

WORKING PAPERS IN ECONOMICS

No. 08/07

SJUR DIDRIK FLÅM

OPTION PRICING BY MATHEMATICAL PROGRAMMING



Department of Economics

UNIVERSITY OF BERGEN

Option Pricing by Mathematical Programming

SJUR DIDRIK FLÅM*

July 10, 2007

Dedicated to Professor Hubertus Th. Jongen on his 60th birthday.

ABSTRACT. Financial options typically incorporate times of exercise. Alternatively, they embody set-up costs or indivisibilities. Such features lead to planning problems with integer decision variables. Provided the sample space be finite, it is shown here that integrality constraints can often be relaxed. In fact, simple mathematical programming, aimed at arbitrage or replication, may find optimal exercise, and bound or identify option prices. When the asset market is incomplete, the bounds stem from nonlinear pricing functionals.

Keywords: asset pricing, arbitrage, options, finite sample space, scenario tree, equivalent martingale measures, bid-ask intervals, incomplete market, linear programming, combinatorial optimization, totally unimodular matrices.

JEL: C61, G12; MSC: 90C05, 90C15, 90C27.

1. INTRODUCTION

Asset pricing is a chief concern in finance. That concern relates intimately to queries about possible *arbitrage*. In fact, much of financial analysis or speculation aims at finding money for free. When such can be found, and portfolios are unrestricted, asset pricing is not a well posed task.

But otherwise, while market values remain finite, it's often a challenge to estimate them. That challenge is especially pressing in case of *options*. These might derive from underlying papers - or, when stemming from other opportunities, they could reside outside the marketable space. In the latter case, replicating them via portfolios becomes impossible and evaluation non-unique.

Additional complexity comes with the fact that many options, notably of American sort, involve random exercise times [1], [19], [23], [25], [27]. The latter may be construed as discrete decision variables. As is well known, presence of such variables renders many a planning problem combinatorial hence "intricate" [22], [30].

The difficulties of combinatorics notwithstanding, this paper recommends optimization - alias mathematical programming - as a convenient vehicle. In finance its advantages are since long well known and quite many [8], [24]. For one: pricing and replication often obtain in a single shot. For others: one may rather easily incorporate

*Economics Department, Bergen University, 5007 Norway; sjur.flaam@econ.uib.no. Thanks are due Finansmarkedsfondet and the Arne Ryde Foundation for support - and a referee for very helpful criticism and comments.

taxes, transaction costs, trade restrictions, portfolio constraints, and bid-ask spreads [6], [19]. Also, prices may evolve in quite general manner, allowing path dependence.

It appears however, less known among financial analysts that discrete variables often cause no inconvenience. Indeed, it happens frequently that integer choice, when and where required, comes automatically - at no extra cost. This has recently been brought out in finance by Pennanen and King (2004) who dealt with American contingent claims. Following their lead this paper considers assets that incorporate constrained exercise times. When the corresponding constraints are linear, with a totally unimodular matrix, the integrality restrictions can safely be relaxed [12], [13], [16], [26].

The material below addresses both readers of finance and optimization, emphasizing strong links between the two fields. On one hand, it deals with standard financial problems of pricing, portfolio choice, replication, and hedging. On the other hand, presuming a finite sample space, all those problems are stated as mathematical programs, often linear. The approach is computational, direct, and simple. It differs somewhat from the customary one of finance by describing decision processes and information flows in manners most familiar to stochastic programmers and game theorists [3], [9], [10], [17], [21].

The *scenario tree* unifies various descriptions. It helps Sections 2-5 to draw parts of finance and optimization technology somewhat closer to one another. Those sections claim no originality but invite stochastic programmers to financial analysis - and financial analysts to stochastic programming. Sections 6&7 offer some novelties and arguments for option pricing by means of continuous optimization.

2. THE SCENARIO TREE AND THE ASSETS

This section prepares the ground. It starts by briefly describing three different structures, each naturally leading to what is commonly called a *scenario tree* [9]. Prices of primitive assets will thereafter be posted along that tree. As in [11], time is discrete. That choice facilitates both analysis and computation.

An information structure: Considered here is an exhaustive, non-empty set \mathbb{S} of possible but mutually exclusive *scenarios* or *states*. Information as to which state is relevant arrives gradually, step by step. At time $t \in \{0, \dots, T\}$ the investor can ascertain to which part $P_t(s)$ in a prescribed *partition* \mathbb{P}_t of \mathbb{S} the true s belongs. Arrival of novel information means that any part $P_{t+1} \in \mathbb{P}_{t+1}$, $t < T$, is contained in a unique *ancestor* $\mathcal{A}(P_{t+1}) = P_t \in \mathbb{P}_t$; see figure below.

Since no conclusive knowledge about s is given ex ante, one naturally posits $\mathbb{P}_0 := \{\mathbb{S}\}$. Similarly, to avoid redundancy, let $\mathbb{P}_T := \{\{s\} : s \in \mathbb{S}\}$. That is, for the sake of having \mathbb{S} small - in fact, minimal - reduce each terminal part $P_T \in \mathbb{P}_T$ to a singleton. For computational reasons, I hesitate not in assuming \mathbb{S} *finite*.

If some concerned party perceives the setting as *risky*, he predicts that s will happen with (objective or subjective) probability $\Pr(s)$. Then, to maintain \mathbb{S} minimal, he should have $\Pr(s) > 0$ for all $s \in \mathbb{S}$. If plagued instead by *uncertainty*, he might

envisage a closed set of probability measures, each non-degenerate; see [14].

A stochastic process: It's often natural to construe \mathbb{S} as comprising all relevant scenarios or trajectories $s = (s_0, \dots, s_T)$ of a stochastic process, with s_0 already specified. Component s_t is unveiled at time t . Then, $s' \in P_t(s) \Leftrightarrow (s'_0, \dots, s'_t) = (s_0, \dots, s_t)$. In this setting it's convenient to let \mathbb{S} be the *sample space*, still assumed finite. If an agent distributes probability across \mathbb{S} , then reasonably, $\Pr(s) > 0$ for each s - as before.

A decision framework: The said probability measure \Pr can be dispensed with. To see how, suppose a planner merely be concerned with various "decision nodes." These constitute a non-empty finite set \mathcal{N} on which is prescribed an antisymmetric, transitive *precedence order* \prec . That order should never recombine "non-aligned" chains, meaning:

$$(n \prec n' \ \& \ \bar{n} \prec n') \Rightarrow (n \prec \bar{n} \ \text{or} \ \bar{n} \preceq n).$$

Further, the past should always provide common connections:

$$\forall n, \bar{n} \in \mathcal{N} \ \exists \hat{n} \in \mathcal{N} \ \text{such that} \ (\hat{n} \preceq n \ \& \ \hat{n} \preceq \bar{n}).$$

Under these assumptions each node $n \in \mathcal{N}$, except *one*, has a unique immediate *ancestor* $\mathcal{A}(n) \in \mathcal{N}$. The exceptional node, called the *root*, has none. If n is the ancestor of $c \in \mathcal{N}$, the latter is declared a *child*, and we write $c \in \mathcal{C}(n)$. Nodes without children are called *leafs*. When a *chain* of immediate successors $n_0 \prec n_1 \prec \dots \prec n_t$ emanates from the root n_0 to reach node n_t , we say the latter is located at *height* t , and write $n_t \in \mathcal{N}_t$. With no loss of generality let all leafs have the same height $T > 0$.

An investor need not entertain a probabilistic perspective. Instead, he might merely hold beliefs about the likelihood or occurrence of various nodes. For the minimality of \mathcal{N} it imports though, that his subjective opinion, say in the form a non-additive measure [5], assigns positive weight to each leaf.

The scenario tree: The three structures just outlined all fit a common form. Indeed, identify parts $P_t \in \mathbb{P}_t$ with nodes $n_t \in \mathcal{N}_t$ such that $P_t = \mathcal{A}(P_{t+1}) \Leftrightarrow n_t = \mathcal{A}(n_{t+1})$. Thus emerges a *tree* with *node set* $\mathcal{N} = \cup_{t=0}^T \mathcal{N}_t$ that features a *directed branch* from n to c iff $c \in \mathcal{C}(n)$.

Calling this construct a tree is justified by letting an oriented branch lead from $P_t \in \mathbb{P}_t$ to $P_{t+1} \in \mathbb{P}_{t+1}$ iff $P_t \supseteq P_{t+1}, t < T$; see figure of (a fallen over) tree below. The pictorial representation thus obtained is a *directed graph* that springs from the root n_0 and stretches via intermediate nodes up to the leafs $n_T \in \mathcal{N}_T$. As in nature, the tree never recombines. Thus, from the root to each subsequent node leads exactly

one directed path.

$$\text{root } n_0 = \left[\begin{array}{c} s \\ s' \\ s'' \end{array} \right] \begin{array}{l} \nearrow \\ \searrow \end{array} \left[\begin{array}{c} s \\ s' \\ s'' \end{array} \right] \begin{array}{l} \nearrow \\ \rightarrow \\ \rightarrow \end{array} \left. \begin{array}{c} [s] \\ [s'] \\ [s''] \end{array} \right\} \text{ leafs}$$

Legend: A tree with 3 stages/states/scenarios and 6 nodes.

The probability distribution \Pr , if any, plays from here on no chief role. It serves merely to identify the set \mathbb{S} as *support*, comprising precisely those states that always carry positive probability (likelihood or belief).

Any (non-degenerate) probability measure μ over \mathbb{S} amounts to have the same over \mathcal{N}_T . On deeper nodes $n \notin \mathcal{N}_T$ it recursively induces probability $\mu(n) := \mu(\mathcal{C}(n)) := \sum_{c \in \mathcal{C}(n)} \mu(c)$. If $\mu(n) > 0$ and $n \notin \mathcal{N}_T$, there is a *transition probability* $\mu(c|n) := \mu(c)/\mu(n)$ from parent node n to child $c \in \mathcal{C}(n)$. To convey that μ and \Pr are positive on the same states we write $\mu \sim \Pr$.

Traded assets: Central and fixed here is a non-empty finite set J of *primitive traded assets*. At node $n \in \mathcal{N}$ a share of asset $j \in J$ commands nominal price p_{jn} (cum dividend if any). No conditions are imposed on these. While merely defined at the nodes, prices can be: driven by multiple factors, strongly affected by the preceding path, and come rather jumpy in nature. The investor has no impact on prices, and he watches their evolution along the tree.

A special paper, henceforth called a "bond" (or *numeraire*), is singled out and labeled $b \in J$. It has $p_{bn} > 0$ at each node n . If p_{bc} remains constant across $c \in \mathcal{C}(n)$, paper b is declared *predictable* (or *previsible*) at n .¹ In terms of the bond define *discount factors* $\delta_n := p_{bn_0}/p_{bn}$.²

Some remarks on filtrations and adapted variables: Most presentations of finance are probabilistic in form or flavor.³ Specifically, let the *field* \mathcal{F}_t comprise all possible unions of parts $P_t \in \mathbb{P}_t$. Progressive acquisition of knowledge reflects in the string

$$\{\emptyset, \mathbb{S}\} := \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_T := 2^{\mathbb{S}},$$

called a *filtration*. Accordingly, no state-dependent price, when realized at time t , should unveil more information than already imbedded in \mathcal{F}_t . In other words: the entities $p_{jt}(s), j \in J$, must all be \mathcal{F}_t -*measurable*, meaning known or knowable at time t . Under that proviso, the price process is declared *adapted*.

Clearly, a variable defined on \mathbb{S} is \mathcal{F}_t -measurable iff constant on each $P_t \in \mathbb{P}_t$, these parts being the *atoms* of \mathcal{F}_t . Therefore, the one-one correspondence $P_t \leftrightarrow n_t$,

¹For any node $n \notin \mathcal{N}_T$ constancy of p_{bc} across $c \in \mathcal{C}(n)$ points to b as *locally riskless*.

²One may interpret p_{bn} as the face value of a zero-coupon bond that matures at node n . Thus the *spot rate* $(p_{bc} - p_{bn})/p_{bn}, c \in \mathcal{C}(n)$, of a predictable bond is perfectly known at node n . The mapping $n \mapsto \delta_n = p_{bn_0}/p_{bn}$ is often called the *term structure*. It may well be random [29].

³See for instance the excellent text [11].

described above, ensures that adapted prices $p_{jt}(s) = p_{jt}(s_0, \dots, s_t) = p_{jn}$ are defined quite simply and merely in terms of nodes.

A stochastic process $t \mapsto \theta_t$, indexed by time, is adapted - or *progressively measurable* - iff θ_t is defined on \mathcal{N}_t . Then, instead of $\theta_t(s)$ we simply write θ_n , tacitly understanding that node $n \in \mathcal{N}_t$ points to part $P_t(s) \in \mathbb{P}_t$. Henceforth we use 0 as shorthand for the root node n_0 .

3. ARBITRAGE

Denote by $\theta_{jn} \in \mathbb{R}$ the number of shares an investor holds in paper $j \in J$ upon leaving node n . Suppose he buys (outgoing) *portfolio* $\theta_n := (\theta_{jn}) \in \mathbb{R}^J$ at $n \neq 0$ and liquidates there the (incoming) portfolio $\theta_{\mathcal{A}(n)}$ bought at the ancestor node. Absent transaction costs, those operations bring him nominal, current *gain*

$$G_n(\theta) := p_n \cdot \theta_{\mathcal{A}(n)} - p_n \cdot \theta_n.$$

(The dot always denotes the standard inner product.) At node $n = 0$ let $G_0(\theta) := -p_0 \cdot \theta_0$. He might naturally ask: *Can the market be milked for money?* That simple question motivates the following

Definition: *The market allows **arbitrage** iff the system*

$$G_n(\theta) \geq 0 \text{ for all } n \text{ and } p_n \cdot \theta_n \geq 0 \text{ for each leaf,} \quad (1)$$

*admits a solution $\theta = (\theta_n)$ with at least one strict inequality. Otherwise the market is declared **arbitrage-free**.*⁴ \square

The fundamental theorem of asset pricing: *The market is arbitrage-free iff there exists a strictly positive probability measure μ on \mathcal{N}_T such that the transition probabilities, induced by μ on \mathcal{N} , satisfy the martingale condition*

$$\delta_n p_n = E_\mu [\delta_c p_c | n] = \sum_{c \in \mathcal{C}(n)} \delta_c p_c \mu(c | n) \text{ for all } n \notin \mathcal{N}_T. \quad (2)$$

In particular, whenever the bond is predictable at some $n \notin \mathcal{N}_T$, and $\delta_{(n)} := \delta_c / \delta_n$, $c \in \mathcal{C}(n)$, denotes the local discount factor there, the corresponding equation in (2) amounts to

$$p_n = \delta_{(n)} E_\mu [p_c | n] = \delta_{(n)} \sum_{c \in \mathcal{C}(n)} p_c \mu(c | n).$$

Proof. Fix any probabilities $\pi_n > 0$ across $n \in \mathcal{N}_T$, and use the induced probabilities π_n at nodes $n \notin \mathcal{N}_T$. Consider the homogeneous linear program

$$\max_{\theta} \sum_n \delta_n \pi_n G_n(\theta) + \sum_{n \in \mathcal{N}_T} \delta_n \pi_n p_n \cdot \theta_n \text{ s.t. (1).} \quad (3)$$

⁴Customary but weaker definitions of arbitrage require that θ be *self-financing* in that $G_n(\theta) = 0$ for all $n \neq 0$ (or for all n); see e.g. [18].

Clearly, the market is arbitrage-free iff the optimal value of (3) is 0. Associate multiplier $\delta_n y_n \geq 0$ to inequality $G_n(\theta) \geq 0$, and $\delta_n Y_n \geq 0$ to leaf constraint $p_n \cdot \theta_n \geq 0$. Maximizing the resulting Lagrangian

$$\begin{aligned} & \sum_n \delta_n (\pi_n + y_n) G_n(\theta) + \sum_{n \in \mathcal{N}_T} \delta_n (\pi_n + Y_n) p_n \cdot \theta_n = \\ & \sum_{n \notin \mathcal{N}_T} \left[\sum_{c \in \mathcal{C}(n)} \delta_c (\pi_c + y_c) p_c - \delta_n (\pi_n + y_n) p_n \right] \cdot \theta_n + \sum_{n \in \mathcal{N}_T} \delta_n (Y_n - y_n) p_n \cdot \theta_n \quad (4) \end{aligned}$$

with respect to the free variable θ we see that the dual of (3) amounts to solve

$$\delta_n (\pi_n + y_n) p_n = \sum_{c \in \mathcal{C}(n)} \delta_c (\pi_c + y_c) p_c \quad \text{for all } n \notin \mathcal{N}_T \text{ with } y \geq 0.$$

Suppose the latter system is indeed solvable. In that case, by LP duality, problem (3) has 0 as optimal value, and there are no arbitrage opportunities. Then consider component b of the last equation to get $\pi_n + y_n = \sum_{c \in \mathcal{C}(n)} (\pi_c + y_c)$. Therefore $\mu(c|n) := (\pi_c + y_c) / (\pi_n + y_n)$ defines strictly positive transition probabilities that satisfy (2).

Conversely, suppose some strictly positive measure μ on \mathcal{N}_T suits (2). In (4) let $\pi = \mu$ and each $y_n, Y_n = 0$ to get

$$\sum_n \delta_n \mu_n G_n(\theta) + \sum_{n \in \mathcal{N}_T} \delta_n \mu_n p_n \cdot \theta_n = \sum_{n \notin \mathcal{N}_T} \left[\sum_{c \in \mathcal{C}(n)} \delta_c \mu_c p_c - \delta_n \mu_n p_n \right] \cdot \theta_n = 0$$

for all θ . Thus arbitrage is impossible. \square

For subsequent reference let M denote the (bounded convex) set of all probability measures $\mu \sim \text{Pr}$ on \mathbb{S} that satisfy (2). Clearly, M depends only on the price process $p = (p_{jn}) \in \mathbb{R}^{J \times \mathcal{N}}$. By the fundamental theorem, $M = \emptyset$ iff there are arbitrage opportunities. Any $\mu \in M$ is called an *equivalent martingale measure*.

When b is predictable at some node $n \notin \mathcal{N}_T$, construe $\delta_{(n)} \mu(c|n)$ as the price there of an elementary Arrow-Debreu like paper that pays 1 unit of account in contingency $c \in \mathcal{C}(n)$ and nil otherwise. Viewed from that perspective, component j of (2) regulates consistent and arbitrage-free prices of paper j at any non-terminal node. In probabilistic jargon: under any $\mu \in M$ the discounted price process becomes a martingale.

4. ASSET PRICING AND SUPER-REPLICATION

Besides the given ensemble J of primary securities, consider next another asset, fully described by its "dividend process" $D = (D_n)_{n \neq 0} \in \mathbb{R}^{\mathcal{N} \setminus 0}$. The latter derives from payouts on some special paper or exogenous project. We assume that introduction of D doesn't affect price process p .

Now, if ownership to D is thinly traded, or not traded at all, one naturally asks: *how much is it worth at node 0?* More precisely, at that node, *above what value \bar{v} should D be sold?*

For an answer, this section recalls some known results, following [18], [19], [23] and [24]. The answer - or the value estimate \bar{v} - typically comes by *super-replicating* D via iterated trades in existing papers $j \in J$. In that business, called *dynamic portfolio choice*, optimization has the great advantage of aiming expressly at net gains or efficient choices.

Specifically, the investor might ask: *what up-front expense $p_0 \cdot \theta_0$ and subsequent portfolio choices θ_n would suffice to gain $G_n(\theta) \geq D_n$ at each node $n \neq 0$ - and avoid terminal debt as well?* That question amounts to state linear problem

$$\bar{\mathbb{P}}[D] : \begin{array}{l} \text{minimize} \quad p_0 \cdot \theta_0 \\ \text{subject to} \quad p_n \cdot \theta_{\mathcal{A}(n)} - p_n \cdot \theta_n \geq D_n \text{ for } n \neq 0 \quad (*) \\ \text{with} \quad p_n \cdot \theta_n \geq 0 \text{ for each leaf } n \in \mathcal{N}_T. \end{array}$$

Any feasible portfolio process $\theta = (\theta_n)$ for problem $\bar{\mathbb{P}}[D]$ is said to *super-replicate* or *super-hedge* D . Such a process represents an information-adapted investment strategy that stays on the upper safe side, eliminating all downside risk. In optimum, if any, one may posit $\theta_n = 0$ at each leaf.

Note that problem $\bar{\mathbb{P}}[D]$ is free of preference and probability. It involves no utility function, no risk aversion parameters, and no ad hoc probability measure.⁵ Let $\bar{v} = \bar{v}(D) = \inf \bar{\mathbb{P}}[D]$ denote the optimal value. Suppose D is demanded for price $\mathcal{V} > \bar{v}$ at node 0. A seller or supplier can there pocket positive profit $\mathcal{V} - \bar{v}$ right away, pay $\bar{v} = p_0 \cdot \theta_0$ for an optimal super-hedge θ of D , and still look forward to net payment process $G(\theta) - D \geq 0$ with no terminal debt.

The instance $D = 0$ is particularly interesting because $\bar{\mathbb{P}}[0]$, being homogeneous, has optimal value $\bar{v}(0) \in \{0, -\infty\}$. Clearly, $\bar{v}(0) = -\infty$ iff the system

$$p_0 \cdot \theta_0 < 0, \quad G_n(\theta) \geq 0 \text{ for } n \neq 0, \quad \text{and } p_n \cdot \theta_n \geq 0 \text{ at each leaf,}$$

is solvable. Absent arbitrage, no solution exists but, by the fundamental theorem, there is least one equivalent martingale measure. This observation motivates what is called

Martingale Pricing: Assign a Lagrange multiplier $\delta_n \mu_n \geq 0$ to restriction (*). Similarly, couple a multiplier $\delta_n \bar{\mu}_n \geq 0$ to the constraint $p_n \cdot \theta_n \geq 0$, $n \in \mathcal{N}_T$. Thereby emerges a real-valued Lagrangian

$$\mathcal{L}(\theta, \mu, \bar{\mu}) := p_0 \cdot \theta_0 + \sum_{n \neq 0} \delta_n \mu_n [D_n + p_n \cdot \theta_n - p_n \cdot \theta_{\mathcal{A}(n)}] - \sum_{n \in \mathcal{N}_T} \delta_n \bar{\mu}_n p_n \cdot \theta_n \quad (5)$$

⁵For a relation to utility maximization see [4] and references therein.

$$= \begin{cases} (\sum_{n \neq 0} \delta_n \mu_n D_n) + [p_0 - \sum_{c \in \mathcal{C}(0)} \delta_c \mu_c p_c] \cdot \theta_0 + \\ \sum_{n \notin 0 \cup \mathcal{N}_T} [\delta_n \mu_n p_n - \sum_{c \in \mathcal{C}(n)} \delta_c \mu_c p_c] \cdot \theta_n + \\ \sum_{n \in \mathcal{N}_T} \delta_n (\mu_n - \bar{\mu}_n) p_n \cdot \theta_n. \end{cases}$$

As customary, for given $\mu \in \mathbb{R}_+^{\wedge 0}$ and $\bar{\mu} \in \mathbb{R}_+^{\mathcal{N}_T}$, minimize $\mathcal{L}(\theta, \mu, \bar{\mu})$ with respect to $\theta = (\theta_n)$. That operation, when finite-valued, associates to $\bar{\mathbb{P}}[D]$ a corresponding *dual problem*, namely:

$$\begin{aligned} & \text{maximize} && \sum_{n \neq 0} \delta_n \mu_n D_n \\ & \text{subject to} && \delta_n \mu_n p_n = \sum_{c \in \mathcal{C}(n)} \delta_c \mu_c p_c \quad \forall n \notin \mathcal{N}_T, \quad \mu \geq 0, \quad \text{and} \quad \mu_0 = 1. \end{aligned}$$

(The vector $\bar{\mu} \in \mathbb{R}_+^{\mathcal{N}_T}$ disappeared here because $\mu_n = \bar{\mu}_n$ for $n \in \mathcal{N}_T$.) Considering component b of the last constraint we get $\mu_n = \sum_{c \in \mathcal{C}(n)} \mu_c$ at each $n \notin \mathcal{N}_T$. Since $\mu_0 = 1$ and $\mu \geq 0$, it follows that μ defines a probability measure over $\mathcal{N}_T \leftrightarrow \mathbb{S}$, and the preceding problem reads

$$\bar{\mathbb{D}}[D] : \boxed{\begin{array}{l} \text{maximize} \quad E_\mu \sum_{n \neq 0} \delta_n D_n \text{ over probability all measures } \mu \\ \text{such that} \quad \delta_n \mu_n p_n = \sum_{c \in \mathcal{C}(n)} \delta_c \mu_c p_c \text{ at each non-terminal node } n. \quad (**) \end{array}}$$

One may regard optimal solutions to $\bar{\mathbb{D}}[D]$ as risk-neutral, equivalent probability distributions that depict the expected present value of process D as favorably as possible. When an optimal μ_n is unique, it supports a standard interpretation. To wit, if the investor withdraws wealth $w_n \approx 0$ at node $n \neq 0$ subject to the reasonable restriction

$$p_n \cdot \theta_{\mathcal{A}(n)} - p_n \cdot \theta_n - w_n \geq D_n,$$

the up-front investment cost increases by $\delta_n \mu_n w_n$. Thus the contingent "shadow price" $\mu_n \geq 0$ reports the value of money made available only at node n .

Clearly, when all $\mu_n > 0$, condition (***) coincides with (2). In case the market is arbitrage-free, and $M \subseteq \hat{M} \subseteq clM$, problem $\bar{\mathbb{D}}[D]$ can be restated as $\sup_{\mu \in \hat{M}} E_\mu \sum_{n \neq 0} \delta_n D_n$.

Proposition (On the ask value \bar{v}). *Suppose the market is arbitrage-free. Then the optimal value $\bar{v} = \inf \bar{\mathbb{P}}[D]$ is finite and attained, and*

$$\bar{v} = \bar{v}(D) := \sup_{\mu \in M} E_\mu \sum_{n \neq 0} \delta_n D_n = \sup_{\mu \in clM} E_\mu \sum_{n \neq 0} \delta_n D_n.$$

The function $D \mapsto \bar{v}(D)$ is convex, piecewise linear and positively homogenous.

Proof. By linear programming duality, $\bar{v}(D) = \sup \bar{\mathbb{D}}[D]$. Positive homogeneity derives directly. The closure clM of M is a non-empty compact polyhedron, having a finite set $ext(clM)$ of extreme points. Therefore $\bar{v}(D)$, being the maximum of linear

functions $\mu \mapsto E_\mu \sum_{n \neq 0} \delta_n D_n$, $\mu \in \text{ext}(clM)$, comes out convex and piecewise linear. \square

Problem $\bar{\mathbb{P}}[D]$ didn't embody the constraint $G_0(\theta) \geq 0$ in (1). Also, its objective differs from (3). Therefore \bar{v} may be finite even when there are arbitrage opportunities.

Absent arbitrage, only non-degenerate measures $\mu \sim \text{Pr}$ need be considered. To argue differently for non-degeneracy, suppose some market agent maximizes a separable "utility" criterion $\sum_{n \in \mathcal{N}} \delta_n u_n(d_n)$ in decision variables $d, \theta \in \mathbb{R}^{\mathcal{N}}$ subject to $-p_0 \cdot \theta_0 \geq d_0$ and the constraints of $\bar{\mathbb{P}}[d]$. Then, in terms of *conjugates* $u_n^{(*)}(\mu_n) := \sup_{d_n} \{u_n(d_n) - \mu_n d_n\}$, the dual problem reads:

$$\text{minimize } \sum_{n \in \mathcal{N}} \delta_n u_n^{(*)}(\mu_n) \text{ s.t. } \mu \geq 0 \text{ and } \delta_n \mu_n p_n = \sum_{c \in \mathcal{C}(n)} \delta_c \mu_c p_c \text{ for all } n \notin \mathcal{N}_T.$$

At each leaf n naturally assume $u_n^{(*)}(\mu_n) = +\infty$ whenever $\mu_n \leq 0$. Consequently, the very last problem is infeasible unless $\mu_n > 0$ at every $n \in \mathcal{N}_T$. This suffices to have all μ_n positive.

5. BID PRICING AND COMPLETENESS

Problem $\bar{\mathbb{P}}[D]$ approximates D within the market, but *from above and as cheaply as possible*. Alternatively, D might be approximated *from below and as expensively as possible*. Thus, a prudent buyer of D , who prefers conservative estimates of present values, could choose to formulate the following problem:

$$\underline{\mathbb{P}}[D] : \begin{array}{l} \text{maximize } p_0 \cdot \theta_0 \\ \text{subject to } D_n \geq p_n \cdot \theta_{\mathcal{A}(n)} - p_n \cdot \theta_n \text{ for } n \neq 0, \\ \text{with } p_n \cdot \theta_n \leq 0 \text{ when } n \in \mathcal{N}_T. \end{array}$$

In optimum, if any, one may posit $\theta_n = 0$ at each leaf. The optimal value $\underline{v} = \underline{v}(D)$ of $\underline{\mathbb{P}}[D]$ is the maximal amount of current cash an investor can extract from the market at node 0 when allowed terminal debt but no gain $G_n(\theta) > D_n$ at any node $n \neq 0$. If process D is offered for value $\mathcal{V} < \underline{v}$ at node 0, a buyer could choose any minimizing θ , cash in $\underline{v} - \mathcal{V} > 0$ at once, and thereafter still enjoy the subsequent payment process $D - G(\theta) \geq 0$.

The Lagrangian of $\underline{\mathbb{P}}[D]$ assumes the same form (5). So, arguing exactly as above, the associated dual now reads:

$$\underline{\mathbb{D}}[D] : \text{minimize } E_\mu \sum_{n \neq 0} \delta_n D_n \text{ over probability all measures } \mu \text{ s.t. } (**).$$

It follows likewise a

Proposition (On the bid value \underline{v}). *Suppose the market is arbitrage-free. Then the optimal value $\underline{v} := \sup \mathbb{P}[D]$ is finite and attained, and*

$$\underline{v} = \underline{v}(D) := \inf_{\mu \in M} E_{\mu} \sum_{n \neq 0} \delta_n D_n = \inf_{\mu \in cl M} E_{\mu} \sum_{n \neq 0} \delta_n D_n.$$

The function $D \mapsto \underline{v}(D)$ is concave, piecewise linear and positively homogenous. Moreover, $\underline{v}(D) = -\bar{v}(-D)$. \square

Proof. Only the very last assertion requires justification. For this, simply note that θ is feasible for $\mathbb{P}[D]$ iff $-\theta$ is feasible for $\bar{\mathbb{P}}[-D]$. \square

Again, because the constraint set and objective of $\mathbb{P}[D]$ differ from counterparts (1),(3), it may happen that \underline{v} is finite even under arbitrage opportunities.

Anyway, the upshot is that two valuation schemes $\underline{v}(\cdot)$, $\bar{v}(\cdot)$ operate on dividend processes. When these schemes differ, both are *nonlinear*. It turns out that *linearity*, when in vigor, relates to what is called

A complete market. Any price outside the *bid-ask interval* $[\underline{v}, \bar{v}]$ creates an arbitrage opportunity. Indeed, D is undervalued at a price $< \underline{v}$ and overvalued at a price $> \bar{v}$. So, when is the appropriate value unique? That is, *when is the interval $[\underline{v}, \bar{v}]$ degenerate?*

The answer comes in terms of the *marketable space* \mathbf{D} , consisting of all dividend processes $\mathcal{D} = (\mathcal{D}_n)_{n \neq 0} \in \mathbb{R}^{\mathcal{N} \setminus 0}$ that satisfy

$$\mathcal{D}_n = p_n \cdot \theta_{\mathcal{A}(n)} - p_n \cdot \theta_n \tag{6}$$

for some portfolio process $\theta = (\theta_n)$ with $\theta_n = 0$ at each leaf. One says that process $\mathcal{D} \in \mathbf{D}$ is *attainable* - or *replicable*, or made *redundant* - by dynamic portfolio choice. Accordingly, the market is declared *complete* iff $\mathbf{D} = \mathbb{R}^{\mathcal{N} \setminus 0}$.

Theorem (On unique values and a complete market). *Suppose the market is arbitrage-free. Then,*

- *a process $D \in \mathbb{R}^{\mathcal{N} \setminus 0}$ has a unique value $v(D) = \underline{v}(D) = \bar{v}(D)$ iff $D \in \mathbf{D}$;*
- *the asset market is complete iff there is only **one** equivalent martingale measure.*

Proof. Suppose $\theta = (\theta_n)$ replicates $D \in \mathbf{D}$ in that (6) holds for all $n \neq 0$ with $\theta_n = 0$ on \mathcal{N}_T . Because θ is feasible for both $\bar{\mathbb{P}}(D)$ and $\underline{\mathbb{P}}(D)$, it follows that

$$\bar{v}(D) \leq p_0 \cdot \theta_0 \leq \underline{v}(D).$$

Also, since M is non-empty, $\bar{v}(D) \geq \underline{v}(D)$. Thus, $\underline{v}(D) = \bar{v}(D) = p_0 \cdot \theta_0$, and this takes care of the necessity in first bullet. For the second bullet, note that each martingale measure μ generates a node-based function $n \mapsto \mu_n$, whence a linear mapping

$$\langle \mu, \mathcal{D} \rangle := E_{\mu} \sum_{n \neq 0} \delta_n \mathcal{D}_n = \sum_{n \neq 0} \mu_n \delta_n \mathcal{D}_n$$

on dividend processes $\mathcal{D} = (\mathcal{D}_n)_{n \neq 0} \in \mathbb{R}^{\mathcal{N} \setminus 0}$. The functional $\langle \mu, \cdot \rangle$ is thus defined on all $\mathbb{R}^{\mathcal{N} \setminus 0}$ - and naturally called the μ -valuation.

If more than one martingale measure exists, then some asset has different martingale-valuations. That asset can't belong to \mathbf{D} , and then the market must be incomplete.

If indeed the market is incomplete, consider any $\mathcal{D} \notin \mathbf{D}$. Let $\bar{\mathcal{D}}$ be its closest approximation (alias projection) in \mathbf{D} , using the inner product $\langle \mathcal{D}^*, \mathcal{D} \rangle := \sum_{n \neq 0} \mathcal{D}_n^* \delta_n \mathcal{D}_n$ already mentioned above. Then $\mathcal{D}^* := \mathcal{D} - \bar{\mathcal{D}}$ is orthogonal on \mathbf{D} , and $\hat{\mu} := \mu + \varepsilon \mathcal{D}^* > 0$ for small enough $\varepsilon > 0$. The particular dividend process $\hat{\mathcal{D}}$ that first and last pays $1/\delta_n$ at leaf n , is marketable; that is, $\hat{\mathcal{D}} \in \mathbf{D}$. Therefore,

$$1 = \sum_{n \in \mathcal{N}_T} \mu_n = \sum_{n \in \mathcal{N}_T} \mu_n \delta_n \hat{\mathcal{D}}_n = \langle \mu, \hat{\mathcal{D}} \rangle = \langle \mu + \varepsilon \mathcal{D}^*, \hat{\mathcal{D}} \rangle = \langle \hat{\mu}, \hat{\mathcal{D}} \rangle = \sum_{n \in \mathcal{N}_T} \hat{\mu}_n.$$

This string of equalities tells that $\hat{\mu}$ also constitutes a non-degenerate probability distribution across the leafs. The orthogonality of \mathcal{D}^* on \mathbf{D} entails that $\hat{\mu}$ generates a martingale measure with $\langle \mu, \cdot \rangle = \langle \hat{\mu}, \cdot \rangle$ on \mathbf{D} . However, because $\langle \mathcal{D}^*, \mathcal{D} \rangle > 0$, we get $E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n < E_{\hat{\mu}} \sum_{n \neq 0} \delta_n \mathcal{D}_n$. \square

Some comments on *extensions* and *reductions* briefly conclude this section. For *extensions*, θ_n could comprise long and short positions $\theta_n^+, \theta_n^- \in \mathbb{R}_+^J$ traded at corresponding prices p_n^+, p_n^- . Also there might be node-dependent transaction costs or taxes $\mathcal{T}_n(\theta_{\mathcal{A}(n)}, \theta_n)$. And clearly, there might be constraints on various branches and nodes.

For *reductions*, consider merely the *sub-tree* that emanates from a non-terminal node $n \neq 0$. Pricing the corresponding part of D only within that sub-tree, as described above, gives values $\underline{v}_n \leq \bar{v}_n$, node n now figuring as root. Finally, for up-front asset pricing, the sub-tree may thereafter be erased, leaving only node n as a new leaf with dividend $D_n + [\underline{v}_n, \bar{v}_n]$.

When $\underline{v}_n = \bar{v}_n$, numerous textbooks illustrate this procedure, in one form or another, under various headings called backward recursion, dynamic programming or portfolio replication. If $\underline{v}_n < \bar{v}_n$, the said sub-tree features incompleteness.

6. PRICING PARTLY MANUFACTURED ASSETS

Many a dividend process is exogenous, meaning totally unaffected by the investor. Often though, he has great impact on its evolution. In particular, such is the case when pay-outs become nil beyond some deliberately chosen stopping time.⁶ More generally, there might be opportunities to revise positions prior to their time of expiration. To model instances of that sort suppose the dividend process is *partly manufactured* - or largely influenced - by the investor in that adapted dividends have the form

$$D_n = \mathcal{D}_n(x_{n_0}, x_{n_1}, \dots, x_n). \quad (7)$$

⁶As an example consider purchase of a physical asset, before or at the expiration of a lease.

Here $x = (x_n)$ is an underlying process, affected by uncertainty, but controlled by the investor who responds to changing opportunities. The said x must reside in a prescribed, non-empty set $X \subseteq \Pi_{n \in \mathcal{N}} X_n$. In (7) n_0, n_1, \dots, n is the unique path of adjacent nodes that leads from the root up to n . To simplify notations write \vec{x}_n for the corresponding part $(x_{n_0}, x_{n_1}, \dots, x_n)$ of x . Denote by $\mathcal{D}(x)$ the resulting dividend process, and assume that $\sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n)$ is finite for each $x \in X$.

The preceding results apply immediately. To wit, for fixed $x \in X$, the *inner problem* $\bar{V}(x) := \inf \bar{\mathbb{P}}[D]$ amounts to find an optimal super-hedge for $D = \mathcal{D}(x)$. The *outer problem* $\bar{v} := \sup_{x \in X} \bar{V}(x)$ first chooses x , to be followed by an optimal super-hedge. Together these decisions generate a max-min saddle problem: $\sup_{x \in X} \inf_{\theta} \bar{\mathbb{P}}[\mathcal{D}(x)]$.

The interpretation is pretty much as before but depends on whether the seller owns \mathcal{D} or not. I consider only the upper problem $\bar{\mathbb{P}}[\mathcal{D}]$. Suppose a buyer offers value $\mathcal{V} > \bar{v}$ for \mathcal{D} and plans to implement $x \in \arg \max \bar{V}$. If the owner of \mathcal{D} accepts the offer, and opts for an optimal solution θ of $\bar{\mathbb{P}}[\mathcal{D}(x)]$, he immediately gets $\mathcal{V} - \bar{v} > 0$, thereafter payoff profile $G(\theta) - \mathcal{D}(x) \geq 0$, and finally no debt. These arrangements leave him no liabilities.

By contrast, consider somebody who doesn't own \mathcal{D} but decides to sell it short. Admittedly, this situation is somewhat more intricate. Recall that any $x \in X$ is a *strategy*, implemented as a contingent plan. To hedge his position, suppose the seller immediately requires, by contractual agreement, that the buyer commits to some strategy $x \in X$, decided by the latter. Next, upon choosing $\theta \in \arg \min \bar{\mathbb{P}}[\mathcal{D}(x)]$ the seller gets $\mathcal{V} - \bar{V}(x) \geq \mathcal{V} - \bar{v} > 0$ up front. Thereafter his gain process $G(\theta)$ covers his liabilities $\mathcal{D}(x)$, and finally he exits without debt.

Selling \mathcal{D} short thus resembles a *Stackelberg game* [21]: The *leader* (alias buyer) first commits a strategy $x \in X$; thereafter the *follower* (alias seller) responds with θ . Plainly, such a setting isn't always satisfactory or convincing. Why should the leader reveal his strategy ex ante, play open loop, and forego all sorts of discretion or opportunism?

Duality again proves useful by pointing to the equivalent optimization problem: $\sup_{x \in X} \sup_{\mu \in M} \bar{\mathbb{D}}[\mathcal{D}(x)]$ in which both variables x, μ are oriented towards supremum. Moreover, these interact only in the objective. The upshot is that saddle problems are avoidable, as described next:

Proposition (On max-min price estimates). *Suppose the market is arbitrage-free.*

- Then the optimal value $\bar{v} := \bar{v}(\mathcal{D}) := \sup_{x \in X} \inf \bar{\mathbb{P}}[\mathcal{D}(x)]$ of the upper price problem equals

$$\sup_{x \in X} \sup_{\mu \in M} \left\{ E_{\mu} \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\}.$$

- Similarly, the optimal value $\underline{v} := \underline{v}(\mathcal{D}) := \sup_{x \in X} \sup \underline{\mathbb{P}}[\mathcal{D}(x)]$ of the lower price

problem equals

$$\sup_{x \in X} \inf_{\mu \in M} \left\{ E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\}.$$

• Provided X be closed convex, and $x \mapsto \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n)$ upper semicontinuous concave, there exists $\bar{\mu} \in clM$ such that

$$\underline{v} = \inf_{\mu \in M} \sup_{x \in X} \left\{ E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\} = \sup_{x \in X} \left\{ E_{\bar{\mu}} \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\}.$$

Proof. Only the last bullet needs justification. Since the objective

$$(\mu, x) \in (clM) \times X \mapsto E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n)$$

is convex in μ and concave in x , the result follows from the lopsided minimax theorem in Chap. 6 of [2]. \square

Clearly, the operation $\sup_{\mu \in M} E_\mu$ amounts to $\max_{\mu \in clM} E_\mu$, and similarly for infimum. Anyway, the preceding proposition underscores the convenience of having a unique martingale measure μ whence a complete market. Problems

$$\sup_{x \in X} \sup_{\mu \in M} E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \quad \text{and} \quad \sup_{\mu \in M} \sup_{x \in X} E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n)$$

amount, in essence, to one and the same. Also, for any $\mu \in M$, it follows from dynamic programming, starting at leafs and proceeding recursively towards deeper nodes, that

$$\sup_{x \in X} E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) = E_\mu \sup_{x \in X} \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n).$$

When solving for \underline{v} or \bar{v} one naturally inquires what curvature the intermediate, reduced objectives might have. To that end, recall that the pointwise supremum of convex functions remains convex. Also, the maximum, if any, of a convex function occurs at an *extreme point* of its domain [20]. (Quite similar properties hold for concave functions, infimum then replacing the role of supremum.) By an extreme point of a set, contained in a real vector space, is understood an element that can't equal a proper convex combination of other set members.

In the special case when $X = \prod_{n \in \mathcal{N}} X_n$, with each X_n convex, extremality of $x = (x_n)$ means that x_n must be an extreme point of $X_n \forall n$. Reverting to scenario $s \in P_t \leftrightarrow n \in \mathcal{N}_t$ we see that $x_n = x_t(s)$ is an extreme point in the \mathcal{F}_t -measurable set $X_n = X_t(s)$.

Proposition (On curvature of objectives, and extreme solutions).

- If $x \mapsto \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n)$ is convex continuous, then so is the reduced upper objective

$$\sup_{\mu \in M} \left\{ E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\}.$$

In that case, with X compact convex, an optimal solution to the upper pricing problem $\sup_{x \in X} \inf \bar{\mathbb{P}}[\mathcal{D}(x)]$ is realized at an extreme point $x \in X$.

- Similarly, if $x \mapsto \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n)$ is concave, then so is the reduced lower objective $\inf_{\mu \in M} \left\{ E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\}$.
- The objective $\mu \mapsto \sup_{x \in X} \left\{ E_\mu \sum_{n \neq 0} \delta_n \mathcal{D}_n(\vec{x}_n) \right\}$ is convex. Consequently, the upper value \bar{v} , if realized, is attained at an extreme point $\mu \in \text{cl}M$. \square

7. PRICING OPTIONS THAT FEATURE DISCRETE DECISIONS

A surprisingly wide class of financial decision problems fit the format of continuous optimization - as exemplified here above. An even larger class comes on stage when some variables are integers; see [22], [30].

Such variables could stem from set-up costs or indivisibilities. Other, more challenging instances correspond discrete interventions at judiciously chosen times [7]. Examples include contingent closure/opening of fishing grounds/seasons - or of oil/mineral deposits. Also covered are rotation problems in forestry. Common to these is presence of at least one *stopping time*, meaning a mapping $\tau : \mathbb{S} \rightarrow \{0, \dots, T\} \cup \{+\infty\}$ such that $\{s : \tau(s) \leq t\} \in \mathcal{F}_t$ for every $t \in \{0, \dots, T\}$.

To accommodate such objects, recall the one-one correspondence between nodes and partitions: $n \leftrightarrow P_t(s)$ - and the resulting identification $x_t(s) \leftrightarrow x_n$. So, to highlight the time aspect, write now $x = (x_t)$ for the decision process, indexed by time, with the tacit understanding that x_t be \mathcal{F}_t -measurable, $t \in \{0, \dots, T\}$. In short, only *adapted* processes are allowed. Further, to elaborate on discrete choices, assume henceforth that each component x_t be purely integral.⁷

Examples of stopping times in finance: Numerous financial derivatives yield dividend of the form $D_t = \mathcal{D}_t(S_0, \dots, S_t, x_0, \dots, x_t)$ at time t . Underlying is then the price process (S_t) on a specified stock. For instance, a *call option*

$$C_t := \max_{\tau \in T(t)} \{S_\tau - K_\tau, 0\} = \max_{\tau \in T(t), x_\tau \in \{0,1\}} \{x_\tau(S_\tau - K_\tau), 0\}$$

or a *put option*

$$P_t := \max_{\tau \in T(t)} \{K_\tau - S_\tau, 0\} = \max_{\tau \in T(t), x_\tau \in \{0,1\}} \{x_\tau(K_\tau - S_\tau), 0\}$$

⁷One may envisage that continuous components, if any in x , have already been optimized away, leaving a reduced objective. The remaining variables could correspond to stopping times.

with strike price K_τ , exercised or not at time $\tau \in \mathbb{T} \subseteq \{0, \dots, T\}$, is named

American if $T(t) = \{t\}$ and $\mathbb{T} = \{0, \dots, T\}$,

Bermudan if $T(t) = \{t\}$ and $\mathbb{T} \subset \{0, \dots, T\}$,

European if $T(t) = \{t\}$ and $\mathbb{T} = \{T\}$,

Russian if $T(t) = \{0, \dots, t\}$ and $\mathbb{T} = \{0, \dots, T\}$.

Each option is exercised at most *once*. Correspondingly, in case of a unique martingale measure μ , one should

$$\text{maximize } E_\mu \sum_{t \geq 1} \delta_t \mathcal{D}_t x_t$$

subject to (x_t) adapted, all $x_t \in \{0, 1\}$,

$$\sum_{t \notin \mathbb{T}} x_t = 0, \quad \text{and} \quad \sum_{t \in \mathbb{T}} x_t \leq 1, \quad (8)$$

with either $\mathcal{D}_t = C_t$ or $\mathcal{D}_t = P_t$ for all t as the case may be. Some comparative statics deserve mention here, briefly limited to call options:

Proposition (On comparative statics of call options).

- *Any first- or second-order increased uncertainty in some stock-price S_t , causes the call option value to increase.*
- *Everything else equal, if δ_t or S_t increases, or if K_t decreases, for some t , then the optimal response x_t cannot decrease for a call option.*

Proof. Because C_t is non-decreasing and convex in S_t the first bullet follows immediately from stochastic dominance [31]. The second bullet derives from supermodularity [28] and

$$\frac{\partial^2}{\partial S_t \partial x_t} C_t \geq 0 \quad \text{and} \quad \frac{\partial^2}{\partial (-K_t) \partial x_t} C_t \geq 0. \quad \square$$

It's natural to probe beyond linear restrictions like (8). This motivates a look at

Linearly constrained instances: Hereafter, for simplicity, consider the convex polyhedron

$$C := \{x = (x_t) : \underline{b} \leq Ax \leq \bar{b}, \underline{x} \leq x \leq \bar{x}\} \quad (9)$$

with bounds $\underline{b} \leq \bar{b}$ and $\underline{x} \leq \bar{x}$. The matrix A is random, lower triangular stochastic, and it comes in the block form:

$$A = \begin{bmatrix} A_{00} & 0 & \cdots & \cdots \\ A_{10} & A_{11} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ A_{T0} & A_{T1} & \cdots & A_{TT} \end{bmatrix}$$

The block entry $s \mapsto A_{t\tau}(s)$ is constant on each part $P_t \in \mathbb{P}_t$ for $t \geq \tau$, and it has exactly as many columns as x_τ has components. The bounding vectors

$$\underline{b} = \begin{bmatrix} \underline{b}_0 \\ \underline{b}_1 \\ \vdots \\ \underline{b}_T \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} \bar{b}_0 \\ \bar{b}_1 \\ \vdots \\ \bar{b}_T \end{bmatrix}, \quad \underline{x} = \begin{bmatrix} \underline{x}_0 \\ \underline{x}_1 \\ \vdots \\ \underline{x}_T \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} \bar{x}_0 \\ \bar{x}_1 \\ \vdots \\ \bar{x}_T \end{bmatrix}$$

are of compatible measurability and size. For example, the part \underline{b}_t is observable at time t , and it has exactly as many rows as $A_{t\tau}$ for $t \geq \tau$. When all these qualifications are in vigor, data $(A, \underline{b}, \bar{b}, \underline{x}, \bar{x})$ is certified as *adapted*. Henceforth suppose $x \in X \Leftrightarrow x$ is adapted and belongs to polyhedron C .

Unimodularity: A matrix A is declared *totally unimodular* if for any integer bounding vectors $\underline{b}, \bar{b}, \underline{x}, \bar{x}$, of suitable dimensions, all extreme points of the convex polyhedron C , defined in (9), become integral. This happens iff every square sub-matrix of A has determinant $-1, 0$ or $+1$. Then, necessarily, every entry of A must be either $-1, 0$ or $+1$; see [16]. Sufficient for total unimodularity of a matrix A , having but entries $-1, 0$ or $+1$, is that

- no more than two nonzero entries appear in each column;
- the rows can be partitioned into *two* subsets such that if a column contains two nonzero elements of the same (opposite) sign, then the corresponding row indices belong to the opposite (respectively, same) subset [13].

Proposition (On integer extreme solutions). *Suppose $(A, \underline{b}, \bar{b}, \underline{x}, \bar{x})$ is adapted and integral. Also suppose that every diagonal block A_{tt} is totally unimodular. Then every adapted, extreme solution x is integer-valued in each component.*

Proof. x_0 must satisfy

$$\underline{b}_0 \leq A_{00}x_0 \leq \bar{b}_0 \quad \text{and} \quad \underline{x}_0 \leq x_0 \leq \bar{x}_0.$$

It follows that every extreme x_0 must be integer. Next, an extreme x_1 must satisfy

$$\underline{b}_1 - A_{10}x_0 \leq A_{11}x_1 \leq \bar{b}_1 - A_{10}x_0 \quad \text{and} \quad \underline{x}_1 \leq x_1 \leq \bar{x}_1$$

hence be integral as well. Continue in this manner to conclude. \square

Proposition (On integer solutions). *Suppose $(A, \underline{b}, \bar{b}, \underline{x}, \bar{x})$ adapted and integral. Further suppose that every diagonal block A_{tt} is totally unimodular. If*

$$x \mapsto \sum_{t \geq 1} \delta_t \mathcal{D}_t(x_0, \dots, x_t)$$

is convex, then so is the reduced upper objective

$$\bar{V}(x) := \sup_{\mu \in M} \left\{ E_{\mu} \sum_{t \geq 1} \delta_t \mathcal{D}_t(x_0, \dots, x_t) \right\}.$$

In that case an optimal solution, if any, to the upper pricing problem $\sup_{x \in X} \inf \bar{\mathbb{P}}[\mathcal{D}(x)]$ is realized at an **integral** extreme point $x \in X$. \square

Proof. $\bar{V}(\cdot)$, being the supremum of convex functions, must itself be convex. Therefore, as argued above, if $\arg \max \bar{V}$ is non-empty, it must intersect the set $extX$ of extreme points in X . By the preceding proposition each $x \in extX$ is integral. \square

Quite often dividends depend on the path. Examples include American, Bermudan and barrier options [1], [19], [23], [25], [27]. Common to these are restrictions on when and how often the investor may exercise contracted rights. Considered in conclusion is

An example with American-like options: Suppose exercise happens at most once. For computational purposes construct a *flow problem* in the following *capacitated network* [26]: Let $ex\mathcal{N} \subseteq \mathcal{N} \setminus 0$ denote the set nodes at which the option can be exercised. For each node $n \in ex\mathcal{N}$ create a duplicate node n' and the directed link (n, n') . From each said duplicate n' - and from each original leaf $n \in \mathcal{N}_T$ - introduce a link towards a common, auxiliary node, called *sink* to reflect its absorbing role. Denote by \mathbb{E} the resulting set of edges; see figure below.

In the oriented network so constructed, let each edge that leads into the sink have *capacity interval* $[0, 1]$. This means that the amount which flows through that edge is bounded below by 0 and above by 1. Other edges impose no upper restrictions; they all have $[0, +\infty)$ as capacity interval.

The root is the only *source* of flow. Endow that special node with integer *supply* $|\mathbb{S}| = |\mathcal{N}_T|$. That same amount is *demanded* at the unique sink. Any other node serves merely for *transshipment*: what flows in there equals precisely what flows out. Let $y = (y_e)$ denote any feasible flow pattern along the directed edges $e \in \mathbb{E}$ in the extended network just laid out.

So far, this was all physical design. Economic data enter next. At the duplicate n' of original node $n \in ex\mathcal{N}$ let $D_{n'}$ equal the dividend if the option is exercised at n . At each leaf $n \in \mathcal{N}_T$ let D_n equal the value of *not* exercising there. The reduced objective

$$y \mapsto \sup_{\mu \in M} E_{\mu} \left\{ \sum_{n \in ex\mathcal{N}} \delta_n D_{n'} y_{(n', sink)} + \sum_{n \in \mathcal{N}_T} \delta_n D_n y_{(n, sink)} \right\}$$

is convex. Thus the upper value \bar{v} , if realized, is attained at an extreme point of the polytope comprised of all feasible flows. Such points have integral coordinates [20].

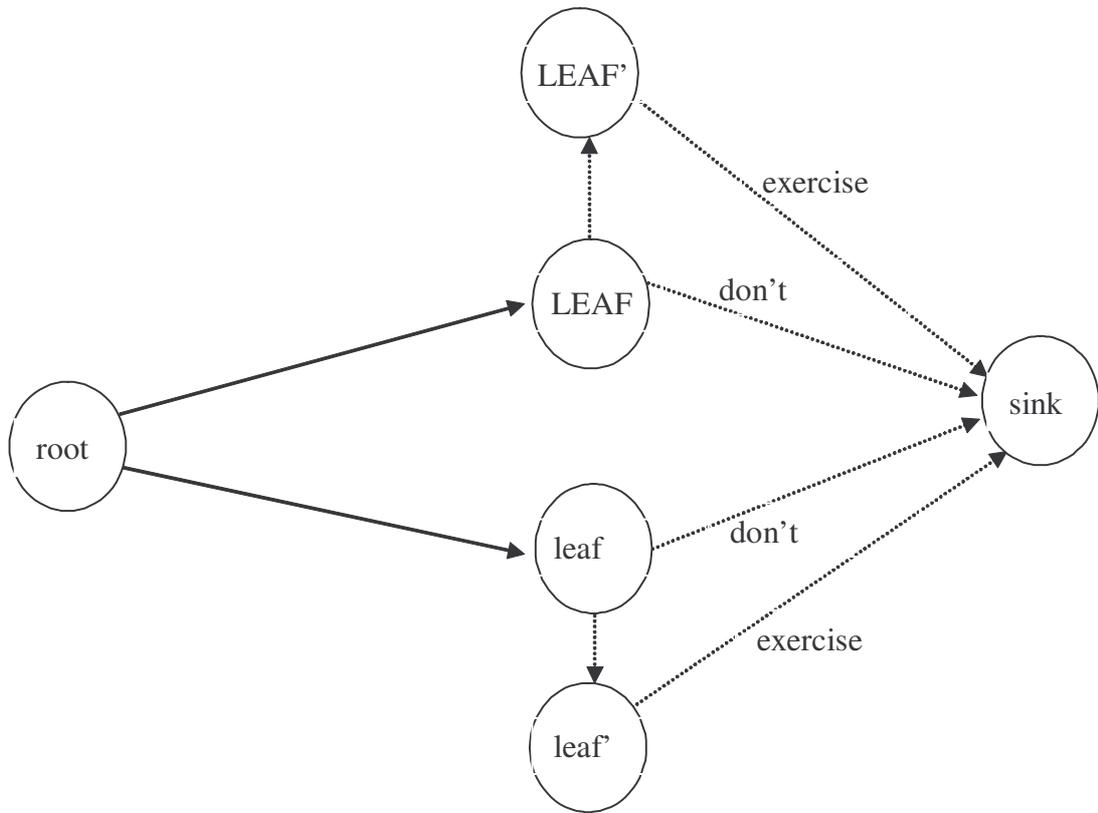


Figure: A capacitated network for a simple American option.

8. CONCLUDING REMARKS

To have something for free - that is, to make *arbitrage* - motivates much of finance. Related activities include asset evaluation and portfolio choice. For such purposes optimization helps a lot and is often indispensable. Yet surprisingly few finance books make much out of optimization technology. And conversely, financial problems are rare or absent in most texts devoted to optimization.

In fact, mathematical finance has very much become a field for specialists in stochastic processes, optimal control, partial differential equations, or numerical analysis. Mathematical programming, although a close relative, stays somewhat at distance. In my opinion that discipline may gain and offer much by connecting closer to finance. Reflecting on this, the present paper has advocated that popular algorithms be brought to bear on some chief financial issues. On a didactical note, it emphasized the great convenience and generality of scenario trees. On a more substantial note, it observed that integral constraints, stemming from indivisibilities or exercise times, can often be relaxed.

REFERENCES

- [1] L. Andersen and M. Broadie, Primal-dual simulation algorithm for pricing multidimensional American options, *Management Science* 50, 9, 1222-1234 (2004).
- [2] J.-P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, J. Wiley & Sons, New York (1984).
- [3] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer, Berlin (1997).
- [4] L. Carassus and M. Rásonyi, Convergence of utility indifference prices to superreplication price, *Math. Meth. Oper. Research* 64, 145-154 (2006).
- [5] D. Denneberg, *Non-Additive Measure and Integral*, Kluwer, Dordrecht (1994).
- [6] J. C. Dermody and R. T. Rockafellar, Cash stream valuation in face of transaction costs and taxes, *Mathematical Finance* 1, 1, 31-54 (1991).
- [7] A. K. Dixit and R. S. Pindyck, *Investment under Uncertainty*, Princeton University Press (1994).
- [8] D. Duffie, *Dynamic Asset Pricing*, Princeton University Press (1992).
- [9] J. Dupačová, G. Consigli, and S. W. Wallace, Scenarios for multistage stochastic programs, *Ann. Oper Res.* 100, 25-53 (2000).
- [10] I. V. Evstigneev and S. D. Flåm, Stochastic programming: non-anticipativity and Lagrange multipliers, in C.A. Floudras and P.M. Pardalos: *Encyclopedia of Optimization*, Volume V, R-Z. Dordrecht: Kluwer (2001).
- [11] H. Föllmer and A. Schied, *Stochastic Finance*, W. de Gruyter, Berlin (2002).
- [12] D. Gale, *Linear Economic Models*, The University of Chicago Press (1960).
- [13] R. S. Garfinkel and G. L. Nemhauser, *Integer Programming*, J. Wiley & Sons, New York (1972).
- [14] I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique prior, *J. Math. Economics* 18, 141-153 (1989).
- [15] M. B. Haugh and L. Kogan, Pricing American options: a duality approach, *Operations Research* 52, 2, 258-270 (2004).
- [16] A. J. Hoffman and J. B. Kruskal, Integral boundary points of convex polyhedra, in H. W. Kuhn and A. W. Tucker (eds.) *Linear Inequalities and Related Systems*, *Ann. of Math. Studies* 38, 233-246 (1956).
- [17] P. Kall and S. W. Wallace, *Stochastic Programming*, J. Wiley&Sons, New York (1994).

- [18] A. J. King, Duality and martingales: a stochastic programming perspective on contingent claims, *Mathematical Programming* 93 (3, Ser. B) 543-562 (2002).
- [19] A. J. King, M. Koivu and T. Pennanen, Calibrated option bounds, Manuscript (2002).
- [20] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, London (1984).
- [21] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*, MIT Press, Cambridge (1994).
- [22] G. R. Parija, S. Ahmed and A. J. King, On bridging the gap between stochastic integer programming and MIP solver technologies, *INFORMS Journal on Computing* 16, 1, 73-83 (2004).
- [23] T. Pennanen and A. J. King, Arbitrage pricing of American contingent claims in incomplete markets - a convex optimization approach, Manuscript (2004).
- [24] S. R. Pliska, *Introduction to Mathematical Finance*, Blackwell, Oxford (1997).
- [25] P. Ritchken and S. Kuo, Option bounds with finite revision opportunities, *Journal of Finance* 43, 301-308, (1988).
- [26] R. T. Rockafellar, *Network Flows and Monotropic Optimization*, J. Wiley & Sons, New York (1984).
- [27] L. C. R. Rogers, Monte Carlo valuation of American options, *Mathematical Finance* 12, 3, 271-286 (2002).
- [28] S. M. Ross, *Introduction to Stochastic Dynamic Programming*, Academic Press, London (1983).
- [29] K. Sandmann, The pricing of options with an uncertain interest rate: a discrete-time approach, *Mathematical Finance* 3 (2), 201-216 (1993).
- [30] H. P. Williams, *Model Building in Mathematical Programming*, J. Wiley & Sons, Chichester (1978).
- [31] E. Wolfstetter, *Topics in Microeconomics*, Cambridge University Press (1993).

Department of Economics
University of Bergen
Fosswinckels gate 6
N-5007 Bergen, Norway
Phone: +47 55 58 92 00
Telefax: +47 55 58 92 10
<http://www.svf.uib.no/econ>