The practice of portfolio replication
A practical overview of forward and inverse problems

Ron Dembo and Dan Rosen

Algorithmics Inc., 185 Spadina Avenue, Toronto, Ontario, Canada M5T 2C6

E-mail: {ron; drosen}@algorithmics.com

Portfolio replication is a powerful tool that has proven in practice its applicability to enterprise-wide risk problems such as static hedging in complete and incomplete markets and markets that gap; strategic asset and capital allocation; benchmark tracking; design of synthetic products; and portfolio compression. In this paper, we revise the basic principles behind this methodology, currently used by financial institutions worldwide, and present several practical examples of its application. We further show how inverse problems in finance can be naturally formulated in this framework. In contrast to mean-variance optimization, the scenario approach allows for general non-normal, discrete and subjective distributions, as well as for the accurate modeling of the full range of nonlinear instruments, such as options. It also provides an intuitive, operational framework for explaining basic financial theory.

1. Introduction

It is now widely recognized by financial institutions and regulatory bodies that an effective risk management function must be organized on an enterprise-wide basis, and must cover all aspects of risk in the financial markets. It is no longer sufficient to manage risk primarily at the local levels. Enterprise-wide risk management (EWRM) refers to the set of policies and procedures put in place to monitor, control and manage all financial risks of an institution in a unified way.

In addition to monitoring risk, an effective risk management function must help the firm understand the sources of its exposures, restructure its risks profile, and obtain optimal risk vs. profit trade-offs, within and across various business lines. *Portfolio replication* is a powerful, robust, methodology that can be used to address these issues generally within an EWRM and portfolio management framework. Its applications include:

- Strategic (firm-wide) and tactical asset and capital allocation.
- Risk restructuring.
- Benchmark tracking.

© J.C. Baltzer AG, Science Publishers

- Selection of efficient investment portfolios.
- Hedging and pricing in complete and incomplete markets.
- Portfolio compression.
- Estimation of implied no-arbitrage parameters and implied views (inverse problems).

Not only does the technique naturally accommodate general (non-normal) distributions and nonlinear instruments (such as options), but it also explicitly models discrete markets that we often observe in practice, where trading may be costly and liquidity limited. The methodology also naturally accommodates transaction costs, liquidity and other specified user constraints, as well as investor preferences. In this paper, we revise the basic principles behind this methodology, currently used by financial institutions worldwide, and present several practical examples of its application. The reader is further referred to Dembo [3,4], Dembo et al. [5,6] for a complete discussion and proofs of some of the results.

We focus on static replication (single stage); i.e. on replication based on a single decision at the start of the horizon, although the replication may cover several time periods (multi-step). For some multi-stage stochastic programming applications, see for example Carino and Ziemba [2], Ziemba and Mulvey [20], and the references therein.

Even traditional areas of finance (where dynamic hedging is customarily used) have shown in the last few years a growing interest in static replication as an effective tool for pricing and hedging (cf. Carr and Chou [1], and Derman et al. [7]). Static portfolio replication can easily be extended to full dynamic, multi-stage strategies, or "semi-dynamic" strategies where only several parametrized strategies are considered in the future. However, the computational effort and complexity required by full multi-stage stochastic dynamic programming solutions grows exponentially. Thus, static replication may be the best, and perhaps the only, practical solution in many problems related to enterprise-wide risk management where the dimensionality and complexity of the problems are very large. In practice, it is not uncommon for financial institutions to handle hundreds of risk factors. Consider, for example, the case of the RiskMetrics™ covariance matrix with over 400 risk factors (see RiskMetrics [17]).

We further discuss inverse portfolio optimization problems. "Implied views" allow portfolio and risk managers to determine whether a portfolio or a strategy is consistent with some existing views on future movements of all the relevant factors. This is particularly useful in many cases when these views are not, or can not be made, explicit in advance, or also when the market leads to some counter-intuitive conclusions. Other interesting inverse problems include the computation of risk-neutral probabilities from option prices, or estimation of zero curves from bonds and forward contracts.

The formulation of implied views in the context of Modern Portfolio Theory (MPT) and Markowitz mean-variance efficient frontiers (Markowitz [16]) has been the subject of previous work in the literature (cf. Litterman [14]). We present the extension of such analysis to a Scenario Optimization framework, which has useful applications for risk management, hedging and pricing. These problems arise naturally and their solutions are found directly from basic duality. Inverse theory can also be applied to obtain implied market parameters from traded instruments. See Dembo et al. [5,6] for discussions on inverse problems in portfolio replication. For a general discussion on estimating risk-neutral probabilities, see for example Derman et al. [8], and Rubinstein [18].

The rest of the paper is organized as follows. Section 2 presents the basic concepts and mathematical formulation of portfolio replication together with some example applications; section 3 introduces inverse problems, points to their scenario affirmation and shows how these concepts can be applied in practice; finally, section 4 presents some concluding remarks.

2. The portfolio replication framework

Consider a single period where only a finite number s of events (scenarios) can occur. Exactly one of these events will occur at the end of the period but, at the start of the period, there is uncertainty as to which one will occur. Furthermore, assume that an investor in the market has a perception of the likelihood p of these scenarios.

The objective is to replicate a *benchmark*, or *target*, portfolio at the end of the period, over all scenarios, with only a finite number n of financial instruments (assets) and generally accounting for transaction costs and liquidity. The term "replicate" is used loosely now, but its precise meaning will be evident below.

Let $q = (q_1, q_2, \ldots, q_n)^T$ be the known prices of the instruments at the start of the period; and D the s by n matrix that gives the values of each instrument in each scenario, i.e. each entry d_{ij} is the value of instrument j at the end of the period, if scenario i were to occur. Similarly, we denote the price of the benchmark at the beginning of the period by c, and its value at the end of the period by the price vector $\tau = (\tau_1, \tau_2, \ldots, \tau_s)^T$. A portfolio is characterized by the vector x, with components x_i denoting the amount it contains of each instrument.

Originally, each scenario may be a single realization of all the relevant risk factors in the market, but must be transformed to a realization of the values of all the instruments and the benchmark. In practice, this is accomplished with a risk engine that captures, through a set of pricing functions, the relationships between the underlying factors and the instruments. For example, in a standard enterprise-wide risk management (EWRM) application, the original risk factors may be the zero curves for every currency in the portfolio, FX spot rates, equity indices, and macro-economic factors. Thus, joint scenarios of all these factors are first created and, then, instruments and benchmark are priced consistently with the engine to generate the scenario values entering the optimization problem $\{d_{ij}\}$, $\{\tau_i\}$.

This framework has two very important characteristics. First, it describes observable market situations quite accurately; and second, since everything about a market is deterministic under a given scenario, it is a straightforward exercise to evaluate the behavior of a portfolio for each scenario. Furthermore, to make the problem operational, and perhaps also realistic, we have assumed that the set of possible future market events is finite and given. However, it is also possible to generalize the problem to the continuous case.

The objective of portfolio replication is to find a replicating portfolio x that behaves "identically" to the benchmark under all scenarios. Such a portfolio is called a *perfect replication*. The deviations from this perfect replication are referred to as *benchmark risk*.

The *tracking error* R of a portfolio is the norm of the difference between the portfolio and the benchmark:

$$R = \| Dx - \tau \|. \tag{1}$$

A portfolio with zero tracking error perfectly matches the benchmark under all scenarios. If a 1-norm is used, then the tracking error is the expectation of these differences under the measure p: $R = p^T | Dx - \tau |$. If an infinity norm is used, then it is

$$R_\infty = \max\{ | (Dx - \tau)_i |, \ i = 1, \ldots, s \}.$$

One useful measure of benchmark risk is given by *regret*: the expected underperformance from the benchmark. The regret R^- of a portfolio is

$$R^- = \| (Dx - \tau)^- \| = \| \min(0, Dx - \tau) \|. \tag{2}$$

Denote the vector of negative deviations from the benchmark by $y^- = (Dx - \tau)^- = \min(0, Dx - \tau)$ and the vector of positive deviations by $y^+ = (Dx - \tau)^+ = \max(0, Dx - \tau)$. Then, $R = p^T(y^+ + y^-)$ and $R^- = p^T y^-$.

In principle, we can use any weighted combination of upside and downside deviations to define benchmark risk as well. Here, we focus exclusively on (downside) regret as a measure of risk. The reader is also referred to Konno and Yamazaki [13] and Zenios and Kang [19] for applications of this type of risk measure in finance.

2.1. Mathematical programming formulation

We are ready to formulate the problem. Similar to the Markowitz framework, (Markowitz [16]), an investor is guided by the trade-off between risk, given in this case by R^-, and expected excess profit over the horizon. However, several differences must be stressed. First, risk is defined as downside risk with respect to a benchmark (which may also be a stochastic variable) as opposed to variance. Hence, we explicitly address more realistic, non-normal, and non-symmetric distributions. Second, we work in the space of profits and losses, as opposed to returns; this allows us to treat efficiently leveraged instruments and derivatives. Finally, by modeling explicitly risk factor distributions and using simulation, we effectively capture instruments' non-linearities, path dependency, expiry, as well as market jumps or specific extreme market conditions.

The efficient frontier can then be constructed using the parametric program

$$MR(k) = \min_x R^-(x) \quad \text{s.t. } E_p\{\Pi(x)\} \geq k; \text{ trading constraints,} \tag{3}$$

where $E_p\{\Pi(x)\}$ denotes the expected excess profit over the benchmark. The trading constraints are discussed below. If the value of the portfolio today equals that of the benchmark, the expected excess profit, expressed in time T value, is

$$E_p\{\Pi(x)\} = p^T(y^+ - y^-). \tag{4}$$

If the values of the portfolio and the benchmark are not equal, the expected excess profit is

$$E_p\{\Pi(x)\} = (c - q^T x) + W_T\, p^T(y^+ - y^-), \tag{5}$$

where the first term on the right represents the profits at the start of the period, and the second term the expected profit at the horizon; the weight of the future profits, W_T, reflects in general the discounting of future gains to today, or investor's preferences for realizing profits at different points in time.

When the 1-norm or the infinity norm are used, equation (3) leads to a linear parametric program that can be readily solved. In this case, the efficient frontier $MR(k) = f(k)$ is a piecewise linear increasing convex function.

If the efficient frontier passes through the origin (zero excess profit for zero regret), the market is said to be benchmark-complete. The market is complete when it is benchmark-complete for all possible benchmarks (when any pay-offs can be replicated). When one can obtain zero regret and still have a positive excess profit, then the benchmark can be arbitraged. Finally, when the frontier does not cross the horizontal axis, the market is incomplete, and there is some residual error.

Once the frontier is found, managers may use different ways to find the best portfolio for them given their risk aversion and goals. For this purpose, a "mean-regret" convex utility function can be used. Alternatively, one may also choose the portfolio with highest return to risk (ratio of excess profit over regret), the scenario equivalent to a Sharpe ratio.

The frontier can also be parameterized in a slightly different way by introducing a set of linear utility functions $U(x; \lambda) = E_p\{\Pi(x)\} - \lambda R^-(x)$, $\lambda \geq 0$. The condition of a positive λ ensures that the utility is decreasing in risk levels. To maximize this utility, we use the parametric program

$$U^*(x;\lambda) = \max_x E_p\{\Pi(x)\} - \lambda R^-(x), \quad \lambda \geq 0 \tag{6}$$
$$\text{s.t. trading constraints.}$$

The parametrization is now in terms of the *risk aversion parameter* λ, instead of the excess profit k. It can be shown through the simple application of Kuhn–Tucker conditions that the two formulations given by expressions (3) and (6) are equivalent; i.e. the set of efficient portfolios generated by both methods is the same (note that the program (6) arises from the Lagrangian of (3), when it is expressed as a maximization problem).

Substituting the expressions for regret and profit yields

$$U(\lambda) = c - q^T x + W_T p^T y^+ - \gamma p^T y^- \quad (\gamma = \lambda + W_T \geq W_T). \tag{7}$$

The efficient frontier is given by the solution to the parametric program

$$\begin{aligned} U^*(\gamma) - c = \max_{x, y^+, y^-} \quad & -q^T x + W_T p^T y^+ - \gamma p^T y^- \\ \text{s.t.} \quad & D x - y^+ + y^- = \tau, \\ & A_L x \leq b_L \quad \text{(liquidity constraints)}, \\ & x \text{ free}, \ y^+, y^- \geq 0. \end{aligned} \tag{8}$$

For completeness, we have introduced a set of general linear "liquidity" constraints explicitly in the formulation. These constraints could represent, for example:

- constraints on the amount one can trade in each instrument (i.e. $L \leq x \leq U$);
- group constraints (e.g. no more than P% in a given asset class, etc.);
- piecewise linear constraints in the amount one can trade for different prices;
- specific constraints representing a portfolio manager's preference or strategy;
- a constraint on the maximum cost of the replicating portfolio, $q^T x \leq c_o$.

In general, the program (8) will lead to arbitrage and an unbounded solution without the proper liquidity constraints, unless there is one scenario where all of the replicating instruments have a zero pay-off (see Dembo et al. [5]).

It is sometimes useful to place an additional set of constraints on the tracking errors y, thus, for example, guaranteeing that the risk never exceeds a certain amount under any circumstance (e.g. $y^- \leq Y^{\max}$).

In the presence of bid–ask spreads, one can further split the instruments into long and short instruments with prices $q_l \geq q_s$, resulting in the split of the decision vector x into a vector of long and short positions $x = (x_l, -x_s)^T$, x_l, $x_s \geq 0$. Furthermore, we can assess the impact of liquidity by comparing the frontiers generated with different liquidity and spread assumptions as shown in figure 1.

The dual formulation of (8) leads to very interesting interpretations. In particular, duality forms the basis for obtaining state price vectors (risk neutral probabilities) and pricing in incomplete markets and markets with frictions, see Dembo [4], Dembo et al. [5], and solving some interesting inverse problems (see section 3).

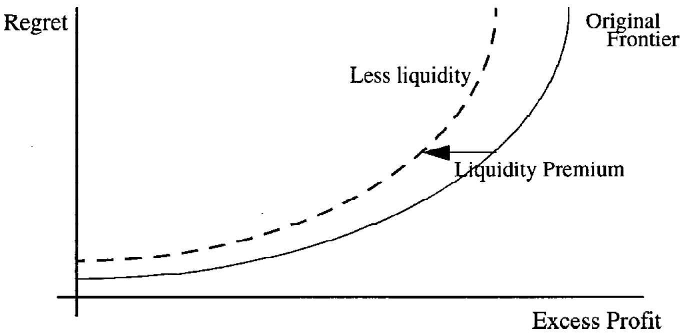

Figure 1. Effects of liquidity on the effiicient frontier.

The optimization problem finds the best possible portfolio that replicates one attribute of the benchmark (usually value, UP&L, CashFlow) at a single rollover date over all the scenarios. The formulation is readily generalized to find portfolios that replicate the benchmark in several attributes and at several points in time. The resulting program is a multi-objective optimization problem.

Consider a multi-attribute optimization problem with m attributes and T time steps t_i. Denote by R_{ij} ($i = 1,\ldots, m$; $j = 1,\ldots, T$) the portfolio regret in the ith attribute at the jth time step; i.e. R_{ij} gives the expected difference in the ith attribute between the portfolio and the benchmark at time t_j. Then, the optimization program can be expressed in terms of a measure of risk given by *the weighted sum of attribute-regrets*. We use the function

$$R = \sum_{i,j} w_{ij} R_{ij}, \tag{9}$$

where $R_{ij} = E_S|\{P(a_i, t_j) - \tau(a_i, t_j)\}_-|$, with $P(a_i, t_j)$ and $\tau(a_i, t_j)$ the values of attribute i for the portfolio and the benchmark, respectively, at the corresponding time step, and the w_{ij} are user-defined weights for each attribute. In general, for each attribute, we can choose separately whether all errors or only negative or positive errors (downside or upside regret) are to be used.

The weights should be chosen to represent individual preferences with respect to the errors in replicating each attribute at the set of time steps, and caution should be exercised to guarantee that these preferences are properly modeled. Only the *relative values* of the weights are important, since the same solution would be obtained if the weights were all scaled by a constant. In practice, special attention must be paid to the units of measurement of each attribute in order for the selection of weights to represent accurately one's preferences.

For example, consider the case where we are interested in tracking the value, duration and convexity of a bond portfolio (portfolio immunization over scenarios), and suppose we are equally interested in matching all attributes. Then, a sensible strategy is to choose *all the weights equal to one*, but to use *monetary duration* and *monetary convexity* as the tracking attributes; this ensures that the attributes are within the same range. Accordingly, a higher weight to the attribute "value", for example, places a higher priority for matching each unit of value. Alternatively, one may choose as relative weights (e.g. w_1/w_2) the ratio of average values for each attribute.

A *prioritized stratified strategy*, is one where one ensures that we first minimize the errors in one attribute before attempting to minimize those of a second attribute. It can be obtained by placing a *very large relative weighting* between the two attributes. This is also a good way to get "*better*" *optimal hedges* when *multiple optimal solutions* exist (e.g. if there are many more bonds to be traded than scenarios, there may be multiple possible hedges that match a target in value; one can then choose to optimize also for the monetary duration with a very small weight, say 0.001). In practice, the prioritized optimization problem, represented by weighting the second attribute by an infinitesimal non-archimidean ε, is more effectively solved, from a numerical perspective, in multiple stages.

Another example is the multi-step cash-flow tracking over scenarios (a scenario "tracking" portfolio dedication). In this case, a reasonable choice may be to assign as weights the average discount factor at each time step, thus minimizing the present value of the cashflow mismatches (or underperformance).

2.2. Examples

2.2.1. Static replication of a barrier option

Consider the replication of an "up-and-out" knock-out call option on an equity index. This option behaves like a standard call if the value of the index does not hit the barrier before maturity, but it pays zero otherwise.

The current level of the index is 900. The option is a 25-day call with strike at 880 and barrier at 930. The option volatility is set at 20%. The theoretical price of the option is $4.25. Figure 2 presents its profile over the following twelve days; it shows the mean value as well as both tails of the distribution at the 66% and 90% levels. Figure 3 gives the full distribution two days and twelve days in the future. For the simulation, we have assumed that the index follows a geometric Brownian motion with a zero drift. The returns of the underlying are normally distributed; however, the distribution of the Barrier option values is far from normal and very skewed.

Previous theoretical work has focussed on a static replication based on defining the best instruments that would theoretically replicate the target's pay-offs (cf. Carr and Chou [1], Derman et al. [7]). We attempt a different approach by using portfolio replication to find the best possible hedge with a pre-defined universe of traded instruments in the market. Also, we may devise the hedge to hold only for a limited period of time.

We perform a static replication using a set of 56 exchange traded, European options with maturities ranging from 28 to 370 days. We choose a horizon of 12 days. The scenario set comprises 50 Monte Carlo paths, plus three extreme scenarios which

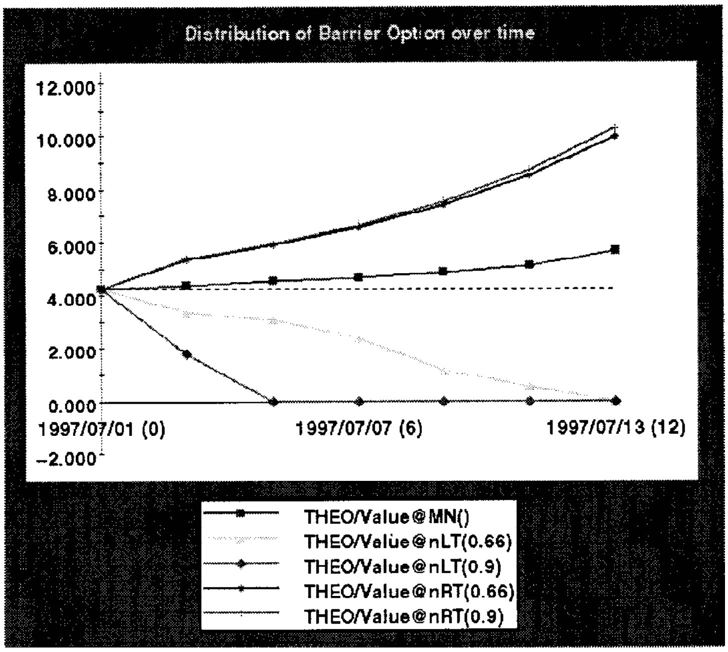

Figure 2. Simulated values of up-and-out option over twelve days.

represent a sudden drop and a sudden jump of about 10%, and a hit in the barrier; see figure 4. The extreme scenarios are weighted about five times higher than the other scenarios; thus, we want to buy some reasonable extra insurance to cover event risk. In practice, a trader would likely perform a more sophisticated joint simulation of both the equity index and a parametrized option volatility surface.

The multi-step portfolio replication model is used to create an efficient frontier that reflects the trade-offs between the cost and the quality of the replication (its regret). This is easily done by placing a zero weight to the future gains ($W_T = 0$). Therefore, the excess profit in this case simply gives the difference between the theoretical cost of the barrier ($4.25) and the cost of the replicating portfolio today. The results are given in figure 5, where the horizontal axis gives the difference in value between the replication and the theoretical cost of the barrier ($4.25); the vertical axis gives the regret.

The frontier shows that perfect replication is not possible and a residual regret of about 10 cents still remains even after spending $6.00 (–1.75 in the graph). This is independent of the bid–ask prices. The quality of the replication does not appreciably improve after spending $5 (–0.75 in the graph). In general, the replicating portfolios contain from 5 to 11 individual positions, out of the 56 options.

We test the robustness of the $5.00 replication over an out of sample simulation of 500 daily paths. The results for the hedge, given by a short position in the barrier

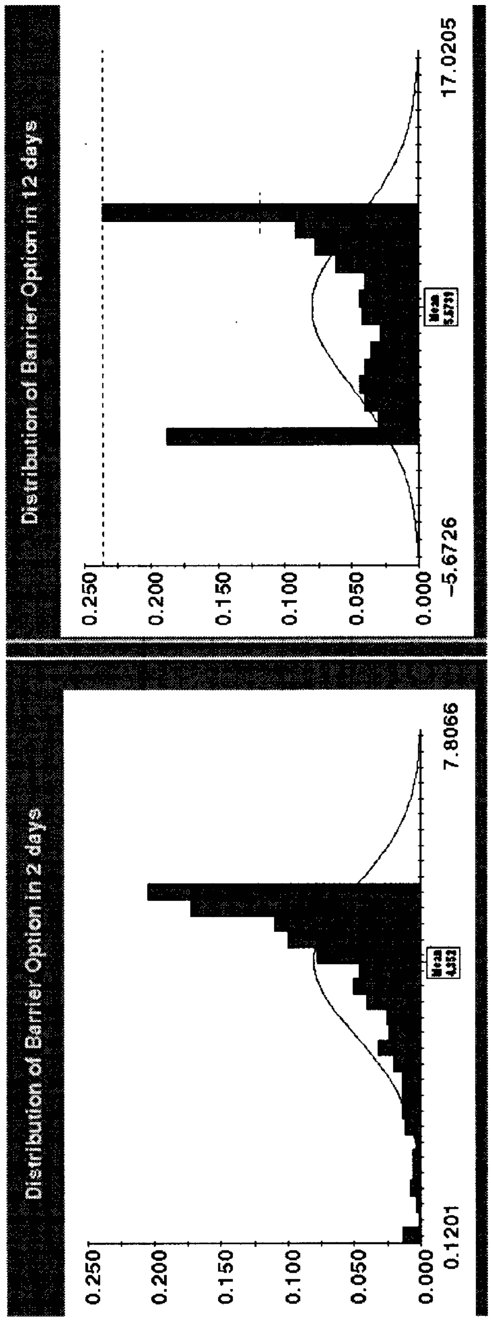

Figure 3. Distribution of knock-out option at two and twelve days.

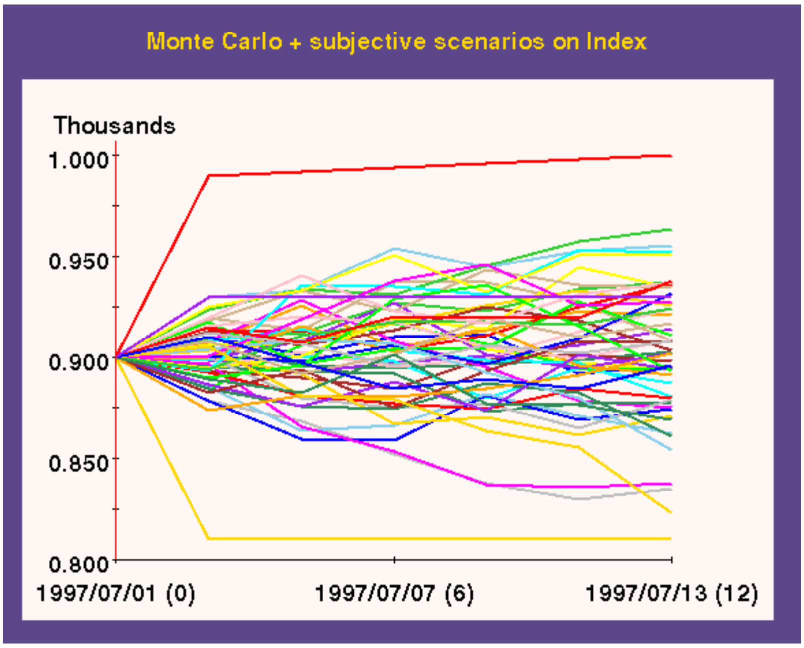

Figure 4. Scenarios for replication.

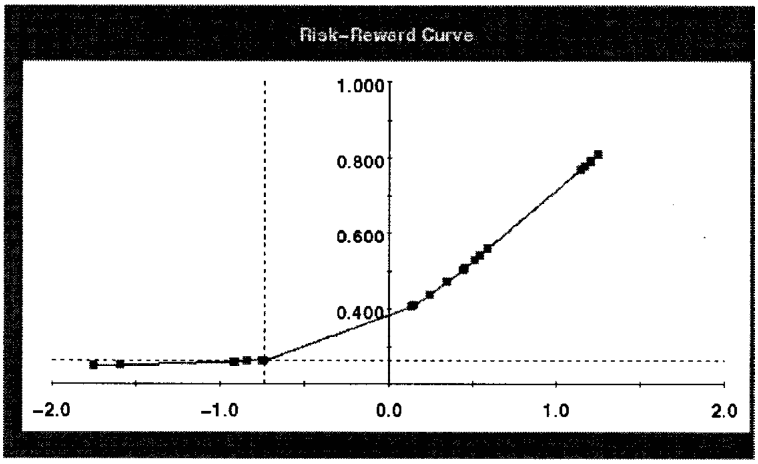

Figure 5. Cost vs. quality of hedge frontier for barrier option.

and a long position in the replicating portfolio, are presented in figure 6. The replication is very stable and never exceeds an error of more than 50 cents over the simulation period. Stress testing of the replication also shows the stability of the hedge over extreme conditions; this was of course expected as a result of the high weights given to the extreme scenarios.

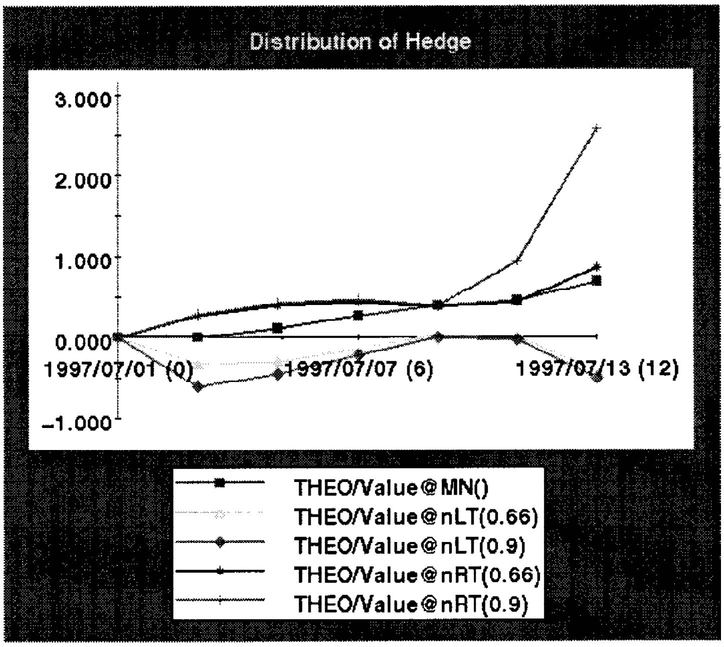

Figure 6. Value of the hedge over out of sample simulation.

In this example, we have used random Monte Carlo scenarios. Other advanced scenario techniques, such as quasi-Monte Carlo methods as described by Joy et al. [12] and stratified sampling as described by Jamshidian and Zhu [11], have been shown to be very powerful for this purpose. These methods generate scenarios that tend to "cover" the risk factor space more uniformly, thus leading to robust estimations.

2.2.2. Risk–reward analysis and enterprise capital allocation

As a second example, we use the portfolio replication model for the capital allocation of a US-based global portfolio with trading desks in New York, Toronto, London, Frankfurt, Paris, Tokyo and Singapore. The portfolio contains about 1000 fixed income, foreign exchange and equity positions with a high concentration in derivatives. The current portfolio weights are

New York	London	Frankfurt	Tokyo	Paris	Toronto	Singapore
59.8%	20.4%	10.8%	7.8%	1.0%	0.1%	0.1%

We use a single-period joint simulation of 9 interest rate curves, 2 equity indices, and 6 FX spot rates over one month. We assume that all risk factors follow a joint log-normal distribution with the RiskMetrics covariance matrix as of 1997/07/01. This matrix is estimated as described in RiskMetrics [17]. For capital allocation purposes, we only allow for assigning the weights in each desk, although the simulation is done for every instrument held in the portfolio.

For simplicity, we use as a benchmark a portfolio gaining the risk-free rate, and we have normalized both the benchmark and the current portfolios to be worth \$100,000. In practice, a more sophisticated benchmark representing the institutions goals must be used.

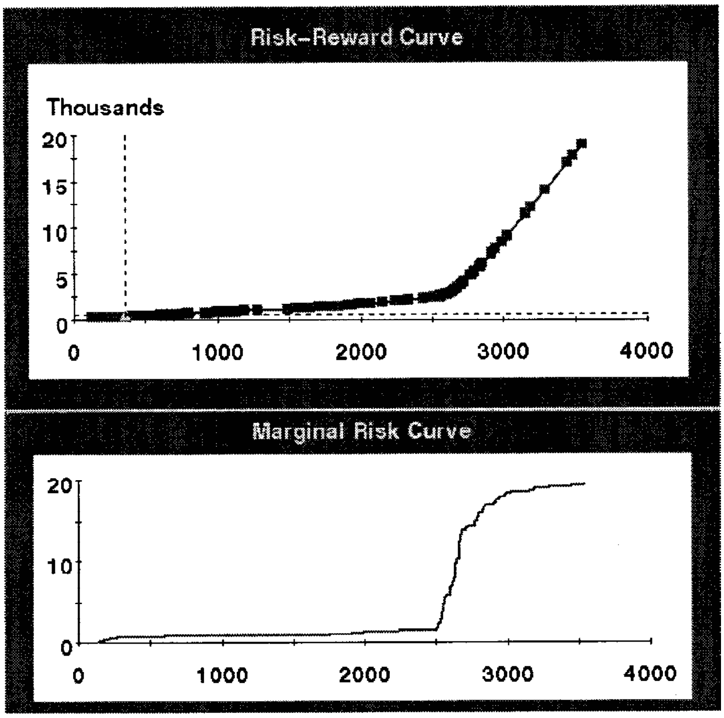

Figure 7. Risk–reward frontier and marginal risk curve for asset allocation.

The risk–reward frontier and the marginal risk curve are given in figure 7. The horizontal axis gives the excess profit and the vertical axis the regret of the portfolio. The marginal risk curve gives the additional regret that one must take to gain an extra dollar of profit; i.e. the slope of the frontier. The sudden change in slope at around 2500 shows that the additional risk one must take to gain one extra dollar of profit jumps from about \$3 to almost \$20. This of course, may be a natural point for investment.

Table 1 shows a summary of the efficient allocations suggested by the model. The current portfolio is close to the efficient frontier but still shows some slight inefficiencies. The portfolio called “projection” is the vertical “radial” projection of the current portfolio on the frontier; i.e an efficient portfolio with the same level of

Table 1

Efficient portfolio allocations.

Desk	Current (%)	Projection (%)	Low risk (%)	Medium risk (%)	Inflection (%)	High risk (%)
Toronto	0.1	0.00	0.0	0.0	0.0	0.0
Singapore	0.1	0.00	0.0	0.0	0.0	0.0
New York	59.8	78.9	93.1	26.6	0.0	0.0
Frankfurt	10.8	6.5	0.0	33.7	5.4	0.0
Paris	1.0	0.4	0.0	2.3	4.8	93.7
Tokyo	7.8	10.1	4.0	28.8	89.8	6.3
London	20.4	4.0	2.9	8.6	0.0	0.0

expected excess profit but with a lower regret (similarly we can get the horizontal projection). It indicates a higher weight on the NY portfolio (which contains a large position in US government bonds), and a slightly higher position on the Tokyo portfolio. This is done at the expense of reducing the positions primarily in London and, less so, in Paris and Frankfurt. The portfolio with lowest regret has a large position in NY, with a large investment in US bonds. The medium risk portfolio points to a more balanced portfolio across the main desks, thus allowing for a more leveraged position. A portfolio with high risk, and excess profit, is mostly weighted in Paris, where highly leveraged positions are held. The "inflection" portfolio has a large weight in Tokyo, which holds a substantial position on equity options with considerable volatility. Shifting the weights from Tokyo to Paris leads to low marginal profits for the risk one must take.

3. Implied views and inverse problems

In general, a *forward portfolio optimization problem* is formulated as follows: given a distribution of market factors (yields, prices, etc.) in the future, some market prices and liquidity constraints, find the set of positions (or weights) in a given, finite, set of instruments (assets) that maximizes (or minimizes) some given objective function. The optimality criteria may be, for example, risk minimization, profit maximization or, more generally, a utility function that combines risk and profit in a meaningful way.

The *inverse portfolio optimization problem* is obtained when we reverse what is known and unknown at the beginning of the problem: "given a chosen portfolio, find the market conditions, its *implied views*, that would make it optimal; i.e. find those conditions under which no other (feasible) portfolio would perform better".

These views are only implied by the portfolio under some behavioral assumptions: e.g. utility maximization. Then, whoever is holding the portfolio prefers more returns and less risk (and weights their trade-offs in a given way). Implied views allow managers to determine whether a portfolio or a strategy is consistent with some existing views on future movements of the relevant factors. In most cases these views are not, or can not be made, explicit in advance.

Inverse problems, in general, do not present a unique solution (cf. Huestis [10]). The goal is the recovery of useful information about the solution while recognizing the fundamental non-uniqueness of the problem. A common example of this non-uniqueness in finance arises when estimating the implied risk neutral distribution (or state prices) from the set of trade options in an incomplete market (cf. Duffie [9]).

In this section, we discuss portfolio implied views in the context of portfolio replication. We refer the reader to Dembo et al. [6] and Litterman[14] for its application to MPT.

The forward portfolio optimization problem introduced in the previous section requires as inputs

- the set of scenarios, and their probabilities (the measure over the set),
- the prices of the instruments and target today,
- the risk aversion parameter λ (or alternatively the desired excess profit K),
- assumptions on liquidity,

and obtains as output a portfolio x that is optimal (or a set of portfolios $x(k)$ or $x(\lambda)$, depending on the parameterization). The most natural inverse problem is to find the *scenario implied views* assuming liquidity and everything else constant. However, in principle, implied views can also be obtained for liquidity, risk aversion, current prices, etc.

The scenario implied views inverse problem can be formulated formally as follows:

> Given a portfolio x_o, find a measure p_o (a set of scenario probabilities), for which x_o is optimal; i.e. where
>
> $$E_{p_o}\{\Pi(x_o)\} - \lambda R^-_{p_o}(x_o) \geq E_{p_o}\{\Pi(x)\} - \lambda R^-_{p_o}(x), \qquad x \quad X, \tag{10}$$
>
> with X the set of feasible portfolios (satisfying liquidity and other constraints).

The subscript p_o denotes what expectation and regret are with respect to this measure.

It is also possible to formulate the problem in terms of near optimality of the portfolio as opposed to full optimality; i.e. to find the set of measures for which

$$E_{p_o}\{\Pi(x_o) - \Pi(x)\} - \lambda[R^-_{p_o}(x_o) - R^-_{p_o}(x)] \leq \varepsilon, \tag{11}$$

with ε usually a small number giving the maximum acceptable difference in utility between the optimal portfolio and the current portfolio.

The solution to (10) is easy to find using linear programming principles and duality theory. The final form of the solution, given by a set of equations, and its derivation are further given in Dembo et al. [6]. The main characteristics are:

- If everything else is fixed, the set of implied measures is given by the solution of a set of linear equations.
- Solutions may be non-unique and hence proper criteria are required to obtain useful solutions and represent the set in a practical way; this is further discussed below.
- Solutions in terms of implied probabilities of scenarios are useful for small (subjective) scenario sets and few factors. However, in many applications, the solution may be more useful if condensed in the proper way (e.g. showing the densities for an asset class, or expressing the solution in terms of volatilities, drifts, correlations of risk factors or asset returns). This is necessarily the case when Monte Carlo scenarios are used.
- Implied simultaneous views on p and λ require the solution of nonlinear equations that can be solved readily.

Since the inverse problem, in general, does not present a unique solution, we must impose some a priori restrictions on the solution to recover useful information for the risk manager. Thus, we look for solutions that optimize various functionals, subject to the data constraints. We summarize below some useful criteria:

- Closest to prior solutions; prior distributions may be the *real* underlying distribution or any other desired benchmark distribution. In general, a closest to prior optimization strategy presents a conservative picture of the deviations from our prior, and probably best, guesses, and can be expressed directly on the probabilities or any other distribution parameters.
- Maximum deviations from priors: while the closest to prior solutions reflect minimum deviation solutions, we may also want to find upper bounds for the deviations.
- Minimum and maximum parameter values: we only search for some form of absolute bounds on the implied views.

3.1. *Example*

Consider the case of a Canadian-based manager holding a portfolio with positions in equities and equity derivatives worldwide, as well as Government of Canada and US Treasury Bonds. Due to the option content, such a portfolio may be highly nonlinear. Figure 8 presents a report with the implied views of the mean portfolio returns. These are the closest to prior views, thus giving the smallest differences. The scenario set comprises a Monte Carlo simulation of the relevant risk factors over the following quarter, with several stress scenarios properly weighted. The period is in 1997, before the Asian crisis.

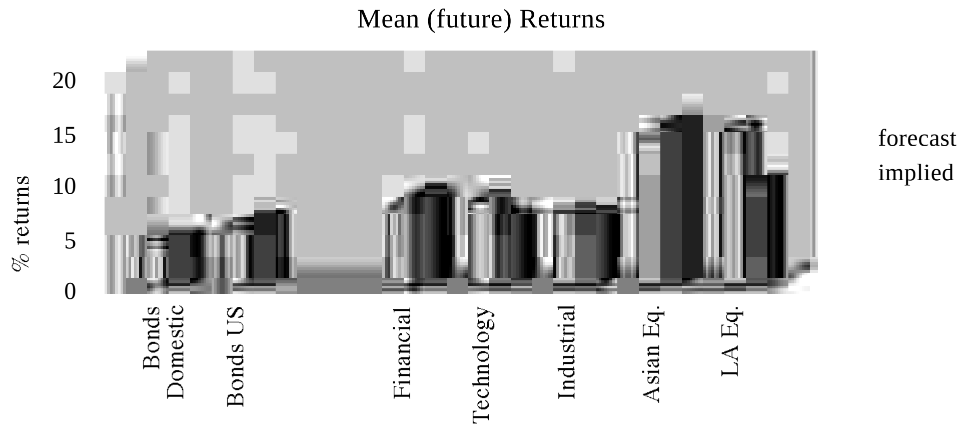

Figure 8. Implied views report of a Canadian asset manager.

The report shows, in particular, a noticeable discrepancy in the Latin American portfolio, with an over 4% difference. This means that the portfolio would be optimal only if the mean returns of the Latin American portfolio were 4% lower than current forecasts and simulations would show. This can be caused by a true difference between the trader's and the manager's expectations (drifts), by the effects of negative correlations or nonlinearity in the simulation. A similar report can be generated in this case for the variances and higher moments to assess the implied risks.

4. Concluding remarks

Portfolio replication is a powerful tool that has proven in practice its applicability to financial problems such as static hedging in complete and incomplete markets and markets that gap; strategic asset and capital allocation; benchmark tracking; design of synthetic products; and portfolio compression. In contrast to mean-variance optimization, the scenario approach allows for general non-normal, discrete and subjective distributions, as well as for the accurate modeling of the full range of nonlinear instruments, such as options. It also provides an intuitive, operational framework for explaining basic financial theory. We further show how inverse problems in finance can be naturally formulated in this framework. From the technical perspective, the formulation and solution of these problems arises directly from basic duality results in mathematical programming.

Acknowledgements

We would like to thank Andrew Aziz for many fruitful discussions and valuable financial modeling in the examples. We are also thankful to Michael Zerbs and Judy Farvolden for their useful comments.

References

[1] P. Carr and A. Chou, Braking barriers, Risk 10(9)(1997)139–146.
[2] D.R. Carino and W.T. Ziemba, Formulation of the Russell–Yasuda Kasai financial planning model, Operations Research 46(4)(1998).
[3] R. Dembo, Scenario optimization, Annals of Operations Research 30(1991)63–80.
[4] R. Dembo, Optimal portfolio replication, Algorithmics Technical Paper series 95-01.
[5] R. Dembo, D. Rosen and D. Saunders, Illiquid neutrality: On implied risk neutral probabilities with transaction costs and liquidity constraints, Algorithmics Working Paper, 1997.
[6] R. Dembo, L. Merkoulovitch and D. Rosen, Images of a portfolio, Algorithmics Technical Paper series 97-02, 1997.
[7] E. Derman, D. Ergener and I. Kani, Static option replication, Goldman Sachs Quantitative Strategies Research Notes, May 1994.
[8] E. Derman and C.N. Kani, Implied trinomial trees of the volatility smile, The Journal of Derivatives (Summer 1996).
[9] D. Duffie, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, 1992.
[10] S.P. Huestis, The use of linear programming in the construction of extremal solutions to linear inverse problems, Classroom Notes, SIAM Review 38 (1996)496–506.
[11] F. Jamshidian and Y. Zhu, Scenario simulation: Theory and methodology, Finance and Stochastics 1(1997)43–67.
[12] C. Joy, K.S. Tan and P.P. Boyle, Quasi-Monte Carlo methods in numerical finance, Management Science 42(1996)926–938.
[13] H. Konno and H. Yamazaki, Mean absolute deviation portfolio optimization model and its applications to Tokyo stock market, Management Science 37(1991)519–531.
[14] R. Litterman, Hot spots and hedges, The Journal of Portfolio Management, Special Issue (1996).
[15] A. MacKay and E.Z. Pressman, Estimating valuator operators in incomplete markets with noises, Working Paper, University of York, 1996.
[16] H.M. Markowitz, Portfolio selection, Journal of Finance 7(1952)77–91.
[17] RiskMetrics^{TM}, *Technical Document*, J.P. Morgan Guaranty Trust Company, New York, 3rd edition, 1995.
[18] M. Rubinstein, Implied binomial trees, Journal of Finance 69(3)(1994).
[19] S.A. Zenios and P. Kang, Mean-absolute deviation portfolio optimization for mortgage-backed securities, Annals of Operations Research 45(1993)443–450.
[20] W.T. Ziemba and J.M. Mulvey (eds.), *Worldwide Asset and Liability Modeling*, Cambridge University Press, 1998.