Shortfall risk minimising strategies in the binomial model: characterisation and convergence

Gino Favero*
Università Commerciale "L. Bocconi"
Istituto di Metodi Quantitativi
viale Isonzo 25, I-20133 Milano – Italy
gino.favero@unibocconi.it

Tiziano Vargiolu
Università degli Studi di Padova
Dipartimento di Matematica Pura ed Applicata
via Belzoni 7, I-35131 Padova – Italy
vargiolu@math.unipd.it

April 19, 2007

Abstract

In this paper we study the dependence on the loss function of the strategy which minimises the expected shortfall risk when dealing with a financial contingent claim in the particular situation of a binomial model. After having characterised the optimal strategies in the particular cases when the loss function is concave, linear or strictly convex, we analyse how optimal strategies change when we approximate a loss function with a sequence of suitable loss functions.
Key Words: shortfall risk minimization, binomial model, Dynamic Programming algorithm, robustness

1 Introduction

Given a market with a vector S of asset prices (one of which is non-risky), let H_N be a liability to be hedged at some future time N. If $V_N(\pi)$ is the value at time N of a portfolio corresponding to a self-financing investment strategy π, the *shortfall risk minimization problem* consists in determining $J_0(S_0, V_0)$, where for every $n = 0, 1, \ldots, N$,

$$J_n(s, v) := \inf_{\pi} \mathrm{E}\left\{\ell\big([H_N - V_N(\pi)]^+\big) \;\middle|\; S_n = s, V_n = v\right\}$$

---

*This author accomplished most of this research as a post-doc fellow in the Pure and Applied Mathematics Dept. of Padua University, which is gratefully acknowledged.

for a suitable "loss function" ℓ, with S_n and V_n the (known) price at time n of the assets and the value at time n of the portfolio. This problem typically arises when an agent does not have enough initial capital to (super)replicate the contingent claim H_N. The *loss function* ℓ is classically considered to be increasing and such that $\ell(0) = 0$. It is also customary in economic models to suppose that ℓ is either concave or convex on $\mathbb{R}_+$, depending on the investor's inclination or aversion (respectively) to taking risks. While the most used loss function is $\ell(x) = x$ (see [2, 3, 4, 9] and the references therein), and the choice $\ell(x) = 1_{x>0}$ corresponds to the problem of maximising the probability of a perfect hedge (see e.g. [1, 8]), the shortfall risk minimisation problem with a convex ℓ is used as a building block when dealing with the so-called *convex measures of risk* (see [10] for details). The major application of the foregoing problem consists in finding the optimal policy π^* and, if possible, to determine it explicitly.

In this paper, we want to analyse the dependence of the optimal policy π^* on the loss function ℓ in the case of the Cox-Ross-Rubinstein "binomial model". Namely, we want to prove that the shortfall risk minimization problem is *robust* with respect to the loss function ℓ, meaning that, if $(\ell_n)_n$ is a sequence of loss functions converging to another loss function ℓ, then the optimal strategies $\pi^{*,k}$ for the "approximating" problems also converge in some sense to the optimal strategy for the limit problem. To this extent, a necessary first step is to characterise the optimal strategies, both in the case of a concave loss function and in the case of convexity. Indeed, for the convex case we find results similar to those cited in the references above, while in the concave (also including the linear) case, we find different results, closer to those of [6, 7, 15, 16]). It is remarkable that, due to the finite state space of our model, we are allowed to drop the constraints $H_T \geqslant 0$, $V_N(\pi) \geqslant 0$ (or some weaker version) which are usually needed when solving the shortfall risk minimisation problem in infinite state spaces (see e.g. [2, 3, 4, 9, 8, 13]). Nevertheless, when the robustness issue is considered, it can be proved that our main results still hold true even if some "admissibility" constraints are taken into consideration.

In the Cox-Ross-Rubinstein "binomial model", the non-risky asset is supposed (by considering forward prices) identically equal to 1, and a single risky asset S is featured, whose price S follows the dynamics

$$S_{n+1} = S_n \omega_{n+1}, \qquad n = 0, \ldots, N-1$$

where $S_0 > 0$ is given and $(\omega_n)_{n=1,\ldots,N}$ are i.i.d. random variables taking values in the set $\{d, u\}$ (with u and d known real numbers such that $0 < d < 1 < u$) with probability law

$$p := P\{\omega_n = u\} = 1 - P\{\omega_n = d\}, \qquad n = 1, \ldots, N.$$

To avoid technicalities, we shall take as our underlying probability space the minimal one for our model; that is, we let $\Omega = \{u, d\}^N$, $\mathcal{F} = \mathcal{P}(\Omega)$.

An *investment strategy* is a sequence $\pi = (\pi_n)_{n=0,\ldots,N-1}$, where $\pi_n = (\alpha_n, \beta_n)$ is the decision to buy at time $n-1$, and hold up to time n, β_n units of the non-risky asset and α_n units of the stock S_n. We shall always suppose our investment strategies to be self-financing, namely that

$$\begin{cases} V_0 = \beta_0 + \alpha_0 S_0 \\ V_n := \beta_n + \alpha_n S_n = \beta_{n+1} + \alpha_{n+1} S_n \end{cases}$$

so that the dynamics of the portfolio value can be written as

$$V_{n+1}^{\alpha} = V_{n+1}^{\alpha}(V_n^{\alpha}, S_n, \alpha_{n+1}, \omega_{n+1}) := V_n^{\alpha} + \alpha_{n+1} S_n(\omega_{n+1} - 1).$$

By the last equality, then, the main concern in the hedging/replication problem is the determination of α_n, and this is why we use the notation V_n^{α} to stress the dependence of the portfolio dynamics upon the choice of α. Moreover, we shall suppose H to be a European contingent claim depending only on the final value S_N of the stock. Hence, the shortfall risk minimization problem can be written, for every $n = 0, \dots, N$, as

$$J_n(S_n, V_n) := \inf_{\alpha} \mathrm{E}\left\{\ell\big([H(S_N) - V_N^{\alpha}]^+\big) \;\big|\; S_n, V_n\right\}. \tag{1.1}$$

The paper is organised as follows. Section 2 contains the definition of the Dynamic Programming Algorithm (DPA) and some interesting results that hold when the DPA is applied to the binomial model. Moreover, in the same section the main results of [6] and [15], which will be a useful reference in the sequel, are recalled. In Section 3, after the determination of the optimal control and optimal value for the case when ℓ is concave in $\mathbb{R}_+$, it will be surprisingly clear that the investor with "linear" loss function behaves in exactly the same way as a risk prone investor. Moreover, Section 3 also explains a way to approach the incomplete information case (i.e., the situation when p is unknown) for a risk prone investor and investigates the robustness of the solutions in the case when a concave function ℓ is approximated with a sequence $(\ell_n)_n$ of concave functions. Section 4 treats the case when the loss function is convex, and the explicit determination of the optimal strategy is made using a Neyman-Pearson technique similar to [9]. Moreover, we study the robustness of the optimal strategy considering the two cases when we approximate a strictly convex loss function ℓ with a sequence $(\ell_n)_n$ of strictly convex functions and the case when the functions in the approximating sequence $(\ell_n)_n$ are strictly convex and ℓ is linear.

Appendix A is dedicated to a self-contained treatment of the optimization in particular data structures that we call "recombining binomial trees" , which are used in the determination of the optimal strategy in Section 3.

2 Preliminary results from the literature: the linear case

In [15], the key tool used for solving the shortfall risk minimization problem is the *Dynamic Programming Algorithm* (DPA for short). Since it will be also used in the present work, we summarize here its main features. Unlike the rest of the paper (and with the exception of Proposition 2.1), this section gathers results that are already known in the literature.

The DPA is computed according to the following backwards recursion:

$$J_N(s, v) := \ell\big([H(s) - v]^+\big)$$
$$J_n(s, v) := \inf_{\alpha} \mathrm{E}\{J_{n+1}(S_{n+1}, V_{n+1}^{\alpha}) \mid S_n = s, V_n^{\alpha} = v\}, \; n = 0, \dots, N-1.$$

Once solved, the DPA allows us to compute the optimal value $J_0(S_0, V_0)$ for our problem; moreover, if the inf operators in the various recursive steps are realized as min, the minimizing α is the optimal strategy.

For the DPA applied to the shortfall risk minimization in the binomial model, there are some remarkable results that do not depend on the choice of the loss function ℓ. First of all, there exists a unique equivalent martingale measure P^* such that

$$p^* = P^*\{\omega_n = u\} := \frac{1-d}{u-d}, \qquad 1 - p^* = P^*\{\omega_n = d\} = \frac{u-1}{u-d}. \tag{2.1}$$

In particular, for every $n = 0, \dots, N$, the arbitrage free price of $H(S_N)$ at time n is $V_n^*(S_n) := \mathrm{E}^*\{H(S_N) \mid S_n\}$ (Cox-Ross-Rubinstein valuation formula). Note that, in particular, $V_N^*(S_N) = H(S_N)$. It is also well known that, if $V_0 \geqslant V_0^*(S_0)$, there is a replicating strategy given by

$$\alpha_{n+1}^* := \frac{V_{n+1}^*(S_n u) - V_{n+1}^*(S_n d)}{S_n(u-d)} \tag{2.2}$$

As a consequence, the shortfall risk minimization problem for this binomial model is non trivial only in the case $V_0 < V_0^*(S_0)$. Henceforth, we shall therefore suppose that we are dealing with this case. The problem has already been solved in the case $\ell(x) = x$ in [15], even when there is incomplete information on the underlying model (see [6] and [15] again, and [12] for a different approach).

Other results on the DPA are gathered in the following proposition, whose proof – a straightforward calculation by backwards induction on n, based on the DPA and on the monotonicity of ℓ – is extensively featured in [7].

2.1 Proposition. *In the notations of the present section and for every $n = 0, \dots, N-1$,*

1. *$J_n(s, v)$ is decreasing in v;*
2. *$J_n(s, v) = 0$ for $v \geqslant V_n^*(s)$;*
3. *there exist $\overline{\alpha} < \widetilde{\alpha}$ such that*

$$J_n(s, v) = \inf_{\alpha \in [\overline{\alpha}, \widetilde{\alpha}]} \mathrm{E}\{J_{n+1}(S_{n+1}, V_{n+1}^\alpha) \mid S_n = s, V_n^\alpha = v\};$$

4. *if $v \leqslant V_n^*(s)$, in (3) above one can choose*

$$\overline{\alpha} = \alpha_{n+1}^{(1)}(s, v) := \frac{V_{n+1}^*(sd) - v}{s(d-1)}, \quad \widetilde{\alpha} = \alpha_{n+1}^{(2)}(s, v) := \frac{V_{n+1}^*(su) - v}{s(u-1)}. \tag{2.3}$$

We now want to review the most important points of [6] and [15], which completely solve the problem in the case $\ell(x) = x$. These results are reported here for the convenience of the reader, as they permit a better understanding of the meaning of the results and remarks contained in the next section.

Here, we shall often use the expressions defined in (2.3) in correspondence of the current values $v = V_n$ and $s = S_n$ for the portfolio and the risky asset respectively. For simplicity of notation, then, we shall write $\alpha_{n+1}^{(i)}$ instead of $\alpha_{n+1}^{(i)}(S_n, V_n)$. The reader should nevertheless keep in mind the dependence of the $\alpha_{n+1}^{(i)}$s not only on the current time (via the Cox-Ross-Rubinstein valuation formula) but also on S_n and V_n.

In the case $\ell(x) = x$, $V_0 < V_0^*(S_0)$, the recursive step in the dynamic programming algorithm defined above reduces, at each time n, to the minimization of a function which is piecewise linear in α on the three disjoint intervals $(-\infty, \alpha_{n+1}^{(1)})$, $(\alpha_{n+1}^{(1)}, \alpha_{n+1}^{(2)})$ and $(\alpha_{n+1}^{(2)}, +\infty)$. As a consequence, the inf is attained as a min either for $\alpha = \alpha_{n+1}^{(1)}$ or for $\alpha = \alpha_{n+1}^{(2)}$. A straightforward calculation (see [15, Theorem 4.1]) allows one to check that the optimal control $\alpha^\dagger$ is calculated at each time n by choosing

$$\alpha_{n+1}^\dagger = \begin{cases} \alpha_{n+1}^{(1)} & \text{if } p < p^*, \\ \alpha_{n+1}^{(2)} & \text{if } p > p^*. \end{cases} \tag{2.4}$$

(The case $p = p^*$ corresponds to S and V being martingales, so that any choice of the control would give the same results in the mean.) Note that the choice between the strategy $\alpha^{(1)}$ and $\alpha^{(2)}$ does not depend on n, so the optimal control is either $\alpha_n^\dagger \equiv \alpha_n^{(1)}$ for $p < p^*$ or $\alpha_n^\dagger \equiv \alpha_n^{(2)}$ for $p > p^*$. In particular, the choice of the optimal control can be made "a priori". The optimal value for the shortfall risk is then $\min\{\frac{p}{p^*}, \frac{1-p}{1-p^*}\}^N (V_0^*(S_0) - V_0)^+$.

In [6], some additional properties of the strategies belonging to the class

$$\Pi := \big\{(\alpha_n)_n \;\big|\; \alpha_n \in \{\alpha_n^{(1)}, \alpha_n^{(2)}\} \text{ for every } n = 1, \dots, N\big\} \tag{2.5}$$

are investigated. In particular, a slight modification of the proof of Theorem 4.1 in [15] shows that for every $\alpha \in \Pi$,

$$\mathrm{E}\{H(S_N) - V_N^\alpha \mid S_0, V_0\} = \Big(\frac{p}{p^*}\Big)^{\lambda_N} \Big(\frac{1-p}{1-p^*}\Big)^{N-\lambda_N} [V_0^*(S_0) - V_0]^+, \tag{2.6}$$

where $\lambda_N := \#\{n \mid \alpha_n = \alpha_n^{(1)}\}$. As a consequence, it is possible to calculate an adaptive strategy on a "step-by-step" basis, namely, choosing $\alpha_n = \alpha_n^{(1)}$ (respectively, $\alpha_n = \alpha_n^{(2)}$) if, given our information at time n, we are lead to believe that $p < p^*$ (respectively, $p > p^*$). Moreover, it is possible to define a one to one correspondence between controls $\alpha \in \Pi$ and "events" $\omega = (\omega_n)_{n=0,\dots,N-1}$ by defining

$$\big(\omega(\alpha)\big)_n := \begin{cases} u & \text{if } \alpha_n = \alpha_n^{(1)} \\ d & \text{if } \alpha_n = \alpha_n^{(2)} \end{cases}. \tag{2.7}$$

In other words, $\omega(\alpha)$ is the only event such that the stock S goes up every time the investor "bets" on it going down on average, and the stock goes down every time the investor bets on it going up on average. Now, it is straightforward to prove that *any strategy* $\alpha \in \Pi$ *gives perfect hedging in all events* $\omega \neq \omega(\alpha)$. This is quite a strong result, which has two main consequences. First of all, Π is a class of "quasi-replicating" controls, in the sense that every $\alpha \in \Pi$ gives perfect hedging with probability close to 1. More in detail, since $\mathrm{E}\{(H(S_N) - V_N^\alpha)^+\} > 0$ and $H(S_N)(\omega) - V_N^\alpha(\omega) = 0$ in all events $\omega \neq \omega(\alpha)$, one has

$$P\{H(S_N) - V_N^\alpha > 0\} = P\{\omega(\alpha)\} = p^{\lambda_N}(1-p)^{1-\lambda_N} \tag{2.8}$$

with λ_N defined as above. Moreover, from (2.6), it follows that in the only "critical" event $\omega(\alpha)$ we have

$$H(S_N(\omega(\alpha))) - V_N^\alpha(\omega(\alpha)) = \Big(\frac{1}{p^*}\Big)^{\lambda_N} \Big(\frac{1}{1-p^*}\Big)^{N-\lambda_N} [V_0^*(S_0) - V_0]^+,$$

which is generally much greater than the initial "lack" of capital.

3 Risk prone investor: the concave case

Throughout this section ℓ will be supposed to be a concave function on $\mathbb{R}_+$. As in Section 2, we use the convention of writing $\alpha_n^{(i)}$ instead of $\alpha_n^{(i)}(S_n, V_n)$.

Before starting the discussion, we need an extension of Proposition 2.1. Its proof, a straightforward calculation by backwards induction, is again featured in [7].

3.1 Proposition. *If ℓ is concave then, in the notations of Section 2, the function $J_n(s,\cdot)_{|(-\infty, V_n^*(s)]}$ is concave for every $n = 0, \ldots, N$.*

The fact that the concavity "propagates" along the DP algorithm allows us to find a straightforward explicit solution for our problem.

3.2 Theorem. *Let ℓ be a concave function and, for every $n = 0, \ldots, N$, let J_n be the function defined in Equation (1.1). Suppose that $V_0 < V_0^*(S_0)$. Then, for every n,*

$$J_n(S_n, V_n) = \min_{k=n,\ldots,N} p^{k-n}(1-p)^{N-k} \ell\left(\frac{V_n^*(S_n) - V_n}{(p^*)^{k-n}(1-p^*)^{N-k}}\right). \tag{3.1}$$

Define, for every $n = 0, \ldots, N-1$,

$$\alpha_{n+1}^{(1)} := \frac{V_{n+1}^*(S_n d) - V_n}{S_n(d-1)}, \qquad \alpha_{n+1}^{(2)} := \frac{V_{n+1}^*(S_n u) - V_n}{S_n(u-1)}.$$

Then the optimal strategy is given by

$$\alpha_{n+1}^* = \begin{cases} \alpha_{n+1}^{(1)} & \text{if } p^{N-n}\ell\left(\frac{V_n^*(S_n)-V_n}{(p^*)^{N-n}}\right) \leqslant (1-p)^{N-n}\ell\left(\frac{V_n^*(S_n)-V_n}{(1-p^*)^{N-n}}\right) \\ \alpha_{n+1}^{(2)} & \text{if } p^{N-n}\ell\left(\frac{V_n^*(S_n)-V_n}{(p^*)^{N-n}}\right) > (1-p)^{N-n}\ell\left(\frac{V_n^*(S_n)-V_n}{(1-p^*)^{N-n}}\right) \end{cases}. \tag{3.2}$$

3.3 Remark. As the proof of this theorem will make clear, we can always restrict the class of admissible controls to the class Π defined in (2.5). In particular, this allows a proof by induction via a straightforward calculation that from $V_0 \leqslant V_0^*(S_0)$ follows $V_n \leqslant V_n^*(S_n)$ almost surely (i.e., for every event ω) for every n. This is the reason why in the expression for $J_n(S_n, V_n)$ there is no need to consider the positive part of the argument of ℓ.

Proof of Theorem 3.2. For every $n = 0, \ldots, N-1$ define j_n^u and j_n^d as in the proof of Proposition 2.1, so that

$$J_n(S_n, V_n) = \inf_\alpha \{j_{n+1}^u(s, v, \alpha) + j_{n+1}^u(s, v, \alpha)\}.$$

By (3) and (4) of Proposition 2.1 we can restrict the computation of the inf to the interval $[\alpha_{n+1}^{(1)}, \alpha_{n+1}^{(2)}]$ where, as already seen in the proof of Proposition 3.1, $j_{n+1}^u + j_{n+1}^d$ is concave. So, since a concave function on an interval can attain its lowest value only at the extremal points, the problem reduces to comparing the values taken by the function at the extrema. Namely, on recalling that for

$\alpha = \alpha^{(1)}_{n+1}$ (respectively, $\alpha = \alpha^{(2)}_{n+1}$) one has $j^d_{n+1} = 0$ (respectively, $j^u_{n+1} = 0$), one finds that

$$\begin{aligned} J_n(S_n, V_n) &= \min\{pJ_{n+1}(S_n u, V_n + \alpha^{(1)}_{n+1} S_n(u-1)), \\ &\qquad (1-p)J_{n+1}(S_n d, V_n + \alpha^{(2)}_{n+1} S_n(d-1))\}, \\ \operatorname{argmin}[\cdots] = \alpha^{(1)}_{n+1} &\iff pJ_{n+1}(S_n u, V_n + \alpha^{(1)}_{n+1} S_n(u-1)) \leqslant \\ &\qquad \leqslant (1-p)J_{n+1}(S_n d, V_n + \alpha^{(2)}_{n+1} S_n(d-1)), \end{aligned}$$

and, as in [6, Section 2] (see also Section 2 here), we are allowed to restrict the class of admissible controls to the class Π defined in (2.5). By recalling that $\mathrm{E}^*\{V^*_{n+1}(S_{n+1}) \mid S_n = s\} = V^*_n(s)$, it is straightforward to verify by backward induction on n that for every $\alpha \in \Pi$, defining

$$\lambda_n := \#\{k > n \mid \alpha_k = \alpha^{(1)}_k\}, \qquad \mu_n := \#\{k > n \mid \alpha_k = \alpha^{(2)}_k\} = N - n - \lambda_n,$$

gives

$$\mathrm{E}_{S_n,V_n}\{\ell([H(S_N) - V^{\alpha}_N]^+)\} = p^{\lambda_n}(1-p)^{\mu_n}\,\ell\left(\frac{V^*_n(S_n) - V_n}{(p^*)^{\lambda_n}(1-p^*)^{\mu_n}}\right). \tag{3.3}$$

Hence it follows that the optimal value is as in (3.1).

To prove that the strategy given in (3.2) is optimal, we start by observing that from (3.3) it follows that, for the purpose of determining the shortfall risk, only the number of times that $\alpha_n = \alpha^{(1)}_n$ matters, and not the particular sequence of choices. The evolution of the shortfall risk with respect to the chosen control from time n on can then be seen as a "recombining binomial tree" of depth $N-n$ (see Definition A.1 in the Appendix for the definition) as follows. For every $m = n+1, \ldots, N$, associate the choice $\alpha_m = \alpha^{(1)}_m$ (respectively, $\alpha_m = \alpha^{(2)}_m$) with the decision of branching left (respectively, right) at depth $m-n-1$, so as to set a one-to-one correspondence between strategies in the class Π and branching sequences for the tree. Formula (3.3) suggests associating with each leaf τ^{N-n}_k the value $p^k(1-p)^{N-n-k}\ell\big(\frac{V^*_n(S_n)-V_n}{(p^*)^k(1-p^*)^{N-n-k}}\big)$, i.e., the shortfall risk associated with any strategy leading to that leaf. The proof is now completed by observing that the proposed strategy is exactly the "optimal branching sequence" of Proposition A.2. □

3.4 Remark. At the end of Section 2 we noted that any strategy in the class Π is "quasi replicating", i.e., it leads to perfect hedging in all events $\omega \neq \omega(\alpha)$ as defined in (2.7). This result does not depend on the function ℓ taken into consideration, so we can conclude that the expected value (3.3) found in the proof to Theorem 3.2 corresponds to a shortfall of $\frac{V^*_n(S_n)-V_n}{(p^*)^{\lambda_n}(1-p^*)^{\mu_n}}$ in the only "critical" non-hedging event $\omega(\alpha)$, whose probability is $P(\omega(\alpha)) = p^{\lambda_n}(1-p)^{\mu_n}$ as seen in (2.8).

3.5 Remark. In the case of "complete information" (i.e., when the probability p is known by the investor), the optimal strategy can be chosen *a priori* instead of on a step-by-step basis. Actually, once the $\overline{k}$ that minimizes $p^k(1-p)^{N-k}\ell\big(\frac{V^*_0(S_0)-V_0}{(p^*)^k(1-p^*)^{N-k}}\big)$ is determined, any strategy choosing $\overline{k}$ times

$\alpha_n = \alpha_n^{(1)}$ is optimal. Note also that all of these strategies have the same probability $p^{\overline{k}}(1-p)^{N-\overline{k}}$ of being non-hedging and the same shortfall $\ell\big(\frac{V_0^*(S_0)-V_0}{(p^*)^{\overline{k}}(1-p^*)^{N-\overline{k}}}\big)$ in the "critical" event, so that the choice of the "preferred" optimal strategy has to be made according to completely subjective criteria.

The expression (3.2) for the optimal control has been chosen because it can easily be adapted in the spirit of [6, Section 3] (summarized in Section 2) to elaborate optimal adaptive controls in the case of incomplete knowledge of the model.

3.6 Remark. In the case $\ell(x) = x$ examined in [6] and [15], a quite remarkable fact is that the (generally unique) optimal strategy consists of choosing either $\alpha \equiv \alpha^{(1)}$ or $\alpha \equiv \alpha^{(2)}$. In other words, in this case the optimal leaf in the tree described in the proof of Theorem 3.2 is either at the extreme "left" or at the extreme "right". This may not happen in the general case, as the following (perhaps quite artificial) example shows even in the very simple case of three possible final outcomes.

3.7 Example. Take $\ell(x) := \sqrt{x} + \sqrt[5]{x}$, $N = 2$, $p = .245$, $u = 3.17$ and $d = .8$ (so that $p^* = .0845$). Consider a contingent claim H such that $V_0^*(S_0) > 170$ and choose V_0 such that $V_0^*(S_0) - V_0 = 170$. The three values associated with the final leaves are then

$$\begin{aligned} p^2\ell\left(\frac{V_0^*(S_0)-V_0}{(p^*)2}\right) &\sim 9.77,\\ p(1-p)\ell\left(\frac{V_0^*(S_0)-V_0}{(p^*)(1-p^*)}\right) &\sim 9.54,\\ (1-p)^2\ell\left(\frac{V_0^*(S_0)-V_0}{(1-p^*)2}\right) &\sim 9.72. \end{aligned}$$

This way the optimal "leaf" is the central one, and two optimal controls can be built by choosing either $\alpha_0 = \alpha_0^{(1)}, \alpha_1 = \alpha_1^{(2)}$ or $\alpha_0 = \alpha_0^{(2)}, \alpha_1 = \alpha_1^{(1)}$. Note also that the policy proposed in Theorem 3.2 will choose $\alpha_1 = \alpha_1^{(2)}$.

3.8 Example. Consider the case when $\ell(x) = x^\kappa$ for some $0 \le \kappa \le 1$. In particular, the case $\kappa = 1$ corresponds to minimising the mean shortfall risk, and, due to the assumption that $\ell(0) = 0$, the case $\kappa = 0$ corresponds to minimising the probability of positive shortfall. These two particular cases have already been considered in [6] and [15]: when $\kappa = 1$, the optimal strategy is the $\alpha^\dagger$ defined in (2.4), and when $\kappa = 0$ the optimal strategy $\alpha^\ddagger$ is calculated by choosing

$$\alpha_n^\ddagger = \begin{cases} \alpha_n^{(1)} & p < 0.5 \\ \alpha_n^{(2)} & p > 0.5 \end{cases}$$

(when $p = 0.5$, all strategies $\alpha \in \Pi$ are optimal).

When $0 < \kappa < 1$, the solution still takes an appearance close to the cited results, namely, the optimal policies always choose either $\alpha \equiv \alpha^{(1)}$ or $\alpha \equiv \alpha^{(2)}$. Actually, in this case, for every $n = 0, \dots, N-1$ the optimal control α^* is computed by choosing

$$\alpha_n^* = \alpha_n^{(1)} \iff \frac{p}{(p^*)^l} \leqslant \frac{1-p}{(1-p^*)^\kappa} \quad \text{i.e.,} \quad \frac{p}{1-p} \leqslant \Big(\frac{p^*}{1-p^*}\Big)^\kappa,$$

independent of n.

Moreover, let $\overline{\kappa} = \frac{\log(p)-\log(1-p)}{\log(p^*)-\log(1-p^*)}$ (namely, the value for which $\frac{p}{1-p} = \left(\frac{p^*}{1-p^*}\right)^{\overline{\kappa}}$). Excluding the "undecidable" cases $p = 0.5$ (all events have the same probability), $p^* = 0.5$ (division by zero in $\overline{\kappa}$) and $p = p^*$ (the stock and the portfolio are martingales under the real world probability measure), the following cases may then occur.

$p^* < p < 0.5$ $(0 < \kappa < 1)$	$\alpha^* = \begin{cases} \alpha^\ddagger = \alpha^{(1)} & \kappa < \overline{\kappa} \\ \alpha^\dagger = \alpha^{(2)} & \kappa > \overline{\kappa} \end{cases}$
$p^* < 0.5$, $p \neq (p^*, 0.5)$ $(\kappa < 0$ or $\kappa > 1)$	$\alpha^* = \alpha^\ddagger = \alpha^\dagger = \alpha^{(1)}$
$0.5 < p < p^*$ $(0 < \kappa < 1)$	$\alpha^* = \begin{cases} \alpha^\ddagger = \alpha^{(2)} & \kappa < \overline{\kappa} \\ \alpha^\dagger = \alpha^{(1)} & \kappa > \overline{\kappa} \end{cases}$
$p^* > 0.5$, $p \neq (0.5, p^*)$ $(\kappa < 0$ or $\kappa > 1)$	$\alpha^* = \alpha^\ddagger = \alpha^\dagger = \alpha^{(2)}$

Note that if $\alpha^\ddagger = \alpha^\dagger$, then the optimal control α^* for the shortfall risk minimization problem coincides with both $\alpha^\ddagger$ and $\alpha^\dagger$ for every κ. On the other hand, if $\alpha^\ddagger \neq \alpha^\dagger$, the exponent $0 < \overline{\kappa} < 1$ is "critical", in the sense that $\alpha^* = \alpha^\ddagger$ for $\kappa \in (0, \overline{\kappa})$ and $\alpha^* = \alpha^\dagger$ for $\kappa \in (\overline{\kappa}, 1)$. In other words, the study of the shortfall risk minimization problem with $\ell(x) = x^\kappa$ can be reduced to the two fundamental problems with $\kappa = 0$ and $\kappa = 1$.

Suppose now that $(\ell_k)_k$ is a given sequence of concave loss functions converging pointwise to a (concave) loss function ℓ. For each function ℓ_k (resp. for ℓ), let $\alpha^{*,k}$ (resp. α^*) be the optimal strategy for the corresponding shortfall risk minimisation problem. It is then a natural question whether the $\alpha^{*,k}$ converge in some sense to α^*.

Though the answer is somewhat complicated by the fact that the optimal solution of a concave problem in general is not unique, it is nevertheless possible to give the following result.

3.9 Theorem. *Let $(\ell_k)_{k\in\mathbb{N}}$ be a sequence of concave functions such that $\ell_k \to \ell$ pointwise and let $\alpha^{*,k}$ be the optimal strategy for the shortfall risk minimisation problem corresponding to ℓ_k for every k. Then there exists a $(k_h)_h$ such that $\alpha_n^{*,k_h} \to \alpha_n^*$ almost surely for all $n = 1, \ldots, N$, where α^* is an optimal strategy for the shortfall risk minimisation problem corresponding to ℓ.*

Proof. Define for every $k \in \mathbb{N}$,

$$\varphi_k^0(S_0, V_0) := p^N \ell_k \left(\frac{V_0^*(S_0) - V_0}{(p^*)^N} \right) - (1-p)^N \ell_k \left(\frac{V_0^*(S_0) - V_0}{(1-p^*)^N} \right),$$

so that according to (3.2) the optimal strategy at time 0 for the problem associated to ℓ_k is to choose $\alpha_1^{*,k} = \alpha_1^{(1)}$ (respectively, $\alpha_1^{*,k} = \alpha_1^{(2)}$) if $\varphi_k^0 \leqslant 0$ (respectively, $\varphi_k^0 > 0$). Note also that, since $\alpha_1^{(1)} = \frac{V_1^*(S_0 d) - V_0}{S_0(d-1)}$, $\alpha_1^{(2)} = \frac{V_1^*(S_0 u) - V_0}{S_0(u-1)}$ (see (2.3) for the definition) and $V_n^*(S_n) = \mathrm{E}\{H(S_N) \mid S_n\}$ is independent of the loss function, $\alpha_1^{(1)}$ and $\alpha_1^{(2)}$ also do not depend on the loss function. We can now distinguish two cases:

- if either $\varphi_k^0 \leqslant 0$ or $\varphi_k^0 > 0$ from some $\overline{k}$ on, then the sequence $(\alpha_1^{*,k})_k$ is constant from $\overline{k}$ on, and thus it converges to a limit α_1^*,

- if φ_k^0 converges to 0 taking both positive and negative values infinitely many times, consider the subsequence $(\varphi^0_{k_{h_0}})_{h_0}$ formed either by the positive or by the non-positive values taken by $(\varphi_k^0)_k$ and the problem reduces to the previous case.

This way, we are dealing with a subsequence $(\ell_{k_{h_0}})_{h_0}$ such that $(\alpha_1^{*,k_{h_0}})_{h_0} \equiv \alpha_1^*$ from some $\overline{h}_0$ on. Note that, since $V_1^\alpha = V_0 + \alpha_1 S_0(\omega_0 - 1)$ only depends on the chosen strategy and not on the loss function ℓ, the optimal portfolios of the problems associated to the $\ell_{k_{h_0}}$s for $h_0 \geqslant \overline{h}_0$ all follow the same evolution in the first time interval.

The existence of the limit strategy α^* now follows by induction on n in a similar way, i.e., defining at each step $n = 1, \ldots, N-1$

$$\varphi_k^n(S_n, V_n) := p^{N-n} \ell_k \left(\frac{V_n^*(S_n) - V_n}{(p^*)^{N-n}} \right) - (1-p)^{N-n} \ell_k \left(\frac{V_n^*(S_n) - V_n}{(1-p^*)^{N-n}} \right)$$

for every $k \in \mathbb{N}$, and extracting from the sequence $(\ell_{k_{h_{n-1}}})_{h_{n-1}}$ a subsequence $(\ell_{k_{h_n}})_{h_n}$ such that $(\alpha_{n+1}^{*,k_{h_n}})_{h_n} \equiv \alpha^*_{n+1}$ from some $\overline{h}_n$ on.

It only remains to show that α^* is an optimal strategy for the problem associated to ℓ. Note that, by Equation (3.1) and the independence between the strategies and the loss functions, the optimal values of the problems associated to the ℓ_ks converge to the shortfall associated to strategy α^* under the loss function ℓ. On the other hand, since for every sequence of functions $(f_n)_n$ converging pointwise to f one has $\liminf f_n \leqslant \inf f$, this limit value necessarily must be the optimal value for the problem associated to ℓ. □

3.10 Remark. Note that the optimal limit strategy α^* is not necessarily unique, i.e., there might be different subsequences of $(\alpha_k^*)_k$ converging to different optimal solutions to the problem (4.2). Note also that one does *not* need *strict* concavity of the ℓ_ks, so that the limit function ℓ might also be linear.

3.11 Remark. The concave problem can also be solved from the "static" point of view defined in the next section. In this case, the optimal modified contingent claim X^* will have the same form as in (4.3) where E is the set of the "critical" ωs corresponding to any optimal branching sequence leading to the optimal "leaf". This way, it is possible to see that any convex combination of the $\binom{N}{k}$ optimal modified claims corresponding to the $\binom{N}{k}$ optimal strategies in Π still gives an optimal solution to the shortfall risk minimization problem.

This allows us to understand that, if the optimal strategy $\alpha^* \in \Pi$ is not unique, then there also exist infinitely many optimal strategies not in Π. Moreover, it is no longer possible to give an analogue to Theorem 3.9 unless the modified contingent claims optimal for the approximating problems are required to be chosen according to some common condition (e.g., all the loss is concentrated on a single ω, or γ constant on E).

4 Risk averse investor: the convex case

Throughout this section ℓ will be assumed to be a convex function on $\mathbb{R}_+$. As in the above section (see Proposition 3.1), it is possible to give an extension of Proposition 2.1, which we cite below. We give the formulation of this extension below but, as the proof is based on rather technical calculations and the proposition itself will not be used in this paper, we invite the interested reader to find it in [7].

4.1 Proposition. *If ℓ is convex then, in the notations of Section 2, the function $J_n(s,\cdot)$ is convex for every $n = 0, \ldots, N$.*

This proposition has as an immediate consequence the existence of an optimal strategy for the convex case. It also follows that, if ℓ is strictly convex, then the optimal strategy is unique. Nevertheless, when trying to determine explicitly the optimal strategy in general form by using dynamic programming arguments one is led to quite complex calculations. For the sake of simplicity, then, from now on we shall shift from a DPA-based approach to another one, with techniques similar to those in [9]. Notice that, for the case when ℓ is strictly convex, we could also use techniques based on convex duality as in [12] or [13]. However, due to the absence of lower bounds on the claim $H(S_N)$ and on the portfolio X, it would not be possible to apply these techniques to the case when $\ell(x) = x$.

4.2 Definition. We define the set of the *modified contingent claims* as

$$\mathcal{X} := \{X \mid X \leqslant H(S_N) \text{ (a.s.)}, \mathrm{E}^*\{X\} \leqslant V_0\}. \tag{4.1}$$

Roughly speaking, $\mathcal{X}$ is the set of all the claims less than $H(S_N)$ which can be replicated with initial capital (less than or equal to) V_0 or, equivalently, the set of all the possible final states of adapted, self-financing strategies starting from initial capital (less than or equal to) V_0.

We can now consider the shortfall risk minimization problem from a "static" point of view:

$$\min_{X\in\mathcal{X}} \mathrm{E}_{S_0}\left\{\ell\big(H(S_N) - X\big)\right\}. \tag{4.2}$$

We want to show that the modified contingent claim that solves (4.2) coincides with the payoff of the optimal portfolio for the shortfall risk minimization problem. We shall start from the linear case ($\ell(x) = x$) and from this case we shall derive the solution for the strictly convex case.

The following lemma, a key tool for the proof of the main Theorem 4.4, shows how this approach, an alternative to the DPA used in Section 2, can be applied to the study of the "mean shortfall risk minimization problem", i.e., the shortfall risk minimization problem in the case $\ell(x) = x$. Note that, in this case, (4.2) reduces to the problem $\max_{x\in\mathcal{X}} \mathrm{E}\{X\}$.

4.3 Lemma. *If $X^* \in \mathcal{X}$ solves the problem $\max_{x\in\mathcal{X}} \mathrm{E}\{X\}$, then the hedging strategy α^* for the claim X^* also solves the mean shortfall risk minimization problem $\min_\alpha \left\{ \mathrm{E}_{S_0,V_0}\{(H(S_N) - V_N^\alpha)^+\}\right\}$. Moreover, define $c_{es} := \min_\omega \frac{\mathrm{d}P}{\mathrm{d}P^*}$, where P and P^* are, respectively, the "real world" and the martingale probability measures. Set $E := \{\omega \mid \frac{\mathrm{d}P}{\mathrm{d}P^*}(\omega) = c_{es}\}$, and*

$$X^* := H(S_N)1_{E^c} + \gamma 1_E$$

where γ is any random variable such that $\mathrm{E}^*\{X^*\} = V_0$. Then X^* solves the problem (4.2).

Proof. This proposition can be proved directly as in [16], but we give a shorter proof that uses the results of Section 2. Consider the strategy $\alpha^\dagger$ defined in (2.4). Since the corresponding portfolio value $V_N^{\alpha^\dagger}$ is the payoff of a self-financing strategy starting from the initial capital V_0 and $V_N^{\alpha^\dagger} \leqslant H(S_N)$, then $V_N^{\alpha^\dagger}$ itself is a modified contingent claim in the sense of (4.1). Moreover, if $X \in \mathcal{X}$, then one has $\mathrm{E}_{S_0,V_0}\{(H(S_N)-X)^+\} = \mathrm{E}_{S_0,V_0}\{H(S_N)\} - \mathrm{E}_{S_0,V_0}\{X\}$, so that the first part of the proposition follows.

As for the second part, if $p = p^*$ then all the modified contingent claims $X \in \mathcal{X}$ are solutions. If $p \neq p^*$, then from Section 2 we also deduce that $V_N^{\alpha^\dagger}$ is equal to $H(S_N)$ on all events except for the "least probable" one, where the payoff is equal to

$$H(S_N) - \frac{1}{\min\{(p^*)^N, (1-p^*)^N\}}(V_0^*(S_0) - V_0).$$

In particular, it is straightforward to check that $\mathrm{E}^*\{V_N^{\alpha^\dagger}\} = V_0$, and thus that $V_N^{\alpha^\dagger}$ can be defined as in the statement above. □

To apply this approach to the convex case, we make the further assumption that ℓ is strictly convex, continuously differentiable and such that $\ell'(0) = 0$. Indeed, these properties appear to be the right ones to state the following results in a reasonable notation, but we believe that a generalization to non-$\mathcal{C}^1$ functions, in terms of the sub-differential, should be straightforward.

4.4 Theorem. *Set $I := (\ell')^{-1}$ and define the modified contingent claim*

$$X^* := H(S_N) - I\left(c^* \frac{\mathrm{d}P^*}{\mathrm{d}P}\right)$$

with $c^* > 0$ chosen in such a way that $E^*\{X^*\} = V_0$. Then X^* solves the "static" problem (4.2).

The proof can be carried out either by using Neyman-Pearson techniques as in [9] (see [7]), or using convex duality techniques (i.e., Lagrange multipliers).

4.5 Remark. Once the optimal modified contingent claim X^* is determined, the optimal strategy is simply the Cox-Ross-Rubinstein replicating strategy for X^* as in (2.2), that is,

$$\alpha_{n+1}^* = \frac{\mathrm{E}^*\{X^*|S_{n+1} = S_n u\} - \mathrm{E}^*\{X^*|S_{n+1} = S_n d\}}{S_n(u-d)}.$$

Now let $(\ell_k)_k$ be a sequence of strictly convex, $\mathcal{C}^1$ loss functions that converge pointwise to a loss function ℓ. For each function ℓ_k (resp. for ℓ), we call X_k^* (resp. X^*) the optimal modified claim which solves the problem in (4.2), and $\alpha^{*,k}$ (resp. α^*) the corresponding optimal strategy. We shall also write

$$\Delta := \mathrm{E}^*\{H(S_N)\} - V_0 = V_0^*(S_0) - V_0.$$

Since the optimal modified claims are very different in the two cases when the limit loss function ℓ is strictly convex and when $\ell(x) = x$, we shall deal separately with these two situations. We start with the first case using a technique similar to [11].

4.6 Theorem. *Let $(\ell_k)_{k\in\mathbb{N}}$ be a sequence of strictly convex, $\mathcal{C}^1$ functions such that $\ell_k'(0) = 0$ for every k. If $\lim_k \ell_k = \ell$ pointwise with ℓ strictly convex, $\mathcal{C}^1$ and such that $\ell'(0) = 0$, then $X_k^* \to X^*$ almost surely.*

Proof. We start by proving that $c_k^* \to c^*$. To do this, define

$$\varphi_k(c) = \mathrm{E}\left\{ I_k\left(c\frac{\mathrm{d}P^*}{\mathrm{d}P} \right) \right\}$$

and note that c^* and the c_k^*s have the property that $\Delta = \varphi_k(c_k^*) = \varphi(c^*)$ (where φ is defined in the obvious way). Since the probability space Ω is finite, it is straightforward to check that $\varphi_k(c) < +\infty$ for all $c \geq 0$ and that φ_k is continuous for all k.

Since $I_k \to I$ pointwise, we have that for every $c \in \mathbb{R}+$

$$I_k\left(c\frac{\mathrm{d}P^*}{\mathrm{d}P}(\omega) \right) \leq \sup_{k\in\mathbb{N}} I_k\left(c\frac{\mathrm{d}P^*}{\mathrm{d}P}(\omega) \right) \leq \sup_{\omega\in\Omega}\sup_{k\in\mathbb{N}} I_k\left(c\frac{\mathrm{d}P^*}{\mathrm{d}P}(\omega) \right) =: K < \infty,$$

i.e., the sequence $I_k(c\frac{\mathrm{d}P^*}{\mathrm{d}P}(\omega))$ is dominated. We know from [11] that if $\ell_k \to \ell$ pointwise then $I_k \to I$ uniformly on compact sets. This means that $\varphi_k(c) \to \varphi(c)$ for all $c \geq 0$. Since φ and the φ_k are continuous and strictly increasing, it follows that $\varphi_k^{-1} \to \varphi^{-1}$ pointwise. Moreover, these functions are continuous, so $c_k^* = \varphi_k^{-1}(\Delta) \to \varphi^{-1}(\Delta) = c^*$.

Now we are able to prove the theorem. First of all notice that by Theorem 4.4

$$X_k^* = H(S_N) - I_k\left(c_k^*\frac{\mathrm{d}P^*}{\mathrm{d}P} \right).$$

Since I_k converges uniformly on compact sets to I, the right hand side converges to $H(S_N) - I(c^*\frac{\mathrm{d}P^*}{\mathrm{d}P})$ almost surely, so that one gets $X_k^* \to X^*$ almost surely. $\square$

Now we consider the case $\ell(x) = x$. For this case, since I cannot be defined, we can no longer use techniques similar to [11], and we have to develop an *ad-hoc* technique. Recall from Lemma 4.3 that an optimal claim in this case is given by

$$X^* = H(S_N) - \frac{\Delta}{P^*(E)} 1_E \tag{4.3}$$

with $E = \left\{ \frac{\mathrm{d}P}{\mathrm{d}P^*} = c_{es} \right\}$.

4.7 Theorem. *Let $(\ell_k)_{k\in\mathbb{N}}$ be a sequence of strictly convex, $\mathcal{C}^1$ functions such that $\ell_k'(0) = 0$ for every k. If $\lim_k \ell_k(x) = x$ pointwise, then $X_k^* \to X^*$ almost surely, where X^* is defined as in (4.3).*

Proof. Since $\ell_k(x) \to x$ pointwise and the ℓ_k are convex, then $\ell_k'(x) \to 1$ uniformly on compact sets of $\mathbb{R}_+$ (see [11] or [14]). Since the ℓ_k' are all increasing functions, it is easy to prove that

$$\lim_{k\to\infty} I_k(x) = \begin{cases} 0 & \text{if } x < 1 \\ +\infty & \text{if } x > 1 \end{cases}$$

The convergence is uniform on compact sets of $(0, 1)$, and also the convergence to $+\infty$ is "uniform" in the sense that for every $\varepsilon > 0$ and $M > 0$ there exists $\overline{k}$ such that $I_k(x) > M$ for all $k > \overline{k}$ and $x > 1 + \varepsilon$.

This convergence implies that $c_k^* \to c_{es}$. Indeed, for every $\varepsilon > 0$ there exists $\bar{k}$ such that $0 < I_k(x) < \varepsilon$ for every $k > \bar{k}$, $x \in (0, 1-\varepsilon)$. Then

$$\Delta = \mathrm{E}^* \left\{ I_k \left(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \right) \right\} \leq \varepsilon + \mathrm{E}^* \left\{ I_k \left(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \right) 1_{\left\{ c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \geqslant 1-\varepsilon \right\}} \right\}$$

so $\mathrm{E}^* \left\{ I_k(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P}) 1_{\{c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \geq 1-\varepsilon\}} \right\} \geq \Delta - \varepsilon$; in particular, $\left\{ c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \geq 1-\varepsilon \right\}$ is not empty for all $k > \bar{k}$, and this means that $c_k^* \max_\omega \frac{\mathrm{d}P^*}{\mathrm{d}P}(\omega) \geq 1-\varepsilon$, so $c_k^* < (1-\varepsilon) \min_\omega \frac{\mathrm{d}P}{\mathrm{d}P^*}(\omega) = (1-\varepsilon) c_{es}$. Conversely, for all $\varepsilon > 0$ and $M > \Delta / P^*(E)$ there exists $\bar{k}$ such that $I_k(x) > M$ for all $k > \bar{k}$, $x > 1+\varepsilon$. This implies that

$$\Delta = \mathrm{E}^* \left\{ I_k \left(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \right) \right\} \geq \mathrm{E}^* \left\{ I_k \left(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \right) 1_E \right\} = I_k \left(\frac{c_k^*}{c_{es}} \right) P^*(E)$$

Thus we must have $\frac{c_k^*}{c_{es}} \leq 1+\varepsilon$. In fact, if $\frac{c_k^*}{c_{es}} > 1+\varepsilon$, we obtain $\Delta \geq M \cdot P^*(E)$, but we took $M > \Delta / P^*(E)$, so this is absurd. In conclusion, we have that for all ε, $(1-\varepsilon) c_{es} \leq c_k^* \leq (1+\varepsilon) c_{es}$ from a certain $\bar{k}$ on, which yields $c_k^* \to c_{es}$.

Now we only have to prove that $I_k(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P}) \to \frac{\Delta}{P^*(E)} 1_E$. Since $c_k^* \to c_{es}$, it follows that $\lim_{k\to\infty} c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} = c_{es} \frac{\mathrm{d}P^*}{\mathrm{d}P}$, which is equal to 1 on E and less than one on E^c. Since $I_k \to 0$ uniformly on compact sets of $(0,1)$, we have that $I_k(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P}) \to 0$ on E^c, and the limit is uniform. Thus for all $\varepsilon > 0$ there exists $\bar{k}$ such that for all $k > \bar{k}$ we have

$$\Delta - \mathrm{E}^* \left\{ I_k \left(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \right) 1_E \right\} = \mathrm{E}^* \left\{ I_k \left(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P} \right) 1_{E^c} \right\} \in (0, \varepsilon)$$

This means that $I_k(\frac{c_k^*}{c_{es}}) P^*(E) \to \Delta$, and finally that $I_k(c_k^* \frac{\mathrm{d}P^*}{\mathrm{d}P}) \to \frac{\Delta}{P^*(E)} 1_E$. This completes the proof. □

4.8 Corollary. *Under the assumptions of Theorems 4.6 or 4.7, we have that* $\alpha_n^{*,k} \to \alpha_n^*$ *almost surely for all* $n = 0, \ldots, N-1$.

Proof. Recall that the optimal strategy for the shortfall risk minimization problem with loss function ℓ_k is

$$\alpha_n^{*,k} = \frac{\mathrm{E}^*\{X_k^* | S_{n-1} u\} - \mathrm{E}^*\{X_k^* | S_{n-1} d\}}{S_{n-1}(u-d)}$$

(see Remark 4.5). Since $X_k^* \to X^*$ almost surely and the probability space Ω is finite, the conditional expectations above converge almost surely to

$$\frac{\mathrm{E}^*\{X^* | S_{n-1} u\} - \mathrm{E}^*\{X^* | S_{n-1} d\}}{S_{n-1}(u-d)},$$

i.e., $\alpha_n^{*,k} \to \alpha_n^*$ almost surely for all $n = 0, \ldots, N-1$. □

A Optimal paths in recombining binomial trees

This appendix contains some technical definitions and results that formalise a “branch-and-bound” type algorithm for the purpose of determining an optimal path in a particular directed graph, which we call a recombining binomial tree. These results constitute a key tool for the proof of Theorem 3.2 concerning the solution of the shortfall risk minimization problem in the case when ℓ is concave.

A.1 Definition. For $n \in \mathbb{N}$, a *recombining binomial tree* (r.b.t. for short) of depth n is a directed graph $T(n)$ with $\frac{(n+1)(n+2)}{2}$ nodes τ_k^i, where $k = 0, \ldots, i$, $i = 0, \ldots, n$. In this way, $T(n)$ features $i+1$ nodes at each "depth" i. In particular, at the depth 0 there is only one node τ_0^0, which will be called the *vertex* of the tree $T(n)$.

For $i = 0, \ldots, n-1$, each node τ_k^i is supposed to be connected with the two nodes τ_k^{i+1} and τ_{k+1}^{i+1}. The nodes at depth n are thus "terminal" nodes, and will be called the *leaves* of the tree. We shall refer to moving from τ_k^i to τ_k^{i+1} (respectively, to τ_{k+1}^{i+1}) with the expression *branching left* (respectively, *branching right*).

Since the graph is directed, the maximum length paths on the graph start from the vertex τ_0^0 and reach a leaf τ_k^n for some k. Note that there is a one-to-one correspondence between maximum length paths and branching sequences $(\vartheta_i)_{i=0}^{n-1} \in \{l, r\}^n$, where $\vartheta_i = l$ (respectively, $\vartheta_i = r$) means the decision to branch left (respectively, right) when passing from depth i to depth $i+1$.

The problem we want to solve is the following. Suppose that each leaf τ_k^n of the r.b.t. $T(n)$ is associated with some value $r_k \in \mathbb{R}$. Then, we want to find a branching sequence so as to reach a leaf associated with the minimum value, that we shall call, respectively, an *optimal branching sequence* (or *strategy*) and an *optimal leaf.* (The problem of reaching the maximum value can trivially be reformulated as a minimum problem by associating the values $-r_k$ to the leaves.)

Note that, although if the correspondence between paths and branching sequences is one to one, in general the strategy leading to a chosen leaf τ_k^n is not unique. Actually, any strategy (starting from the vertex) which branches left $n-k$ times and right k times in any order will end in leaf τ_k^n, and it is clear that there are $\binom{n}{k}$ such strategies. Thus, the choice of the optimal branching strategy is not a consequence of the determination of the optimal leaf.

The algorithm we propose below solves the problem of choosing an optimal strategy by deciding the branching sequence while scanning the leaves to determine the minimum value.

A.2 Proposition. *Let $n \in \mathbb{N}$, and $T(n)$ be the recombining binomial tree of depth n, with values r_k, $k = 0, \ldots, n$ associated with the leaves.*

Define

$$\vartheta_0 = \begin{cases} l & \text{if } r_0 \leqslant r_n \\ r & \text{if } r_0 > r_n \end{cases}$$

and then, recursively for every $i = 1, \ldots, n$,

$$k_{\vartheta,i} := \#\{j < i \mid \vartheta_j = r\},$$
$$\vartheta_i = \begin{cases} l & \text{if } r_{k_{\vartheta,i}} \leqslant r_{n-i+k_{\vartheta,i}} \\ r & \text{if } r_{k_{\vartheta,i}} > r_{n-i+k_{\vartheta,i}}. \end{cases}$$

Then the leaf $\tau_{k_{\vartheta,n}}^n$ is associated with the minimum value, and $\vartheta = (\vartheta_i)_{i=1}^n$ is an optimal branching sequence.

Proof. Induction on n.

($n = 1$). In this case, there are only the two leaves τ_0^1 and τ_1^1 associated with the values r_0 and r_1 respectively. The thesis translates into the obvious decision to branch left and reach τ_0^1 if r_0 is the minimum value, and to branch right and reach τ_1^1 otherwise.

($n > 1$). To determine the branching at depth 0 the following considerations can be made. Note that the two outmost leaves τ_0^n and τ_n^n are both reached from the vertex by a unique branching strategy, namely, τ_0^n can only be reached by branching always left and τ_n^n can only be reached by branching always right. Any other leaf can be reached with a suitable branching sequence whatever ϑ_0 is. Thus, the proposed strategy decides to branch so as to make unreachable the outmost leaf which is associated with the higher (and, thus, non-optimal) value and reach the node $\tau^1_{k_{\vartheta,1}}$.

One can now consider the"sub-tree" T_1 of depth $n-1$ with vertex $\tau^1_{k_{\vartheta,1}}$ and leaves τ_k^n, $k = k_{\vartheta,1}, \ldots, n-1+k_{\vartheta,1}$. Note that the minimum value has to be associated with one of the leaves of T_1, since the value associated with the "discarded" leaf is greater than a value associated with a leaf of T_1. The recursive step then corresponds to the first step in this sub-tree, and so the proposition follows by recursion. □

References

[1] Cvitanic J (1998) Methods of partial hedging. Asia-Pacific Financial Markets 00: 1–28

[2] Cvitanic J (2000) Minimizing expected loss of hedging in incomplete and constrained markets. SIAM J. Control and Optimization 38 (4): 1050–1066

[3] Cvitanic J, Karatzas I (1999) On dynamic measure of risk. Finance and Stochastics 3 (4): 451–482

[4] Cvitanic J, Karatzas I (2001) Generalized Neyman-Pearson lemma via convex duality. Bernoulli 7 (1): 79–97

[5] Duffie D (1999) Dynamic Asset Pricing Theory. Princeton University Press, Princeton, NJ

[6] Favero G (2001) Shortfall risk minimization under model uncertainty in the binomial case: adaptive and robust approaches. Math. Meth. Oper. Research 53: 493–503

[7] Favero G, Vargiolu T (2002) Robustness of shortfall risk minimising strategies in the binomial model. Università degli studi di Padova, Dipartimento di matematica pura e applicata, Preprint n. 14 - 2002. http://www.math.unipd.it/~favero/workpage.html.

[8] Föllmer H, Leukert P (2000) Quantile hedging. Finance and Stochastics 3: 251–277

[9] Föllmer H, Leukert P (2000) Efficient hedging: cost versus shortfall risk. Finance and Stochastics 4: 117–146

[10] Föllmer H, Schied A (2002) Convex measures of risk and trading constraints. Finance and Stochastics 6: 429–447

[11] Jouini E, Napp C (2004) Convergence of utilityu functions and convergence of optimal strategies. Finance and Stochastics 8 (1): 133–144

[12] Kirch M (2001) Efficient hedging in incomplete markets under model uncertainty. Ph.D. Thesis, Humboldt University, Berlin.

[13] Pham H (2000) Dynamic L^p-hedging in discrete time under cone constraints. SIAM J. Control and Opt. 38: 665-682

[14] Rockafellar RT (1970) Convex Analysis. Princeton University Press, Princeton, NJ

[15] Runggaldier WJ, Trivellato B, Vargiolu T (2002) A Bayesian Adaptive Control Approach to Risk Management in a Binomial Model. In: Dalang RC, Dozzi M, Russo F (ed.) Proceedings of the Ascona '99 Seminar on Stochastic Analysis, Random Fields and Applications. Birkhauser, pp. 243–258

[16] Schulmerich M, Trautmann S (2003) Local expected shortfall-hedging in discrete time. Eur. Finance Rev. 7 (1): 75–102