Centre for Computational Finance and Economic Agents

WP014-07

Working Paper Series

Dietmar Maringer

Risk Preferences and Loss Aversion in Portfolio Optimization

November 2007

University of Essex

CCFEA

www.ccfea.net

Risk Preferences and Loss Aversion in Portfolio Optimization

Dietmar Maringer*

November 2007

Abstract

Traditionally, portfolio optimization is associated with finding the ideal trade-off between return and risk by maximizing the expected utility. Investor's preferences are commonly assumed to follow a quadratic or power utility function, and asset returns are often assumed to follow a Gaussian distribution. Investment analysis has therefore long been focusing on the first two moments of the distribution, mean and variance. However, empirical asset returns are not normally distributed, and neither utility function captures investors' true attitudes towards losses. This contribution investigates the impact of these specification errors under realistic assumptions. As traditional optimization techniques cannot deal reliably with the extended problem, we suggest the use of a heuristic approach. We find that loss aversion has a substantial impact on what investors consider to be an efficient portfolio and that mean-variance analysis alone can be utterly misguiding.

Keywords. Loss aversion, risk aversion, portfolio optimization, heuristics, differential evolution.

*Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. `dmaring@essex.ac.uk`

1 Assets and Asset Selection

1.1 Properties of Asset Returns

Portfolio optimization and understanding of investment decisions have come a long way over the last decades. Following the seminal work of Markowitz (1952), the standard paradigm in financial modeling has long been that the first two moments, mean and variance, are sufficient to describe asset returns and capture the preferences of a rational risk averse investor. In this framework the investment decision is based on the trade-off between higher mean and higher variance of the returns. The underlying assumption for this approach is that either the returns are normally distributed, or that investors have a quadratic utility function. Unfortunately, neither seems to hold for the real world.

One of the salient properties of normal distribution is that any observation from minus to plus infinity is assigned a positive probability. Negative log returns up to minus infinity require that losses up to the amount of initial investment can occur, but not more so. Hence it does apply, e.g., to bonds, stocks, or options when held in a long position, but not to futures or stocks and options when held in a short position. Arbitrarily large positive returns can be achieved for long positions when prices can increase without any limit. This is true (at least in principle) for assets such as stocks, futures or call options. Put options already have a theoretical upper limit on their value (though this bound is unlikely to be reached anyway unless the underlying defaults). For assets such as bonds, however, there is a much clearer upper bound on potential profits: a long position in a simple bond will earn the yield to maturity if it does not default or less if it does; additional deviations due to changes in the general interest rate structure might cause slight, but no dramatic deviations from this. Hence, the assumption of symmetric deviations from the expected value which make positive and negative deviations of the same magnitude equally likely, will by definition apply only to a small set of assets such as stocks. And even for those assets, skewed empirical returns are not uncommon.

The even bigger problem with the normality assumption is that is does not capture extreme events appropriately. Looking at the Dow Jones Industrial Average's daily log returns, there were two days with losses bigger than 5% within the 1684 days considered for the computational study in this chapter; for the observed mean and variance, the normal distribution predicts that it ought to take about 1.5 million days for these two extreme events to happen. For individual stocks, excess kurtosis and the gaps between predicted and actually observed frequencies are typically even larger. Hence, looking only at the first two moments of return distributions leaves out crucial information. As a consequence, other approaches have become more popular, such as including student's t-distribution (which, in particular with low degrees of freedom, does exhibit some excess kurtosis) or, more recently, extreme value theory (Gilli and Këllezi (2006)); alternatively, empirical distributions and historical sampling methods are widely used and sanctioned by regulatory authorities as are alternative risk measures (Basel Committee on Banking Supervision (2003), Gilli, Këllezi and Hysi (2006)).

1.2 Investor Preferences

At the same time, the usual assumptions about risk preferences imply that investors actually might prefer deviations from a normal distribution. A common assumption in modeling an investor's preferences is diminishing marginal utility of wealth; this implies that larger deviations on the positive side are required to offset her for losses. In other words, investors are assumed to prefer positively skewed returns. For the same reason, variance and excess kurtosis are not desirable: both measure deviations from the mean but ignore the sign; with marginally diminishing utility, losses lower the utility more than profits of the same magnitude would increase it. As a consequence, higher variance and kurtosis, respectively, are accepted only if they are rewarded with an increase in the mean payoff. All things considered, the representative risk averse rational investor should find an investment that optimizes the trade-off between its expected return and expected deviations from it.

More recent work in behavioral finance suggests that utility functions should also capture a more human property: investors do not only dislike losses, sometimes they also put more emphasis on it than pure statistics or a traditional framework of rational investors would suggest. In their seminal studies, Kahneman and Tversky (1979) found that losses play a more prominent part in making decisions than traditional utility analysis would predict. Hence, decisions are merely driven by loss aversion and the prospect of ending up with a lower than the current wealth, and not just by risk aversion which captures any deviation – positive and negative alike – from the expected wealth. Prospect theory, which arose from these findings, assumes that investors put additional weight on losses, meaning that they either overestimate the likelihood or the magnitude of losses. Empirical work found these effects not only for ex ante decision problems (Thaler, Tversky, Kahneman and Schwartz (1997)), but also for dynamic portfolio revisions where investors tend to sell winners and to keep losers (Odean (1998)). These findings have also coincided with (and, arguably, supported) the advent of new risk measures that measure only losses, such as Value at Risk and conditional Value at Risk, or take the asymmetries between profits and losses into account, such as the Omega risk measure (Gilli et al. (2006)).

1.3 Consequences for Asset Selection

If neither the asset returns nor the decision maker's preferences agree with the assumptions of traditional optimization problems, disagreement on what is an optimal solution can be expected. Behavioral finance has made progress in identifying more realistic models of preference and choice, there exists little evidence of how this would (or should) affect the actual investment selection process in a realistic market. This contribution aims to answer to this question by analyzing the consequences from deviating preferences and properties. Based on an empirical computational study for Dow Jones stock data, it is investigated how the choice of utility function, level of risk aversion and, in addition, level of loss aversion affects the investment process. The results show that these differences in preferences not only lead to different ideas of "efficient" portfolios, but also that investors' portfolios will differ substantially in their stylized facts and properties.

The remainder of this chapter is organized as follows. Section 2 presents the financial background in more and formalizes the optimization problems, Section 3 offers a method that can solve them. Section 4 presents an empirical study, and Section 5 concludes.

2 Portfolio Optimization under Loss Aversion

2.1 Loss aversion, Risk Aversion and Utility Analysis

In myopic portfolio optimization, a popular choice for modeling investors' preferences is the quadratic utility function,

$$\mathcal{U}_q(w) = w - \frac{b}{2}w^2, \quad 0 < b < 1/w \tag{1}$$

where w is the wealth and b is the level of risk aversion. With this function, higher moments are either ignored or assumed to not exist – as is the case under a normal distribution. One peculiarity of this function is that it is not defined for wealth levels $w < 1/b$: the function's first derivative, i.e., the marginal utility of wealth, would become negative which clearly contradicts the usual assumption of non-satiation and investors preferring more wealth to less. A simple remedy is to assume zero marginal utility for wealth beyond this bliss point and set the utility to $\mathcal{U}_q(w) = 1/(2b)\ \forall\, w \geq 1/b$.

Alternatively, utility functions such as the power utility

$$\mathcal{U}_p(w) = \left(w^{1-\gamma}\right)\cdot(1-\gamma)^{-1}, \quad \gamma > 0 \tag{2}$$

implicitly do take higher moments into account. $(1-\gamma)$ measures the risk aversion, with γ also being the constant coefficient of relative risk aversion, $-w \cdot \mathcal{U}_p''(w)/\mathcal{U}'(w)$. For $\gamma = 1$, the function turns into the log utility function (or Kelly criterion) where the log of the wealth, $\mathcal{U}_l(w) = \ln(w)$, is considered.

However, individuals tend to be more sensitive to falling below a certain level of wealth than exceeding it (Samuelson (1963)). Typically, this reference point is the current level of wealth, w_0, which makes the decision maker loss averse. A simple way to account for this phenomenon is to use modified future levels of wealth, m, that enhance the losses:

$$\begin{aligned} m &= w_0 + \underbrace{(w - w_0)\Im_{w \geq w_0}}_{\text{profit}} + \lambda \underbrace{(w - w_0)\Im_{w < w_0}}_{\text{loss}} \\ &= w + (\lambda - 1)(w - w_0)\Im_{w < w_0} \end{aligned}$$

where $\lambda \geq 1$ is the degree of loss aversion; the indicator function $\Im_{w<w_0}$ ($\Im_{w \geq w_0}$) is 1 (0) if a loss is occurred and $w < w_0$, and 0 (1) otherwise. When using the quadratic or power utility function, the utility curves remain concave, but have a kink at w_0 and are steeper left of it than their "traditional" counterpart. Alternatively, the levels of wealth can be left unchanged but the probabilities for incurring losses and profits are altered. In either way, the left-hand part of the curve is over-emphasized when the expected utility is computed.

In their seminal experiments, Kahneman and Tversky (1979) (KT) find that individuals are not only loss averse, but that the preference actually form an S-shaped value curve of changes in wealth, $x = w - w_0$,

$$v(x) = \begin{cases} x^{\alpha} & x \geq 0 \\ -\lambda(-x)^{\beta} & x < 0 \end{cases} \tag{3}$$

Furthermore, they assume decision makers use decision weights, π_i, which are nonlinear transforms of the probabilities of outcome i, p_i. According to their *prospect theory*, the prospective utility of a gamble G can then be modeled as

$$V(G) = \sum_i \pi_i v(x_i). \tag{4}$$

They find for their experiments that the parameters for the value function are $\alpha = \beta = 0.88$ and $\lambda = 2.25$.

2.2 Operational Aspects of Portfolio Optimization

Investors are assumed to maximize their expected utility. In perfect markets with well behaved return distributions, investors prefer a combination of a (universal) risky portfolio with a risk free asset. Depending on their level of risk aversion, they choose different fractions of wealth to invest in either of these (Tobin (1958)). In the absence of a risk free asset, however, and with distributions that do not follow the usual parametric distributions, solutions become less straightforward. The problem becomes even more demanding when practical constraints are considered. While aspects such as transaction costs, taxes, liquidity risk, etc. can be approximated by modifying the (distribution of) returns, constraints on the asset weights (i.e., the assets' fractions within the risky portfolio) can cause severe challenges to numerical approaches. Not uncommonly, weights for individual assets can have upper limits, $x_i \leq x^u$ in order to encourage diversification and reduce excessive exposure to a single asset's risk. At the same time, there can also be minimum weights for included assets, $x_i = 0 \vee x_i \geq x^\ell$, to avoid over-fragmentation and keep costs to a reasonable level.

In the presence of such weight constraints, the portfolio optimization problem for a maximizer of expected utility can be stated as follows:

$$\max_{\mathbf{x}} E(\mathcal{U}(w, \lambda)) \tag{5}$$

subject to

$$
\begin{aligned}
& w_i = \sum_{a=1}^{N} w_0 \exp(r_{a,i}) x_a \\
& m = w + (\lambda - 1)(w - w_0)\Im_{w<w_0} \qquad (6) \\
& E(\mathcal{U}(w,\lambda)) = \sum_{i=1}^{I} p_i \mathcal{U}(w_i,\lambda) \\
& \mathcal{U}(w,\lambda) = \begin{cases} \text{quadratic utility:} & \mathcal{U}_q(m) = \begin{cases} m - \frac{b}{2}m^2 & b < m/2 \\ m/2 & \text{otherwise} \end{cases} \\ \text{power utility:} & \mathcal{U}_p(m) = \begin{cases} \frac{m^{1-\gamma}}{1-\gamma} & \gamma \neq 1 \\ \ln(m) & \text{otherwise} \end{cases} \\ \text{prospect theory:} & \mathcal{U}_{PT}(w) = (w-w_0)^{\alpha}\Im_{w\geq w_0} - \lambda(w_0 - w)^{\beta}\Im_{w<w_0} \end{cases} \\
& \sum_i x_i = 1 \\
& x_i : \begin{cases} x_i = 0 \\ x^{\ell} \leqslant x_i \leqslant x^u \end{cases} \qquad \forall i \\
& \Im_{w<w_0} = \begin{cases} 1 & w < w_0 \\ 0 & \text{otherwise} \end{cases}, \quad \Im_{w\geq w_0} = 1 - \Im_{w<w_0}
\end{aligned}
$$

As with many financial optimization problems, this problem, too, is non-convex, and standard non-linear optimization methods cannot be guaranteed to find the global optimum. Hence, alternative methods such as heuristics or non-deterministic approaches have gained popularity since they can deal with non-convexity and discontinuities.

3 Heuristic Methods for Portfolio Optimization

3.1 A Brief History of Heuristics in Finance

Heuristic methods have become an increasingly popular alternative to traditional optimization methods. One salient feature of these methods is the inclusion of non-deterministic elements, another the occasional acceptance of impairments (despite the obvious preference for improvements). Because of these two principles, the search process is no longer deterministic – which can be very beneficial: local optima can be overcome more easily. Also, restarts do not necessarily produce the same reported result; if the search converges to an inferior solution once, another run can identify another optimum – ideally the global one. Restarts will therefore increase the likelihood of identifying the true optimum, which is not the case for traditional deterministic search methods. Furthermore, heuristics are often designed as general principles (meta-heuristics), that can easily be adapted to all sorts of problems and constraints. An introduction to the use of heuristics in finance can be found in Gilli, Maringer and Winker (forthcoming). For applications of deterministic methods, see, e.g., Konno (2005).

A common traditional approach is to find a suitable trajectory through the search space that leads to the optimum. Gradient search, e.g., takes its steps from first order conditions and, in the case of a maximization problem, moves the current solution towards the steepest ascend. This simple deterministic rule, however, has difficulties coping with rough search spaces that have several optima; often, the search ends in the local optimum closest to the starting point which is not necessarily the global optimum. A simple yet effective remedy is the inclusion of noise: adding noise with expected value zero to the gradient will lead to diversions in the trajectory, and local optima can be circumvent and left. Parpas and Rustem (2006) and Maringer and Parpas (forthcoming) apply this approach to a portfolio selection problem, where updates of the current solutions contain deterministic (i.e., the gradient) and non-deterministic (i.e., noise) elements and where the magnitude of noise is gradually reduced and replacement decisions are based on improvements only.

Introduced by Kirkpatrick, Gelatt and Vecchi (1983), simulated annealing (SA) is a heuristic that abandons the deterministic ingredient all together. Suggested changes are randomly picked based on a neighborhood definition (i.e., adding noise to the current solution), and acceptance is based on a stochastic rule. This rule takes into account the magnitude of the change in the objective function and how progressed the search is: Other things being equal, larger impairments are less likely to be accepted than smaller ones or improvements. The more iterations have passed, the algorithm should have moved to a favorable region of the search space (due to its preference for improvements over impairments), and it becomes less tolerant in accepting impairments. Hence, in the course of multiple iterations the search process becomes more and more like a greedy uphill search.

Dueck and Scheuer (1990) suggest threshold acceptance (TA) where SA's random generation of new solutions is kept, but the stochastic acceptance principle is replaced with a deterministic rule and any improvement is accepted as well as any impairment that does not exceed a given threshold. This threshold is lowered in the course of iterations which, again, makes it more intolerant to impairments. More on TA and applications to economics and econometrics can be found in Winker (2001) and Winker and Maringer (2007); applications of SA and TA in finance are presented in Maringer (2005).

TA has been successfully applied to different portfolio optimization problems. For example, Gilli and Këllezi (2002) show how to use it for index tracking problems with demanding constraints, and Gilli et al. (2006) offer a significant and helpful extension to the usability and practical application of TA by suggesting a data driven approach for tuning one of the ingredients, the threshold sequence, and showing how to tackle investment decisions under different downside risk measures. In these contributions, Gilli and colleagues managed not only successfully answer challenging and demanding problems that would be unsolvable with traditional methods; they also helped to establish the use of heuristics in finance as well as econometrics.

An alternative class of (meta-) heuristics is based on evolutionary principle and natural evolvement. Evolutionary strategies (Rechenberg (1965, 1973)) and the more popular genetic algorithms (Holland (1975)) are typical and widely used examples. The main principles are mimicking natural processes: New solutions are produced by slightly modifying existing ones ("mutation") and, in the case of "populations" of candidate solutions, combining them into new ones ("cross-over"). For both approaches, randomness is the key element: which part of the solution is altered to which extent is decided non-deterministically, and so is which "parent" solutions contributes what to their offspring's. Following the "survival of the fittest principles," new solutions will replace current ones if they outperform them – yet this decision, too, can be subject to a tournament which adds another stochastic element to the process. Fogel (2001) offers a detailed account of evolutionary methods and general applications, while financial applications of evolutionary methods can be found, e.g., in Brabazon and O'Neill (2006) or Maringer (2005).

3.2 Differential Evolution

A recent addition to the class of evolutionary heuristics is a method called differential evolution (DE). Suggested by Storn and Price (1995, 1997), DE uses a population of P vectors, $\mathbf{v}_p$, $p = 1 \ldots P$, where the N real valued elements of the vectors represent the objective variables. The basic idea is to produce a new solution for each current vector $\mathbf{v}_{p_0}$, where the new solution is a combination of four distinct current solutions. This involves the following steps: First, three different vectors are randomly chosen from the current population. One vector, $\mathbf{v}_{p_1}$, is used as the base vector to which the weighted difference of two other vectors $\mathbf{v}_{p_2}$ and $\mathbf{v}_{p_3}$ is added. The combined solution is then $\mathbf{v}_k = \mathbf{v}_{p_1} + F \cdot (\mathbf{v}_{p_2} - \mathbf{v}_{p_3})$. Finally, this combined solution is crossed-over with a fourth solution, $\mathbf{v}_{p_0}$. This procedure of producing a new solution is repeated for each member p_0 of the current population. Once all the crossed-over solutions have been generated, they replace their respective parent, $\mathbf{v}_{p_0}$, if the objective value is better.

As long as different vectors do not "agree" on the values of the elements, the difference vector $(\mathbf{v}_{p_2} - \mathbf{v}_{p_3})$ will have non-zero elements, and genuinely new solutions will be produced. If, however, one or more of the elements of the difference vector are small or even zero, the new solution will inherit the (more or less) unchanged corresponding values of the base vector. The new solutions will then move towards what they consider the best values for the elements; these reference points can change in the course of multiple iterations as the solutions are steadily improved. Eventually, they will agree, however, and (graphically speaking) flock at or around a point in the solution space which they consider the global optimum.

Diversity within the population is salient to avoid premature convergence. For example, additional perturbation can be achieved by adding some noise to the difference vector, by adding a second difference vector, etc. Alternatively, to reinforce good solutions, the best solution found so far, the so-called elitist, can be chosen as the base vector or as the vector $\mathbf{v}_n$ is crossed over with. A detailed presentation of variants and how to apply these and further extension can be found in Price, Storn and Lampinen (2005). In Maringer and Parpas (forthcoming), DE is applied to a portfolio optimization problem and compared against an alternative method, Stochastic Differential Equations, while Maringer and Meyer (forthcoming) compare it against Threshold Accepting in an application to model selection. Both studies find that DE is a competitive method capable of solving demanding optimization problems.

One of the advantages of DE is that the number of technical parameters is rather low. In its typical version with the difference vector consisting of two current elements and a cross-over with a fourth current solution, only the population size, the scaling factor F and the cross-over probability need to be prespecified. If noise is used to add diversity, it is common to add normally distributed random values with expected value zero and small standard deviation. Also it is claimed to be easy to use as it needs little or no parameter tuning.

These favorable property was also experienced in this application; a more detailed discussion of calibrating DE can be found in Maringer (forthcoming).

DE's way of generating new solutions has been designed for search in a continuous space. In order to deal with the discontinuities of the search space due to the constraints on the weights, a repair function is introduced that maps the candidate solutions $\mathbf{v}$ into a feasible solution of weights, $\mathbf{x}$. Following the approach suggested in Maringer and Oyewumi (2007), this repair function first assigns the minimum positive weight, x^{ℓ}, to all assets where the values of the corresponding elements of $\mathbf{v}$ exceed this value. If this would leave fewer than $n^{\min} = 1/x^{u}$ assets, then the $n^{\min}$ assets with the highest values in $\mathbf{v}$ are picked. Next, the weights of the included assets are increased in proportion to the values in $\mathbf{v}$ until the weights add up to one and no asset weights violates the upper limit. In preliminary experiments, this repair function was tested against several alternatives. It was found that the function itself is computationally cheap as in most of the cases the mechanism simply picks the largest positive elements and scales them such that the consumption constraint is met. Furthermore, it does not encourage premature convergence: while the values for the included elements tend to be similar to their repaired counterparts, those of the other elements tend to converge either to zero or to some arbitrary negative value. Given the production algorithm for new candidate solutions, the latter case can cause that these elements become positive again in the candidate solution and the corresponding assets are included in the portfolio, and local optima can be escaped. In addition to upper and lower limits, Maringer and Oyewumi (2007) also consider a cardinality constraint which limits the number of different assets in a portfolio with, say, k. Other things equal, the repair mechanism then picks only the k elements with the biggest values if there are more than k positive ones.

For the generation of new solution, the inclusion of noise was found helpful. More specifically, the i-th element of the new vector was computed according to $v_{n,i} = v_{p_1,i} + (F + z_1) \cdot (v_{p_2,i} - v_{p_3,i} + z_2)$ where z_1 and z_2 are either zero (with a probabilities of 1 and 2 percent, respectively) or normally distributed random variables with expected values of zero and standard deviations of 0.02. Typically, the population size was set to 50, and the number of iterations to 2500. Preliminary experiments suggest that the results did not differ substantially for other values of the technical parameters. For all problems, up to 50 restarts were performed and the best of the reported results were used for the analysis in the following section. Algorithm 1 summarizes the pseudocode.

Algorithm 1: Pseudocode for utility maximization with Differential Evolution

randomly initialize population of vectors $\mathbf{v}_p, p = 1 \ldots P$;
repeat
  %% `generate new solutions` $\tilde{\mathbf{v}}_p$;
  for *current solutions* $\mathbf{v}_p, p = 1 \ldots P$ **do**
    randomly pick $c_1 \neq c_2 \neq c_3 \neq p$;
    for *all elements* i *in the solution vector* **do**
      with probability $\pi_1 : z_1[i] \leftarrow N(0, \sigma_1)$ else $z_1[i] \leftarrow 0$;
      with probability $\pi_2 : z_2[i] \leftarrow N(0, \sigma_2)$ else $z_2[i] \leftarrow 0$;
      randomly pick $u[i] \sim U(0,1)$;
      if $u[i] < \pi$ **then**
        $\tilde{v}_p[i] \leftarrow v_p[i]$;
      else
        $\tilde{v}_p[i] \leftarrow v_{c_1}[i] + (F + z_1[i]) \cdot (v_{c_2}[i] - v_{c_3}[i] + z_2[i])$;
  %% `select new population`: replace if improvement;
  for *current solutions* $\mathbf{v}_p, p = 1 \ldots P$ **do**
    map solutions into asset weights, $\mathbf{v}_p \rightarrow \mathbf{x}_p$, $\tilde{\mathbf{v}}_p \rightarrow \tilde{\mathbf{x}}_p$;
    if $E(\mathcal{U}(w(\tilde{\mathbf{x}}_p), \lambda)) > E(\mathcal{U}(w(\mathbf{x}_p), \lambda))$ **then**
      $\mathbf{v}_p \leftarrow \tilde{\mathbf{v}}_p$;
until *halting criterion met* ;

4 Empirical Study

4.1 Data

The empirical study is based on the stocks included in the Dow Jones Industrial Average (DJIA). Using adjusted daily prices downloaded from `finance.yahoo.com` for 2 March 2000 to 17 November 2006, 1684 log returns where computed. Figure 1 contains scatter plots of the moments of the asset returns; points on the same vertical level refer to the same asset. Some of the assets were not included in any of the optimized portfolios and are represented by dots; circles depict assets that are included in at least one of the optimized portfolios. It is noteworthy that it is not the assets with the highest volatility or kurtosis or most negative

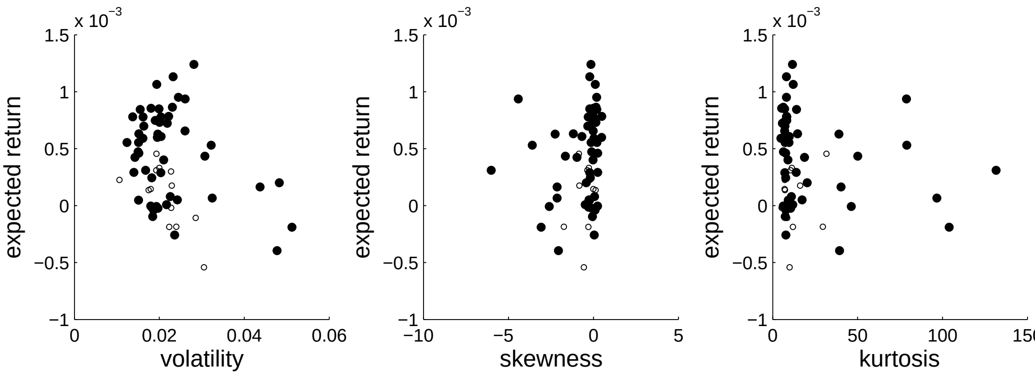


Figure 1: Assets in the mean-volatility, mean-skewness and mean-kurtosis space. Dots: assets included in at least on portfolio; circles: assets never included

skewness that are avoided; also, in some cases assets that are dominated in the mean-volatility space are included. The main reason for this is the structure of the (higher) co-moments.

4.2 Portfolios under Risk Aversion and Loss Aversion

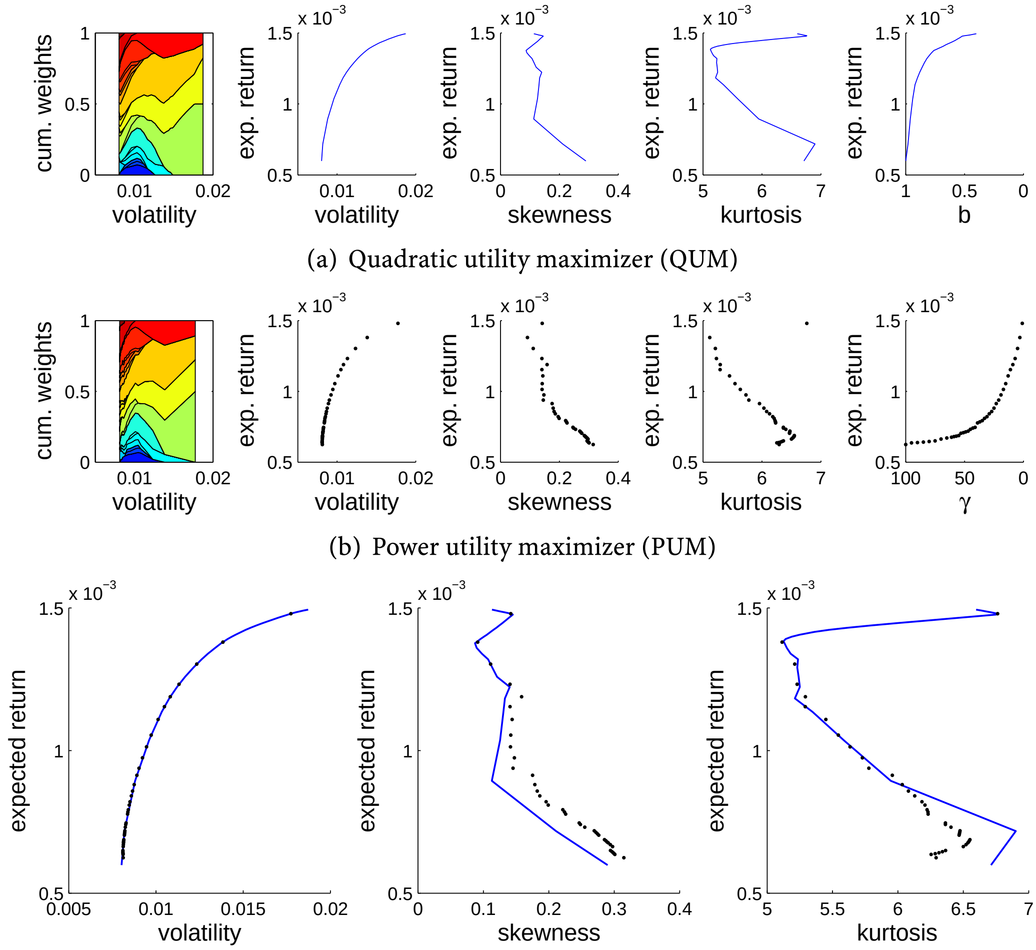
In traditional utility analysis, an investor's utility of a (future) level of wealth, w is irrespective of his initial wealth w_0. In this case, the investor is concerned with the expected value of future wealth and its expected distribution, but there is no explicit distinction whether or not a certain level of wealth is a loss or a gain. Hence, the investor is risk averse, but there is no additional loss aversion. In terms of optimization problem (5), this implies a parameter of loss aversion of $\lambda = 1$, and no correction is required to estimated the expected utility.

Figure 2 depicts the moments of the arithmetic returns of optimized portfolios for a quadratic utility (Figure 2(a), line) and a power utility (Figure 2(b), dots) maximizer with different parameters of risk aversion. As expected, decreasing risk aversion (i.e., decreasing parameters b and γ, respectively) leads to portfolios with higher expected return, yet also higher volatility. In this region, fewer, but more profitable and riskier assets are included, as can be seen from the cumulated weights. For both utility functions, the efficient sets are the usual curves in mean-volatility space with slight kinks due to the weight constraints.

Quadratic utility is oblivious to skewness and kurtosis; power utility, on the other hand, is affected by higher moments. Not surprisingly, the power utility maximizer therefore prefers portfolios with higher skewness and/or lower kurtosis. The graphs in Figure 2(c) show that for cases with a rather low level of risk aversion (high return portfolios), the two investors hardly differ in their investment choice. The higher the risk aversion (low return portfolios), however, the more pronounced the deviations become. Though hardly noticeable, when the level of risk aversion is very high, the power utility maximizer (PUM) is even prepared to accept a slightly higher volatility than the quadratic utility maximizer (QUM), if this comes with more favorable values for the higher moments. When returns are perfectly normally distributed, these deviations will not manifest themselves. However, in the presence of higher order moments PUM and QUM might consider each other's optimized portfolios as inefficient; a look beyond the usual mean-volatility framework, however, shows that neither of them behaves irrationally – the differences are due to different, yet perfectly rational, preferences and, consequently, they have different notions of what an efficiency means.

These differences become even more apparent when investors are not only risk averse, but also loss averse. When the initial wealth, w_0 influences the utility of a future level of wealth, then the same future level of wealth w will give the investor a higher utility if it comes with a gain than if it actually comes with a loss: $\mathcal{U}(w|w_0) > \mathcal{U}(w|w_0') \; \forall w_0 < w < w_0'$. With concave utility functions, decision makers will then be even more keen on avoiding losses. Primarily, this means a reduction of risk in terms of volatility. If, however, the utility function can look beyond the second moment, than investors will try to avoid not just any, but specifically negative deviations from the expected value; i.e., they will try to increase the (positive) skewness. Also, they want to avoid extreme events and will therefore aim to reduce the kurtosis of their portfolios.

Figure 3 depicts the moments of optimized portfolios for different utility functions, levels of risk aversion and levels of loss aversion. As can be seen immediately from the graphs for the mean-volatility space (left), the high return / high volatility regions of the previous graph are missing, and both QUM and PUM are now in regions which, in the absence of loss aversion, would have been chosen only with a substantially higher level of risk aversion. This effect is the stronger, the higher the loss aversion. Since the modified wealth levels (equation (6)) include the level of loss aversion and emphasize the losses accordingly, the efficient lines for quadratic utility maximizers now deviate from the one without loss aversion. Again, the constraints on minimum and maximum weights for included assets cause kinks in the efficient lines (here: in a mean-volatility space); the actual effects of these constraints, however, are not too big. Looking

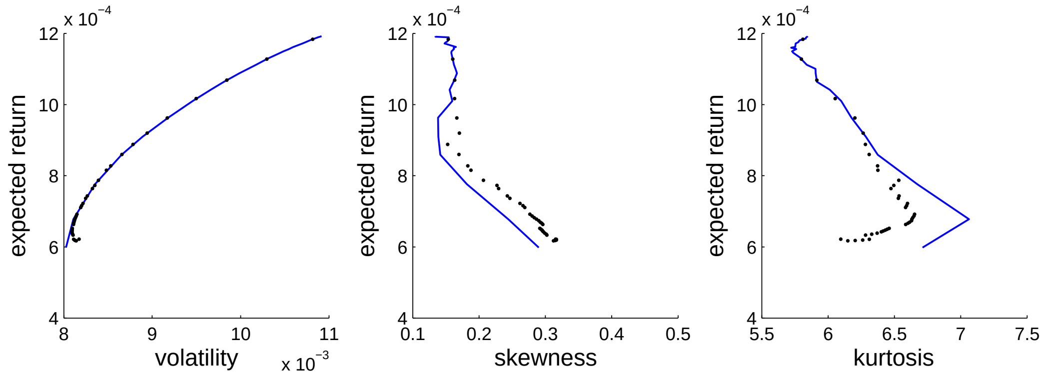


(a) Quadratic utility maximizer (QUM)

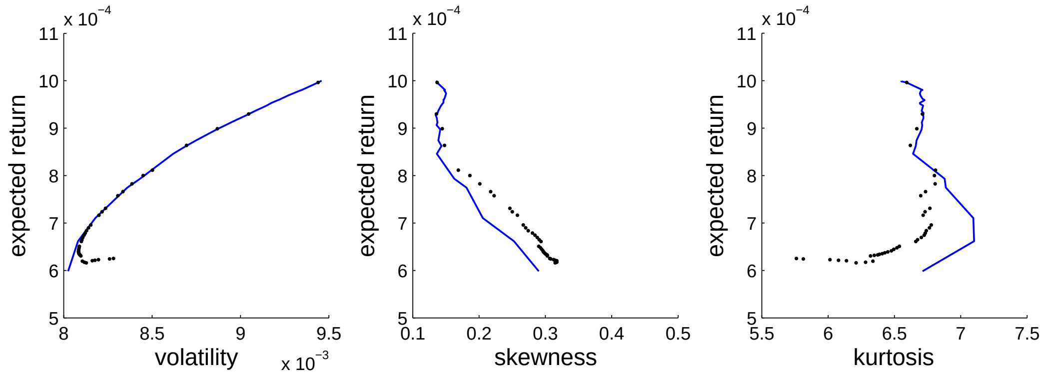
(b) Power utility maximizer (PUM)

(c) Comparison of quadratic and power utility maximizer

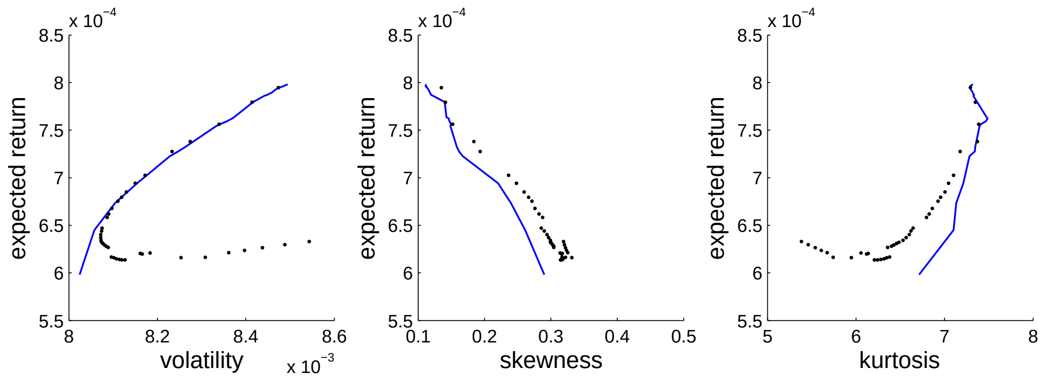
Figure 2: Portfolios optimized under quadratic utility (line) and power utility (dots) without loss aversion ($\lambda = 1$), projected into the mean-volatility, mean-skewness and mean-kurtosis space



(a) loss aversion: $\lambda = 1.25$



(b) loss aversion: $\lambda = 1.5$



(c) loss aversion: $\lambda = 2$

Figure 3: Moments of portfolios optimized under quadratic utility (line) and power utility (dots) and different levels of loss aversion, λ, projected into the mean-volatility, mean-skewness and mean-kurtosis space

at the higher moments, however, shows QUM's lack of concern with deviations from the expected value that are not captured by volatility: By definition, higher risk aversion demands lower volatility, but not necessarily higher skewness or lower kurtosis.

For power utility maximizers, the picture looks different. PUM with higher risk aversion and higher loss aversion, too, prefer lower volatility. However, they care more about higher moments than QUM. Notably, for any given mean return, their portfolios never have lower skewness and higher kurtosis than the respective portfolios of QUM. However, they are prepared to accept higher volatility than QUM because the benefits of the higher skewness and the reduced kurtosis offsets.

Eventually, this concern with higher moments leads to the effect that PUM with increasing risk aversion start increasing their portfolio's volatility. Looking at the mean-volatility diagram alone might suggest irrationality or inefficiency. Looking at the higher moments, however, shows that this is perfectly rational. With higher risk aversion, PUM are inclined to avoid (extreme) losses. Loss aversion emphasizes the losses, and the way the modified (or perceived) wealth levels, m, are computed (equation (6)) adds negative skewness. In the optimization process, these distortions are to be counterbalanced, and a potential increase in skewness and reduction of kurtosis outweighs a mere volatility reduction. To a small degree, this can be observed already in the results for the case without loss aversion but a very high level of risk aversion; when there is loss aversion in addition, this effect can become prevailing.

As can be seen from the results, the effects of loss aversion are different for different risk preferences. Obviously, when some of the investors look at different (higher) moments of return distributions, they will have different views about efficient and optimal solutions. When only looking at the mean and volatility, power utility maximizers with loss aversion might look as if they become risk lovers, while in actual facts they become even more risk averse but are increasingly concerned with the higher moments of their portfolios. Volatility is no longer sufficient to describe all aspects of risk, and basing the definition of "effectiveness" on this risk measure alone might lead to flawed conclusions.

It must be emphasized that all the empirical cases considered so far are for investors with strictly convex utility functions who are always risk averse and rational. What the consequences for optimal portfolios and their higher moments are when, as suggested by prospect theory, decision makers become risk seeking when occurring losses, is discussed in the next subsection.

4.3 Portfolios under Prospect Theory

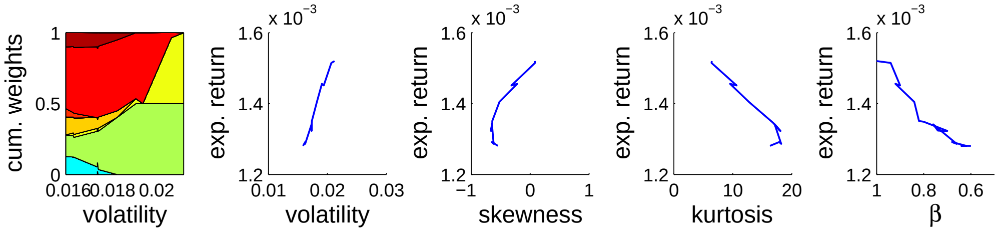
When investors no longer have a decreasing marginal utility for any level of wealth, then their utility curve becomes convex in certain regions – which characterizes risk lovers. In Kahneman and Tversky (1979), this is found for losses: investors dislike bigger losses, yet the additional disutility is decreasing. In addition, however, investors are more sensitive to losses than they are to profits of the same magnitude; their utility of losses is therefore weighted with a loss aversion coefficient λ. The resulting utility curve is S-shaped with a kink at the origin, and, assuming the convexity for losses is equal to the concavity is profits ($\alpha = \beta$), it can be modeled by

$$\begin{aligned}\mathcal{U}_{PT}(w) &= (w - w_0)^{\alpha} \Im_{w \geq w_0} - \lambda(w_0 - w)^{\beta} \Im_{w < w_0} \\ &= (|w - w_0|)^{\beta} \left(1 - (1+\lambda)\Im_{w<w_0}\right)\end{aligned}$$

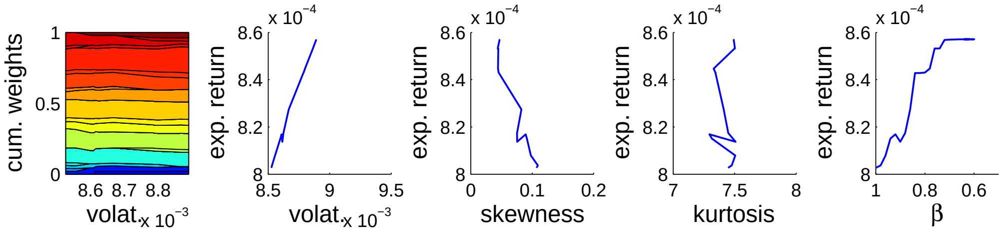
Furthermore, it is assumed that the decision weights are equal to the probabilities of the different states, i.e., $\pi_i = p_i$.

If $\lambda = 1$ then the investor is not loss averse. When the parameter β is equal to one, the utility function is a straight line, and the investor will simply maximize the expected wealth. However, when β is lowered the curvature on either side increases. This has two effects: on the right hand side, bigger profits earn less and less additional utility and the investor will accept lower expected returns if the distribution is more favorable. On the left hand side, however, bigger losses cause less and less additional concern, and, other things equal, investors with low β are less concerned with lower or even negative skewness. Also, kurtosis is now less important: extreme events provide virtually the same (dis-)utility as events that are not quite so extreme, hence increasing the kurtosis does has no substantial effect if none of the other moments is affected. Figure 4(a) illustrates these effects. Note that the lower and upper limits on the weights of included assets cause discontinuities.

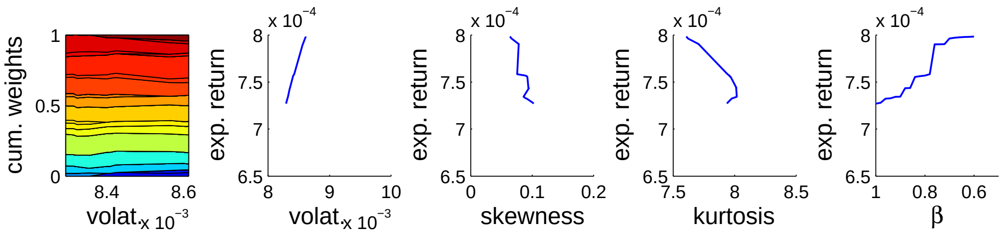
When loss aversion is introduced, however, the situation changes; this can be seen from Figures 4(b) and 4(c). When they are loss averse, investors concentrate on low risk portfolios, and the actual value of the parameter β, governing the curvature of the utility function, looses importance: with the same level of loss aversion, an investor with high risk aversion (low β) will choose a portfolio quite similar to that of an investor who is risk neutral. While in the absence of loss aversion, the risk neutral aims for the highest returns (and accepts higher volatility), it is now investors with stronger curvature who are prepared to accept



(a) No loss aversion: $\lambda = 1$



(b) loss aversion: $\lambda = 2$



(c) loss aversion: $\lambda = 3$

Figure 4: Moments of portfolios optimized under prospect theory without and with λ, projected into the mean-volatility, mean-skewness and mean-kurtosis space

higher risk. Hence, even though parts of their utility curve is concave indicating they are risk seeking, their loss aversion prevents them from choosing more risky portfolios, and they select better diversified portfolios than they would in the absence of loss aversion.

5 Conclusion

Over the last decades, the importance of behavioral aspects in finance have widely been recognized. One of the insights of this stream of research is that simple utility functions are not sufficient to capture the actual preferences of investors. In particular, they do not capture that investors take their current level of wealth into account when evaluating the utility of future levels of wealth; that moving from profits to losses triggers a higher level of sensitivity; and that eventually they cease to care about even bigger losses as much as they cease to care about even bigger profits. The latter property is addressed in prospect theory where the profits are evaluated with the concave utility function of a risk averse, while losses are subject to a convex utility function of a risk loving investor. The former property can be considered by not only looking at the risk aversion, but also at the loss aversion of a decision maker.

In this chapter, the consequences of different attitudes towards risk and losses are investigated for portfolio optimization where additional practical constraints on the asset weights where introduced. Based on an empirical study, different utility functions and levels of loss aversions are tested and the properties of the resulting portfolios are investigated. It is found that higher order moments are needed to explain the choices: Even when investors are strictly risk averse, a higher level of risk aversion might lead to an investment with more, not less volatility. What seems irrational at first sight, is actually perfectly rational as under loss aversion, investors are more sensitive towards skewness and kurtosis, and increasing the desirable positive skewness and decreasing kurtosis might prevail a mere volatility reduction strategy once the level of risk aversion is big enough. Loss aversion even outbalances some of the preferences of investors who become risk seeking when they face losses.

The results found in this chapter help to understand seemingly irrational behavior and underline the inadequacy of the mean-volatility framework for many real life investment problems. Furthermore, it is shown that demanding optimization problems can be solved with heuristic optimization techniques. However, the results also open some new questions. A broader computational study for different markets, data sets shall help to generalize the findings, while additional constraints on asset selection can analyze the practical implications. Furthermore, the inclusion of other types of assets such as bonds and derivatives can provide additional insights on the consequences for financial engineering. Also, the workings and convergence behavior of the heuristic search method should be investigated in more detail. All this, however, has to be left to future research.

Acknowledgement

The author is grateful for numerous invaluable discussions with Manfred Gilli on heuristics, optimization, econometrics and financial modeling, for encouragement and support.

Financial support from the EU Commission through MRTN-CT-2006-034270 COMISEF is gratefully acknowledged.

References

Basel Committee on Banking Supervision: 2003, Consultative document: The new basel capital accord, bank for international settlement, available from: `www.bis.org`.

Brabazon, A. and O'Neill, M.: 2006, *Biologically inspired algorithms for financial modeling*, Springer, New York.

Dueck, G. and Scheuer, T.: 1990, Threshold accepting: A general purpose algorithm appearing superior to simulated annealing, *Journal of Computational Physics* **90**, 161–175.

Fogel, D. B.: 2001, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 2nd edn, IEEE Press, New York, NY.

Gilli, M. and Këllezi, E.: 2002, The threshold accepting heuristic for index tracking, *in* P. Pardalos and V. Tsitsiringos (eds), *Financial engineering, E-commerce, and supply chain*, Kluwer Academic Press, Bosten, Mass.

Gilli, M. and Këllezi, E.: 2006, An application of extreme value theory for measuring financial risk, *Computational Economics* **27**(2-3), 207–228.

Gilli, M., Këllezi, E. and Hysi, H.: 2006, A data-driven optimization heuristic for downside risk minimization, *The Journal of Risk* **8**(3), 1–18.

Gilli, M., Maringer, D. and Winker, P.: forthcoming, Applications of heuristics in finance, *in* D. Seese, C. Weinhardt and F. Schlottmann (eds), *IT and Finance*, Springer.

Holland, J. H.: 1975, *Adaption in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, MI.

Kahneman, D. and Tversky, A.: 1979, Prospect theory: An analysis of decision under risk, *Econometrica* **47**, 263–291.

Kirkpatrick, S., Gelatt, C. and Vecchi, M.: 1983, Optimization by simulated annealing, *Science* **220**(4598), 671–680.

Konno, H.: 2005, *Essays and Surveys in Global Optimization*, Springer, chapter Applications of Global Optimization to Portfolio Analysis, pp. 195–210.

Maringer, D.: 2005, *Portfolio Management with Heuristic Optimization*, Springer-Verlag.

Maringer, D.: forthcoming, Constrained index tracking under loss aversion using differential evolution, *in* A. Brabazon and M. O'Neill (eds), *Natural Computing in Computational Economics and Finance*, Springer, chapter Constrained Index Tracking under Loss Aversion using Differential Evolution.

Maringer, D. and Meyer, M.: forthcoming, Smooth transition autoregressive models – new approaches to the model selection process, *Studies in Nonlinear Dynamics and Econometrics* .

Maringer, D. and Oyewumi, O.: 2007, Index tracking with constrained portfolios, *Intelligent Systems in Accounting, Finance and Management* **15**, 57–71.

Maringer, D. and Parpas, P.: forthcoming, Global optimization of higher moments in portfolio selection, *Journal of Global Optimization* .

Markowitz, H. M.: 1952, Portfolio selection, *The Journal of Finance* **7**(1), 77–91.

Odean, T.: 1998, Are investors reluctant to realize their losses?, *Journal of Finance* **53**, 1775–1798.

Parpas, P. and Rustem, B.: 2006, Global optimization of the scenario generation and portfolio selection problems, *in* M. Gavrilova, O. Gervasi, V. Kumar, C. Tan, A. Taniar, D.; Laganà, Y. Mun and H. Choo (eds), *Computational Science and Its Applications – ICCSA 2006, Proceedings Part III*, Vol. 3982 of *Lecture Notes in Computer Science*.

Price, K., Storn, R. and Lampinen, J.: 2005, *Differential Evolution: A Practical Approach to Global Optimization*, Springer.

Rechenberg, I.: 1965, Cybernetic solution path of an experimental problem, *Library Translation 1122*, Royal Aircraft Establishment.

Rechenberg, I.: 1973, *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Fromman-Holzboog Verlag, Stuttgart.

Samuelson, P. A.: 1963, Risk and uncertainty: A fallacy of large numbers, *Scientia* **XCVIII**, 108–113.

Storn, R. and Price, K.: 1995, Differential evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces, *Technical report*, International Computer Science Institute, Berkeley.

Storn, R. and Price, K.: 1997, Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization* **11**(4), 341–359.

Thaler, R. H., Tversky, A., Kahneman, D. and Schwartz, A.: 1997, The effect of myopia and loss aversion on risk taking: An experimental test, *Quaterly Journal of Economics* **112**, 647–661.

Tobin, J.: 1958, Liquidity preference as behavior towards risk, *Review of Economic Studies* **26**(1), 65–86.

Winker, P.: 2001, *Optimization Heuristics in Econometrics. Applications of Threshold Accepting*, John Wiley & Sons, ltd., Chichester et al.

Winker, P. and Maringer, D.: 2007, The threshold accepting optimization algorithm in economics and statistics, *in* E. J. Kontoghiorges and C. Gatu (eds), *Advances in Computational Economics, Finance and Management Science*, Kluwer, pp. 107–125.