

Randomized stopping times and American option pricing with transaction costs

Prasad Chalasani¹, Somesh Jha²

¹ Arizona State University, e-mail: chal@asu.edu

² Carnegie Mellon University, e-mail: sjha@cs.cmu.edu

Abstract In a general discrete-time market model with proportional transaction costs, we derive new expectation representations of the range of arbitrage-free prices of an arbitrary American option. The upper bound of this range is called the upper hedging price, and is the smallest initial wealth needed to construct a self-financing portfolio whose value dominates the option payoff at all times. A surprising feature of our upper hedging price representation is that it requires the use of randomized stopping times (Baxter and Chacon, 1977), just as ordinary stopping times are needed in the absence of transaction costs. We also represent the upper hedging price as the optimum value of a variety of optimization problems. In addition we show a two-player game where at Nash equilibrium, the value to both players is the upper hedging price, and one of the players must in general choose a mixture of stopping times. We derive similar representations for the lower hedging price as well. Our results make use of strong duality in linear programming.

Key words American options – transaction costs – randomized stopping times – martingale measures – super-replication – hedging – pricing – arbitrage – incomplete markets – linear programming – strong duality – complementary slackness – game theory – Nash equilibrium – mixed strategy.

1 Introduction

This paper examines the no-arbitrage pricing of American options in a discrete-time market model where stock trades are subject to proportional transaction costs. We characterize the range of option prices that is consistent with no-arbitrage, and find necessary and sufficient conditions for the absence of arbitrage in the model. Similar results are known for European options, but not for American options. The American case turns out to be significantly more challenging because of the added complication of there being an exercise strategy at the disposal of the option bearer. To help put this paper in context, we review below the related literature on the topic of transaction costs.

The classical option valuation framework of Black and Scholes (1973) and Merton (1973) is based on a replication argument, in a continuous trading model with no frictions,

i.e., unrestricted short-sales and no transaction costs. In the scenario of these papers, when the payoff from an option can be perfectly replicated by trading in the underlying stock(s) and bond, there is a unique arbitrage-free option price, which is given by the initial value of the replicating portfolio. Moreover, as has been made explicit by Harrison and Kreps (1979) and by Harrison and Pliska (1981), in markets where there is a unique equivalent martingale measure the no-arbitrage price of a European option can be expressed as the expected discounted terminal payoff from the option, under this martingale measure. For American options one takes the supremum over all stopping times τ (which represent possible exercise strategies of the option bearer) of the expected discounted payoff at τ , under the martingale measure.

The appeal of the arbitrage-free approach is that it is independent of buyers' and sellers' preferences, attitude towards risk, or probability beliefs. In the presence of transaction costs, however, the replication argument in continuous time breaks down, because perfect replication may require that the portfolio be rebalanced infinitely often, incurring infinite transaction costs. Thus perfect replication is impossible, the market model is incomplete, and there may no longer be a unique arbitrage-free option price.

As a result, several recent papers in the transaction costs area are concerned with defining and characterizing the *range* of option prices that are required by the absence of arbitrage. This work has for the most part focused exclusively on *European* options, and American options have not received any serious treatment, even in the discrete-time case. In the absence of arbitrage, the highest possible price h_{up} of a European option is the smallest initial wealth required in order to *super-replicate* the option payoff at expiration, via a self-financing trading strategy. Thus h_{up} , called the *upper hedging price*, is the minimum price the seller of the option needs to receive in order to construct a portfolio that enables him to meet his obligation to the option buyer at expiration without risk. Similarly, the *lower hedging price* h_{low} is the largest amount that the option buyer can borrow (in the form of money from a bank or stocks for short-selling) to buy the option, so that by means of a self-financing trading strategy he is able to pay off his debt using the option payoff at expiration, without risk. The range $[h_{low}, h_{up}]$ is precisely the arbitrage-free range for the option price; any price outside this range leads to an arbitrage, and no price in the range does.

For European options, a number of authors have derived *expectation representations* of the upper and lower hedging prices, in both discrete and continuous-time settings. The continuous-time papers include those of Jouini and Kallal (1995) (who call these prices *arbitrage bounds*), Cvitanić and Karatzas (1996) (who only consider the upper hedging price and call it the *hedging price*), Cvitanic et al. (1999), (who show an explicit form for the expectation defining the upper hedging price), and Karoui and Quenez (1995) (who call these price-bounds the *selling price* and *buying price*). Papers showing similar results in discrete-time include those of Koehl et al. (1999), Jaschke (1996), and Ortu (1996). Our use of the terms upper and lower hedging price, and some other terms and definitions, is borrowed from Karatzas and Kou (1998), who give expectation expressions for the no-arbitrage price-range of an American option with constraints on hedging portfolios.

These results have the following flavor. The upper (lower) hedging price can be expressed as the supremum (respectively, infimum) of the expected discounted terminal payoff from the option, over all probability measures that make the discounted stock price “nearly” a martingale, in a certain precise sense. For example in Jouini and Kallal (1995) the “near-martingale” measures are those under which some process (appropriately discounted) that lies between the bid and ask prices of the stock becomes a martingale. Note

that the presence of proportional transaction costs on stock purchases and sales is equivalent to assuming that the stock has a bid price process and an ask price process that is never below the bid price process. (This is made clearer later in the paper.)

In *continuous time*, the above arbitrage-based price bounds are principally of theoretical interest, and yield trivial results. For instance, a result conjectured by Davis and Clark (1994) and later proved by Soner et al. (1995), and Levental and Skorohod (1997), states that for a European call option, the upper hedging price is equal to the initial stock price, and the minimal super-replicating strategy is the trivial buy-and-hold strategy: buy one share of the stock and hold it until the option expires. This upper hedging price is typically too high to be used as a reasonable option price in practice, suggesting that something more than no-arbitrage is needed to obtain a satisfactory pricing theory for options in the presence of frictions. Despite this negative result, as Cvitanic et al. (1999) argue, the study of super-replicating strategies can be helpful in solving other problems involving transaction costs, and also in finding a more realistic, utility-based price of an option. Examples of the utility-maximization approach to pricing options under transaction costs include the work of Davis et al. (1993) and Cvitanic and Karatzas (1996) on European options, and that of Davis and Zariphopoulou (1995) on American options.

In contrast to the continuous-time situation, it is possible to obtain non-trivial results in discrete time, as the aforementioned papers on European options show.

Expectation representations for American option prices in incomplete but frictionless markets have been shown before. For instance Harrison and Kreps (1979) show that in their frictionless model, if \mathcal{P} is the class of equivalent martingale measures, \mathcal{T} the class of stopping times, and R the discounting factor, then the upper hedging price of an American option with payoff process G is given by

$$h_{\text{up}} = \sup_{\tau \in \mathcal{T}} \sup_{\mathbf{P} \in \mathcal{P}} \mathbb{E}^{\mathbf{P}} (G_{\tau} R^{-\tau}). \quad (1.1)$$

Similarly the lower hedging price is

$$h_{\text{low}} = \sup_{\tau \in \mathcal{T}} \inf_{\mathbf{P} \in \mathcal{P}} \mathbb{E}^{\mathbf{P}} (G_{\tau} R^{-\tau}). \quad (1.2)$$

The chief contribution of our paper is to show various representations for the upper and lower hedging price of American options in the presence of proportional transaction costs. Some of these representations are similar in form to (1.1) and (1.2) above.

1.1 Outline of the paper

In Sections 2 and 3 we introduce our market model and related definitions. We introduce in Section 4 a central notion of this paper: *randomized stopping times*. In our discrete market model a randomized stopping time is a non-negative adapted process X_0, X_1, \dots, X_n (n being the time horizon of the model) whose sum on any path is 1. Such stopping times, also called *fuzzy* stopping times, have been studied in other contexts (Baxter and Chacon, 1977; Dalang et al., 1988). It turns out that they play an essential role in our expectation representations of the upper hedging price of American options with proportional transaction costs. In fact the bulk of the paper consists of showing a variety of representations for the upper hedging price.

Section 5 collects, for later reference, several useful properties of randomized stopping times and related notions. Section 6 introduces another basic concept, that of an *approximate martingale measure*.

Sections 7 defines self-financing portfolios. In Section 8 we borrow notation from Karatzas and Kou (1998) to define the upper and lower hedging prices ((8.2) and (8.4)) for American options under transaction costs, and define what constitutes an arbitrage opportunity. In contrast to the European case, the situations of the buyer and seller here are fundamentally asymmetric, a point that has been made clearly in Karatzas and Kou (1998): The seller needs to hedge against *any* exercise strategy of the buyer, whereas the buyer need only hedge against *some* exercise strategy since he has the choice of it.

Section 9 contains results showing a variety of representations of the upper hedging price of an American option with proportional transaction costs. Not surprisingly, one of these representations, namely (9.2), is similar in form to the above expression (1.1) with zero transaction costs. The differences are that in the case of (9.2) the outer maximization is over randomized stopping times, and the inner maximization is over approximate martingale measures.

We find it surprising that the upper hedging price representation requires randomized exercise strategies. We present two other results that provide some insight into why randomized strategies appear to be necessary to represent the upper hedging price when there are transaction costs. The first involves Nash equilibria in game theory, where *mixed strategies* are often necessary. In Section 10 we define a natural game between a seller and a “devil” where it turns out that at Nash equilibrium, the seller’s “cost” and the devil’s “utility” are both equal to the upper hedging price, and the devil must in general choose a *mixture* of stopping times (a notion closely related to randomized stopping times, see Section 4). The second result is presented in Section 11 where we make a distinction between an *oblivious hedge* (a single portfolio that hedges against all exercise strategies) and an *adaptive hedge* (a possibly different portfolio to hedge against each exercise strategy). With zero transaction costs, the minimum initial wealth needed for an adaptive hedge also suffices for an oblivious hedge. But with non-zero transaction costs, an oblivious hedge may require more initial wealth.

In Section 12 we show that the lower hedging price for an American option with proportional transaction costs can be represented by an expression analogous to (1.2) above. The difference is that with non-zero transaction costs, the inner minimization is over approximate martingale measures. We prove in Section 13 various conditions that are equivalent to the absence of arbitrage in our model. Section 14 contains a proof that our results imply those of Koehl et al. (1999) (and others mentioned above) for European options with proportional transaction costs. In Section 15 we show examples that illustrate our main results. Section 16 concludes with a discussion of possible future research directions.

The techniques we use to derive our expectation representations of h_{up} and h_{low} are elementary, and depend only on *strong duality of linear programming*. In the case of h_{up} we observe that its definition is a linear programming problem. Strong duality assures us that the dual optimum equals the primal optimum, so after some algebraic manipulation of the dual linear program, we are able to interpret its variables in terms of an approximate martingale measure and a randomized exercise strategy. In the case of h_{low} our approach is much the same, except that h_{low} is a linear programming problem only for a fixed stopping time τ . We are also able to show that the existence of a certain approximate martingale measure is necessary and sufficient for the absence of arbitrage in our market model (Theorem 13.1).

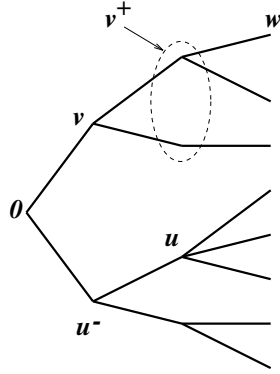


Fig. 2.1 Illustrating notation on an event tree with $n = 3$. For instance on this tree, $w \in T$ and $t(u) = 2$

The form of our expectation formulas is closest to the discrete-time results of Koehl et al. (1999) for European options. However they appear to use much more complex methods than ours. We show in Section 14 that results very similar to theirs follow from our results for American options. Jaschke (1996), Naik (1995) and Ortu (1996) obtain results similar to ours via linear programming, but only for European options. Other authors have used linear programming to study hedging strategies for European options with transaction costs. These include Edirisinghe et al. (1993), who also show efficient approximations to the linear program defining an optimal super-replication strategy; and Bensaid et al. (1992), who were one of the first to point out that a super-replicating hedge strategy may have less initial cost than a perfectly replicating one.

2 Market model

We model discrete time by means of a finite **event tree** which contains nodes corresponding to **trading dates** (or **time-steps**) $0, 1, \dots, n$, where n is the expiration date of the option under consideration. Nodes will be identified by letters such as u, v, w . There is exactly one node at time-step 0, called the **root node**, and it is denoted 0; this node represents the present time. The set of nodes at time n is denoted T ; nodes in T are called **terminal nodes** or **leaves**. Each non-terminal node u has a *non-empty* set of **immediate successors** denoted u^+ . If u is a terminal node, u^+ will be treated as the empty set. The unique **immediate predecessor** of a node $u \neq 0$ is denoted u^- . We use the term **path** to refer to a sequence of nodes from the root to a leaf, where each node but the last is the predecessor of the next node in the sequence. Paths will be denoted by letters such as ω, ω' , etc. The node at time i on path ω is denoted ω_i . We write $u \in \omega$ to say that node u is on the path ω . The **time** corresponding to a node is denoted $t(u)$. Thus, $t(0) = 0$ and for any $u \in T$, $t(u) = n$. If u and v are nodes on a given path, and $t(v) > t(u)$ we write for brevity $v > u$ or $u < v$. We also write $u \leq v$ to indicate that u is either equal to v or a predecessor of v . Figure 2.1 illustrates some of this notation. For ease of reference a summary of the notation used in this paper is provided in Appendix B.

All random variables in this paper will be defined with respect to the measurable space (Ω, \mathcal{F}) where Ω is the set of all paths in the tree from time 0 to time n , and \mathcal{F} is the σ -algebra consisting of all subsets of Ω . Note that each node in the tree represents the set of all tree paths that contain that node. This allows us to define a natural *filtration*

$F_0, F_1, \dots, F_n = F$ (denoted $\{F_k\}$ for brevity) on the space Ω , where F_k is the σ -algebra generated by the path-sets represented by all the time- k nodes. In other words, F_k represents the information known at time k : at time k , market participants only know which atom of F_k the tree path belongs to.

For our purposes, any real-valued function F of the tree nodes defines a *random process* $\{F_0, F_1, \dots, F_n\}$. Here F_k is the random variable that assigns to any path ω in Ω , the value of the process F at ω_k . We will denote the process either by $\{F_k\}$ or simply by F . By definition, such a process is *adapted* to the filtration $\{F_k\}$. If u is a node in the event-tree, then we write $F(u)$ for the value of the process F at u , whereas if k is a non-negative integer and $\omega \in \Omega$ is a path, then we write either $F_k(\omega)$ or $F(\omega_k)$ for the value of F at time k on path ω .

To avoid notational clutter, we will only consider a market consisting of a single riskless security called the *bond* and a single risky security called the *stock*. Our results can be easily extended to the case of multiple risky assets. The stock price process is denoted by S , which means S_k is the stock price random variable at time k , and $S(u)$ is the stock price at a specific node u . We assume trading in the stock is subject to *proportional* transaction costs. In particular, to buy a stock whose market price is S dollars, one pays $(1 + \lambda)S$ dollars, and upon selling a share at price S , one receives $(1 - \mu)S$ dollars, where $\lambda \geq 0$, and μ lies in the interval $[0, 1]$. In other words, λ and μ are the transaction costs per dollar of share bought and sold respectively. Another way to say this is that the **bid price** of the stock at node u is $S(u)(1 - \mu)$ and its **ask price** is $S(u)(1 + \lambda)$. We assume that the *risk-free interest rate* is r , and that one dollar in bonds at time k will grow to $R = 1 + r$ dollars at time $k + 1$. For any adapted process Z , the corresponding **discounted process** will be denoted Z^* , or in other words, $Z_k^* := Z_k R^{-k}$. (Here and elsewhere, we follow Williams (1991) in using the notation “:=” to stand for “is defined as”, and “=:” to stand for “is the definition of”).

An **American option** on the stock is specified by a *non-negative* adapted process G , where $G(u)$ is the payoff from exercising the option at node u in the tree. The function $G(u)$ may be any arbitrary path-dependent function, and not just a function of the stock price $S(u)$ at u . An American option can be exercised at any time until it expires. The exercise strategy of the holder of an American option can be described by a **stopping time**. In our model a stopping time τ is a random variable that maps each path $\omega \in \Omega$ to a number in $\{0, 1, \dots, n\}$, with the restriction that for any k , the indicator-random variable $I_{\tau=k}$ is F_k -measurable. In other words, if there is some path ω with $\tau(\omega) = k$ and $\omega_k = u$, then for every path ω' containing u , we must have $\tau(\omega') = k$. For any path ω , there is thus a unique **stopping point** $\omega_{\tau(\omega)}$. For any adapted process $\{F_k\}$, we use the standard notation F_τ to denote the random variable that maps a path $\omega \in \Omega$ to $F_{\tau(\omega)}(\omega)$. The set I_τ of all stopping points of a stopping time τ is called the **stopping boundary** of τ . The set of all stopping times τ is denoted T .

3 Measures and Node-measures

A non-negative node-function \mathbf{P} is called a **probability measure** (or simply “measure” for short) on the event tree if $\mathbf{P}(0) = 1$, and for any node $u \notin T$, the sum of \mathbf{P} over all immediate successors of u equals $\mathbf{P}(u)$, i.e.,

$$\sum_{v \in u^+} \mathbf{P}(v) = \mathbf{P}(u). \quad (3.1)$$

We can think of $\mathbf{P}(u)$ as the probability of reaching u , or as the total probability of all paths containing node u . Note that a non-terminal node with positive probability must have at least one successor with positive probability. For any path $\omega \in \Omega$ we define the **probability $\mathbf{P}(\omega)$ of the path ω** to be the probability of its last node, i.e., $\mathbf{P}(\omega) := \mathbf{P}(\omega_n)$. When we say a certain relationship among random variables (defined below) holds **\mathbf{P} -almost surely**, (abbreviated **\mathbf{P} -a.s.**) we mean that it holds on every path with positive probability under \mathbf{P} . The expectation of a random variable under \mathbf{P} is denoted $\mathbb{E}^{\mathbf{P}}(Z)$, and the conditional expectation $\mathbb{E}^{\mathbf{P}}(Z|F_k)$ is abbreviated $\mathbb{E}_k^{\mathbf{P}}(Z)$. Note that $\mathbb{E}_k^{\mathbf{P}}(Z)$ is a random variable, and its value on a particular path $\omega \in \Omega$ is therefore $\mathbb{E}_k^{\mathbf{P}}(Z)(\omega)$; if $\omega_k = u$ then for brevity we denote this value by $\mathbb{E}_u^{\mathbf{P}}(Z)$.

We will also find it convenient to use the notion of a **node probability measure** (or just “node-measure” for short), which is simply a non-negative node-function q such that $\sum_u q(u) = 1$. (Here and elsewhere in the paper, a summation sign, subscripted by a node-name, such as u , stands for a summation over *all* nodes in the event tree.) The **support** of a node-measure q is the set of nodes u such that $q(u) > 0$; we write $u \in q$ if u is in the support of q . The set of all possible node-measures is denoted Q .

Node-measures whose support is on the stopping boundary a stopping time will be called **simple** node-measures. In other words q is a simple node-measure if and only if there do *not* exist nodes u, v such that $u < v$, $q(u) > 0$ and $q(v) > 0$, i.e., on every path in the event-tree there is at most one node where q is strictly positive. We write Q^τ to denote the set of simple node-measures whose support lies on the boundary of a stopping time τ .

4 Randomized Stopping Times and Mixed Strategies: Definitions

An ordinary exercise strategy is specified by a stopping time τ , which specifies exactly one exercise point on every event-tree path. A natural generalization is to allow *randomized* (or “fractional” or “fuzzy”) stopping times; these will play a central role in this paper.

Definition 4.1 A **randomized stopping time** (or **randomized exercise strategy**) is a non-negative adapted process (i.e. node-function) X with the property that on every path ω in the event tree,

$$\sum_{i=0}^n X(\omega_i) = 1,$$

i.e., the sum of the random variables X_0, X_1, \dots, X_n is 1 on every path (see Figure 4.1). We denote the collection of all randomized stopping times by X . The **support** of a randomized stopping time $X \in X$ is the set of nodes u such that $X(u) > 0$; we write $u \in X$ if u is in the support of X .

Randomized stopping times have been studied, among others, by Baxter and Chacon (1977), Dalang et al. (1988), Edgar et al. (1981) and Mazziotto and Millet (1986), primarily as an aid to show the existence and properties of optimal ordinary stopping times in certain optimal control problems. In certain optimization problems, randomized stopping times can be viewed as a natural linear *relaxation* (Grötschel et al., 1993) of ordinary stopping times. When a randomized stopping time X is used to describe an exercise strategy, we can think of $X(u)$ as the probability of exercising the option at u given that the node u has been reached.

Remark 4.2 Note that an ordinary stopping time τ corresponds to the degenerate randomized stopping time X^τ whose values are restricted to lie in the set $\{0, 1\}$, and defined as follows for any $\omega \in \Omega$ and $k = 0, 1, \dots, n$:

$$X^\tau(\omega_k) := \begin{cases} 1 & \text{if } \tau(\omega) = k, \\ 0 & \text{otherwise.} \end{cases} \quad (4.1)$$

We will sometimes abuse notation and use τ to mean X^τ .

It is easy to see that ordinary (or pure) stopping times are extreme points of the convex set of randomized stopping times (Baxter and Chacon, 1977). Thus a randomized stopping time can always be written as a convex combination, or *mixture*, of pure stopping times (see Fig. 4.1), in a manner reminiscent of mixed strategies in game theory (Fudenberg and Tirole, 1998). Section 10 explores this connection in detail. A mixture of stopping times is formally defined as follows.

Definition 4.3 A **mixture** σ of stopping times is a probability assignment $\sigma(\tau)$ for each $\tau \in T$ such that

$$\sigma(\tau) \geq 0 \quad \forall \tau \in T; \quad \sum_{\tau \in T} \sigma(\tau) = 1.$$

The collection of all mixtures of stopping times will be denoted \overline{T} . The **support** of a mixture σ is defined as the set of stopping times $\tau \in T$ such that $\sigma(\tau) > 0$; we write $\tau \in \sigma$ if τ is in the support of σ . We write τ to denote the mixture with support at just one stopping time τ .

Note that a given mixture σ corresponds to the unique randomized exercise strategy X^σ defined by:

$$X^\sigma := \sum_{\tau \in \sigma} \sigma(\tau) X^\tau. \quad (4.2)$$

However, the converse is not true, i.e., there may be more than one way to express a randomized exercise strategy $X \in X$ as a mixture of pure stopping times.

4.1 Average values

There are many natural ways of defining the “average” value of an adapted process Z on the event tree. First, for a randomized stopping time $X \in X$, we define the **time- X value of Z** as the random variable

$$Z_X := \sum_{i=0}^n (X_i Z_i). \quad (4.3)$$

Note that if $X = X^\tau$ for some pure stopping time τ , then Z_X is the familiar random variable Z_τ . For a specific path $\omega \in \Omega$, it is clear that

$$Z_X(\omega) = \sum_{u \in \omega} Z(u) X(u). \quad (4.4)$$

For a randomized stopping time $X \in X$, we define the **maximum time- X value of Z** as

$$\widehat{Z}_X := \max_{\omega \in \Omega} Z_X(\omega). \quad (4.5)$$

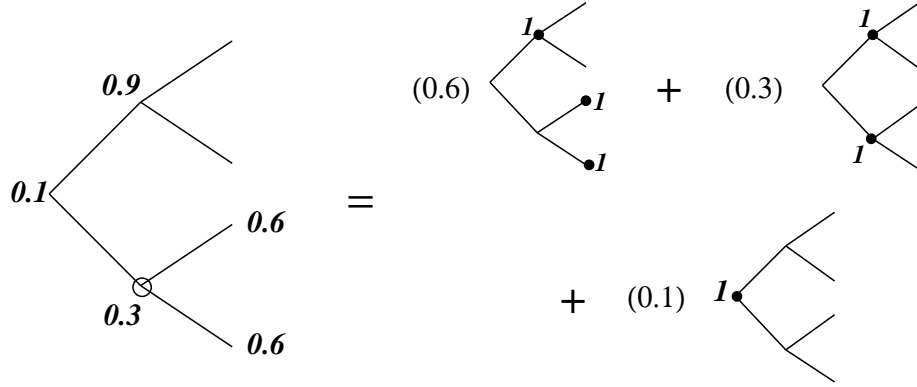


Fig. 4.1 Illustrating how a randomized exercise strategy X on an event tree with $n = 2$ can be viewed as a weighted average of three “pure” stopping times. The numbers on the leftmost tree show the value of X alongside the nodes. Only the positive X values are shown; all others are 0. Note that along any path, the sum of the X values is 1. The residual X -exercise probability at the circled node is 0.6. Each of the three trees on the right shows a “pure” stopping time represented as a 0 – 1 valued randomized stopping time; again only the 1 values are shown. Alongside each of these trees the weight of the corresponding stopping time used to produce X is indicated in parentheses.

For a (pure) stopping time $\tau \in \mathcal{T}$, we define the **maximum value of Z at τ** as

$$\hat{Z}_\tau := \max_{\omega \in \Omega} Z_\tau(\omega), \quad (4.6)$$

and for a mixture $\sigma \in \overline{\mathcal{T}}$ we define the **σ -expected maximum value of Z** as

$$\hat{Z}_\sigma := \sum_{\tau \in \sigma} \sigma(\tau) \hat{Z}_\tau. \quad (4.7)$$

Lastly, for any node-measure $q \in \mathcal{Q}$, we let \bar{Z}_q denote the **q -weighted average of Z** over the entire event-tree:

$$\bar{Z}_q := \sum_u q(u) Z(u). \quad (4.8)$$

It turns out that a useful equivalence relation can be defined between measure-strategy pairs (\mathbf{P}, X) and node-measures.

Definition 4.4 Suppose \mathbf{P} is a probability measure, X is a randomized exercise strategy, and q is a node-probability measure. We say that q is **equivalent** to (\mathbf{P}, X) if $\mathbf{P}(u)X(u) = q(u)$ holds at every node u (See Figure 4.2). We will use the notation $q \equiv (\mathbf{P}, X)$ to mean that q is a node-measure and (\mathbf{P}, X) is an equivalent measure-strategy pair.

Theorem 5.4 shows that for any node-measure q there always exists an (essentially unique) equivalent measure-strategy pair (\mathbf{P}, X) , and vice versa.

4.2 Future average values

We will often want to refer to the future average value of an adapted process at some arbitrary node u . We do this by defining appropriate generalizations of \bar{Z}_q and Z_X . First

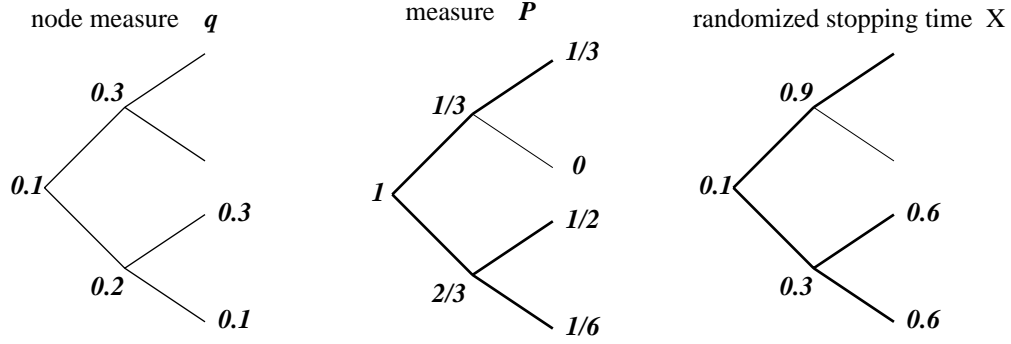


Fig. 4.2 A node-probability measure q and an equivalent measure-strategy pair (\mathbf{P}, X) . Only the positive q and X values are shown. Note that $q(u) = \mathbf{P}(u)X(u)$ holds at every node u . Starting from the (\mathbf{P}, X) values shown on the right, the q values on the left are obtained by simply setting $q(u) = \mathbf{P}(u)X(u)$ at every u . Starting from the q values on the left, the \mathbf{P}, X values on the right are computed as in Theorem 5.4.

we need a notation for the remaining future X -value or q -weight. More precisely, for a randomized exercise strategy X , we define the random variable

$$X_k^+ := \sum_{i=k+1}^n X_i = 1 - \sum_{i=0}^k X_i, \quad k \leq n, \quad (4.9)$$

where the first summation is interpreted as 0 if $k = n$. The second expression above follows from the definition of a randomized strategy, and implies that X_k^+ is F_k -measurable. When X represents a randomized option exercise strategy, for example, X_k^+ can be viewed as the probability that at time k , the option has not yet been exercised at time k or earlier (see Figure 4.1).

Remark 4.5 If X corresponds to an ordinary stopping time τ , then $X_k^+ = 1$ for $k < \tau$, and $X_k^+ = 0$ for $k \geq \tau$. The node-function corresponding to the process X^+ is given by $X^+(v) = 1 - \sum_{u \leq v} X(u)$ for any node v . Note also that if $v \in u^+$ then $X^+(u) = X(v) + X^+(v)$.

For a node-measure q , the node-function analogous to X_k^+ is

$$q^+(u) := \sum_{v > u} q(v), \quad \forall u, \quad (4.10)$$

the sum being interpreted as 0 if u is a terminal node. We write $\bar{q}(u)$ to denote $q(u) + q^+(u)$. Notice an important difference between the node-functions $q^+(u)$ and $X^+(u)$: the former is the sum of $q(v)$ over *all* nodes v that are successors of u , while the latter is the sum of $X(v)$ over all successors v of u that lie on a *specific* path containing u (the choice of path is arbitrary since this sum must be the same on every such path).

We can now define the following generalizations of \bar{Z}_q and Z_X . For any integer $k \in [0, n]$, if $X_k^+ > 0$, the random variable

$$Z_{X|k} := \frac{\sum_{i=k+1}^n (X_i Z_i)}{X_k^+} \quad (4.11)$$

represents the **future time- X value of Z at time k** . Note that if \mathbf{P} is a probability measure on the event-tree, then $\mathbb{E}_k^{\mathbf{P}} Z_{X|k}$ is the *expected* future time- X value of Z at time k ; for brevity we

will write this expectation as $\mathbb{E}_k^{\mathbf{P}} Z_X$. Also for a node u such that $\omega_k = u$ we write $(\mathbb{E}_k^{\mathbf{P}} Z_X)(\omega)$ more compactly as $\mathbb{E}_u^{\mathbf{P}} Z_X$; this quantity is well-defined only if $\mathbf{P}(u) > 0$ and $X^+(u) > 0$. It is important to bear in mind that in the expression $\mathbb{E}_k^{\mathbf{P}} Z_X$, the Z_X stands for $\sum_{i=k+1}^n (Z_i X_i)$, i.e., the term $(Z_k X_k)$ is *not* included in this summation.

If u is a node with $q^+(u) > 0$, we define the q -**weighted average future value of Z** at node u as

$$\bar{Z}_q(u) := \frac{\sum_{v>u} q(v) Z(v)}{q^+(u)}. \quad (4.12)$$

Theorem 5.6 shows that if $q \equiv (\mathbf{P}, X)$ then $\bar{Z}_q = \mathbb{E}^{\mathbf{P}} Z_X$, and $\bar{Z}_q(u) = \mathbb{E}_u^{\mathbf{P}} Z_X$ for all u such that $q^+(u) > 0$.

5 Useful properties

We are now ready to show several useful properties that relate to node-measures, randomized stopping times, and mixtures.

Proposition 5.1 *Suppose $q \equiv (\mathbf{P}, X)$, and let $\sigma \in \bar{\mathcal{T}}$ be a mixture such that $X = X^\sigma$. Then for any adapted process Z ,*

$$\bar{Z}_q \leq \hat{Z}_X \leq \hat{Z}_\sigma.$$

Proof By definition of equivalence, we have $q(u) = \mathbf{P}(u)X(u)$ for all u . Recalling that $\mathbf{P}(\omega)$ represents $\mathbf{P}(\omega_n)$, the probability of the last node on ω , we can then write

$$\begin{aligned} \bar{Z}_q &:= \sum_u q(u) Z(u) = \sum_u Z(u) X(u) \mathbf{P}(u) \\ &= \sum_u \left[Z(u) X(u) \sum_{\omega \in \Omega: \omega \ni u} \mathbf{P}(\omega) \right] \\ &= \sum_{\omega \in \Omega} \mathbf{P}(\omega) \sum_{u \in \omega} Z(u) X(u) \\ &= \sum_{\omega \in \Omega} \mathbf{P}(\omega) Z_X(\omega) \\ &\leq \max_{\omega \in \Omega} Z_X(\omega) =: \hat{Z}_X. \end{aligned}$$

To show the second inequality, note that for any fixed $\omega \in \Omega$ and $\tau \in \sigma$ we have

$$Z_\tau(\omega) \leq \max_{\omega \in \Omega} Z_\tau(\omega) =: \hat{Z}_\tau,$$

and taking the σ -weighted average of both sides over all $\tau \in \sigma$,

$$Z_X(\omega) := \sum_{\tau \in \sigma} \sigma(\tau) Z_\tau(\omega) \leq \sum_{\tau \in \sigma} \hat{Z}_\tau =: \hat{Z}_\sigma,$$

whence taking the maximum of the left hand side over all $\omega \in \Omega$,

$$\hat{Z}_X := \max_{\omega \in \Omega} Z_X(\omega) \leq \hat{Z}_\sigma.$$

■

An easy consequence of the above proposition is:

Proposition 5.2 *Let \mathcal{C} be any closed set of adapted processes. Then*

$$\max_{q \in \mathcal{Q}} \min_{Z \in \mathcal{C}} \bar{Z}_q \leq \max_{X \in \mathcal{X}} \min_{Z \in \mathcal{C}} \hat{Z}_X \leq \max_{\sigma \in \mathcal{T}} \min_{Z \in \mathcal{C}} \hat{Z}_\sigma \leq \min_{Z \in \mathcal{C}} \max_u Z(u),$$

where the last \max_u above is taken over all nodes u in the event tree.

Proof Let q' be an optimal node-measure in the first expression above. By Theorem 5.4 there exists a pair (\mathbf{P}', X') equivalent to q' . Suppose $Z = Z'$ minimizes $\hat{Z}_{X'}$. Then by Proposition 5.1, $\hat{Z}'_{X'} \geq \bar{Z}'_{q'}$, and so

$$\min_{Z \in \mathcal{C}} \hat{Z}_{X'} = \hat{Z}'_{X'} \geq \bar{Z}'_{q'} \geq \min_{Z \in \mathcal{C}} \bar{Z}_{q'} = \max_{q \in \mathcal{Q}} \min_{Z \in \mathcal{C}} \bar{Z}_q,$$

from which the first inequality follows by noting that the leftmost expression above is at most $\max_{X \in \mathcal{X}} \min_{Z \in \mathcal{C}} \hat{Z}_X$. A similar argument proves the second inequality. To show the third inequality, note first that

$$\max_{\sigma \in \mathcal{T}} \min_{Z \in \mathcal{C}} \hat{Z}_\sigma \leq \min_{Z \in \mathcal{C}} \max_{\sigma \in \mathcal{T}} \hat{Z}_\sigma.$$

For a fixed $Z \in \mathcal{C}$, the value of \hat{Z}_σ is maximized by picking σ to be a pure stopping time τ whose stopping boundary contains some node u where $Z(\cdot)$ is maximum. Thus the right hand side above equals $\min_{Z \in \mathcal{C}} \max_u Z(u)$, and this proves the third inequality. ■

Equivalence between a node-measure q and a measure-strategy pair (\mathbf{P}, X) has the following useful consequence.

Proposition 5.3 *Suppose $q \equiv (\mathbf{P}, X)$. Then for any node v we have*

$$\mathbf{P}(v)X^+(v) = q^+(v). \quad (5.1)$$

Proof By definition of an exercise strategy, if v is a time- k non-terminal node, then

$$\sum_{i=k+1}^n X_i(\omega) = 1 - \sum_{u \leq v} X(u)$$

for every path ω containing v , so we can write

$$X^+(v) = 1 - \sum_{u \leq v} X(u) = Y(v), \quad (5.2)$$

where Y is the random variable defined by

$$Y := \mathbb{E}_k^{\mathbf{P}} \left(\sum_{i=k+1}^n X_i \right).$$

Now multiplying (5.2) by $\mathbf{P}(v)$ we get

$$\mathbf{P}(v)X^+(v) = \sum_{w > v} \mathbf{P}(w)X(w) = \sum_{w > v} q(w) = q^+(v).$$

■

Now we turn to the question of constructing an equivalent node-measure given a measure-strategy pair, and vice versa. Not surprisingly, given a node-measure pair (\mathbf{P}, X) , the unique equivalent node-measure q is given by $q(u) = \mathbf{P}(u)X(u)$ for all u . Conversely, given a node-measure q , there is a natural way to define a measure \mathbf{P} , and it turns out that the node-function $X(u) = q(u)/\mathbf{P}(u)$ (with a modified definition if $\mathbf{P}(u) = 0$) is actually a randomized stopping time. Also, the (\mathbf{P}, X) pair defined in this way is “effectively unique”. This is formalized below:

Theorem 5.4 (Existence of equivalent (\mathbf{P}, X) and q) *If (\mathbf{P}, X) is a measure-strategy pair, then the node-function q defined by $q(u) = \mathbf{P}(u)X(u)$ at all u is the unique equivalent node-measure. Conversely, suppose q is a node-measure on the event tree. Then there always exists a measure-strategy pair (\mathbf{P}, X) equivalent to q , where \mathbf{P} is uniquely defined at nodes u where $\bar{q}(u) > 0$, and X is uniquely defined at nodes u where $q(u) > 0$.*

Proof To prove the first statement note that the unique node-function q defined at all u by $q(u) = \mathbf{P}(u)X(u)$, is a node-probability measure: Clearly q is non-negative everywhere, and

$$\sum_u q(u) = \sum_u \mathbf{P}(u)X(u) = \sum_{i=0}^n \mathbb{E}^{\mathbf{P}}(X_i) = \mathbb{E}^{\mathbf{P}} \sum_{i=0}^n X_i = 1.$$

To show the converse, consider some node-measure q . Define the node function (which will turn out to be a probability measure) \mathbf{P} forward-inductively as follows: Set $\mathbf{P}(0) = \bar{q}(0) = 1$ (Recall that $\bar{q}(u) = q(u) + q^+(u)$). For any non-terminal node u , if $q^+(u) = 0$ then set $\mathbf{P}(v) = \mathbf{P}(u)$ for an arbitrary $v \in u^+$ and set $\mathbf{P}(v) = 0$ for all other $v \in u^+$. (The reason for doing this is that if \mathbf{P} is a probability measure then any non-terminal node u with $\mathbf{P}(u) > 0$ must have at least one successor v with $\mathbf{P}(v) > 0$.) If $q^+(u) > 0$, then for each $v \in u^+$, we will define $\mathbf{P}(v)$ so that the ratio $\mathbf{P}(v)/\mathbf{P}(u)$ equals the *proportion* of $q^+(u)$ that is on the subtree starting at v :

$$\mathbf{P}(v) = \mathbf{P}(u) \frac{\bar{q}(v)}{q^+(u)} \quad \forall v \in u^+. \quad (5.3)$$

The \mathbf{P} values defined above are clearly non-negative at all nodes since q is non-negative. For any non-terminal u , if $q^+(u) = 0$, then one of the immediate successors v of u has $\mathbf{P}(v) = \mathbf{P}(u)$ and all other immediate successors v have $\mathbf{P}(v) = 0$, so $\sum_{v \in u^+} \mathbf{P}(v) = \mathbf{P}(u)$. On the other hand, if $q^+(u) > 0$ for some u , then for each $v \in u^+$,

$$\sum_{v \in u^+} \mathbf{P}(v) = \frac{\mathbf{P}(u)}{q^+(u)} \sum_{v \in u^+} \bar{q}(v) = \mathbf{P}(u),$$

Thus \mathbf{P} satisfies the requirements (3.1) of a probability measure on the event tree.

Next define the node function X (which will turn out to be a randomized stopping time) forward-inductively in terms of q and the above-defined \mathbf{P} as follows:

$$X(v) = \begin{cases} q(v)/\mathbf{P}(v) & \text{if } \mathbf{P}(v) > 0, \\ q^+(v^-)/\mathbf{P}(v^-) & \text{if } \mathbf{P}(v^-) > 0, \mathbf{P}(v) = 0, \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

We now argue that X is a randomized stopping time. Consider any non-terminal node u with $\mathbf{P}(u) > 0$. By definition of a measure, u has an immediate successor u' with $\mathbf{P}(u') > 0$. Then we have

$$\begin{aligned} X(u) &= \frac{q(u)}{\mathbf{P}(u)} = \frac{q(u) + q^+(u)}{\mathbf{P}(u)} - \frac{q^+(u)}{\mathbf{P}(u)} \\ &= \frac{\bar{q}(u)}{\mathbf{P}(u)} - \frac{\bar{q}(u')}{\mathbf{P}(u')} \quad (\text{definition of } \mathbf{P}(u')). \end{aligned}$$

Therefore if v is a non-terminal node with $\mathbf{P}(v) > 0$, and v' is an immediate successor of v with $\mathbf{P}(v') > 0$ (there must be at least one such $v' \in v^+$ by definition of a measure), we have

$$\sum_{u \leq v} X(u) = \frac{\bar{q}(0)}{\mathbf{P}(0)} - \frac{\bar{q}(v')}{\mathbf{P}(v')} = 1 - \frac{\bar{q}(v')}{\mathbf{P}(v')}.$$

If v' is a terminal node, then $\bar{q}(v') = q(v')$ and $X(v') = q(v')/\mathbf{P}(v')$, so we have by the first case of the definition (5.4) of X ,

$$\sum_{u \leq v} X(u) = 1 - \frac{q(v')}{\mathbf{P}(v')} = 1 - X(v'),$$

which means the sum of the X values on any positive-probability path on the event-tree is 1. Now suppose that in addition to an immediate successor v' with $\mathbf{P}(v') > 0$, v has another immediate successor w with $\mathbf{P}(w) = 0$. Then by definition of \mathbf{P} we have

$$\sum_{u \leq v} X(u) = 1 - \frac{\bar{q}(v')}{\mathbf{P}(v')} = 1 - \frac{q^+(v)}{\mathbf{P}(v)},$$

which equals $1 - X(w)$ by the second case of the definition (5.4) of X . Note that the X values on all successors of w will be zero, by the third case of the definition of X above. Thus the sum of the X values on a zero-probability path is also 1. Clearly X is non-negative everywhere, so X is indeed a randomized stopping time.

Now we verify that $q \equiv (\mathbf{P}, X)$. From the definition of \mathbf{P} we can see that for any v , if $\mathbf{P}(v) = 0$ then $q(v) = 0$. This fact and the definition of X imply that $\mathbf{P}(v)X(v) = q(v)$ holds at every v .

To see the uniqueness properties, suppose (\mathbf{P}, X) is a measure-strategy pair equivalent to the node-measure q . By definition $\mathbf{P}(0)$ is necessarily 1 for any such measure \mathbf{P} . Let us consider any node $v \neq 0$ where $\bar{q}(v) > 0$, and let $u = v^-$ be its unique predecessor. This means $q^+(u) > 0$ and by Proposition 5.3 we have $q^+(u) > 0$, and

$$\mathbf{P}(u) \frac{\bar{q}(v)}{q^+(u)} = \mathbf{P}(u) \frac{\mathbf{P}(v)X(v) + \mathbf{P}(v)X^+(v)}{\mathbf{P}(u)X^+(u)} = \mathbf{P}(v) \frac{X(v) + X^+(v)}{X^+(u)} = \mathbf{P}(v).$$

Thus the equation (5.3) necessarily holds at all v where $\bar{q}(v) > 0$. It is then easy to see by induction that $\mathbf{P}(v)$ is unique at such v . In particular $\mathbf{P}(v)$ is unique at nodes v where $q(v) > 0$, and at such v , $X(v)$ is therefore also unique, since it must satisfy $\mathbf{P}(v)X(v) = q(v)$. ■

This Theorem immediately implies a corresponding result for simple node-measures:

Corollary 5.5 *If for some pure stopping time τ , (\mathbf{P}, X^τ) is a measure-strategy pair; then the unique equivalent node-measure q is a simple node-measure in Q^τ , defined by $q(u) = \mathbf{P}(u)$ for every u that is on the stopping boundary of τ , and $q(u) = 0$ for all other u . Conversely, if $q \in Q^\tau$ is a simple node-measure, then (\mathbf{P}, X^τ) is the unique equivalent measure-strategy pair, where \mathbf{P} is defined by $\mathbf{P}(u) = q(u)$ whenever $q(u) > 0$.*

The term “equivalent” is justified by the following Theorem.

Theorem 5.6 *Suppose $q \equiv (\mathbf{P}, X)$. Then for any adapted process Z ,*

$$\mathbb{E}^{\mathbf{P}} Z_X = \bar{Z}_q, \quad (5.5)$$

and expected future values are identical under q and (\mathbf{P}, X) , i.e., $\bar{Z}_q(u)$ is defined exactly when $\mathbb{E}_u^{\mathbf{P}} Z_X$ is defined (i.e., $q^+(u) > 0$ if and only if both $\mathbf{P}(u) > 0$ and $X^+(u) > 0$), and

$$\bar{Z}_q(u) = \mathbb{E}_u^{\mathbf{P}} Z_X, \quad \forall u : q^+(u) > 0. \quad (5.6)$$

Proof Proposition 5.3 implies that $q^+(u) > 0$ if and only if both $\mathbf{P}(u) > 0$ and $X^+(u) > 0$. Note that

$$\mathbb{E}_u^{\mathbf{P}} Z_X = \frac{\sum_{v>u} X(v) \mathbf{P}(v) Z(v)}{\mathbf{P}(u) X^+(u)}.$$

Applying the equalities $\mathbf{P}(u) X(u) = q(u)$ and $\mathbf{P}(u) X^+(u) = q^+(u)$ to expression (4.12) shows that $\bar{Z}_q(u)$ equals the above expression for $\mathbb{E}_u^{\mathbf{P}} Z_X$. This proves (5.6) above, and a similar argument proves 5.5. ■

6 Approximate martingale measures

Now we have enough vocabulary to define “approximate martingale measures”. Roughly speaking, given an exercise strategy X , a measure \mathbf{P} is an approximate martingale measure if at any time, the discounted expected future exercise value of the stock price is not too far from the current stock price.

Definition 6.1 *For any probability measure \mathbf{P} and randomized strategy X , we say that \mathbf{P} is a (λ, μ, X) -approximate martingale measure, if \mathbf{P} -almost surely,*

$$S_k^*(1 - \mu) \leq \mathbb{E}_k^{\mathbf{P}} S_X^* \leq S_k^*(1 + \lambda), \quad \forall k < n : X_k^+ > 0. \quad (6.1)$$

The set of all measures \mathbf{P} that are (λ, μ, X) -approximate martingale measures is denoted by $P(\lambda, \mu, X)$.

(Recall that S_k^* stands for the *discounted* stock price, i.e., $S_k^* = S_k R^{-k}$. It is worth noting some special cases of this definition. From the above remarks it follows that $\mathbf{P} \in P(\lambda, \mu, \tau)$ if and only if, \mathbf{P} -almost surely,

$$S_k^*(1 - \mu) \leq \mathbb{E}_k^{\mathbf{P}} S_\tau^* \leq S_k^*(1 + \lambda), \quad \forall k < \tau.$$

Clearly the process

$$Y := \mathbb{E}_k^{\mathbf{P}} S_\tau^*$$

is a \mathbf{P} -martingale up to the stopping time τ , i.e., for $k \leq \tau$. Thus $\mathbf{P} \in P(\lambda, \mu, \tau)$ implies that there is some process S' lying between the bid and ask price processes of the stock, such that the discounted process S'_k/R^k is a \mathbf{P} -martingale, up to the stopping time τ . If $\mathbf{P} \in P(\lambda, \mu, \tau)$ for the special stopping time $\tau = n$, then this means that there is some process S' lying between the bid and ask price processes, such that the discounted process S'_k/R_k is a \mathbf{P} -martingale. (This type of interpretation has been made, for example, by Jouini and Kallal (1995) for European option prices under transaction costs.) In this case we say simply that \mathbf{P} is a (λ, μ) -**approximate martingale measure**, and denote the set of such measures by $P(\lambda, \mu)$. Cvitanic and Karatzas (1996) and Koehl et al. (1999) define a very similar class of measures in the context of pricing European options with transaction costs. Clearly an ordinary martingale measure \mathbf{P} belongs to the set $P(0, 0, \tau)$ where $\tau = n$ on every path; we denote the set of martingale measures by P .

From the above remarks it is easy to see that the following holds:

Proposition 6.2 *The following relationship holds among the various sets of probability measures on the event tree:*

$$P \subseteq P(\lambda, \mu) \subseteq \bigcup_{\tau \in T} P(\lambda, \mu, \tau) \subseteq \bigcup_{X \in \mathcal{X}} P(\lambda, \mu, X).$$

We can define a notion of (λ, μ) -approximate martingale node-measures, analogous to (λ, μ, X) -approximate martingale measures:

Definition 6.3 *We say that a node-measure q is a (λ, μ) -approximate martingale node-measure if*

$$S^*(u)(1 - \mu) \leq \overline{S}_q^*(u) \leq S^*(u)(1 + \lambda), \quad \forall u: q^+(u) > 0.$$

The collection of (λ, μ) -approximate martingale node-measures will be denoted by $Q(\lambda, \mu)$. For brevity we use the notation

$$Q(\lambda, \mu, \tau) := Q^\tau \cap Q(\lambda, \mu).$$

Remark 6.4 Note that $Q(\lambda, \mu)$ is never empty: the trivial node-measure with $q(0) = 1$ and $q(u) = 0$ for all nodes $u \neq 0$ is a (λ, μ) -approximate martingale node-measure.

From Theorem 5.6 it is clear that:

Corollary 6.5 *If $q \equiv (\mathbf{P}, X)$ then $q \in Q(\lambda, \mu)$ if and only if $\mathbf{P} \in P(\lambda, \mu, X)$. In particular, if $q \equiv (\mathbf{P}, \tau)$ then $q \in Q(\lambda, \mu, \tau)$ if and only if $\mathbf{P} \in P(\lambda, \mu, \tau)$.*

This implies a result that we will frequently appeal to in this paper:

Theorem 6.6 *For any adapted process Z ,*

$$\max_{q \in Q(\lambda, \mu)} \overline{Z}_q = \max_{X \in \mathcal{X}} \max_{\mathbf{P} \in P(\lambda, \mu, X)} \mathbb{E}^{\mathbf{P}} Z_X, \quad (6.2)$$

$$\max_{q \in Q(\lambda, \mu, \tau)} \overline{Z}_q = \max_{\mathbf{P} \in P(\lambda, \mu, \tau)} \mathbb{E}^{\mathbf{P}} Z_\tau, \quad \forall \tau \in T, \quad (6.3)$$

and the corresponding statements with “max” replaced by “min” also hold.

Proof The equality (6.2) (either with “max” everywhere or “min” everywhere) follows from: Theorem 5.4 which says that for every $q \in Q$ there is an equivalent measure-strategy pair (\mathbf{P}, X) and vice versa, Corollary 6.5 which says that if $q \equiv (\mathbf{P}, X)$ then $q \in Q(\lambda, \mu)$ if and only if $\mathbf{P} \in P(\lambda, \mu, X)$, and Theorem 5.6 which says that if $q \equiv (\mathbf{P}, X)$ then $\bar{Z}_q = \mathbb{E}^{\mathbf{P}} Z_X$. The argument for the second equality is the same, except that we use Corollary 5.5 in place of Theorem 5.4. ■

Clearly the set of node-measures $Q(\lambda, \mu)$ is convex, since it is defined by linear constraints. If there are no transaction costs, the Theorem below shows that all the *extreme points* of this set are *simple* node-measures. Recall from Section 3 that a node-measure q is said to be simple if its support lies on the boundary of some stopping time.

Theorem 6.7 (Extreme points of $Q(0, 0)$) *The extreme points of the set of node-measures $Q(0, 0)$ are simple node-measures.*

Proof Recall that a point x is an extreme point of a convex set S if it cannot be expressed as a strict convex combination of distinct points of S (see Bertsekas (1995)). Consider some non-simple node-measure $q \in Q(0, 0)$. We will construct two new node-measures a and c that belong to $Q(0, 0)$ such that $q = \frac{1}{2}(a + c)$, and this would imply that no non-simple node-measure can be an extreme point of $Q(0, 0)$. Recall that for any node measure p , $\bar{p}(u) \equiv p(u) + p^+(u)$ for any node u .

Now recall from Section 3 that if q is not simple, there is a node u such that $q(u) > 0$ and $q^+(u) > 0$. Fix such a node u . Fix some strictly positive ε such that $\varepsilon < q(v)$ for all $v \geq u$ with $q(v) > 0$. Define a node-measure a by “redistributing” the node-measure q as follows: reduce $q(u)$ by an amount ε , and for every $v > u$, increase $q(v)$ by a fraction of ε that is equal to value of $q(v)$ relative to $q^+(u)$. More precisely, a is identical to q everywhere except that : $a(u) = q(u) - \varepsilon$, and

$$a(v) = q(v) \left(1 + \frac{\varepsilon}{q^+(u)} \right), \quad \forall v > u. \quad (6.4)$$

Notice that the total amount by which q is increased on all the successors of u matches the amount by which q is decreased at u , i.e., a is just a redistribution of q , and so is also a node-measure. Note also that

$$\bar{q}(v) = \bar{a}(v), \quad \forall v \leq u, \quad q^+(v) = a^+(v), \quad \forall v < u. \quad (6.5)$$

The above statements also hold for the node-function c constructed as a but with $-\varepsilon$ instead of ε . It is easy to see that $q = \frac{1}{2}(a + c)$.

We now argue that $a \in Q(0, 0)$; the argument for c is similar. To see this we first note that if a node-measure p is in $Q(0, 0)$ then the “martingale-like” constraints hold:

$$\begin{aligned} S(v)p^+(v) &= \sum_{w>v} S(w)p(w) = \sum_{v' \in v^+} \left(S(v')p(v') + \sum_{w>v'} S(w)p(w) \right) \\ &= \sum_{v' \in v^+} (S(v')p(v') + S(v')p^+(v')), \end{aligned}$$

or

$$S(v)p^+(v) = \sum_{v' \in v^+} S(v')\bar{p}(v') \quad (6.6)$$

In fact it is easy to see that a node-measure p is in $Q(0,0)$ if and only if (6.6) holds at all $v \notin T$. We know (6.6) holds for $p = q$, and since for any $v > u$, $a(v)$ is a fixed multiple of $q(v)$ (definition (6.4)), this equality also holds for $p = a$ and any $v \geq u$. From (6.5), the equality (6.6) holds for $p = a$ and any $v < u$. Clearly for nodes v that are not on a path containing u , $q(v) = a(v)$, so the equality (6.6) holds for such v and $p = a$. Thus $a \in Q(0,0)$. ■

7 Portfolios

A **portfolio process** is defined by a triple (x, Δ, B) where x is the **initial wealth**, and Δ_k and B_k are adapted processes representing the **stock and bond positions** at time k respectively, for $k = 0, 1, \dots, n-1$. (Note that the portfolio process is not defined at time n). In particular for a non-terminal node u , $\Delta(u)$ is the number of shares of stock held in the state represented by node u , from time $t(u)$ to time $t(u) + 1$, and $B(u)$ is the number of dollars in bonds held at node u , between times $t(u)$ and $t(u) + 1$. Consider some node $v \neq 0$. Recall that v^- is the unique predecessor of v . At node v , just before it is re-balanced, the portfolio consists of $\Delta(v^-)S(v)$ dollars in stocks and $B(v^-)R$ dollars in bonds (or just x dollars in bonds if $v = 0$). We therefore define the (pre-rebalancing) **portfolio value** at any node v as

$$V^{x, \Delta, B}(v) = \begin{cases} x & \text{if } v = 0 \\ \Delta(v^-)S(v) + B(v^-)R & \text{otherwise.} \end{cases} \quad (7.1)$$

Note that if we define $\Delta(0^-) := 0$ and $B(0^-) := x/R$ then we can omit the first case in the above definition. We will therefore assume these definitions throughout. If $v \notin T$, after re-balancing, the portfolio at v consists of $\Delta(v)$ shares and $B(v)$ dollars. This rebalancing involves either buying or selling of shares at node v , *but not both* (it can never help to both buy and sell shares at the same node, since this will only incur more transaction costs). If $\Delta(v) > \Delta(v^-)$, then $b(v) = (\Delta(v) - \Delta(v^-))$ shares must be bought. Otherwise, $s(v) = (\Delta(v^-) - \Delta(v))$ shares must be sold. Clearly $b(v)$ and $s(v)$ are uniquely defined by the constraints

$$b(v) - s(v) = \Delta(v) - \Delta(v^-), \quad b(v)s(v) = 0, \quad b(v) \geq 0, \quad s(v) \geq 0. \quad (7.2)$$

The portfolio (x, Δ, B) is said to be **self-financing** if the cost of (or proceeds from) re-balancing the stock position is offset by a change in the bond position. In other words, at every node v , in case $b(v) > 0$ the cost of shares bought is financed by a reduction in the bond holdings, and in case $s(v) > 0$ the proceeds from selling shares is offset by an increase in bond holdings. Thus for a self-financing portfolio (x, Δ, B) , at any node $v \notin T$ we must have

$$b(v)S(v)(1 + \lambda) \leq B(v^-)R - B(v) \quad \text{if } b(v) > 0,$$

and

$$s(v)S(v)(1 - \mu) \leq B(v) - B(v^-)R \quad \text{if } s(v) > 0.$$

Noting that $b(v)$ and $s(v)$ can never both be positive, these conditions can be combined into

$$-b(v)S(v)(1 + \lambda) + s(v)S(v)(1 - \mu) + B(v^-)R - B(v) \geq 0 \quad \forall v \notin T, \quad (7.3)$$

where $b(v)$ and $s(v)$ are given by (7.2). A similar formulation appears in Edirisinghe et al. (1993).

We write Φ to denote the set of all self-financing portfolios (x, Δ, B) . We will sometimes abuse notation and write $\Phi(x)$ to mean the set of all pairs of adapted processes (Δ, B) such that $(x, \Delta, B) \in \Phi$. Since the constraints defining self-financing portfolios are all linear, it is clear that Φ is *convex* and *closed*.

8 Hedging and the pricing problem

In order to characterize the range of prices for an American option that are consistent with the absence of arbitrage, we first need to understand the situation of the buyer and seller of the option. In Karatzas and Kou (1998), the authors point out eloquently that these two situations are fundamentally asymmetric. The treatment in this section closely follows theirs.

The buyer and seller are essentially entering into the following agreement at time 0: the seller agrees to pay the buyer a random amount $G_\tau \geq 0$ at time τ , where τ is a stopping time *chosen by the buyer*. In return for this commitment, the buyer agrees to pay an amount $x \geq 0$ to the seller at time 0. We consider the question: what is the “fair” price for the buyer to pay the seller at time 0, for his obligation to deliver the amount $G_\tau \geq 0$ at any stopping time τ picked by the buyer?

Let us examine the seller’s objective. When he receives x from the buyer, he wants to find a portfolio $(x, \Delta, B) \in \Phi$ that enables him to meet his obligation without risk (i.e. on all paths) and for *any* stopping time τ chosen by the buyer, i.e.,

$$V_\tau^{x, \Delta, B} \geq G_\tau, \quad \forall \tau \in \mathcal{T}. \quad (8.1)$$

Note that we are implicitly making the following **assumptions**: (a) when it is exercised, the option is *settled in cash*, and (b) *there are no transaction costs when a portfolio is liquidated* to settle the option. If there were transaction costs at this time, we would have to consider the liquidated portfolio value, which would be smaller than the pre-rebalancing portfolio value. (See Musiela and Rutkowski (1998, page 56) for a discussion of these assumptions.) We make these assumptions for simplicity; removing these assumptions only complicates the notation without adding much insight.

The smallest initial wealth $x \geq 0$ that allows the seller to meet his obligation as described above, is called the **upper hedging price** for the American option:

$$h_{\text{up}}(\lambda, \mu) := \min\{x \geq 0 \mid \exists (\Delta, B) \in \Phi(x) \text{ s.t. (8.1) holds}\}. \quad (8.2)$$

We will refer to the optimization problem in (8.2) as the **min-cost super-replication problem**, and an optimal solution (x', Δ', B') will be called an **optimal super-replicating portfolio**. Clearly in any optimal super-replicating portfolio (x', Δ', B') , we have $x' = h_{\text{up}}(\lambda, \mu)$ by definition.

We will need the following definition in Section 9.

Definition 8.1 *The discounted deficit process $D^{x, \Delta, B}$ of the portfolio (x, Δ, B) is the discounted amount by which the portfolio value falls short of the payoff:*

$$D_k^{x, \Delta, B} := (G_k - V_k^{x, \Delta, B})R^{-k}.$$

In particular at a given node u ,

$$D^{x, \Delta, B}(u) = (G(u) - V^{x, \Delta, B}(u))R^{-t(u)}.$$

Now let us consider the buyer's objective: he starts out by borrowing the amount $x \geq 0$ (i.e. he starts with wealth $-x$) at time 0 to pay the seller for the option, and picks a stopping time $\tau \in T$ and portfolio $(\Delta, B) \in \Phi(-x)$ such that, by exercising his option at time τ , the payment that he receives enables him to repay without risk, the debt incurred at time 0 when purchasing the option, i.e.,

$$V_{\tau}^{-x, \Delta, B} + G_{\tau} \geq 0. \quad (8.3)$$

Note that again we are assuming that (a) a debt must be repaid *in cash*, and (b) *there are no transaction costs when a portfolio is liquidated to settle a debt*.

The largest amount $x \geq 0$ that enables the buyer to repay his debt in this way, is called the **lower hedging price** of the option:

$$h_{\text{low}}(\lambda, \mu) := \max\{x \geq 0 \mid \exists \tau \in T, (\Delta, B) \in \Phi(-x), \text{ s.t. (8.3) holds.}\} \quad (8.4)$$

Notice the asymmetry between the seller and the buyer: the former needs to hedge against *any* stopping time $\tau \in T$, while the latter need only hedge for *some* stopping time $\tau \in T$. It is important to point out that h_{low} could be bigger than h_{up} . However, in the absence of arbitrage, $h_{\text{low}} \leq h_{\text{up}}$ always holds, as we show in Theorem 8.3. We now define the notion of arbitrage.

Suppose the price of the American option G is y . We say that the pair (G, y) **admits an arbitrage opportunity** in the market if there exists either

1. a portfolio $(\Delta, B) \in \Phi(x)$ that satisfies

$$V_{\tau}^{x, \Delta, B} \geq G_{\tau}, \quad \forall \tau \in T, \quad (8.5)$$

for some $0 < x < y$; or

2. a stopping time $\tau \in T$ and a portfolio $(\Delta, B) \in \Phi(-x)$ such that

$$V_{\tau}^{-x, \Delta, B} + G_{\tau} \geq 0, \quad (8.6)$$

for some $x > y$.

In the first case, the option price y is too high because the seller is able to construct a portfolio using just $x < y$ that allows him to meet his obligations to the option buyer without risk, so he can pocket the difference $y - x$, which represents a risk-less profit, i.e., an arbitrage. In the second case, the option price y is too low because the buyer can borrow $x > y$, and pay only y for the option, pocketing the difference $x - y$, and construct a portfolio that allows him to repay his initial debt without risk at some stopping time τ that he chooses. Again this is an arbitrage.

We say that the market is **arbitrage-free** if for any American option G there is some price y such that (G, y) does not admit an arbitrage opportunity.

The following theorem is adapted from Karatzas and Kou (1998), and is immediate from the above definitions.

Theorem 8.2 *Suppose G is an American option. Then (G, y) admits an arbitrage opportunity if and only if $y < h_{\text{low}}(\lambda, \mu)$ or $y > h_{\text{up}}(\lambda, \mu)$.*

Proof Suppose $y < h_{\text{low}}(\lambda, \mu)$. By definition (8.4), $h_{\text{low}}(\lambda, \mu)$ is the largest amount x that allows the option bearer to construct a portfolio $(\Delta, B) \in \Phi(-x)$ such that (8.6) holds for some stopping time τ . Note that any amount x smaller than $h_{\text{low}}(\lambda, \mu)$, and in particular any x in the interval $(y, h_{\text{low}}(\lambda, \mu))$, will also allow the option bearer to do this. Thus y satisfies the condition 2 of an arbitrage opportunity as defined above. On the other hand suppose $y > h_{\text{up}}(\lambda, \mu)$. By definition (8.2), $h_{\text{up}}(\lambda, \mu)$ is the smallest initial wealth x that allows the option seller to construct a portfolio $(\Delta, B) \in \Phi(x)$ such that (8.5) holds. Note that any amount x larger than $h_{\text{up}}(\lambda, \mu)$, and in particular any x in the interval $(h_{\text{up}}(\lambda, \mu), y)$, will also allow the option seller to do this. Thus y satisfies the condition 1 of an arbitrage opportunity defined above.

On the other hand, suppose that $h_{\text{low}}(\lambda, \mu) \leq y \leq h_{\text{up}}(\lambda, \mu)$ and that the conditions of (1) in the definition of arbitrage opportunity above are satisfied, for some positive $x < y$. Since by definition (8.2), $h_{\text{up}}(\lambda, \mu)$ is the smallest $x \geq 0$ such that these conditions hold, this implies that $h_{\text{up}}(\lambda, \mu) \leq x < y$, which is a contradiction. The argument is similar if the conditions of (2) are satisfied. ■

An immediate corollary to this is:

Corollary 8.3 *The market is arbitrage-free if and only if $h_{\text{low}}(\lambda, \mu) \leq h_{\text{up}}(\lambda, \mu)$ holds for every American option G , and in this case the interval $[h_{\text{low}}(\lambda, \mu), h_{\text{up}}(\lambda, \mu)]$ is called the **arbitrage-free interval**.*

In Theorem 13.1 we show some other conditions that are equivalent to the absence of arbitrage. The remainder of the paper is concerned with characterizing the upper and lower hedging prices $h_{\text{up}}(\lambda, \mu)$ and $h_{\text{low}}(\lambda, \mu)$.

9 Representations of the Upper Hedging Price

The goal of this section is to represent the upper hedging price $h_{\text{up}}(\lambda, \mu)$ in terms of several optimization problems. Consider the following seven optimization problems.

$$h(Q(\lambda, \mu)) := \max_{q \in Q(\lambda, \mu)} \overline{G^*}_q, \quad (9.1)$$

$$h(X) := \max_{X \in \mathcal{X}} \max_{\mathbf{P} \in \mathcal{P}(\lambda, \mu, X)} \mathbb{E}^{\mathbf{P}} G_X^*, \quad (9.2)$$

$$h(Q, \Phi) := \max_{q \in Q} \min_{(x, \Delta, B) \in \Phi} (x + \overline{D}_q^{x, \Delta, B}), \quad (9.3)$$

$$h(X, \Phi) := \max_{X \in \mathcal{X}} \min_{(x, \Delta, B) \in \Phi} (x + \widehat{D}_X^{x, \Delta, B}), \quad (9.4)$$

$$h(\overline{T}, \Phi) := \max_{\sigma \in \overline{T}} \min_{(x, \Delta, B) \in \Phi} (x + \widehat{D}_\sigma^{x, \Delta, B}), \quad (9.5)$$

$$h(\Phi) := \min_{(x, \Delta, B) \in \Phi} \max_u (x + D^{x, \Delta, B}(u)), \quad (9.6)$$

$$h(Q(\lambda, \mu, \tau)) := \max_{q \in Q(\lambda, \mu, \tau)} \overline{G^*}_q, \quad \forall \tau \in \overline{T}. \quad (9.7)$$

We have used a mnemonic scheme to name each of the optimum values above. For example we chose the name $h(X, \Phi)$ in (9.4) because the domains of the max and min in the corresponding optimization problem are \mathcal{X} and Φ respectively. Thus the “arguments” of h in each name above should not be treated literally as arguments of a function.

Some explanation of each of the above problems might be helpful here. Also the reader may wish to review the definitions of the various averages in Section 4.1. Expression (9.1) defines the “maximum discounted q -weighted payoff, over all (λ, μ) -approximate martingale node-measures q ”. Expression (9.2) for $h(X)$ is a generalization of expression (1.1), the difference being that the outer maximization is now performed over all *randomized* stopping times $X \in \mathcal{X}$ and the inner maximization is over all (λ, μ, X) -approximate martingale measures. Note that unlike in (1.1), the domain of the inner “max” in (9.2) depends on which X is fixed in the domain of the outer “max”. In Theorem 11.1 we show that with zero transaction costs this expression reduces to (1.1) if there exists a martingale measure with positive measure on every path. Expression (9.3) is essentially the dual of the linear program corresponding to (9.1). For a fixed $q \in \mathcal{Q}$ the inner “min” represents the “best hedge” achievable by self-financing portfolios with respect to q , in the sense of minimizing the sum of initial wealth and q -weighted discounted deficit. Expression (9.4) is similar, except that the outer maximization is over randomized stopping times $X \in \mathcal{T}$, and “best hedge” is defined in terms of the maximum X -value of the discounted deficit. Expression (9.5) is yet another variant where the outer maximization is over mixtures σ of stopping times and “best hedge” is defined in terms of the σ -expected maximum value of the discounted deficit. Expression (9.6) is obtained by interchanging the “max” and “min” in (9.3) and simplifying; for a fixed self-financing portfolio $(x, \Delta, B) \in \Phi$ the inner “max” represents the maximum value of x plus the discounted deficit, over all nodes u . Finally, expression (9.7) is similar to (9.1) except that the maximization is restricted to simple node-measures whose support lies on the stopping boundary of some stopping time τ .

Theorems 9.1 and 9.2 below show that all but the last of the above expressions equals the upper hedging price $h_{\text{up}}(\lambda, \mu)$, as defined by (8.2). The principal tool for our proofs is Strong Duality of Linear Programming (see Appendix A). Example 15.1 in Section 15 illustrates the computation of $h_{\text{up}}(\lambda, \mu)$.

It is also instructive to consider the problem of hedging against a particular exercise strategy τ , i.e., the problem of finding the minimum initial wealth needed to be able to construct a self-financing portfolio that can hedge against τ . More precisely, consider

$$h_{\text{up}}^{\tau} := \min\{x \geq 0 \mid \exists (\Delta, B) \in \Phi(x) \text{ s.t. } V_{\tau}^{x, \Delta, B} \geq G_{\tau}\}. \quad (9.8)$$

Theorem 9.3 shows that $h_{\text{up}}^{\tau} = h(Q(\lambda, \mu, \tau))$.

Theorem 9.1 (Upper hedging price) *The minimum in the definition (8.2) of the upper hedging price is finite, and is given by:*

$$h_{\text{up}}(\lambda, \mu) = h(Q(\lambda, \mu)) = h(X). \quad (9.9)$$

Proof The plan of the proof is to first observe that the optimization problem implicit in the definition (8.2) of the upper hedging price, is a linear programming problem. We then write the dual of this linear program (LP) (see Appendix A), and by making a judicious choice of dual variables observe that the dual LP is equivalent to the definition of $h(Q(\lambda, \mu))$. The equality of $h(Q(\lambda, \mu))$ and $h(X)$ then follows from Theorem 6.6. Thus we only need to show the first equality in (9.9).

Let us then write the optimization in the definition (8.2) of $h_{\text{up}}(\lambda, \mu)$ as an LP. This defines $h_{\text{up}}(\lambda, \mu)$ as the smallest x such that $V^{x, \Delta, B}(u) \geq G(u)$ holds at every node u , subject to the self-financing constraints (7.2) and (7.3). First observe that since it is always sub-optimal to simultaneously buy and sell shares of the stock, for any v the variables $b(v)$ and $s(v)$ cannot both be positive at an optimal solution. Therefore we can dispense with

the product constraints $b(v)s(v) = 0$ appearing in (7.2). Thus we can write the LP in the variables $x, b(\cdot), s(\cdot), \Delta(\cdot), B(\cdot)$:

$$h_{\text{up}}(\lambda, \mu) = \min x$$

subject to the following constraints (the dual variable corresponding to each constraint is shown on the left):

the self-financing constraints (7.2) and (7.3) for $v = 0$:

$$z(0) : \quad -b(0)(1 + \lambda)S(0) + s(0)(1 - \mu)S(0) - B(0) + x \geq 0, \quad (9.10)$$

$$y(0) : \quad -\Delta(0) + b(0) - s(0) = 0, \quad (9.11)$$

self-financing constraints (7.2) and (7.3) for each node $v \notin T \cup \{0\}$:

$$z(v) : \quad -b(v)(1 + \lambda)S(v) + s(v)(1 - \mu)S(v) - B(v) + B(v^-)R \geq 0, \quad (9.12)$$

$$y(v) : \quad -\Delta(v) + \Delta(v^-) + b(v) - s(v) = 0, \quad (9.13)$$

the initial domination constraint:

$$z'(0) : \quad x \geq G(0), \quad (9.14)$$

the domination constraints for every $v \neq 0$ (see definition (7.1) of $V^{x, \Delta, B}$):

$$z'(v) : \quad \Delta(v^-)S(v) + B(v^-)R \geq G(v) \quad (9.15)$$

and the non-negativity constraints

$$b(u) \geq 0, \quad s(u) \geq 0, \quad \forall u \notin T. \quad (9.16)$$

We need not add the non-negativity constraint $x \geq 0$ since the option payoff $G(\cdot)$ is non-negative and so the initial domination constraint $x \geq G(0)$ is a stronger constraint. Clearly the LP above is feasible since for a sufficiently large initial wealth x one can always construct a self-financing portfolio whose value dominates the payoff everywhere. Furthermore, the initial domination constraint ensures that the solution to the linear program is bounded from below. Therefore the LP is feasible and has a bounded optimum value, and by strong duality of linear programming (Schrijver, 1986) (see Appendix, Theorem A.1), the dual LP also has a bounded optimum value equal to that of the primal LP.

Let us now write down the dual of the above LP. We have defined three families of dual variables: $z(u), y(u)$ for each node $u \notin T$, and $z'(u)$ for each node u . Thus we can express $h_{\text{up}}(\lambda, \mu)$ in terms of the dual LP as

$$h_{\text{up}}(\lambda, \mu) = \max \sum_u z'(u)G(u) \quad (9.17)$$

subject to the following constraints, with the corresponding primal variables shown on the left. (The appendix describes how the dual LP is written in general.)

$$B(u) : \quad -z(u) + R \sum_{v \in u^+} (z(v) + z'(v)) = 0, \quad \forall u : u^+ \not\subset T, \quad (9.18)$$

$$B(u) : \quad -z(u) + R \sum_{v \in u^+} z'(v) = 0, \quad \forall u : u^+ \subset T, \quad (9.19)$$

$$\Delta(u) : -y(u) + \sum_{v \in u^+} (y(v) + z'(v)S(v)) = 0, \quad \forall u : u^+ \not\subset T, \quad (9.20)$$

$$\Delta(u) : \quad -y(u) + \sum_{v \in u^+} z'(v)S(v) = 0, \quad \forall u : u^+ \subset T, \quad (9.21)$$

$$x : \quad z(0) + z'(0) = 1, \quad (9.22)$$

$$b(u) : \quad -z(u)(1 + \lambda)S(u) + y(u) \leq 0, \quad \forall u \notin T, \quad (9.23)$$

$$s(u) : \quad z(u)(1 - \mu)S(u) - y(u) \leq 0, \quad \forall u \notin T, \quad (9.24)$$

and the non-negativity constraints

$$z(u) \geq 0, \quad \forall u \notin T; \quad z'(u) \geq 0, \quad \forall u. \quad (9.25)$$

The dual constraints (9.18)-(9.24) can be re-written in a more informative way as follows:

$$z(u) = R^{-t(u)} \sum_{v > u} z'(v)R^{t(v)}, \quad \forall u \notin T, \quad (9.26)$$

$$y(u) = \sum_{v > u} z'(v)S(v), \quad \forall u \notin T, \quad (9.27)$$

$$\sum_u z'(u)R^{t(u)} = 1, \quad (9.28)$$

$$(1 - \mu)z(u)S(u) \leq y(u) \leq (1 + \lambda)z(u)S(u), \quad \forall u \notin T. \quad (9.29)$$

Now if we define the node function $q(u) := z'(u)R^{t(u)}$ for all u , the dual LP is equivalent to:

$$h_{\text{up}}(\lambda, \mu) = \max \sum_u q(u)R^{-t(u)}G(u) \quad (9.30)$$

subject to

$$\sum_u q(u) = 1, \quad (9.31)$$

$$\begin{aligned} S(u)(1 - \mu) \sum_{v > u} q(v) &\leq R^{t(u)} \sum_{v > u} q(v)R^{-t(v)}S(v) \leq \\ &S(u)(1 + \lambda) \sum_{v > u} q(v), \quad \forall u \notin T, \end{aligned} \quad (9.32)$$

$$q(u) \geq 0, \quad \forall u, \quad (9.33)$$

which is equivalent to $h_{\text{up}}(\lambda, \mu) = \max_{q \in Q(\lambda, \mu)} \overline{G^*} q$. ■

Theorem 9.2 (Upper hedging price) *For the optimization problems defined above,*

$$h(Q, \Phi) = h(X, \Phi) = h(\overline{T}, \Phi) = h(\Phi) = h_{\text{up}}(\lambda, \mu)$$

Moreover, the order of the “max” and “min” can be reversed in the definitions of $h(Q, \Phi)$, $h(X, \Phi)$ and $h(\overline{T}, \Phi)$ without changing the optimum value.

Proof Recall that Φ is a convex, closed set, so we can apply Prop. 5.2 to infer that

$$h(Q, \Phi) \leq h(X, \Phi) \leq h(\overline{T}, \Phi) \leq h(\Phi).$$

Also, if we reverse the order of the “max” and “min” in the definition of $h(Q, \Phi)$, the resulting expression cannot be smaller than $h(Q, \Phi)$, and by an argument similar to the one in the proof of Prop. 5.2, it equals $h(\Phi)$. The same is true if we reverse the “max” and “min” in the expressions defining $h(X, \Phi)$ and $h(\overline{T}, \Phi)$. Now suppose (x, Δ, B) is any optimal super-replicating portfolio (see (8.2)). This means that starting with wealth $x = h_{\text{up}}(\lambda, \mu)$, there is a self-financing portfolio (Δ, B) whose value dominates the payoff everywhere, i.e., the (discounted) deficit at every node is ≤ 0 . Therefore for this portfolio,

$$\max_u (x + D^{x, \Delta, B}(u)) \leq h_{\text{up}}(\lambda, \mu),$$

which means that $h(\Phi) \leq h_{\text{up}}(\lambda, \mu)$.

Thus to prove our Theorem it suffices to show that $h(Q, \Phi) = h_{\text{up}}(\lambda, \mu)$. We show this using Strong Duality of Linear Programming. For brevity we will drop the superscript “ x, Δ, B ” on V and D in what follows. We first write $h(Q, \Phi)$ as

$$h(Q, \Phi) = \max_{q \in Q} h(q),$$

where

$$\begin{aligned} h(q) &:= \min_{(x, \Delta, B) \in \Phi} (x + \overline{D}_q) = \min_{(x, \Delta, B) \in \Phi} (x + \overline{G}_q^* - \overline{V}_q^*) \\ &= \overline{G}_q^* + \min_{(x, \Delta, B) \in \Phi} (x - \overline{V}_q^*). \end{aligned} \quad (9.34)$$

The optimization problem in the last “min” expression above is a linear program:

$$\alpha = \min x - \sum_u q(u) V(u) R^{-t(u)}$$

subject to the self-financing constraints (9.10)-(9.13) in the proof of Theorem 9.1. Recall that $V(u)$ above is given by expression (7.1), which is linear in (x, Δ, B) . This LP and is very similar to the primal LP in the proof of Theorem 9.1. The only difference is that in the present LP, each $V(u)$ term appears in the objective function with coefficient $-q(u)R^{-t(u)}$, whereas in the proof of Theorem 9.1 there was a constraint $V(u) \geq G(u)$ for each node u . Not surprisingly, the dual of the LP above is also similar to the dual LP in that Theorem; it is

$$\alpha = \max 0$$

subject to the constraints (9.18)-(9.24) with each dual variable $z^l(v)$ replaced by the constant $q(v)R^{-t(v)}$. Recall that in the proof of Theorem 9.1 we *defined* $q(v)$ to be $z^l(v)R^{t(v)}$. Thus we can carry out exactly the same algebraic manipulations as in that Theorem, and write the dual of our LP as: $\max 0$ subject to the constraints (9.31)-(9.33). Since the node-measure q is fixed in the present Theorem, these constraints merely enforce the requirement that $q \in Q(\lambda, \mu)$, i.e., that q is a (λ, μ) -approximate martingale node-measure. Thus if $q \in Q(\lambda, \mu)$ the dual LP is feasible and its optimum value, as well as that of the primal LP (i.e. the “min” expression in (9.34)) is 0, and so $h(q) = \overline{G}_q^*$. If $q \notin Q(\lambda, \mu)$, the dual LP is not feasible and

therefore the primal LP is unbounded (since it is clearly feasible) and $h(q) = -\infty$. Since we know $Q(\lambda, \mu)$ is not empty (see Remark 6.4), $h(Q, \Phi)$ can be written as

$$h(Q, \Phi) = \max_{q \in Q} h(q) = \max_{q \in Q(\lambda, \mu)} \overline{G^*}_q =: h(Q(\lambda, \mu)),$$

which by Theorem 9.1 equals $h_{\text{up}}(\lambda, \mu)$. ■

Theorem 9.3 (Seller's hedge against a given τ) *For any stopping time $\tau \in T$,*

$$h_{\text{up}}^\tau = h(Q(\lambda, \mu, \tau)).$$

Proof The proof closely mirrors that of Theorem 9.1. Recall that I_τ is the stopping boundary of τ . We also define the set

$$J_\tau := \{u : \exists v > u \text{ s.t. } v \in I_\tau\} \quad (9.35)$$

of nodes that are *inside* the stopping boundary.

Consider the definition (9.8) of h_{up}^τ . This is a linear program much like the definition (8.2) of $h_{\text{up}}(\lambda, \mu)$. The difference is that in the case of h_{up}^τ , the domination constraints only need to hold at the stopping boundary I_τ . Thus in the LP for h_{up}^τ , the optimum does not change if we restrict the domain of optimization to portfolios that satisfy the self-financing constraints (7.3) only for nodes $v \in J_\tau$. Without writing down the LP explicitly, we can see that it can be written as

$$h_{\text{up}}^\tau = \max x$$

subject to the following constraints: (a) the initial self-financing constraints (9.10) and (9.11) if $0 \in J_\tau$, (b) self-financing constraints (9.12) and (9.13) for every $v \in J_\tau$, (c) the initial domination constraint (9.14) if $0 \in I_\tau$, (d) the domination constraints (9.15) for every $v \in I_\tau$, and (e) the non-negativity constraints (9.16) for every $u \in J_\tau$.

As in the proof of Theorem 9.1 we define the dual variables $z(\cdot), y(\cdot), z'(\cdot)$ but this time we only need dual variables $z(u), y(u)$ for each node $u \in J_\tau$, and $z'(u)$ for each $u \in I_\tau$. Then the dual LP is

$$h_{\text{up}}^\tau = \max \sum_{u \in I_\tau} z'(u) G(u) \quad (9.36)$$

subject to these constraints: (a) constraints (9.18) for all $u \in J_\tau$ such that $u^+ \not\subset T$, (b) constraints (9.19) for all u such that $u^+ \subset T$, (c) constraints (9.20) for all $u \in J_\tau$ such that $u^+ \not\subset T$, (d) constraints (9.21) for all u such that $u^+ \subset T$, (e) the constraint $z'(0) = 1$ if $0 \in I_\tau$, or the constraint $z(0) = 1$ if $0 \notin I_\tau$, (f) constraints (9.23) and (9.24) for every $u \in J_\tau$, and (g) the non-negativity constraints

$$z(u) \geq 0, \quad \forall u \in J_\tau; \quad z'(u) \geq 0, \quad \forall u \in I_\tau.$$

These dual constraints can be re-written as (9.26)-(9.29), except that now we only need (9.26), (9.27), and (9.29) for all $u \in J_\tau$, and the summation in (9.28) should only be over $u \in I_\tau$. As in the proof of Theorem 9.1, if we define the node function $q(u) \equiv z'(u) R^{l(u)}$ for all $u \in I_\tau$, the dual LP is equivalent to

$$h_{\text{up}}^\tau = \max_{u \in I_\tau} q(u) G^*(u)$$

subject to the following constraints: (a) (9.31) with the summation only over nodes $u \in I_\tau$, (b) the constraints (9.32) for all $u \in J_\tau$, with $q(v)$ interpreted as 0 for $v \notin I_\tau$, and (c) the non-negativity constraints (9.33) for $u \in I_\tau$. The constraints (a) and (c) require q to be a simple node-measure in Q^τ , and the constraints (b) require that $q \in Q(\lambda, \mu)$. Thus the constraints together require that $q \in Q^\tau \cap Q(\lambda, \mu) =: Q(\lambda, \mu, \tau)$, so we can write $h_{\text{up}}^\tau = \max_{q \in Q(\lambda, \mu, \tau)} \overline{G}^* q$. ■

10 Randomized Strategies and Nash Equilibria

In the last Section we saw that some optimization problems involved in the representation of the upper hedging price $h_{\text{up}}(\lambda, \mu)$ appear to require the use of mixtures of stopping times (or randomized stopping times). In game theory it is often the case that at a Nash equilibrium, one or more players must choose a *mixture* of strategies. It is thus natural to seek a game-theoretic interpretation of some of our results. We give here such an interpretation of the max-min expressions $h(Q, \Phi)$ and $h(\overline{T}, \Phi)$ defined in the previous Section.

We would like to set up a simultaneous-move game between the *seller* and the *devil*. We will consider two variants of such a game. The **node game** is defined as follows:

- The seller picks a self-financing portfolio $(x, \Delta, B) \in \Phi$,
- The devil picks a node-measure $q \in Q$, i.e., a “mixture” of nodes (in the sense that choosing a single node constitutes a pure strategy, and choosing a node-measure with support at more than one node constitutes a non-degenerate mixed strategy).
- For given choices (x, Δ, B) and q , the seller’s *cost* and the devil’s *utility* are both equal to $x + \overline{D}_q^{x, \Delta, B}$. (Recall from Section 4.1 that for any adapted process Z , \overline{Z}_q is the q -weighted average value of Z over the entire event-tree.)

The **stopping game** is defined similarly except that the strategy space of the devil is the set \overline{T} of all mixtures of stopping times.

- The seller again picks a self-financing portfolio $(x, \Delta, B) \in \Phi$,
- The devil picks a mixture $\sigma \in \overline{T}$ of stopping times,
- For given choices (x, Δ, B) and σ , the seller’s *cost* and the devil’s *utility* are both equal to $x + \widehat{D}_\sigma^{x, \Delta, B}$.

In both games, the seller would like to minimize her cost, while the devil would like to maximize his utility. (We are using the term “devil” rather than “buyer” because the devil’s objective in this game is solely to force the seller’s cost to be as large as possible; the option buyer on the other hand would presumably try to maximize his own payoff from exercising the option). In the node game, a pair of choices $q', (x', \Delta', B')$ is said to constitute a *Nash equilibrium* (Fudenberg and Tirole, 1998) if the seller’s choice of (x', Δ', B') minimizes her cost given the devil’s choice of q' , and the devil’s choice of q' maximizes his utility given the seller’s choice of (x', Δ', B') . A Nash equilibrium in the stopping game is defined similarly.

Theorems 10.2 and 10.3 below characterize Nash equilibria in the node game and stopping game respectively. See Example 15.3 for an illustration of these Theorems. We first need the following definition.

Definition 10.1 We say that a portfolio (x', Δ', B') is **equivalent** to a portfolio (x, Δ, B) if the portfolio (x', Δ', B') is identical to (x, Δ, B) except that (x', Δ', B') starts with some

additional initial wealth α (could be negative) that is simply held in the bank (or in bonds) and allowed to grow at the risk-free interest rate, or in other words $x' = x + \alpha$, and for every node u , $\Delta'(u) = \Delta(u)$, and $B'(u) = B(u) + \alpha R^{t(u)}$.

Clearly if (x, Δ, B) is equivalent to (x', Δ', B') with $x' = x + \alpha$, then at every node, the discounted portfolio values differ by exactly α , i.e.,

$$V^{x', \Delta', B'}(u) R^{-t(u)} = V^{x, \Delta, B}(u) R^{-t(u)} + \alpha, \quad \forall u.$$

Theorem 10.2 (Node game) A node-measure $q' \in Q$ and a self-financing portfolio $(x', \Delta', B') \in \Phi$ constitute a Nash equilibrium in the node game if and only if: (a) q' is an optimal node-measure for expression (9.1) defining $h(Q(\lambda, \mu))$, and (b) (x', Δ', B') is equivalent to an optimal super-replicating portfolio (defined by (8.2)). In particular the seller's cost and devil's utility at any Nash equilibrium is $h_{\text{up}}(\lambda, \mu)$.

Proof Recall from the proof of Theorem 9.1 that the minimization problem (8.2) and the maximization problem in (9.1) are linear programs that are dual to each other. In particular, each variable $q(u)$ in the first LP corresponds to a *domination constraint* in the second LP. *Complementary slackness* (see Appendix A) tells us that $q' \in Q$ is optimal for (9.1) and (x', Δ', B') is optimal for (8.2) if and only if for every node u with $q'(u) > 0$, the domination constraint at u is *tight* for the portfolio (x', Δ', B') , i.e., $D^{x', \Delta', B'}(u) = 0$. For brevity we will denote the process $D^{x', \Delta', B'}$ by D in what follows.

To prove the “if” part of the Theorem, consider any $q' \in Q$ optimal for (9.1), and any $(x', \Delta', B') \in \Phi$ optimal for (8.2). The complementary slackness property stated above implies that $\overline{D}_{q'} = 0$, so $x' + \overline{D}_{q'} = h_{\text{up}}(\lambda, \mu)$. Suppose the seller chooses the portfolio (x', Δ', B') . Since this is a super-replicating portfolio, the deficit at every node is ≤ 0 , so the devil cannot hope to choose a $q'' \in Q$ with $\overline{D}_{q''} > 0$. Thus the devil's choice of q' is optimal, given that the seller has chosen (x', Δ', B') . Now suppose the devil has chosen q' . If the seller chooses (x', Δ', B') , her cost would be $x' + \overline{D}_{q'} = h_{\text{up}}(\lambda, \mu)$. This is the smallest possible cost the seller can hope for, given the devil's choice of q' . This follows from the duality argument in the last paragraph of the proof of Theorem 9.2, where it can be seen that

$$h(q') := \min_{(x, \Delta, B) \in \Phi} (x + \overline{D}_{q'}) = \overline{G}_{q'} = h_{\text{up}}(\lambda, \mu).$$

In fact any portfolio (x'', Δ'', B'') that is equivalent to (x', Δ', B') would also be an optimal choice for the seller.

To prove the “only if” part, consider a Nash equilibrium where the seller's choice is $(x', \Delta', B') \in \Phi$ and the devil's choice is q' . Since this is a Nash equilibrium, the devil must be *indifferent* among all nodes in the support of q' , i.e., the discounted deficit $D(u)$ of the portfolio must equal $\alpha = \overline{D}_{q'}$ for every $u \in q'$ (otherwise the devil could drop some nodes from the support of q' and improve his utility). Moreover, there is no node in the tree where the deficit exceeds α . Therefore we can construct a new portfolio (x'', Δ'', B'') with initial wealth $x'' = x' + \alpha$ that is equivalent to (x', Δ', B') , and the discounted deficit of this portfolio would be ≤ 0 everywhere, and exactly 0 at every node $u \in q'$. In other words (x', Δ', B') is a super-replicating portfolio whose value exactly matches the option payoff at every node in the support of q' . By the complementary slackness property stated above, this implies both that q' is optimal for (9.1) and that (x'', Δ'', B'') is optimal for (8.2). ■

We now prove a necessary condition for a Nash equilibrium in the stopping game.

Theorem 10.3 (Stopping game) *Suppose a mixture $\sigma' \in \overline{T}$ and a self-financing portfolio $(x', \Delta', B') \in \Phi$ constitute a Nash equilibrium in the stopping game. Then σ' is an optimal mixture in the definition (9.3) of $h(\overline{T}, \Phi)$, and (x', Δ', B') is equivalent to an optimal super-replicating portfolio (defined by (8.2)). In particular the seller's cost and devil's utility at any Nash equilibrium is $h_{\text{up}}(\lambda, \mu)$.*

Proof Consider a Nash equilibrium where the seller chooses $(x', \Delta', B') \in \Phi$ and the devil chooses $\sigma' \in \overline{T}$. For brevity we will denote the discounted deficit process $D^{x', \Delta', B'}$ by D . Let $\delta = \widehat{D}_{\sigma'}$, so that the seller's cost and devil's utility is $(x' + \delta)$. Since this is a Nash equilibrium, the devil must be *indifferent* among all $\tau \in \sigma'$, and so $\widehat{D}_{\tau} = \delta$ for all $\tau \in \sigma'$. Also, there is no $\tau \in \overline{T}$ with $\widehat{D}_{\tau}^{x', \Delta', B'} > \delta$. Thus the discounted deficit $D(u)$ of the portfolio (x', Δ', B') is at most δ at every node u . This means we can construct a self-financing portfolio equivalent to (x', Δ', B') by starting with wealth $x'' = x' + \delta$, and the deficit of this portfolio would be ≤ 0 everywhere, i.e., the portfolio value dominates the option payoff everywhere. Thus this is a super-replicating portfolio with initial wealth $x' + \delta$, and by definition of $h_{\text{up}}(\lambda, \mu)$ we must have $x' + \delta \geq h_{\text{up}}(\lambda, \mu)$. However we know that for any mixture σ and any optimal super-replicating portfolio (x, Δ, B) we have $x + \widehat{D}_{\sigma}^{x, \Delta, B} \leq h_{\text{up}}(\lambda, \mu)$, and since the seller must make an optimal choice, her cost (as well as the devil's utility) must be $x' + \delta = h_{\text{up}}(\lambda, \mu)$. Thus the seller's choice is a portfolio equivalent to an optimal super-replicating portfolio. In particular, for the devil's choice of σ' , the seller's minimum cost is $h_{\text{up}}(\lambda, \mu)$. From the definition (9.5) of $h(\overline{T}, \Phi)$ and the fact that $h(\overline{T}, \Phi) = h_{\text{up}}(\lambda, \mu)$ (Theorem 9.2) it follows that the devil's choice of σ' optimizes (9.3). ■

11 The case of zero transaction costs

The previous sections showed that randomized (or mixed) exercise strategies are an essential feature of the upper-hedging price representations. By contrast, with zero transaction costs, it turns out to be sufficient to restrict attention to only pure exercise strategies in all of the previous results. In this section we will examine this difference between the zero and non-zero transaction costs cases in some detail.

Recall that the upper hedging price h_{up} is defined in terms of a super-replication problem: h_{up} is the smallest initial wealth needed so that the option seller is able to hedge against any exercise strategy τ of the buyer's choosing (we will drop the arguments λ, μ for brevity here). In other words, starting with h_{up} , the seller can construct a *single* hedging portfolio that will hedge against any exercise strategy τ the buyer chooses. We call such a hedge an **oblivious hedge**, because it does not depend on any specific exercise strategy. Thus h_{up} is the smallest initial wealth needed to construct an oblivious hedge.

Now it is natural to ask what is the maximum value of h_{up}^{τ} over all τ , i.e., what is the initial wealth needed to hedge against the “toughest” exercise strategy? That is, we want to compute

$$h_{\text{up}}^* := \max_{\tau \in \overline{T}} h_{\text{up}}^{\tau}. \quad (11.1)$$

Notice that h_{up}^* is the smallest initial wealth needed to hedge against any *pre-announced* exercise strategy τ . In other words, if the option buyer announces his exercise strategy τ in advance (and sticks to it), then the seller, starting with an initial wealth h_{up}^* , can construct

a hedge that *depends* on τ , and dominates the payoff at time τ . We call such a hedge an **adaptive hedge**, because the specific portfolio can vary depending on the pre-announced exercise strategy. Thus h_{up}^* is the smallest initial wealth needed for an adaptive hedge. A reasonable question is:

Does h_{up}^* suffice to construct an oblivious hedge, i.e., a *single* hedge portfolio that hedges against *all* exercise strategies?

Here is the subtle difference between the case of zero and no transaction costs. With zero transaction costs, the answer to the above question is always “yes”, i.e., $h_{\text{up}}^* = h_{\text{up}}$, as Theorem 11.2 states. However with non-zero transaction costs, the answer is in general “no”, i.e., h_{up}^* may be smaller than h_{up} ; Section 15 shows an example of this situation.

First we will show in Theorem 11.1 that if $\lambda = \mu = 0$ then the outer maximization in the definition (9.2) of $h(X)$ can be restricted to pure stopping times without affecting the value of the optimum. If in addition there exists a martingale measure with positive measure on every path, then the inner maximization can be restricted to the set P of martingale measures. Thus the representation (1.1) in Harrison and Kreps (1979) follows from our results.

Theorem 11.1 (Upper hedging price with zero transaction costs) *If $\lambda = \mu = 0$ in our model, the upper hedging price is given by*

$$h_{\text{up}}(0, 0) = \max_{\tau \in T} \max_{\mathbf{P} \in P(0, 0, \tau)} \mathbb{E}^{\mathbf{P}} G_{\tau}^*. \quad (11.2)$$

If there exists a martingale measure $\mathbf{Q} \in P$ with $\mathbf{Q}(\omega) > 0$ for every $\omega \in \Omega$, then the upper hedging price is

$$h_{\text{up}}(0, 0) = \max_{\tau \in T} \max_{\mathbf{P} \in P} \mathbb{E}^{\mathbf{P}} G_{\tau}^*. \quad (11.3)$$

Proof Consider the expression (9.1) defining $h(Q(\lambda, \mu))$ (which by Theorem 9.1 equals $h_{\text{up}}(\lambda, \mu)$), with $\lambda = \mu = 0$:

$$h_{\text{up}}(0, 0) = \max_{q \in Q(0, 0)} \sum_u q(u) G^*(u).$$

Theorem 6.7 says that the extreme points of the convex set $Q(0, 0)$ are simple node-measures, and since the maximum of a linear function over a convex set must occur at one of its extreme points (Bertsekas, 1995), it suffices to do the above maximization over simple node-measures in $Q(0, 0)$:

$$h_{\text{up}}(0, 0) = \max_{\tau \in T} \max_{q \in Q(0, 0, \tau)} \overline{G}_q^*, \quad (11.4)$$

and by Theorem 6.6 we can write this as (11.2).

Now suppose there exists a martingale measure \mathbf{Q} with positive measure on every path. Then we claim that the domain of the inner maximization in (11.2) can be restricted to the set P of all martingale measures without affecting the value of the optimum. This follows from two facts. First, any measure $\mathbf{P} \in P$ also belongs to $P(0, 0, \tau)$. Second, any measure $\mathbf{P} \in P(0, 0, \tau)$ can be redefined to be a martingale measure \mathbf{P}' such that $\mathbb{E}^{\mathbf{P}'} G_{\tau}^* = \mathbb{E}^{\mathbf{P}} G_{\tau}^*$, as follows (forward inductively):

$$\mathbf{P}'(u) = \begin{cases} \mathbf{P}(u), & \text{if } u \in I_{\tau} \cup J_{\tau}, \\ \mathbf{P}'(u^-) \mathbf{Q}(u) / \mathbf{Q}(u^-), & \text{otherwise,} \end{cases}$$

where the notation I_τ, J_τ is defined in the proof of Theorem 9.3. This new measure is identical to \mathbf{P} at or before the stopping boundary of τ , and for any node u beyond this boundary, the ratio $\mathbf{P}'(u)/\mathbf{P}'(u^-)$ equals $\mathbf{Q}(u)/\mathbf{Q}(u^-)$ (whenever $\mathbf{P}'(u^-) > 0$). Thus it is easy to verify that \mathbf{P}' is a martingale measure. This establishes our claim, and the Theorem follows. ■

Theorem 11.2 (Adaptive and oblivious hedge with zero transaction costs) *With $\lambda = \mu = 0$, the minimum initial wealth needed to hedge against the “toughest” stopping time also suffices to hedge everywhere. In other words an oblivious hedge requires no more initial wealth than an adaptive hedge, or:*

$$h_{up}^* = h_{up}(0, 0).$$

Proof From Theorem 9.3 it follows that for any $\tau \in T$,

$$h_{up}^\tau = h(Q(0, 0, \tau)) := \max_{q \in Q(0, 0, \tau)} \overline{G}_q^*,$$

and so $\max_{\tau \in T} h_{up}^\tau$ equals expression (11.4) in the proof of Theorem 11.1, and we showed there that it equals $h_{up}(0, 0)$. ■

The above result implies that in the absence of transaction costs, the optimization problems in the definitions of $h(X, \Phi)$ and $h(\overline{T}, \Phi)$, have optimal solutions that involve only pure stopping times. It also implies that in the absence of transaction costs, there is a Nash equilibrium in the stopping game where the devil picks a pure stopping time. This is stated in the following Theorem.

Theorem 11.3 (Representations of $h_{up}(0, 0)$) *Suppose $\lambda = \mu = 0$. Let $\tau \in T$ be a stopping time that maximizes h_{up}^τ . Let $(x, \Delta, B) \in \Phi$ be any optimal super-replicating portfolio. Then:*

1. *The randomized stopping time X^τ and the portfolio (x, Δ, B) are optimal for expression (9.4) defining $h(X, \Phi)$,*
2. *the degenerate mixture τ , and the portfolio (x, Δ, B) are optimal for expression (9.5) defining $h(\overline{T}, \Phi)$, and*
3. *in the stopping game, there is a Nash equilibrium where the devil chooses the pure stopping time τ , and the seller chooses the portfolio (x, Δ, B) .*

Proof For brevity let us denote $h_{up}(0, 0)$ by h_{up} . Note that if $X = X^\tau$ then $\delta := \widehat{D}_X^{x, \Delta, B} = \widehat{D}_\tau^{x, \Delta, B}$. Since (x, Δ, B) is an optimal super-replicating portfolio, $x + \delta \leq h_{up}$. We claim there cannot exist a portfolio (x', Δ', B') with $x' + \widehat{D}_\tau^{x', \Delta', B'} < h_{up}$. This is because if there were such a portfolio, we could construct a new portfolio equivalent to (x', Δ', B') that starts with initial wealth $x' + \delta' < h_{up}$. This would be a self-financing portfolio that starts with wealth less than h_{up} and dominates the option payoff at time τ . This however would contradict Theorem 11.2, since τ is a “toughest” stopping time. Thus we have $x + \widehat{D}_X^{x, \Delta, B} = h_{up}$, and so (x, Δ, B) is optimal for the inner “min” in (9.4) and (9.5), for $X = X^\tau$ and $\sigma = \tau$ respectively. By Theorem 9.2 this in turn implies that $X = X^\tau$ and $\sigma = \tau$ are optimal for the outer “max” in the respective expressions. This proves the first two statements of the Theorem.

Consider now the stopping game. The observations of the previous paragraph imply that the seller’s choice of portfolio (x, Δ, B) is optimal, given that the devil has chosen the pure strategy τ ; the seller’s cost is $x + \widehat{D}_\tau^{x, \Delta, B} = h_{up}$. Also, since (x, Δ, B) is a super-replicating portfolio, the devil cannot hope to pick a τ' for which $\widehat{D}_{\tau'}^{x, \Delta, B} > 0$. Thus, given that the seller has chosen (x, Δ, B) , if the devil chooses τ then his utility would be $x + \widehat{D}_\tau^{x, \Delta, B} = h_{up}$, and this would be highest possible utility he can get. This proves the third statement. ■

12 Representations of the Lower Hedging Price

Just as in Section 9 we will represent the lower hedging price $h_{\text{low}}(\lambda, \mu)$ in terms of some optimization problems:

$$g(T, Q) := \max_{\tau \in T} \min_{q \in Q(\lambda, \mu, \tau)} \overline{G^*}_q \quad (12.1)$$

$$g(T) := \max_{\tau \in T} \min_{\mathbf{P} \in P(\lambda, \mu, \tau)} \mathbb{E}^{\mathbf{P}} G^*_\tau \quad (12.2)$$

$$h_{\text{low}}^\tau := \max\{x \geq 0 \mid \exists(\Delta, B) \in \Phi(-x), \text{ s.t. } V_\tau^{-x, \Delta, B} + G_\tau \geq 0\}, \quad \forall \tau \in T. \quad (12.3)$$

A few remarks about the above expressions might be helpful. In the definition (12.1) of $g(T, Q)$, for a fixed $\tau \in T$, the inner “min” represents the smallest q -weighted discounted option payoff, over all (λ, μ, τ) -approximate martingale node-measures q . Expression (12.2) for $g(T)$ is analogous to (9.2) except that it involves only ordinary stopping times. Theorem 12.2 asserts that $g(T, Q)$ and $g(T)$ equal the lower hedging price $h_{\text{low}}(\lambda, \mu)$, as defined in (8.4). Example 15.1 in Section 15 illustrates the computation of $h_{\text{low}}(\lambda, \mu)$.

It is possible to derive representations analogous to (9.3)-(9.6) for $h_{\text{low}}(\lambda, \mu)$, but we do not do so in this paper. Definition (12.3) of h_{low}^τ is analogous to definition (9.8) of h_{up}^τ ; it is the maximum amount of cash the option buyer can borrow to purchase the option and be able to repay the loan by exercising the option according to τ . Theorem 12.1 shows a representation of h_{low}^τ which is analogous to that shown in Theorem 9.3 for h_{up}^τ , with the “max” replaced by a “min”.

Theorem 12.1 (Buyer’s hedge against a given τ)

$$h_{\text{low}}^\tau = \min_{q \in Q(\lambda, \mu, \tau)} \overline{G^*}_q.$$

Proof The proof is almost identical to that of Theorem 9.3. ■

Theorem 12.2 (Lower hedging price) *In our market model, the lower hedging price $h_{\text{low}}(\lambda, \mu)$ is given by*

$$h_{\text{low}}(\lambda, \mu) = g(T, Q) = g(T), \quad (12.4)$$

where the inner “min” is interpreted as $+\infty$ if $P(\lambda, \mu, \tau)$ is empty.

Proof Clearly $h_{\text{low}}(\lambda, \mu)$ defined in (8.4) can be written as

$$h_{\text{low}}(\lambda, \mu) = \max_{\tau \in T} h_{\text{low}}^\tau,$$

and the Theorem follows from Theorem 12.1 and Theorem 6.6. ■

Remark 12.3 It is possible to show using duality arguments that

$$h_{\text{low}}(\lambda, \mu) = \max_{X \in \mathcal{X}} \min_{\mathbf{P} \in P(\lambda, \mu, X)} \mathbb{E}^{\mathbf{P}} G_X^*, \quad (12.5)$$

where the right hand side is a “relaxation” of the expression (12.2) defining $g(T)$. “Relaxation” is a commonly used term in mathematical programming, to refer to the program that results when integrality restrictions are removed; such relaxed programs are easier to compute numerically. In our case expression (12.2) involves ordinary stopping times, which can be modeled by integer variables, and removing the integer restrictions gives rise to the expression on the right in (12.5) involving randomized stopping times. For reasons of space we do not exhibit the proof of (12.5) in this paper.

Unlike the upper hedging price $h_{\text{up}}(\lambda, \mu)$, the lower hedging price $h_{\text{low}}(\lambda, \mu)$ may be unboundedly large. We show in Theorem 13.1 however that $h_{\text{low}}(\lambda, \mu)$ is finite if and only if the market is arbitrage-free. The intuition behind this difference between the upper and lower hedging prices is the following. Regardless of whether or not the market is arbitrage-free, the definition (8.2) of $h_{\text{up}}(\lambda, \mu)$ implies that it cannot be smaller than 0 since the option payoff function G is non-negative. However if there is a buyer's arbitrage (defined by (8.6)), then the buyer can borrow an unboundedly large amount of cash and still pay it back by exercising the option appropriately.

With zero transaction costs we show that the expression (1.2) of Harrison and Kreps (1979) follows from our results:

Theorem 12.4 (Lower hedging price with zero transaction costs) *If $\lambda = \mu = 0$ in our model, and if there exists a martingale measure $\mathbf{Q} \in \mathcal{P}$ with $\mathbf{Q}(\omega) > 0$ for all $\omega \in \Omega$, then for an American option G the lower hedging price is given by*

$$h_{\text{low}}(0, 0) = \max_{\tau \in T} \min_{\mathbf{P} \in \mathcal{P}} \mathbb{E}^{\mathbf{P}} G_{\tau}^*. \quad (12.6)$$

Proof From Theorem 12.1 it follows that

$$h_{\text{low}}(0, 0) = \max_{\tau \in T} \min_{\mathbf{P} \in \mathcal{P}(0, 0, \tau)} \mathbb{E}^{\mathbf{P}} G_{\tau}^*,$$

where the domain of the inner minimization is the set of measures \mathbf{P} in $\mathcal{P}(0, 0, \tau)$. By the same reasoning as in the proof of Theorem 11.1 we can replace this domain by the set \mathcal{P} of all martingale measures. ■

13 Conditions equivalent to the absence of arbitrage

Theorem 13.1 *The following statements are equivalent:*

1. *The lower hedging price $h_{\text{low}}(\lambda, \mu)$ is finite for every American option G ,*
2. *There exists a (λ, μ) -approximate martingale measure \mathbf{P} , i.e., $\mathcal{P}(\lambda, \mu) \neq \emptyset$,*
3. *$h_{\text{low}}(\lambda, \mu) \leq h_{\text{up}}(\lambda, \mu)$ for every American option G ,*
4. *The market is arbitrage-free.*

Proof We first show 1 and 2 are equivalent. From Theorem 12.2 it follows that (regardless of the American option G) $h_{\text{low}}(\lambda, \mu)$ is finite if and only if for every stopping time τ (and in particular the stopping time $\tau = n$), the inner “min” in the definition (12.2) of $g(T)$ is finite, which means there exists some $\mathbf{P} \in \mathcal{P}(\lambda, \mu, \tau)$, which means $\mathcal{P}(\lambda, \mu) \neq \emptyset$. The equivalence of 1 and 3 follows from the Strong Duality Theorem of Linear Programming (Theorem A.1) and the fact (Theorem 9.1) that $h_{\text{up}}(\lambda, \mu)$ is always finite. The equivalence of 3 and 4 was already stated as Corollary 8.3. ■

We now show that as transaction costs increase, the arbitrage-free interval “expands”, in the sense that the interval with higher transaction costs contains the interval with smaller transaction costs.

Theorem 13.2 *If $0 \leq \lambda \leq \lambda' < 1$ and $0 \leq \mu \leq \mu' < 1$ and $\mathcal{P}(\lambda, \mu) \neq \emptyset$, then*

$$h_{\text{low}}(\lambda', \mu') \leq h_{\text{low}}(\lambda, \mu) \leq h_{\text{up}}(\lambda, \mu) \leq h_{\text{up}}(\lambda', \mu') < \infty.$$

Proof This easily follows from the observation that

$$\bigcup_{X \in \mathcal{X}} P(\lambda, \mu, X) \subseteq \bigcup_{X \in \mathcal{X}} P(\lambda', \mu', X).$$

■

An easy corollary to this Theorem is:

Corollary 13.3 *If there exists a martingale measure in our market model, i.e., if $P \neq \emptyset$, then the range of arbitrage-free option prices without transaction costs, is contained in the range $[h_{low}(\lambda, \mu), h_{up}(\lambda, \mu)]$, and the limits of this range are finite.*

14 European options with transaction costs

As mentioned in the Introduction, this case has been studied in several papers. In particular for European options, our results imply an expectation representation that is very similar to the one shown by Koehl et al. (1999).

Theorem 14.1 *For a European option G in our model, if the market is arbitrage-free, the upper and lower hedging prices are given by*

$$h_{up}(\lambda, \mu) = \max_{\mathbf{P} \in P(\lambda, \mu)} \mathbb{E}^{\mathbf{P}} G_n^*, \quad (14.1)$$

$$h_{low}(\lambda, \mu) = \min_{\mathbf{P} \in P(\lambda, \mu)} \mathbb{E}^{\mathbf{P}} G_n^*. \quad (14.2)$$

Proof A European option G is merely a special case of an American option with payoff $G_k = 0$ for all $k < n$. Therefore in the $\max_{X \in \mathcal{X}}$ in expression (9.2) for $h_{up}(\lambda, \mu)$, we need only consider the special randomized exercise strategy X with $X_n = 1$ (and all other $X_i = 0$ for $i < n$ by definition). This means that the inner max in that expression can be restricted to $\mathbf{P} \in P(\lambda, \mu)$, i.e., the (λ, μ) -approximate martingale measures. This implies expression (14.1) above. The argument for expression (14.2) is similar. ■

15 Examples

In Section 9 we showed several representations for the upper hedging price $h_{up}(\lambda, \mu)$ (defined by (8.2)). In particular the expression (9.2) defining $h(X)$ involves a maximization over randomized stopping times $X \in \mathcal{X}$. A natural question is, if we restrict this maximization to just the pure stopping times $\tau \in \mathcal{T}$, is the optimum the same? In Section 11 we saw that with zero transaction costs the answer to this question is “yes”. The following example shows that in the presence of transaction costs the answer to this question is “no”: (non-degenerate) randomized strategies are essential in the representation of $h_{up}(\lambda, \mu)$. Some more insight into the need for randomized strategies is provided in Example 15.3 later in this section. The example below also illustrates the computation of the lower hedging price.

Example 15.1 *(Need for randomized strategies.)* We will show an example where no optimal simple node-measure exists for the optimization problem in the definition (9.1) of $h(Q(\lambda, \mu))$. Proposition 5.5 then implies that the optimal X in expression (9.2) for $h(X)$

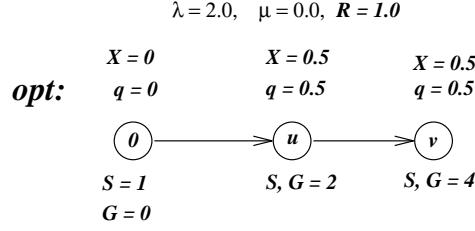


Fig. 15.1 A simple event tree showing that with non-zero transaction costs, it is necessary to use a randomized exercise strategy X to optimize the second expression in (9.9). Node names are shown inside the nodes, and the stock price S and payoff G are indicated under each node. The optimal node-measure q and strategy X are indicated above the nodes.

cannot correspond to a pure stopping time. A non-degenerate randomized exercise strategy is thus necessary in this example.

Consider the simple 3-node event tree of Figure 15.1, with the parameters shown there. The optimization problem corresponding to the expression (9.1) defining $h(Q(\lambda, \mu))$ (and therefore $h_{\text{up}}(\lambda, \mu)$) is:

$$h_{\text{up}}(\lambda, \mu) = \max q(0)G(0) + q(u)G(u) + q(v)G(v) \quad (15.1)$$

subject to

$$q(0) + q(u) + q(v) = 1, \quad (15.2)$$

$$q(v)S(u)(1 - \mu) \leq q(v)S(v) \leq q(v)S(u)(1 + \lambda), \quad (15.3)$$

$$(q(u) + q(v))S(0)(1 - \mu) \leq q(u)S(u) + q(v)S(v) \leq (q(u) + q(v))S(0)(1 + \lambda), \quad (15.4)$$

$$q(0) \geq 0, \quad q(u) \geq 0, \quad q(v) \geq 0. \quad (15.5)$$

We will write $q = (q(0), q(u), q(v))$ for the vector of unknowns, i.e., the node-measure. Constraint (15.3) always holds independently of the q values, since $S(u) < S(v)$ and $S(v) = 4 < S(u)(1 + \lambda) = 2 \times 3.0 = 6$. Similarly, the left inequality in (15.4) always holds since $S(0) < S(u)$ and the middle term is always at least $S(u)$. Since $S(0)(1 + \lambda)$ lies between $S(u)$ and $S(v)$, we can choose $q(0) = 0$, and choose $q(u)$ so that we maximize

$$q(u)G(u) + (1 - q(u))G(v)$$

subject to

$$q(u)S(u) + (1 - q(u))S(v) \leq S(0)(1 + \lambda); \quad 0 \leq q(u) \leq 1.$$

Noting that G and S are equal at u and v , the optimal $q(u)$ is given by

$$2q(u) + 4(1 - q(u)) = 3, \text{ or } q(u) = 0.5,$$

which means the optimal solution is $q = (0, 0.5, 0.5)$. The corresponding value of the objective function is thus

$$h_{\text{up}}(\lambda, \mu) = 3.$$

It is easy to verify that this objective value cannot be achieved with any simple node-measure: there are only two simple node-measures that are (λ, μ) -approximate martingale node-measures: the node-measure $q = (1, 0, 0)$, which yields an objective function value

of 0, and $q = (0, 1, 0)$, which yields an objective value of 2. Thus there is no simple node-measure that optimizes expression (9.1) defining $h(Q(\lambda, \mu))$.

Note that on this simple event tree there can only be one measure \mathbf{P} , namely the one given by $\mathbf{P}(0) = \mathbf{P}(u) = \mathbf{P}(v) = 1$. Thus for any node-measure q on this tree, the unique equivalent measure-strategy pair is (\mathbf{P}, X) , where $X(\cdot) = q(\cdot)$ everywhere. Therefore the optimal X in expression (9.2) defining $h(X) = h_{\text{up}}(\lambda, \mu)$ is the one where $X(0) = 0$ and $X(u) = X(v) = 0.5$. Again it is easy to see that there is no pure stopping time that optimizes that expression, for this example.

Now let us turn to the computation of the lower hedging price $h_{\text{low}}(\lambda, \mu)$ using the max-min expression in (12.4):

$$h_{\text{up}}(\lambda, \mu) = \max_{\tau \in T} \min_{\mathbf{P} \in P(\lambda, \mu, \tau)} \mathbb{E}^{\mathbf{P}} G_{\tau}^*.$$

As noted above, there is only one possible probability measure \mathbf{P} on this event tree, and so for any fixed τ , the inner “min” equals $\mathbb{E}^{\mathbf{P}} G_{\tau}^*$ if $\mathbf{P} \in P(\lambda, \mu, \tau)$, and equals $+\infty$ otherwise. There is only one stopping time τ for which $\mathbf{P} \in P(\lambda, \mu, \tau)$, namely $\tau = 0$ (i.e. stop at node 0). Since we want to maximize the inner “min” above with respect to τ , the optimal τ to pick is either $\tau = 1$ or $\tau = 2$, and in both cases the value of the inner “min” is $+\infty$. Thus the lower hedging price in this example is unboundedly large; the American option in this case admits an arbitrage opportunity. This is not surprising in light of Theorem 13.1 where we showed that the market is arbitrage-free if and only if there is a (λ, μ) -approximate martingale measure, i.e., $P(\lambda, \mu) \neq \emptyset$. In this example the only possible measure \mathbf{P} is *not* a (λ, μ) -approximate martingale measure.

Let us consider this example with zero transaction costs, i.e., $\lambda = \mu = 0$. It is easy to see that the only $(0, 0)$ -approximate martingale node-measure is the one with $q(0) = 1, q(u) = 0, q(v) = 0$, and the corresponding objective function value is $h_{\text{up}}(0, 0) = 0$. The lower hedging price can be seen to be $h_{\text{low}}(0, 0) = +\infty$. ■

The following example illustrates the difference between an adaptive and oblivious hedge.

Example 15.2 (Hedging) Let us consider the previous example from the perspective of adaptive and oblivious hedging (as defined in Section 11). There are just three pure stopping times on this event-tree: we denote the stopping time $\tau = i$ by τ_i , for $i = 0, 1, 2$. Consider the quantity $h_{\text{up}}^{\tau_i}$ defined in (9.8). For brevity we denote $h_{\text{up}}^{\tau_i}$ by h_{up}^i . First take the case of non-zero transaction costs, with $\lambda = 2.0, \mu = 0$. From Theorem 9.3 we know that

$$h_{\text{up}}^i = \max_{q \in Q(\lambda, \mu, \tau_i)} \overline{G^*}_q.$$

In this example the right hand side has a particularly simple form: if the unique node-measure q corresponding to τ_i is a (λ, μ) -approximate martingale node-measure, then the “max” equals $\overline{G^*}_q$, otherwise it equals $-\infty$. Thus $h_{\text{up}}^0 = 0, h_{\text{up}}^1 = 2$ and $h_{\text{up}}^2 = -\infty$, and so the “toughest” exercise strategy is $\tau = 1$, and an initial wealth of

$$h^* = 2$$

suffices to construct a self-financing portfolio that hedges against any pre-announced exercise strategy, i.e., for an *adaptive* hedge. However, since $h_{\text{up}}(\lambda, \mu) = 3$, we know that an *oblivious* hedge requires initial wealth 3. Thus, any portfolio that starts with wealth 2 will fail to dominate the payoff at at least one node of the tree.

By contrast, if $\lambda = \mu = 0$, we know from the previous example that $h^* = h_{\text{up}}(0, 0) = 0$. As noted above, the only $(0, 0)$ -approximate martingale node-measure is $q = (1, 0, 0)$, so that $h_{\text{up}}^0 = 0$ and $h_{\text{up}}^1 = h_{\text{up}}^2 = -\infty$. Thus the “toughest” exercise strategy is $\tau = 0$, and the smallest initial wealth needed to hedge against it is $h^* = 0$. Since $h^* = h_{\text{up}}(0, 0)$, we know that this initial wealth also suffices to construct an oblivious hedge. ■

Nash equilibria in the stopping game and the node game are illustrated in the next example.

Example 15.3 (Nash Equilibrium) Now for the same example above, let us consider the stopping game defined in Section 10. Since there is only one path in our event tree, the stopping game and node game are in fact equivalent. First let us examine the case of non-zero transaction costs, with $\lambda = 2$, $\mu = 0$. Suppose the devil simply picks the “toughest” exercise strategy $\tau = 1$, and the seller picks the smallest-initial-wealth portfolio that dominates the payoff at time 1 (i.e., node u). Such a portfolio would match the payoff exactly at node u , and so the seller’s cost and devil’s utility would be $h_{\text{up}}^1 = 2$. However these choices would *not* constitute a Nash equilibrium, for the following reason. We know that the seller’s portfolio value must fall below the option payoff at node 0 or v , so the devil could switch to a different stopping time and improve his utility. However if the devil chooses a *mixture* of the stopping times $\tau = 1$ and $\tau = 0.5$, each with weight 0.5, and the seller picks an optimal super-replicating portfolio, then by Theorem 10.2 this would be a Nash equilibrium, where the seller’s cost is $h_{\text{up}}(\lambda, \mu) = 3$. ■

Example 15.4 (American call option) A more realistic example where a non-degenerate randomized stopping time is required is an American call option, as shown in Figure 15.2. The node-measure q that optimizes (9.1) for this example is not a simple node-measure. This means that the optimum X in expression (9.2) (which equals $h_{\text{up}}(\lambda, \mu)$) is a non-degenerate randomized strategy. Note that if there were no transaction costs, the optimal X would correspond to a pure stopping time – it is well known (Duffie, 1996) that under the martingale measure the optimal exercise strategy for an American option (on a non-dividend-paying stock) is $\tau = n$, i.e., exercise at expiration. However the example in Fig. 15.2 shows that with non-zero transaction costs, the optimal randomized strategy may have some non-zero “probability” of exercise before expiration. It is not immediately clear what the interpretation of this example ought to be: what is the optimal or “rational” exercise policy for the option holder?

One difficulty with addressing this question is that in expression (9.2), the set of measures in the domain of the inner “max” *depends* on which X is picked in the outer “max”. If this were not the case, standard results in linear programming would imply that the optimum of (9.2) must occur at a pure stopping time. This state of affairs appears to confirm a remark of Duffie (1996): “The real difficulties with analyzing American securities begin with incomplete markets. In that case, the choice of exercise policy may play a role in determining the marketed subspace, and therefore a role in pricing securities. If the state-price-deflator depends on the exercise policy, it could even turn out that the notion of a rational exercise policy is not well defined.” ■

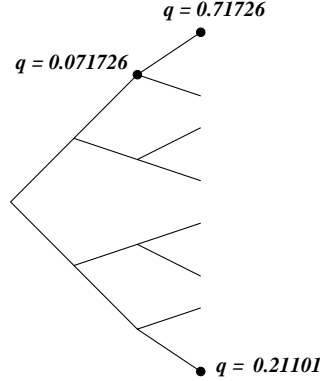


Fig. 15.2 A binomial event tree showing a non-simple node-measure q that optimizes expression (9.1) defining $h(Q(\lambda, \mu))$, for an American call option with the following parameters: $n = 3$, initial stock price $S_0 = 10$, strike price $K = 8$, stock price “up factor” (factor by which stock price moves up or down) $u = 1.1$, $\lambda = \mu = 0.2$, and $R = 1.0$. Only the positive q values are shown.

16 Concluding Remarks

This paper considered the following question. In the presence of proportional transaction costs, what is the permissible range of prices of an American option in the absence of arbitrage? This range is defined by the upper and lower hedging prices. We examined this pricing problem in a general discrete-time event-tree model. Our goal was to generalize the expectation representations (1.1) and (1.2) that were shown by Harrison and Kreps (1979) for a frictionless but incomplete market where there may be multiple martingale measures. We found that in the presence of proportional transaction costs, expressions analogous to these can be written, provided we generalize both the notion of a “martingale” and that of an “exercise strategy”. In particular we need the notion of an approximate martingale measure. More interestingly and surprisingly, we need to expand the class of stopping times to include *randomized* stopping times. We showed several perspectives from which to understand why randomized strategies are needed when there are transaction costs, including one based on Nash equilibria in a two-player game. We derived most of our results using elementary methods, requiring only the Strong Duality Theorem of linear programming.

One of the main contributions of this paper has been to show that randomized stopping times are necessary to arrive at an elegant expectation representation of the upper hedging price of an American option with proportional transaction costs, in discrete time. We believe it would be interesting to consider similar representations in continuous time, or consider the continuous-time limit of our representations. Randomized stopping times can also be defined in continuous time; in fact Baxter and Chacon (1977) introduce this concept in the continuous-time setting.

We have noted in Section 11 that the value of the expression (1.1) for the case of zero transaction costs remains unchanged even if the outer maximization is over randomized exercise strategies. This raises the possibility that randomized stopping times could have *computational* applications as a *linear relaxation* of ordinary stopping times. For instance, even in the absence of transaction costs, is it possible to use randomized stopping times to design efficient Monte Carlo algorithms for pricing American options? A key difficulty in using Monte Carlo simulation to price American options is that the optimal exercise boundary is unknown. Perhaps one way to get around this difficulty is to simulate not

just the stock price paths but also the randomized exercise strategy along each path. An intriguing possibility is whether randomized stopping times can help in the quest for *closed form* formulas for say an American call option (even in the absence of transaction costs).

The computation of the expressions for the upper and lower hedging prices appears to be non-trivial. It would be useful to design efficient algorithms for approximating the values of these expressions

It would also be interesting to show expressions representing upper and lower hedging prices for American options when there are capital gains *taxes*. Karatzas and Kou (1998) have shown expressions for upper and lower hedging prices with *portfolio constraints* but without transaction costs. It would of significant interest to extend their results to the case of non-zero transaction costs; it is conceivable that randomized strategies will play a role here as well, at least in a discrete-time framework.

More works needs to be done to define an appropriate notion of “rational” exercise policy for the holder of an American option, when there are transaction costs. As discussed in Example 15.4, it is not clear what the economic implications are of the fact that (non-degenerate) randomized exercise strategies appear to turn up at the optima of the various upper-hedging price representations.

In Section 10 we showed a game between the option seller and a “devil” from which the upper hedging price emerges fairly naturally. A more economically appealing result would be to define a game between the option seller and the *buyer* which naturally gives rise to the upper and lower hedging prices. This is easy to set up without transaction costs, but is challenging when there are transaction costs.

Acknowledgements We would like to thank Steven Shreve for encouraging us to work on this paper.

A Appendix: Linear programming and duality

Let A be an $m \times n$ (m rows, n columns) matrix with real-valued entries. Let b be a column vector in \mathbb{R}^m , and let c be a column vector in \mathbb{R}^n . Let $x \in \mathbb{R}^n$ be a column vectors of variables. For any matrix Z , Z^T denotes its transpose, $Z(i, \cdot)$ denotes its i 'th row, and $Z(\cdot, j)$ denotes its j 'th column. For any column vector z , z_i refers to its i 'th entry. For two matrices X, Y , XY denotes the usual matrix product.

The following optimization problem is a *linear program* (LP) in the variable vector x :

$$\max\{c^T x : x \geq 0, Ax \leq b\}, \quad (\text{A.1})$$

or more explicitly:

$$\max \sum_{i=1}^n c_i x_i \quad (\text{A.2})$$

subject to

$$x_i \geq 0, \quad i = 1, 2, \dots, n, \quad (\text{A.3})$$

$$A(i, \cdot)x \leq b_i, \quad i = 1, 2, \dots, m. \quad (\text{A.4})$$

The function (A.2) is called the *objective* of the LP. The inequalities (A.3) are the *non-negativity* constraints on the variables x_i . The inequalities (A.4) are referred to simply as the *constraints*. A vector x that satisfies (A.3) and (A.4) is said to be *feasible*, and the set of such vectors is called the *feasible set* of the LP. If the feasible set is empty, the LP is said to be *infeasible*. If the LP has an unbounded optimum (i.e., $+\infty$ or $-\infty$) then the LP is said to be *unbounded*.

Since we are starting with this LP, it is called the *primal*, and the corresponding *dual* LP is

$$\min\{y^T b : y \geq 0, y^T A \geq c^T\}, \quad (\text{A.5})$$

or more explicitly:

$$\min \sum_{i=1}^m y_i b_i \quad (\text{A.6})$$

subject to

$$y_i \geq 0, \quad i = 1, 2, \dots, m, \quad (\text{A.7})$$

$$y^T A(\cdot, i) \geq c_i, \quad i = 1, 2, \dots, n. \quad (\text{A.8})$$

Notice that for $i = 1, 2, \dots, n$, the primal *variable* x_i corresponds to the dual *constraint* $y^T A(\cdot, i) \geq c_i$. Also for $i = 1, 2, \dots, m$, the primal *constraint* $A(i, \cdot)x \leq b_i$ corresponds to the dual *variable* y_i .

In general the primal LP may have *free variables*, i.e., those which are not constrained to be non-negative. There may also be *equality constraints* of the form $A(i, \cdot)x = b_i$. (If there are inequalities such as $x_i \leq 0$ or constraints of the form $A(i, \cdot)x \geq b_i$, the LP can easily be re-written so that all variables are non-negative or free, and all constraints are equalities or of the form (A.4).) To write the dual of a general LP, we modify the above dual specification as follows:

1. If primal variable x_i is *free*, the corresponding dual constraint will be an *equality* constraint: $y^T A(:, i) = c_i$.
2. For a primal *equality* constraint $A(i, :)x = b_i$, the corresponding dual variable y_i will be *free*.

The dual of a maximization LP problem can also easily be written using the above rules. It is therefore clear that the *dual of the dual LP is the primal LP*.

The following relationship between the primal and dual optima is called the **Strong Duality Theorem**, and John von Newmann is credited with stating it in privately circulated notes as early as 1947. (See also Schrijver (1986) for more details.)

Theorem A.1 (Strong Duality of LP) *Consider an LP P in general form, and its dual P' . Then exactly one of the following cases occurs:*

1. P, P' are both infeasible,
2. One of P, P' is unbounded, the other is infeasible,
3. P, P' are both feasible, and their optima are equal.

A useful optimality condition for LP problems is the following.

Theorem A.2 (Complementary Slackness) *Suppose u is a feasible solution for the primal LP, and v is a feasible solution for its dual. Then a necessary and sufficient condition for u and v to be optimal for the primal and dual LP respectively, is: for every i , if $u_i > 0$ then the corresponding dual constraint is tight at v , i.e., holds with equality.*

Let us apply this condition to the primal LP (A.1) and its dual (A.5). If u is primal feasible and v is dual feasible, then u is optimal for the primal and v is optimal for the dual, if and only if:

$$u_i > 0 \implies v^T A(:, i) = c_i, \quad i = 1, 2, \dots, n,$$

or equivalently,

$$v^T A(:, i) > c_i \implies u_i = 0, \quad i = 1, 2, \dots, n.$$

B Notation

For easy reference we give here a summary of notation used in this paper. For each symbol we show where it is defined and a brief definition.

Symbol	Where defined	Brief definition
General event-tree and process notation.		
u^+	Sec. 2	Set of immediate successors of node u .
u^-	Sec. 2	Immediate predecessor of node u .
$u < v$	Sec. 2	v is a successor of u on the event tree.
$u \leq v$	Sec. 2	$u = v$ or $u < v$.
$t(u)$	Sec. 2	Time-period corresponding to node u .
0	Sec. 2	Stands for initial node in event-tree.
T	Sec. 2	Set of terminal nodes in event-tree.
ω_k	Sec. 2	Time k node on path ω .
S_k	Sec. 2	Stock price at time k .
R	Sec. 2	Equals $1 + r$, where r is the risk-free interest rate per time-period.
G_k	Sec. 2	Payoff of American option at time k .
$Z(u)$	Sec. 2	Value of a process $\{Z_k\}$ at node u .
Z^*	Sec. 2	Discounted process corresponding to Z , i.e., $Z_k^* = Z_k R^{-k}$.
I_τ	Sec. 2	Set of nodes on stopping boundary of τ .
J_τ	Eq. (9.35)	Set of nodes inside the stopping boundary of τ .
Node-measures		
q	Sec. 3	A node probability measure. $\sum_u q(u) = 1$, and $q(u) \geq 0$ for all u .
Q	Sec. 3	Set of all possible node-measures q .
Q^τ	Sec. 3	Set of all simple node-measures q corresponding to τ .
$q^+(u)$	Sec. 4.2	$\sum_{v>u} q(v)$.
$\bar{q}(u)$	Sec. 4.2	$q(u) + q^+(u)$.
\bar{Z}_q	Eq. (4.8)	q -weighted average value of Z .
$\bar{Z}_q(u)$	Eq. (4.12)	q -weighted average future value of Z at node u .
$q \equiv (\mathbf{P}, X)$	Def. 4.4	q is equivalent to (\mathbf{P}, X) , i.e., $q(u) = \mathbf{P}(u)X(u)$ for all u .
Randomized stopping times		
X	Def. 4.1	A randomized stopping time: $\sum_{k=0}^n X_k = 1$, $X_k \geq 0$ for all k .
X_k^+	Def. 4.9	Remaining value of X at time k , or $\sum_{i=k+1}^n X_i$.
X^τ	Eq. (4.1)	Randomized stopping time corresponding to pure stopping time τ .
σ	Def. 4.3	A mixture of pure stopping times.
X^σ	Eq. (4.2)	Randomized stopping time corresponding to mixture σ .
Z_X	Eq. (4.3)	Time- X value of Z ; analogous to Z_τ .
$Z_{X k}$	Eq. (4.11)	Future time- X value of Z at time k .
$\mathbb{E}_k^{\mathbf{P}}(Z)$	Sec. 3	Shorthand for the conditional expectation $\mathbb{E}^{\mathbf{P}}(Z F_k)$.
$\mathbb{E}_k^{\mathbf{P}} Z_X$	Sec. 4.2	Shorthand for $\mathbb{E}_k^{\mathbf{P}} Z_{X k}$.
$\mathbb{E}_u^{\mathbf{P}} Z_X$	Sec. 4.2	Shorthand for $\mathbb{E}_k^{\mathbf{P}} Z_X(\omega)$ when $\omega_k = u$.
\hat{Z}_X	Eq. (4.5)	Maximum time- X value of Z .
\hat{Z}_τ	Eq. (4.6)	Maximum value of Z at τ .
\hat{Z}_σ	Eq. (4.7)	σ -expected maximum value of Z .

Symbol	Where defined	Brief definition
Approximate martingale measures		
$P(\lambda, \mu, X)$	Def. 6.1	Set of (λ, μ, X) -approximate martingale measures.
$P(\lambda, \mu, \tau)$	Sec. 6	$P(\lambda, \mu, X)$ with $X = X^\tau$.
$P(\lambda, \mu)$	Sec. 6	$P(\lambda, \mu, \tau)$ with $\tau(\omega) = n$ for all ω .
\mathcal{P}	Sec. 6	The set of all martingale measures.
$Q(\lambda, \mu)$	Def. 6.3	The set of all (λ, μ) -approximate martingale node-measures.
$Q(\lambda, \mu, \tau)$	Def. 6.3	$Q^\tau \cap Q(\lambda, \mu)$.
Portfolios and hedging		
λ	Sec. 2	Transaction cost per dollar of shares bought; $\lambda \geq 0$.
μ	Sec. 2	Transaction cost per dollar of shares sold; $0 \leq \mu \leq 1$.
Δ_k	Sec. 7	Number of shares held at time k .
B_k	Sec. 7	Number of dollars (in bonds) held at time k .
x	Sec. 7	Initial wealth of a portfolio.
Φ	Sec. 7	Set of self-financing portfolios (x, Δ, B) .
$\Phi(x)$	Sec. 7	Set of (Δ, B) such that $(x, \Delta, B) \in \Phi$.
$V^{x, \Delta, B}(v)$	Eq. (7.1)	Value of portfolio (x, Δ, B) at node v .
$D_k^{x, \Delta, B}$	Def. 8.1	Discounted deficit of portfolio (x, Δ, B) at time k .
$h_{\text{up}}(\lambda, \mu)$	Eq. (8.2)	Upper hedging price.
$h_{\text{low}}(\lambda, \mu)$	Eq. (8.4)	Lower hedging price.
h_{up}^τ	Eq. (9.8)	Min. wealth needed for seller to hedge against τ .
h_{up}^*	Eq. (11.1)	Min. wealth needed for seller to hedge against toughest τ .

References

- Baxter, S. and Chacon, R. (1977). Compactness of stopping times. *Z. Wahrsch. Verw. Gebiete*, 40:169–181.
- Bensaid, B., Lesne, J., Pagès, H., and Scheinkman, J. (1992). Derivative asset pricing with transaction costs. *Mathematical Finance*, 2(2):63–86.
- Bertsekas, D. (1995). *Nonlinear Programming*. Athena Scientific, Belmont, MA.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Political Economy*, 81:637–654.
- Cvitanic, J. and Karatzas, I. (1996). Hedging and portfolio optimization under transaction costs: a martingale approach. *Mathematical Finance*, 6:133–165.
- Cvitanic, J., Pham, H., and Touzi, N. (1999). A closed-form solution to the problem of super-replication under transaction costs. *Finance and Stochastics*, 3(1):35–54. Preprint.
- Dalang, R., Trotter, L., and de Werra, D. (1988). On randomized stopping points and perfect graphs. *J. Combinatorial Theory*, 45(3):320–344.
- Davis, M. and Clark, J. (1994). A note on super-replicating strategies. *Phil. Trans. Roy. Soc. London A*, 347.
- Davis, M., Panas, V., and Zariphopoulou, T. (1993). European option pricing with transaction costs. *SIAM J. Control Optim.*, 31:470–493.
- Davis, M. and Zariphopoulou, T. (1995). American options and transaction fees. In et al., M. D., editor, *Mathematical Finance*, volume 65 of *The IMA volumes in mathematics and its applications*, pages 47–62. Springer-Verlag.
- Duffie, D. (1996). *Dynamic Asset Pricing Theory*. Princeton University Press, 2 edition.

- Edgar, G., Millet, A., and Sucheston, L. (1981). On compactness and optimality of stopping times. In *Martingale Theory in Harmonic Analysis and Banach Spaces*, volume 939 of *Lect. Notes in Math.*, pages 36–61. Springer-Verlag, Berlin.
- Edirisinghe, C., Naik, V., and Uppal, R. (1993). Optimal replication of options with transaction costs and trading restrictions. *J. Fin. and Quant. Anal.*, 28:117–138.
- Fudenberg, D. and Tirole, J. (1998). *Game Theory*. MIT Press.
- Grötschel, M., Lovász, L., and Schrijver, A. (1993). *Geometric algorithms and combinatorial optimization*. Springer-Verlag, 2nd corrected ed. edition.
- Harrison, J. and Kreps, D. (1979). Martingales and arbitrage in multiperiod security markets. *J. Econ. Theory*, 20.
- Harrison, J. and Pliska, S. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic proc. and their applications*, 11:215–260.
- Jaschke, S. (1996). Super-hedging and arbitrage pricing in markets with transaction costs and trading constraints. Technical report, Institut für Mathematik.
- Jouini, E. and Kallal, H. (1995). Martingales and arbitrage in securities markets with transaction costs. *J. econ. theory*, 66:178–197.
- Karatzas, I. and Kou, S. (1998). Hedging american contingent claims with constrained portfolios. *Finance and Stochastics*, 2:215–258.
- Karoui, N. E. and Quenez, M.-C. (1995). Dynamic programming and the pricing of contingent claims in an incomplete market. *SIAM J. Control*, 33(1):29–66.
- Koehl, P.-F., Pham, H., and Touzi, N. (1999). Hedging in discrete-time under transaction costs and continuous-time limit. To appear in *J. App. Prob.*
- Levental, S. and Skorohod, A. V. (1997). On the possibility of hedging options in the presence of transaction costs. *Annals of appl. probab.*, 7(2):410–443.
- Mazziotto, G. and Millet, A. (1986). Points, lignes, et systèmes d'arrêt flous et problème d'arrêt optimal. In *Séminaire de Probabilités XII*, volume 649 of *Lect. Notes in Math.*, pages 81–94. Springer-Verlag, Berlin.
- Merton, R. (1973). Theory of rational option pricing. *Bell J. Econ. Manag. Sci.*, 4:141–183.
- Musiela, M. and Rutkowski, M. (1998). *Martingale methods in financial modelling*. Number 36 in Applications of Mathematics: Stochastic Modelling and Applied Probability. Springer.
- Naik, V. (1995). *Finite state securities market models and arbitrage*, chapter 2, pages 31–64. Handbooks in OR and MS, Vol 9: Finance. Elsevier Science B.V.
- Ortu, F. (1996). Arbitrage and martingales in securities markets with transaction costs: a linear programming approach. Technical report, Columbia University Business School.
- Schrijver, A. (1986). *Theory of Linear and Integer Programming*. Wiley.
- Soner, H., Shreve, S., and Cvitanic, J. (1995). There is no nontrivial hedging portfolio for option pricing with transaction costs. *Ann. Appl. Prob.*, 5:327–355.
- Williams, D. (1991). *Probability with Martingales*. Cambridge University Press.