/2 & \text{frac}(x) < 1/2 
\end{cases}
\]

and \( C_{K=1/2} = \{ x : \text{frac}(x) < 1/2 \} \). Then \( C_{K=0} \) and \( C_{K=1/2} \) are disjoint.

This extreme example has been carefully constructed to facilitate computations. However, the intuition remains true with examples with more realistic features. Take \( U \) to be an increasing, concave, piecewise linear function. Suppose there we are close to the option maturity (small time to go) and that it is reasonable to consider \( X \) to be a Brownian motion with a small positive drift. If the values of strike and wealth put us at or near a kink, then the downside risk dominates and we prefer to stop. If on the other hand, parameters mean we are in a linear part of the function, away from kinks, the positive drift means that we would continue. We can now perturb the strike \( K \) in such a way that the stopping and continuation regions are reversed, and thus are not nested.

Our final result in this section tells us when we can expect continuation regions to be nested for different dividend rates. Intuition from Black Scholes informs us that a higher level of dividends reduces the American call boundary and hence the continuation region. In contrast to Theorem 3.7, we will need some assumptions on the price \( X \). We make:

**Assumption 3.8** (i) The price \( X \) follows a diffusion model:

\[
dX_s = (rX_s - X_s q(X_s, s)) ds + X_s \sigma(X_s, s) dB_s \tag{7}
\]

where \( r > 0 \) is a constant interest rate, \( q(X_s, s) \) is a non-negative dividend rate, \( \sigma(X_s, s) \) is the volatility and \( r, q, \sigma \) are such that zero is an absorbing point, and that (7) is unique in law.

(ii) Define \( M = (M_s)_{s \geq t} \) by

\[
M_s = X_s \exp \left( -\int_t^s (r - q(X_u, u)) du \right) \tag{8}
\]

Then \( (M_s)_{s \geq t} \) is a true martingale.

Note that in (8) \( dM_s = \sigma(X_s, s) M_s dB_s \) so that \( M \) is automatically a local martingale. Various tests (e.g. Novikov) exist to ensure \( M \) is a martingale. A simple sufficient condition is that \( \sigma(X_s, s) \) is bounded.

**Theorem 3.9** Suppose Hypothesis 3.1 holds. Suppose \( \hat{X} \) solves

\[
dX_s = X_s (r - \hat{q}(X_s, s)) ds + X_s \sigma(X_s, s) d\hat{B}_s
\]

subject to \( X_s = x \), and that Assumption 3.8(i) holds for \( \hat{X} \). Suppose \( \tilde{X} \) solves

\[
dX_s = X_s (r - \tilde{q}_s) ds + X_s \sigma(X_s, s) d\tilde{B}_s
\]

with \( \hat{X}_s = x \). Suppose \( \tilde{q}_s \geq \hat{q}(x, s) \). Then \( \hat{C} \subseteq \tilde{C} \).
Proof:
There exists a coupling such that $\tilde{X}_s \geq \hat{X}_s \forall s$. For any $V \forall t \leq \tau \leq T$, let $\hat{I}^{(t)}(\tau) = (\hat{X}_\tau - K)^+_{I_{s \geq \tau}}$ and $\tilde{I}^{(t)}(\tau) = (\tilde{X}_\tau - K)^+_{I_{s \geq \tau}}$. Then $d\hat{I}^{(t)}(\tau) = (\hat{X}_\tau - K)_{I_{V \forall t \leq \tau \leq T}}$ and similarly for $d\tilde{I}^{(t)}(\tau)$. It follows from Hypothesis 3.1 that for any such $\tau$, $\mathcal{V}(\hat{I}^{(t)}(\tau)) \geq \mathcal{V}(\tilde{I}^{(t)}(\tau))$ and optimizing over $\tau$ we conclude $\hat{A}(x, t; K, T, V) \geq \tilde{A}(x, t; K, T, V)$. Hence if $(x, t) \in \tilde{C} = \{(y, s) : \hat{A}(y, s; K, T, V) > \tilde{A}(y, s; K, s, V)\}$ then $(x, t) \in \hat{C}$. □

Leung and Sircar [28] show via a pde comparison result that this result is true in a CARA utility indifference model whilst Carpenter, Stanton and Wallace [7] generalize their result to other utilities. In both, prices follow exponential Brownian motions. Theorem 3.9 is much stronger - weakening preferences and assumptions on the asset price process. We revisit this result in Section 5 in the context of utility indifference pricing.

4 Characterizing the Exercise Boundary of the American Call

In this section we will temporarily add more structure to our set-up. This will (i) allow us to give some results on the form of exercise boundaries for single American calls, (ii) demonstrate how complicated the boundaries can be even just for a single option, and (iii) motivate our main result in Section 6 - where we will remove the structure and generalize Theorem 3.7 to option portfolios.

Unless otherwise stated, for this section, we assume that value $V$ is given by the discounted expectation under a martingale measure, and in particular, the risk-neutral expectation. We also assume that the riskless bond pays a positive constant interest rate $r$, although the results generalize to deterministic interest rates. Then the value of income stream $I^{(t)}$ at current time $t$ is given by:

$$\mathcal{V}(I^{(t)}(\tau)) = E \left[ \int_{s \geq t} e^{-r(s-t)} dI^{(t)}_s \right]$$

and

$$A(x, t; K, T, V) = \sup_{\tau : V \forall t \leq \tau \leq T} \mathcal{V}(I^{(t)}(\tau)) = \sup_{V \forall t \leq \tau \leq T} E[e^{-r(\tau-t)}(X_\tau - K)^+_{|X_t = x}] \quad (9)$$

where stopping times $\tau$ are defined relative to the filtration generated by $(X_s)_{s \geq t}$. Note we are not claiming that this is a good assumption for employee stock options (recall employee’s cannot hedge perfectly and thus face an incomplete market), but rather we will develop some results first in a risk neutral setting to demonstrate the complexities involved in characterizing boundaries and thus exercise strategies even in this simple setting.

American option pricing (in complete markets) dates back to McKean [29], Merton [31] and Van Moerbeke [34] who recognize the value of an American call on a stock paying continuous dividends as a free boundary problem. Hedging arguments were given by Bensoussan [3] and Karatzas [26]. Myneni [32] surveys the development of American option pricing.

For this section we make Assumption 3.8(i) and (ii).

We are first interested in when the optimal exercise and continuation regions can be separated by a single price level at each point in time.
**Definition 4.1** We say the exercise region is of single threshold form if:

\[ \mathcal{E} = \{(x, s) : (s \geq V) \text{ and } x \geq x^*(s) \} \cup \{(x, s) : s = T\} \]

for some function of time \( x^*(s) > K \).

**Theorem 4.2** Suppose the price \( X \) satisfies Assumption 3.8 and that \( xq(x, s) \) is increasing in \( x \) for each \( s \). Suppose the American call is priced according to (9). Then the optimal exercise region is of single threshold form.

**Proof:** Suppose \((x, s) \in \mathcal{E}, x > K \) and \( s \geq V \). We want to show that for \( y > x \), \((y, s) \in \mathcal{E}\). We write \( X^{x, s}_t \) to denote that we start from \( x \) at time \( s \). Then we can couple \((X^{y, s}_t)_{u \geq s} \) and \((X^{x, s}_t)_{u \geq s} \) such that \( X^{y, s}_u(\omega) \geq X^{x, s}_u(\omega) \) for all \( u \geq s \) and all \( \omega \). Then for any \( \tau \geq s \)

\[ e^{-r\tau}(X^{y, s}_\tau - K)^+ \leq e^{-r\tau}(X^{x, s}_\tau - K)^+ + e^{-r\tau}(X^{y, s}_\tau - X^{x, s}_\tau). \]

Let \( Z_u = e^{-r(u-s)}(X^{y, s}_u - X^{x, s}_u) \). Then we write \( Z_u = N_u - A_u \) where

\[ N_u = \int_s^u (X^{y, s}_v \sigma(X^{y, s}_v, v) - X^{x, s}_v \sigma(X^{x, s}_v, v)) e^{-r(v-s)} dB_v + (y - x) \]

and

\[ A_u = \int_s^u e^{-r(v-s)}(X^{y, s}_v q(X^{y, s}_v, v) - X^{x, s}_v q(X^{x, s}_v, v)) dv. \]

For the coupled processes \( X^{y, s}_v \geq X^{x, s}_v \) and then by the monotonicity of \( xq(x, v) \) we have that \( A \) is increasing. Then the local martingale \( N \) satisfies \( N_u \geq Z_u \geq 0 \) so is a supermartingale. Hence \( Z \) is a supermartingale and

\[ \sup_{\tau \geq s} \mathbb{E}Z_\tau = Z_s = y - x. \]

Then since \((x, s) \in \mathcal{E}\), we have

\[ \sup_{s \leq \tau \leq T} \mathbb{E} e^{-r(\tau-s)}(X^{y, s}_\tau - K)^+ \leq \sup_{s \leq \tau \leq T} \mathbb{E} e^{-r(\tau-s)}(X^{x, s}_\tau - K)^+ + \sup_{s \leq \tau \leq T} \mathbb{E} e^{-r(\tau-s)}(X^{y, s}_\tau - X^{x, s}_\tau) = (x - K) + (y - x) = (y - K). \]

\( \Box \)

Jacka [24] (under exponential Brownian motion) and Babilua, Bokuchava, Dochviri and Shashiashvili [1] (under time homogenous diffusions) proved similar results for American puts without dividends. Götsche and Vellekoop [15] study the impact of discrete dividends on the boundary in an exponential Brownian motion model. Finally, Vellekoop and Nieuwenhuis [35] prove results in a semimartingale model under certain conditions (which play the role of our condition on \( q(x, t) \) and the no-crossing property for diffusions). However their conditions would seem difficult to verify in any model outside the diffusion framework above.

Theorem 4.2 can also be deduced via convexity of the American call option value in \( x \), under the (stronger) assumption of proportional dividends. The convexity property will follow a lemma which says the call cannot be worth more than the current stock price.
Lemma 4.3 Suppose $X$ satisfies Assumption 3.8 and the American call is priced according to (9). Then $\mathcal{A}(x, t, K, T, V) \leq x$.

Proof:
\[
\mathbb{E}[e^{-r(\tau-t)}(X_\tau - K)^+|X_t = x] \leq \mathbb{E}[e^{-r(\tau-t)}X_\tau|X_t = x] \\
\leq \mathbb{E}[e^{-\int_t^\tau (r-q(x,u))du}X_\tau|X_t = x] = \mathbb{E}[M_\tau|X_t = x] = x
\]
since $M$ is a martingale by Assumption 3.8. Hence $\mathcal{A}(x, t, K, T, V) = \sup_{V \wedge t \leq \tau \leq T} \mathbb{E}[e^{-r(\tau-t)}(X_\tau - K)^+|X_t = x] \leq x$.  

An easy counterexample to the lemma is to consider non risk-neutral expectation pricing where $X$ has a large positive drift, eg. pricing by expectation under the real world measure. Then $M$ is a submartingale and the option is worth more than the current stock price.

Proposition 4.4 Suppose the price $X$ satisfies Assumption 3.8 and dividends are proportional so
\[
dX_s = (r-q)X_sds + \sigma(X_s)sdB_s
\]
Suppose the American call is priced according to (9). Then $\mathcal{A}(x, t; K, T, V)$ is convex in $x$.

Proof:
Recall if $M_s^{x,t} = e^{-(r-q)(s-t)}X_s^x$ then $dM_s^{x,t} = M_s^{x,t}\sigma(X_s^x, s)dB_s = M_s^{x,t}\tilde{\sigma}(M_s^{x,t}, s)dB_s$ (where $\tilde{\sigma}(m, s) = \sigma(e^{(r-q)(s-t)}m, s)$) and by Assumption 3.8, $M_s^{x,t}$ is a martingale.
Define $w(m, u) = (e^{-q(u-t)}m - e^{-r(u-t)}K)^+$. Then $w$ is convex in $m$ and $w(M_u^{x,t}, u) = e^{-r(u-t)}(X_u^{x,t} - K)^+$.
For $s \geq t$ define
\[
v(m, s) = \sup_{\tau \geq s} \mathbb{E}[w(M_{\tau}^{m,s}, \tau)].
\]
Then $v$ is a supermartingale and a martingale for $s \leq \tau^*$ the optimal strategy.
Let $0 < z < y < u$ and for independent Brownian motions’s $\alpha, \beta, \gamma$ define processes $M^{u,t}, M^{y,t}, M^{z,t}$ via
\[
dM_s^{u,t} = \tilde{\sigma}(M_s^{u,t}, s)d\alpha_s; \quad M_t^{u,t} = u \\
dM_s^{y,t} = \tilde{\sigma}(M_s^{y,t}, s)d\beta_s; \quad M_t^{y,t} = y \\
dM_s^{z,t} = \tilde{\sigma}(M_s^{z,t}, s)d\gamma_s; \quad M_t^{z,t} = z.
\]
Let $\tau^y$ be the optimal exercise time for the option with price $M^{y,t}$ started at $y$. Then $V \leq \tau^y \leq T$.
Define $H_u = \inf\{s : M_s^{u,t} = M_s^{y,t}\} = \inf\{s : M_s^{y,t} = M_s^{u,t}\}$ and $\tau = H_u \wedge H_z \wedge \tau^y$.
On $\tau = \tau^y$, $v(M_{\tau}^{u,t}, \tau) = w(M_{\tau}^{u,t}, \tau) = (e^{-q(\tau-t)}M_{\tau}^{y,t} - Ke^{-r(\tau-t)})$ and by convexity,
\[
(M_{\tau}^{u,t} - M_{\tau}^{z,t})v(M_{\tau}^{y,t}, \tau) \leq (M_{\tau}^{y,t} - M_{\tau}^{z,t})w(M_{\tau}^{u,t}, \tau) + (M_{\tau}^{u,t} - M_{\tau}^{y,t})w(M_{\tau}^{z,t}, \tau) \\
\leq (M_{\tau}^{y,t} - M_{\tau}^{z,t})v(M_{\tau}^{u,t}, \tau) + (M_{\tau}^{u,t} - M_{\tau}^{y,t})v(M_{\tau}^{z,t}, \tau).
\]
On $\tau = H_u$, by symmetry
\[
(M_{\tau}^{u,t} - M_{\tau}^{z,t})v(M_{\tau}^{y,t}, \tau) = (M_{\tau}^{y,t} - M_{\tau}^{z,t})v(M_{\tau}^{u,t}, \tau)
\]
13
and \((M^u_t - M^y_t)v(M^z_t, \tau) = 0\). Similarly, on \(\tau = H_z\),
\[
(M^u_t - M^z_t)v(M^y_t, \tau) = (M^y_t - M^y_t)v(M^z_t, \tau).
\]

Hence we always have
\[
(M^u_t - M^z_t)v(M^y_t, \tau) \leq (M^y_t - M^z_t)v(M^y_t, \tau) + (M^u_t - M^y_t)v(M^z_t, \tau).
\]

Taking expectations and by independence,
\[
(u - z)\mathbb{E}v(M^y_t, \tau) \leq (y - z)\mathbb{E}v(M^u_t, \tau) + (u - y)\mathbb{E}v(M^z_t, \tau).
\]

Now \(v(M^y_s, s)\) is a martingale on \(t \leq s \leq \tau \leq \tau^y\) so \(\mathbb{E}v(M^y_t, \tau) = v(y, t)\); and \(v(M^u_t, s), v(M^z_t, s)\) are supermartingales, so the convexity property for \(v(\cdot, t)\) follows. □

This proof generalizes option price convexity results by Bergman, Grundy and Wiener [4] for European options (using analysis of the partial differential equation), and El Karoui, Jeanblanc-Pique and Shreve [12] (using stochastic flows), Hobson [22] (using coupling) (also Ekström [11]) for American options without dividends. In particular, our proof extends the coupling proof of Hobson [22].

**Alternative proof of Theorem 4.2:**

Fix \(K, T, V, t\). From Proposition 4.4, \(A(x, t)\) is convex in \(x\). Since from Lemmas 3.6 and 4.3, \((x - K)^+ \leq A(x, t) \leq x\), then if for any \(x > 0\) we have \(A(x, t) = (x - K)\), then, for all \(y > x\)
\[A(y, t) = (y - K)\]. Thus if \((x, t) \in \mathcal{E}_{K,T}\) then \((y, t) \in \mathcal{E}_{K,T}\). □

**Counterexample: Disconnected Exercise Regions**

We adopt the Black Scholes framework. Suppose \(X\) follows exponential Brownian motion with constant proportional dividends and suppose pricing is risk-neutral as in (9). Under this standard model, the American call has a well studied exercise boundary, which by our definition is a single threshold (see Figure 1). Now we show that if \(x^*(x, s)\) is not monotonic increasing, then the exercise boundary need not be a single threshold. Suppose that dividends are state-dependent and
\[
q = \begin{cases} 
q_{\text{high}}; & X_s \leq x_q \\
q_{\text{low}}; & X_s > x_q 
\end{cases}
\]

for some constant \(x_q\). This violates the assumption in Theorem 4.2. Figure 2 takes \(q_{\text{high}} = 0.1, x_q = 7\) and assumes \(V = 0\). In panel (a), we take \(q_{\text{low}} = 0\), whilst in panel (b), \(q_{\text{low}} = 0.01\). In each panel, the solid lines delineate the exercise and continuation regions for a call with \(K = 1, T = 10\). We first describe the graph in panel (a). Far from maturity, it is optimal to continue at all price levels. There is some prospect of prices reaching the zero dividend region, and zero dividend American calls are not exercised early (Merton [31]). As maturity approaches, the chance of reaching the zero-dividend region becomes less likely, and there is a wedge of price levels between which it is optimal to exercise. This exercise wedge becomes larger as maturity approaches, and close to maturity it is optimal to exercise for all prices in-the-money. The presence of dividends (for \(X > x_q\)) in panel (b) induces a disconnected exercise region at high prices, and the option is exercised above the higher solid line.
Figure 2: The figures show the optimal exercise and continuation regions of a single option in a Black-Scholes model with state-dependent dividends. We take \( x_q = 7 \), \( q_{\text{high}} = 0.1 \). In both panels, the solid line encloses an exercise region for a call with \( K = 1, T = 10 \). In panel (b), there is also a disconnected exercise region at high price levels. In both panels, the dashed line encloses the exercise region for a call with \( K = 2, T = 11 \). Other parameters are: \( r = 0.1, \sigma = 0.15, V = 0 \).

The dashed lines on Figure 2 display the exercise and continuation regions for a call with \( (K = 2, T = 11) \). The exercise region for this call is nested within the region for the call with lower strike and maturity, which is consistent with Theorem 3.2. An agent with the \( (K = 1, T = 10) \) option will exercise before an identical agent with the \( (K = 2, T = 11) \) option. In addition, since this example assumes risk-neutral pricing, we can also conclude that an agent with a portfolio of these two options would exercise the lower strike/maturity option first.

Our next task is to recall some known results concerning the monotonicity of the boundary in time. We need some preliminaries.

**Proposition 4.5** Assume the price \( X \) satisfies Assumption 3.8 and in addition, that the dynamics are time-homogeneous, ie. \( q(x, s) = q(x) \), \( \sigma(x, s) = \sigma(x) \). Assume the American call is priced according to (9). Fix \( K, T, x \). Then

(i) \( \mathcal{A}(x, t; K, T) = \mathcal{A}(x, 0; K, T - t) \);

(ii) \( \mathcal{A}(x, t; K, T) \) non-increasing in \( t \).

Recall Theorem 3.2(i) says the option value is increasing in maturity whereas Proposition 4.5 (ii) requires time-homogeneity. In contrast to Proposition 4.5 (ii), the European call option value is not non-increasing with \( t \).

**Proof:**

(i) Let \( \tilde{X}_u = X_{u+t} \) for \( 0 \leq u \leq T - t \). The filtration \( (\mathcal{F}_s)_{t \leq s \leq T} \) where \( (\mathcal{F}_s) = \sigma((X_v)_{t \leq v \leq s}) \) can be identified with \( (\mathcal{F}_s)_{0 \leq s \leq T-t} \) where \( (\mathcal{F}_s) = \sigma((\tilde{X}_v)_{0 \leq v \leq s-t}) \). Let \( \tilde{\tau} = \tau - t \). Then

\[
e^{-r(\tau-t)}(X_\tau - K)^+ = e^{-r\tilde{\tau}}(\tilde{X}_{\tilde{\tau}} - K)^+.
\]

Since the dynamics of \( X \) are time-homogeneous, \( \tilde{X} \) solves the same SDE as \( X \), and

\[
\sup_{t \leq \tau \leq T} \mathbb{E}^{x, t}[e^{-r(\tau-t)}(X_\tau - K)^+] = \sup_{0 \leq \tilde{\tau} \leq T-t} \mathbb{E}^{x, 0}[e^{-r\tilde{\tau}}(\tilde{X}_{\tilde{\tau}} - K)^+] = \sup_{0 \leq \tau \leq T-t} \mathbb{E}^{x, 0}[e^{-r\tau}(X_\tau - K)^+]
\]
and \( A(\kappa, t; K, T) = A(\kappa, 0; K, T - t) \).

(ii) Let \( t < t' \). From Theorem 3.2 (i), \( A(\kappa, 0; K, T - t) \geq A(\kappa, 0; K, T - t') \), so that from part (i) of this theorem, \( A(\kappa, t; K, T) \geq A(\kappa, t'; K, T) \). \( \square \)

**Theorem 4.6** Suppose \( X \) satisfies Assumption 3.8 and in addition that the dynamics are time homogeneous, i.e. \( q(X_s, s) = q(X_\kappa), \sigma(X_s, s) = \sigma(X_s) \). Suppose the American call is priced according to (9). Assume \( \kappa(x) \) increasing in \( x \) so (by Theorem 4.2) the exercise region takes the form:

\[
\mathcal{E}_{K,T} = \{(x, s) : x \geq x^*(s)\} \cup \{(x, s) : s = T\}
\]

Then

(i) \( x^*(s) \) is non-increasing; 
(ii) \( x^*(s) = f(T - s) \) for some function \( f \) with \( f(u) \) non-decreasing in \( u \), \( f(u) \) does not depend on \( T \), but does depend on \( K \) and \( f(0) \geq K \).

**Proof:**

(i) Fix \( K, T \). Since by Proposition 4.5(ii), \( A \) non-increasing in \( s \), we have that if \( (x, s) \in \mathcal{C}_{K,T} \) then \( (x, s') \in \mathcal{C}_{K,T} \forall s' < s \) since \( A(x, s'; K, T) \geq A(x, s; K, T) > (x - K)^+ \). Define \( \tilde{H}(x) = 
\tilde{H}(x; K, T) = \sup\{u : A(x, u; K, T) > (x - K)^+\}. \) Then \( x^* = \tilde{H}^{-1} \) and \( \tilde{H} \) is non-increasing.

(ii) Fix \( K \). Define \( H(x) = \inf\{u : A(x, 0; K, u) > (x - K)^+\}. \) Since by Theorem 3.2(i), \( A \) increasing in \( u \), for all \( u > H(x) \) we have \( A(x, 0; K, u) > (x - K)^+ \). Fix \( T \). Then by Proposition 4.5(i) and time-homogeneity,

\[
\tilde{H}(x) = \sup\{u : A(x, 0; K, T - u) > (x - K)^+\} = T - \inf\{v : A(x, 0; K, v) > (x - K)^+\} = T - H(x)
\]

Hence \( H(x) = T - \tilde{H}(x) \) is non-decreasing and has an inverse which we write as \( f = H^{-1} \). Then \( x^*(s) = x \iff \tilde{H}(x) = s \iff H(x) = T - s \iff f(T - s) = x \) and hence \( x^*(s) = f(T - s) \). \( \square \)

The intuition is that since the option value is decreasing in time (for a fixed stock price) and the payoff for immediate exercise is time-independent, the exercise boundary will decrease with time. We include Theorem 4.6 to draw a contrast to the boundaries calculated via utility indifference pricing in the next section, where the above intuition breaks down. Further, despite this result being known for many years, it has not always been stated correctly in the literature. Kim [27] gives this result without the condition of time-homogeneity. His Proposition 1 mis-applies the result in Theorem 3.2 (i), which, as Cox and Rubinstein [8] caution, compares the values on a given calendar date of two calls with different maturities. It does not tell us anything about how the value of a call changes with time. Indeed, it is only equivalent to comparing different time-to-expiries with fixed expiry date under the additional assumption of time-homogeneity of the model.

**Counterexample: Non-monotone exercise boundary**

Again we take the Black Scholes setup - risk neutral pricing and exponential Brownian motion with constant proportional dividends. Suppose volatility takes two distinct values \( \sigma = \sigma_{\text{low}}, s \leq T_\sigma \) and \( \sigma = \sigma_{\text{high}}, s > T_\sigma \) for some fixed \( T_\sigma \leq T \). Figure 3 gives the resulting non-monotonic exercise boundary for a call with \( (K = 1, T = 10) \) (the solid line). As \( s \to T \), the boundary approaches the limit \( \frac{s}{q}K = 1.66 \). As \( s \to -\infty \), the boundary approaches the perpetual limit.
The dashed line gives the exercise boundary for a single option with \((K = 2, T = 11)\). Despite the lack of time homogeneity, and resulting non-monotonic boundaries, we see the boundaries are consistent with Theorem 3.7 and the continuation regions are nested. Since the example is under risk-neutral pricing, an agent with a portfolio of both options would exercise the \((K = 1, T = 10)\) option first. Another situation where the exercise boundary is not monotone in time (under risk neutral pricing) is found in Götsche and Vellekoop [15] where dividends are paid discretely.

![Figure 3: The figure shows the optimal exercise threshold of a single American call option under risk neutral pricing with a Black-Scholes price with constant proportional dividends. For \(s \leq T_\sigma = 5\), \(\sigma_{low} = 0.2\), and for \(s > T_\sigma = 5\), \(\sigma_{high} = 0.4\). The solid line depicts the exercise threshold of an option with \(K = 1, T = 10\). The dashed line depicts the exercise threshold of an option with \(K = 2, T = 11\). Other parameters are: \(r = 0.05\), \(q = 0.03\).](image)

5 Single American Options - Utility Indifference Pricing

In this section we consider the exercise of single American options under the utility indifference framework. As we described in Section 2, utility based models have become a standard framework in which to value employee stock options due to the restrictions on the employee’s ability to hedge.

The agent has initial wealth \(w\), increasing concave utility function \(U\), and (for now) a single American call option with strike \(K\), maturity \(T\), vest date \(V\). In addition to the riskless bond (with positive constant interest rate \(r\)) and the underlying stock \(X\) (which cannot be traded), there is also a market asset \(M\) in which the employee may partially hedge her risk.

Prices follow

\[
\frac{dX}{X} = (\nu - q)dt + \sigma dB
\]

\[
\frac{dM}{M} = \mu^M dt + \sigma^M dZ
\]

where standard Brownian motions \(B\) and \(Z\) with constant instantaneous correlation \(\rho \in (-1, 1)\) are defined on a probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_u\}_{u \geq t}, \mathbb{P})\) and where \(\mathcal{F}_u\) is the augmented \(\sigma\)-algebra.
generated by \( \{B_u, Z_u; 0 \leq u \leq t \} \). The volatility of stock returns \( \sigma \), expected return on the stock \( \nu \), constant proportional dividend yield \( q > 0 \), and expected return \( \mu^M \) and volatility of the market returns \( \sigma^M \) are all constants. The mean stock return \( \nu \) is equal to the CAPM return for the stock, given its correlation with the market\(^3\):

\[
\nu = r + \beta(\mu^M - r); \quad \beta = \rho \sigma / \sigma^M
\]

Since \( \rho \in (-1, 1) \), the restricted employee faces some unhedgeable risk through her option position. Allowing the executive to trade in the market asset enables her to partially hedge this risk. Suppose she holds a cash amount \( \theta_s \) in \( M \) at time \( s \) (satisfying the integrability condition \( \mathbb{E} \int_t^T \theta_u^2 du < \infty \)) and invests the remainder of her wealth at the riskless rate \( r \). The employee’s wealth \( W = W^{\theta, I(t)} \) consists of two parts - her trading wealth from investment in the market asset and a cashflow \( I(t) \) which later will represent the cash income from option exercise. We have

\[
dW_u = (rW_u + \theta_u(\mu^M - r))du + \theta_u\sigma^M dZ_u + dI_u; \quad W_t = w
\]

The employee’s goal is to maximize expected utility of terminal wealth at some future date \( T \) (which we take to be the maturity of the option), and maximization is taken over the choice of adapted hedging strategies \( \theta = (\theta_u)_{t \leq u \leq T} \). Define

\[
\bar{U}(w) = \bar{U}(w; t) = \sup_{\theta} \mathbb{E}[U(W_{T})|W_t = w]
\]

which is the value function or indirect utility for the terminal wealth obtainable without any income \( (I(t) = 0) \). Then the certainty equivalent value of the income \( V(I(t)) \) is given by the solution to

\[
\bar{U}(w + V(I(t))) = \sup_{\theta} \mathbb{E}[U(W_{T}^{\theta, I(t)})|W_t = w].
\]

and thus

\[
V(I(t)) = \bar{U}^{-1}(\sup_{\theta} \mathbb{E}[U(W_{T}^{\theta, I(t)})|W_t = w]) - w
\]

**Proposition 5.1** For \( V \) given in (13), Hypothesis 3.5 holds.

**Proof:**

It is sufficient to show that if \( J^{(t)}_u \geq J^{(t)}_v \) \( \forall u \) then for each \( \theta \), \( W_{T}^{\theta, I(t)} \geq W_{T}^{\theta, J(t)} \). Then for increasing \( U \), we have \( V(I(t)(\tau)) \geq V(J(t)(\tau)) \) from (13) as required. In fact, \( W_{T}^{\theta, I(t)} - W_{T}^{\theta, J(t)} = \int_t^T r(W_u^{\theta, I(t)} - W_u^{\theta, J(t)})du + I(t) - J(t) \) so that \( W_{T}^{\theta, I(t)} \geq W_{T}^{\theta, J(t)} \) for each \( s \).

\(^3\)Given the stock is a traded asset, if the CAPM relation did not hold, we would have arbitrage possibilities when \( \rho = \pm 1 \). See Davis [10]. The CAPM drift choice is natural as it leads to monotonicity of the (CARA) utility indifference price in \( |\rho| \) (see Henderson [17], Grasselli and Henderson [16] and Frei and Schweizer [14] in a non-Markovian model with stochastic correlation). Leung and Sircar [28] do not require the CAPM drift to hold, and find an asymmetry in the prices and exercise boundaries for positive and negative correlations of the same magnitude. As argued by Carpenter, Stanton and Wallace [7], “this is because they hold the mean return on the stock fixed as they vary correlation, so that the stock has an abnormal return with respect to the hedging instrument, which is larger the smaller or more negative, the correlation”.

18
Now assume \( I^{(t)}_s(\tau) = (X_s - K)^+ + I_{s \geq \tau} \) is the income from option exercise at \( V \forall t \leq \tau \leq T \). Optimizing over the exercise time, the utility indifference value of the option is

\[
A = A(w, x, t) = \sup_{\tau} V(I^{(t)}_s) = \bar{U}^{-1} \left( \sup_{\theta} \sup_{\tau} E[U(W\theta^{(t)}_t)] | W_t = w, X_t = x \right) - w \quad (14)
\]

First, we return to the dividend comparison result of Theorem 3.9 and show it holds in the context of utility indifference pricing. This example recovers the result of Carpenter, Stanton and Wallace [7].

**Proposition 5.2**. Suppose \( \hat{X} \) solves (10) with \( \hat{q} \) and \( \bar{X} \) solves (10) with \( \bar{q} \), and suppose \( \bar{q} \geq \hat{q} \). For \( V \) given in (13), the conclusion of Theorem 3.9 holds: \( \bar{C} \subseteq \hat{C} \).

**Proof:**
Since \( \hat{X}_s \geq \bar{X}_s \forall s \), \( \hat{I}^{(t)}_s(\tau) \geq \bar{I}^{(t)}_s(\tau) \) as in Theorem 3.9. Then, Proposition 5.1 says Hypothesis 3.5 holds, hence Hypothesis 3.1 holds and by Theorem 3.9, \( \bar{C} \subseteq \hat{C} \). \( \square \)

We now consider two specifications for \( U \) and solve the dynamic programming problem resulting from (14). Note that since (14) is treating a single option, although we display thresholds for different option characteristics, these should be interpreted to be for stand alone or individual options, rather than portfolios.

**CARA utility**
We specialize to take \( U(x) = -e^{-\gamma x}; \gamma > 0 \) and illustrate the optimal exercise boundaries in Figure 4. The free boundary problem resulting from (14) is as stated in Leung and Sircar [28] (and Grasselli and Henderson [16] for the perpetual version). We see the exercise and continuation region are separated by a single threshold, as was the case under the risk-neutral setting in Theorem 4.2. In contrast to Theorem 4.6, the threshold does not have to be monotone in time (see also Rogers and Scheinkman [33]).

Since \( V \) is monetary, we expect the continuation regions for different strikes and maturities to be nested according to Theorem 3.7. Panel (a) shows that this is indeed the case. Of course, if strikes and maturities are not co-monotonic then the thresholds may intersect, as shown in Panel (b).

We can compare the exercise thresholds to the equivalent Black Scholes complete market thresholds in Figure 1. As we would anticipate, the thresholds from the utility indifference model are lower than those from the risk neutral model. The agent is not willing to wait for the price to rise as high before exercising because she is risk averse.

Our next example (extended from Carpenter, Stanton and Wallace [7]) shows that in contrast to the example with CARA preferences, we need not have a single threshold separating the continuation and exercise regions. This is an important observation from the perspective of employee stock option modeling, as there is a strand of literature (see for example Cvitanic, Wiener and Zapatero [9]) which assumes an exogenous form for a boundary of threshold form, rather than deriving the threshold(s) endogenously via a utility (or other) model. It is therefore important to recognize the limitations of such an exogenous specification.

**Disconnected continuation regions**
We again solve the dynamic programming problem from (14) with \( U(W) = \frac{W^{1-A}}{1-A} + cW \). We
simplify the calculations by setting $\theta = 0$, i.e. we do not allow hedging in the asset $M$. This does not alter the observations we wish to make.

The resulting thresholds are displayed in Figure 5. Consider the dashed lines which give boundaries for a call with $K = 1.5, T = 15$. We first describe the graph. Starting from low prices and working upwards to high prices - there is a continuation region for low prices, then an exercise region, then a second continuation region, and then a second exercise region at high prices. The intuition is as follows. For low prices/low wealths, $U$ is like CRRA utility, so exercise occurs when prices become too large relative to wealth and the position is too risky. This explains the presence of the lowest exercise threshold. The continuation region at higher price levels arises from the influence of the risk neutral part of the utility. Dividends are very small (close to zero) and hence a risk neutral agent has an incentive to continue. Indeed, if dividends were zero, this would be the whole story as the agent would continue at high prices everywhere. However, the presence of small dividends induces an exercise region at very high prices. The set of solid lines on the figure give boundaries for an American call option with $K = 1, T = 10$. We see that the continuation region for higher prices does not exist during the whole life of the option and is not present as we approach maturity. We stress that the two options have been treated independently and thus the figure does not give any information about a portfolio of the two options. Observe that the conclusions of Theorem 3.7 hold and the continuation regions are nested (i.e. an agent with a low strike/maturity option would exercise before an identical agent with a high strike/maturity option) even though the hypotheses of Theorem 3.7 are not satisfied as the valuation is not monetary. Thus although counterexamples to the nesting of continuation regions exist, such nesting remains typical behavior.

Carpenter, Stanton and Wallace [7] give a result concerning the existence of a boundary of threshold form for utility models - specifically, when prices follow exponential Brownian motions they give conditions which are always satisfied by CARA, but only sometimes satisfied by CRRA.
Figure 5: The optimal exercise region(s) when pricing according to (14) with \( U(W) = \frac{W^{1-A}}{1-A} + cW \). The solid lines give boundaries for an American call with \( (K = 1, T = 10) \). The dashed lines give boundaries for an American call with \( (K = 1.5, T = 15) \). Note the two options are being treated independently and thus the boundaries do not apply to a portfolio of both options. Other parameters are: \( \nu = r = 0.05, q = 0.01, \sigma = 0.6, A = 10, c = 0.0001, w = 1.2 \).

6 Portfolios of Options

In this section we consider the exercise of portfolios of American call options. We have in mind the situation described earlier, where an employee is granted options periodically and thus builds up an inventory of American call options with varying strikes, vest dates and maturities. Of course, if we were content to assume risk neutrality, a portfolio can be considered as a set of independent options each with its own exercise strategy and value. In this case, Theorem 3.7 can be applied. However, we are interested in incomplete market situations whereby perfect hedging is not available. An important consequence is that we cannot decompose a portfolio into a set of independent options - rather the agent’s strategy concerning an option will be altered by the existence of other options in her portfolio. This was shown to be the case in Grasselli and Henderson [16] and Leung and Sircar [28] where portfolios of options with identical characteristics were considered under CARA preferences. Each option was shown to be exercised at a different threshold level as the agent chooses to unwind risk over time. Note that under risk neutrality, a set of identical options would all be exercised at a single threshold.

We aim to develop an exercise ordering result for portfolios, similar in spirit to the earlier Theorem 3.7, which holds for general preferences, including under utility indifference pricing. The previous section showed exercise and continuation regions from utility indifference pricing may be complex, and that it is difficult to say much in generality. This guides us away from analyzing boundaries and towards finding a model-free approach. In this section we make no assumptions on the price process \( X \) and return to the setup of Section 3. Then, given a portfolio of American call options with co-monotonic strikes, maturities and vest dates, we prove that optimal behavior always involves exercising the shorter-dated, lower-strike option first. This can be shown using a coupling construction where it is shown that an agent who exercises in order of
increasing strike/maturity always generates an amount of wealth that dominates that generated by an agent who follows any other exercise strategy.

Consider an agent with a portfolio of American-style call options. The portfolio consists of \( n \) options with characteristics \( (K^{(i)}, T^{(i)}, V^{(i)})_{1 \leq i \leq n} \) with vesting dates such that \( t \lor V^{(i)} \leq T^{(i)} \). If the option with label \( i \) has no vesting, or has already vested, then it is equivalent to suppose that the vesting date \( V^{(i)} \) is the current time \( t \). We consider an exercise strategy \( S \) to be a collection of \( m \) stopping times (with \( 0 \leq m \leq n \)), and associated labels so that \( S = \{(\tau_i, \ell_i)_{1 \leq i \leq m}\} \) where the sequence \( (\tau_i)_{1 \leq i \leq m} \) is non-decreasing, and \( \ell_i \) denotes label of the option which is exercised at \( \tau_i \). In this section, we allow agents to exercise options which are out-of-the-money (ie. pay \( K^{(i)} > X^{\tau_i} \) to receive cash value \( X^{\tau_i} \)). However as we soon argue, such a strategy is clearly sub-optimal. Furthermore, we allow agents to follow a strategy which involves never exercising an option, in which case it is discarded unexercised. This explains why we may have \( m < n \). We say that such an exercise strategy is feasible if it respects the vesting and maturity requirements, so that if \( \ell_i = l \) then \( t \lor V^{(l)} \leq \tau_i \leq T^{(l)} \).

Now we discuss the optimality, or otherwise, of various exercise strategies. The first result is completely natural. The second depends on Hypothesis 3.5 (prefer more to less and cash now rather than later), and on an ordering property for the characteristics of the options.

**Proposition 6.1** Suppose Hypothesis 3.1 holds. Any exercise strategy which involves paying \( K^{(i)} \) to receive \( X^{\tau_i} \) when \( K^{(i)} > X^{\tau_i} \) is sub-optimal. Any exercise strategy which leaves unexercised options which expire in the money is sub-optimal.

Consider a portfolio of options with characteristics \( (K^{(i)}, V^{(i)}, T^{(i)})_{1 \leq i \leq n} \). We say that the strike, vesting date and maturity are co-monotonic if there is a relabeling of the options, represented by a permutation \( \sigma \) of the labels, such that

\[
K^{(\sigma(i))} \leq K^{(\sigma(j))}; \quad T^{(\sigma(i))} \leq T^{(\sigma(j))}; \quad V^{(\sigma(i))} \leq V^{(\sigma(j))}.
\]

**Theorem 6.2** Suppose Hypothesis 3.5 holds. Suppose that the characteristics of the options are such that the strike, vesting date and maturity are co-monotonic.

Then, for any exercise strategy, there is a modified exercise strategy in which the options are exercised in order of increasing strike (and then also maturity and vesting date) for which the value of the option portfolio is at least as large as the value under the original exercise strategy.

**Corollary 6.3** Suppose Hypothesis 3.5 holds, and that the characteristics of the options are such that the strike, vesting date and maturity are co-monotonic.

Then, in searching for optimal strategies, it is sufficient to look in the class of strategies for which no option is exercised unless it is in the money, every option is exercised if it has reached maturity (and is discarded unless it is in the money); and options are exercised in order of increasing strike.

The proof of the Theorem shows that any exercise strategy may be improved upon if on the same set of exercise dates, the options are instead exercised in order of increasing strike/maturity/vest date. Such an ordered strategy generates at least as much wealth as the original strategy - we show this by showing the cumulative amount spent on strikes at each exercise date is never
more (and could be less) for the ordered strategy. The key point is to show that for the ordered strategy, it is always feasible to exercise the relevant option - that is, that it has vested and has not expired. This is where the co-monotonicity is important.

**Proof of Theorem 6.2**
Without loss of generality we may assume that the options are labelled such that $\sigma$ is the identity permutation and then for $i < j$,

$$K^{(i)} \leq K^{(j)} \quad T^{(i)} \leq T^{(j)} \quad V^{(i)} \leq V^{(j)}.$$

Fix an element of the sample space $\omega$. Consider a first (male) agent, and suppose he follows the strategy $S^M = \{((\tau_j^M, \ell_j^M))_{1 \leq j \leq m}\}$, with resultant cashflow

$$I_s^{(i), M} = I_s^{(i), M}((\tau_j^M, \ell_j^M)) = \sum_{j: \tau_j^M \leq s} (X_{\tau_j^M} - K^{(\ell_j^M)}).$$

Now consider a second agent (who we take to be female). Suppose this agent exercises options on the same dates as the male agent. In each case she exercises the option with lowest label which has vested, but not yet expired. (We argue below that this set is non-empty, so that there is an option she can exercise). Write her strategy as $S^F = \{((\tau_i^F, \ell_i^F))_{1 \leq i \leq m}\}$; then $\tau_i^F = \tau_i^M$, though the labels $(\ell_i^M, \ell_i^F)$ may be different. Her cashflow is

$$I_s^{(i), F} = I_s^{(i), F}((\tau_i^F, \ell_i^F)) = \sum_{i: \tau_i^F \leq s} (X_{\tau_i^F} - K^{(\ell_i^F)}).$$

Suppose that we can show that for each $j$,

$$\sum_{i \leq j} K^{(\ell_i^M)} = \sum_{i \leq j} K^{(\ell_i^F)}.$$  \hfill (15)

Then if $j = \sup\{k : \tau_k^M \leq s\}$,

$$I_s^{(i), M} = \sum_{i \leq j} (X_{\tau_i^M} - K^{(\ell_i^M)})$$

$$\leq \sum_{i \leq j} X_{\tau_i^F} - \sum_{i \leq j} K^{(\ell_i^F)} = I_s^{(i), F}.$$

Now suppose that the two agents have identical preferences, and give identical valuations to cashflows, or rather suppose that we are considering the choices of a single agent between cashflows. Then, given the cashflows are ordered, under Hypothesis 3.5 we have $\mathcal{V}(I^{(i), M}) \leq \mathcal{V}(I^{(i), F})$, and that the agent does no worse, and may do better by exercising the options in order of increasing strike.

It remains to prove (15).

Given an $m$-tuple $\gamma = (\gamma_1, \ldots, \gamma_m)$ of distinct labels $(\gamma_i \in \{1, \ldots, n\})$ we can define the ordered $m$-tuple $\tilde{\gamma} = (\gamma_{(1)}, \ldots, \gamma_{(m)})$ where

$$\gamma_{(1)} = \min\{\gamma_1, \ldots, \gamma_m\} \quad \gamma_{(k)} = \min\{k \in \{\gamma_1, \ldots, \gamma_m\} \setminus \{\gamma_{(1)}, \ldots, \gamma_{(k-1)}\}\}$$

Now, given two $m$-tuples $\gamma, \delta$ we can define a partial order via $\gamma \preceq_m \delta$ if $\gamma_{(i)} \leq \delta_{(i)}$ for each $i \leq m$. We want to show that for each $\ell, \ell^F := (\ell_{i}^F)_{1 \leq i \leq j} \preceq (\ell_{i}^M)_{1 \leq i \leq j} :=: \ell^M$, then (15) follows easily from the monotonicity of the sequence $K^{(i)}$. 

23
Note that, by construction, the elements $\ell_i^F$ are increasing, so that $\tilde{\ell}_i^F = \ell_i^F$. (If that were not the case, then we would have $\ell_k^F < \ell_i^F$ for some $k > i$. Then the option with label $\ell_k^F$ would have vested by the date $\tau_k^F$, since the the option with label $\ell_k^F$ has vested; moreover this option cannot have matured, since it has still not matured by date $\tau_k^F$. Hence this option was available to be chosen by the female agent on date $\tau_i^F$, contradicting the assumption that the female agent exercises the option with smallest label.)

Fix $j \leq m$, and recall $\tilde{\ell}_i^M$ is the ordered set of labels corresponding to $(\ell_i^M)_{1 \leq i \leq j}$. For each $k \leq j$ we want to show that $\tilde{\ell}_k^M \equiv \ell_k^M(\omega) \geq \ell_k^F$. Suppose this is true for $k \leq r - 1$. If we can show that $\tilde{\ell}_r^M > \ell_r^F$, and that $V(\ell_r^M) \leq \tau_r^F \leq T(\ell_r^M)$, then the option with label $\tilde{\ell}_r^M$ has vested, but not yet matured on date $\tau_r^F$, and has not yet been exercised by the female agent. Hence, this option is available to be exercised by the female agent, and since she exercises the option with smallest label, $\ell_r^F \leq \tilde{\ell}_r^M$ as required. But $\tilde{\ell}_r^M > \ell_r^F \geq \ell_r^F - 1$ by the inductive hypothesis. Moreover, note that for any subset $\mathcal{L} \subseteq \{\ell_i^M : 1 \leq i \leq j\}$ of size $r$, $\max\{l \in \mathcal{L}\} \geq \max\{\tilde{\ell}_i^M : 1 \leq i \leq r\} = \tilde{\ell}_r^M$. Then, since the male agent has exercised all the options with labels $\ell_i^M : 1 \leq i \leq r$ by date $\tau_r^M = \tau_r^F$,

$$\tau_r^F \geq \max\{V(l) : l \in \{\ell_i^M : 1 \leq i \leq r\}\} \geq \max\{V(l) : l \in \{\tilde{\ell}_i^M : 1 \leq i \leq r\}\} = V(\tilde{\ell}_r^M)$$

and the option with label $\tilde{\ell}_r^M$ has vested. Similarly,

$$\tau_r^F \leq \min\{T(l) : l \in \{\ell_i^M : r \leq i \leq j\}\} \leq \min\{T(l) : l \in \{\tilde{\ell}_i^M : r \leq i \leq j\}\} = T(\tilde{\ell}_r^M)$$

and the option with label $\tilde{\ell}_r^M$ has not yet matured.

Although Theorem 6.2 is stated for fixed maturities, since the proof fixes $\omega$, it also holds for ordered but random maturities $T^{(1)}(\omega) \leq \ldots \leq T^{(i-1)}(\omega) \leq T^{(i)}(\omega) \leq \ldots \leq T^{(n)}(\omega)$. This is useful in the context of employment termination. In this case we could set the maturities to be $T^{(i)} = T^{(i)}(\omega) \wedge \tau^\gamma$, where $\tau^\gamma$ is an independent random time (e.g. exponentially distributed with intensity $\gamma$) representing the termination time. Employees with stock options may either leave their employment either voluntarily or non-voluntarily, retire from their position, or die. In each of these cases, although the legal terms may differ, typically, non-vested options are cancelled, vested out-of-the-money options are cancelled, and vested in-the-money options must be exercised, perhaps within a short window of time. A second interpretation of $\tau^\gamma$ could be the time of an exogenous income shock which forces the agent to exercise her portfolio for liquidity reasons. Again, the conclusion is that optimal strategies are associated with exercising in label order of increasing characteristics.

**Utility Indifference pricing**

We return to the model given in Section 5, here the employee’s goal is to maximize expected utility of terminal wealth at some future date $T$ ($T \geq \max(T^{(i)})$). Given Proposition 5.1, it is immediate that the conclusions of Theorem 6.2 hold. Thus, in the utility indifference model, under co-monotonicity of strikes and maturities/vest dates, we know it is optimal to exercise in order of increasing strike.

Now assume that the cashflow $I^{(i)}$ is the income from exercise of options in the portfolio. For an exercise strategy $\mathcal{S} = (\tau_i, l_i)_{1 \leq i}$ we write $W = W^\theta, \mathcal{S} = W^\theta, I^{(i)}(\mathcal{S})$ where

$$I_u^{(i)}(\mathcal{S}) = \sum_{\tau_i \leq u} (X_{\tau_i} - K^{(i)})$$

24
Figure 6: Exercise thresholds for a portfolio of two options calculated from (16) with $U(x) = -e^{-\gamma x}$ via dynamic programming. The solid line gives the threshold for call with $(K^{(1)} = 10, T^{(1)} = 5, V^{(1)} = 0)$. The dashed line in each panel is the threshold for the call with $(T^{(2)} = 6, V^{(2)} = 0)$ and different strike in each panel as indicated. Other parameter values are: $\gamma = 0.1, \rho = 0, r = 0.05, q = 0.02, \sigma = 0.4, T = 6 = \max(T^{(1)}, T^{(2)})$.

provided $V^{(l_i)} \leq \tau_i \leq T^{(l_i)}$ and $S$ is feasible.

The utility value of the stream of income $I^{(t)}$ is given by

$$V(S) = V(w, S) = V(w, I^{(t)}_S) = \bar{U}^{-1}(\sup_\theta \mathbb{E}[U(W^\theta, S)|W_t = w, X_t = x]) - w.$$ 

Optimizing over exercise strategies, the utility value of the portfolio of American options is

$$A = A(w, x, t) = \bar{U}^{-1}(\sup_\theta \sup_S \mathbb{E}[U(W^\theta, S)|W_t = w, X_t = x]) - w. \quad (16)$$

These definitions are analogous to those given earlier for the single option and this portfolio value can be calculated by dynamic programming.

The results of Theorem 6.2 greatly simplify the analysis of this problem. In general, if there is no co-monotonicity, we need to consider each label order $(l_1, l_2, \ldots)$. Once we have solved for the optimal hedge and the exercise times $\tau_1, \tau_2, \ldots$ we can do a final optimization over the labels (exercise order). This involves solving up to $n!$ separate optimization problems (each of which itself involves optimizing over hedging strategies and exercise times, conditional on exercising in a particular order) and then finally optimizing over label order. However, in contrast, if $K, V, T$ are co-monotonic, then we know it is optimal to exercise in label order, reducing the problem to a single dynamic programming exercise. Firms often grant employee stock options which are at-the-money and with a fixed vest date of say three years. If such options are granted in a bull market, then their characteristics will indeed be co-monotonic.

**Example - CARA utility**

We assume utility is CARA, $U(x) = -e^{-\gamma x}; \gamma > 0$. We will solve the dynamic programming problem resulting from the valuation in (16) and illustrate the optimal exercise boundaries for a portfolio of two options in Figure 6.

In each panel in the figure, the solid line is the exercise threshold for the call with characteristics $(K^{(1)} = 10, T^{(1)} = 5, V^{(1)} = 0)$. In panel (a), the dashed line is the exercise threshold
for call with \((K^{(2)} = 11, T^{(2)} = 6, V^{(2)} = 0)\), so the characteristics are co-monotonic. We see that as Theorem 6.2 predicts, the low strike/maturity option is exercised first, as it has the lower threshold. Comparing to the stand-alone thresholds in panel (a) of Figure 4, we see that when treated as a portfolio, the threshold for the shorter maturity option has moved downwards, shrinking the continuation region for the shorter maturity option. This is because the presence of the longer maturity option in the agent’s portfolio causes her to exercise the shorter maturity option at lower price levels, in order to unwind some risk. The longer maturity threshold remains the same as the stand-alone threshold in Figure 4 - because once the shorter maturity option is exercised the agent only has the longer maturity option in her portfolio.

In panel (b), the strikes and maturities are not co-monotonic. The dashed line is the threshold for the call with \((K^{(2)} = 8.5, T^{(2)} = 6, V^{(2)} = 0)\). The circle (at around time 4.8 years) indicates the time at which the first option to be exercised switches from the longer maturity to the shorter maturity option. Starting below the lower boundary, one possibility is that we hit the dashed threshold (before 4.8 years) and exercise the \(T = 6\) option first. If so, the relevant boundary for the \(T = 5\) year option is the solid line (which continues to 5 years and is exactly the stand-alone boundary for the \(T = 5\) option given in Figure 4). If, instead, the price hits the solid piece of the lower boundary (which only exists between 4.8 and 5 years), then the \(T = 5\) option is exercised first and the higher solid line over the period (4.8,5) is redundant. Then the relevant boundary for the \(T = 6\) option is the dashed line between 4.8 and 6 years. Again, this is the stand-alone boundary for the \(T = 6\) option, since once the shorter maturity option is exercised, only one option remains.

This example extends the literature on utility based employee stock option models, which previously studied such problems for single (or identical) options. As we see in panel (b), if the strikes and maturities are not co-monotonic, either option could be exercised first, and there are many factors which will influence the precise positioning of the boundaries, and in particular, where they intersect.

Even more complex is the situation with CRRA utility where the computations involve an additional wealth dimension. Carpenter, Stanton and Wallace [7] studied American call exercise in the CRRA utility indifference model for a single option. Given the additional wealth dimension, their problem was already numerically challenging with three state variables (time, stock price, wealth). Again, of course, our exercise ordering result holds and will reduce the computational burden for portfolios.

7 Conclusions

The paper aims to study the optimal exercise of American options in a setting with minimal assumptions on the agent’s preferences, valuation methodology and prices. Our main result tells us that a portfolio of American calls with co-monotonic strikes, maturities and vest dates should be exercised in order of increasing strike. Since employees often receive regular grants of American call options, they should exercise the lower strike options with the least time-to-go before the higher strike options with more time-to-go. Although we concentrate on the American call, similar results will hold for the put. A version of Theorem 6.2 will hold with the proviso that strikes and maturities/vesting dates are counter co-monotonic, and puts will be exercised
in increasing order of maturity but decreasing strike.

We illustrate the portfolio exercise result in a standard CARA utility indifference model and thus give the first treatment of portfolios of options with different characteristics in this setting. Interestingly, our portfolio exercise ordering result can be thought of as a generalization of the single option result of Theorem 3.7. However, the portfolio result does not require that the valuation methodology has the monetary property, and thus holds more widely.

The strength of our results is that we require very few assumptions on preferences and prices. Rather than attempting to solve numerically for thresholds (which depend on preferences and the model for the asset price), we instead ask what can be shown concerning exercise ordering in the absence of such assumptions? This is advantageous for several reasons. It is empirically difficult to ascertain what the true underlying process is, and hence it is useful to develop results which are robust to different price specifications. Although some of our results rely on a diffusion assumption, Theorem 6.2 does not, and thus holds, for example, for models with jumps in prices such as Lévy processes. Second, as Section 4 shows, even under the assumption of risk neutral pricing, exercise and continuation regions can be complex. In incomplete markets (for example, under utility indifference) the regions can be even more complicated, and difficult to characterize, hence the more we can say about how regions “nest” or equivalently, the exercise ordering of a set of options - without having to construct explicitly the thresholds - the better. Finally, there is no consensus on the most appropriate preferences to model agent’s decisions concerning risky outcomes, and our results stand under most popular choices. Although our examples focus on utility indifference pricing, our requirements on preferences are much weaker, and are valid for the S shaped function of prospect theory (Kahneman and Tversky [25]), or valuation using hyperbolic discounting.
References


