

INSTABILITY OF PORTFOLIO OPTIMIZATION UNDER COHERENT RISK MEASURES

Imre Kondor

Collegium Budapest – Institute for Advanced Study, Szentháromság u. 2,
H-1014 Budapest, Hungary

Department of Physics of Complex Systems, Eötvös University, Pázmány
Péter sétány 1/A, H-1117 Budapest, Hungary

Parmenides Center for the Study of Thinking, Kardinal Faulhaber Strasse
14a, Munich, D-80333, Germany

email: kondor@colbud.hu

István Varga-Haszonits

Department of Physics of Complex Systems, Eötvös University, Pázmány
Péter sétány 1/A, H-1117 Budapest, Hungary

Analytics Department of Fixed Income Division, Morgan Stanley Hungary
Analytics, Deák Ferenc u. 15, H-1052 Budapest, Hungary

email: Istvan.Varga-Haszonits@morganstanley.com

Abstract

It is shown that the axioms for coherent risk measures imply that whenever there is a pair of portfolios such that one of them dominates the other one in a given sample (which happens with finite probability even for large samples), then there is no optimal portfolio under any coherent measure on that sample, and the risk measure diverges to minus infinity. This instability was first discovered on the special example of Expected Shortfall which is used here both as an illustration and as a springboard for generalization.

Key Words: Coherent Risk Measures, Portfolio Optimization, Expected Shortfall, Estimation of Risk

JEL Classification: G11, C13, D81

1 Introduction

The proper mathematical characterization of risk is of central importance in finance: in the introductory chapter of a recent volume entirely dedicated to risk measures, the upsurge of interest in risk measures from 1997 onwards was called the third major revolution in finance theory [1]. The choice of the adequate risk measure is a highly nontrivial task that, in principle, involves deep considerations concerning the attitudes of market players, the structure of markets, the instruments in question, etc. No wonder that there is no universally accepted risk measure, the ones in use range from rather ad hoc characteristics to axiomatic constructs. Recently, value at risk (VaR) [2, 3] has gained widespread use, in practice as well as in regulation. VaR has, however, been criticized by a number of academics [4, 5, 6], because as a quantile it has no reason to be convex, and indeed it is easy to construct portfolios for which VaR seriously violates convexity. The shortcomings of VaR and other considerations led [7, 4] to the introduction of coherent risk measures. Independently and essentially simultaneously similar considerations were put forward in various contexts by a number of researchers, such as [8, 9, 10]. The seminal paper of [4] triggered a rich branch of research with the number of papers inspired by it ranging in the hundreds. Expected Shortfall (ES), a particularly appealing risk measure, was quickly demonstrated to be coherent [11, 12], and [13] went on to construct a whole class of coherent risk measures called spectral measures of which ES was a special case. ES is essentially the conditional average loss above a high threshold. As such, it is easily related to VaR that can be identified with this threshold (although for ES this cutoff is defined in probability, not in money). This explains why ES is often identified with conditional value at risk CVaR, although for discrete distributions there is a subtle but significant difference between the two (see [12] for a careful analysis of the relationship between them). As a conditional average, ES is also a more plausible measure than VaR, it is easy to determine from empirical distributions, and as shown by [14] the optimization under ES can be reduced to a linear programming task. All these desirable features have made ES very popular among academics.

Therefore it came as a complete surprise to us when we recognized a few years ago that the optimization of some portfolios, especially the large ones, was not always feasible under ES: in some of the samples the risk measure became unbounded from below [15]. This is a probabilistic phenomenon: it depends on the sample. It is natural to ask whether this is a special pathology of ES or whether other risk measures also display a similar instability. The purpose of the present paper is to answer this question.

The rest of the paper is organized as follows. In Section 2 we recapitulate

some of the results related to the instability of ES. We also display the simple financial content of this mathematical phenomenon, which will immediately make it obvious that the instability must be a common property of every coherent measure. A formal proof of this is provided in Section 3. The paper ends on a short Summary.

2 The Instability of Expected Shortfall

Let us begin with a special case of ES corresponding to the limit where the cutoff α beyond which the conditional average is calculated goes to unity. Then only the single worst loss will contribute and the optimization over the portfolio weights under this risk measure consists in finding the best linear combination of the worst losses. This risk measure is sometimes called the Maximal Loss (ML) or Minimax measure [16]. For this particular measure, and assuming that the returns follow an elliptical distribution with zero mean, [15] showed that it is possible to exactly determine the probability of having a solution as a function of the portfolio size N and the sample size T with the following result:

$$(1) \quad p(N, T) = \Theta(T - N) \cdot \frac{1}{2^{T-1}} \sum_{k=N-1}^{T-1} C_{T-1}^k$$

where $\Theta(x)$ is the step function (equal to 1 if $x \geq 0$ and equal to 0, if $x < 0$), and the symbol C_n^k represents the binomial coefficient.

As shown by (1), for short time series $T < N$ the optimization problem never has a solution, but this is rather natural: if we do not have enough data, then it is impossible to find an optimum. Other, more conventional risk measures (variance and mean absolute deviation or MAD, [17]) also fail to allow for a solution in that parameter range. The remarkable fact about (1) is that for finite N and T the probability of having a solution is less than unity even for $T > N$. In the limit N, T going to infinity with their ratio fixed, the probability distribution becomes a step function: for $N/T < 1/2$ it goes to 1, for $N/T > 1/2$ to zero.

For the case of ES with a cutoff parameter α and again in the limit N, T going to infinity, we find that the critical value of the ratio N/T separating the region where the problem always has a solution from the one where it does not varies with α . This way, we get a kind of phase diagram separating the region where the optimization problem is feasible from that where it is not. This curve is illustrated in Fig. 1. It has been numerically determined in [15]. Subsequently [18] succeeded in calculating this phase diagram exactly

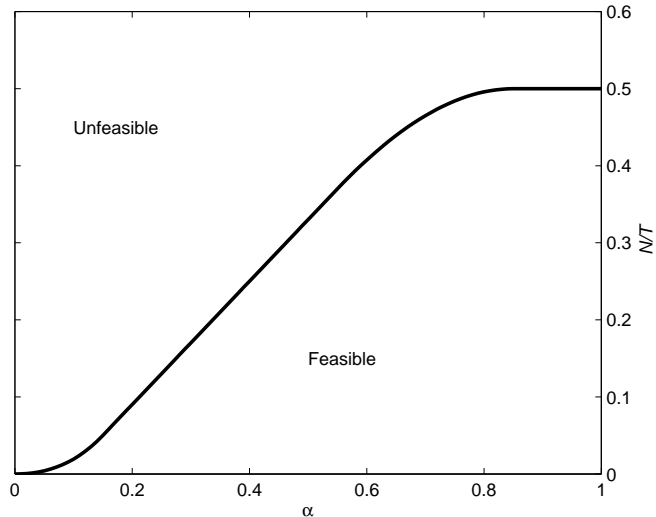


Figure 1: The phase diagram of Expected Shortfall. In the region called 'Unfeasible' in the figure the optimization under ES does not have a solution. The transition across the line separating the feasible and unfeasible regions is sharp if N, T go to infinity with their ratio kept fixed.

via the method of replicas, borrowed from the theory of random systems. The calculation shows that the curve behaves in a non-analytic fashion near $\alpha = 1$, the drop from $1/2$ being exponentially small for α 's in the vicinity of 1. Since the α 's of practical interest are fairly close to 1 (such as .95 or .99) we can say that the critical value of N/T is very close to $1/2$ for any practically interesting case.

It may be worthwhile pointing to a counterintuitive feature of the phase diagram: As can be seen in Fig. 1, the curve is sloping down from right to left, so the critical value of N/T where the optimization task becomes unfeasible is *decreasing* with the cutoff α . Decreasing α means we are taking into account more and more data, so one might have expected that the probability of finding a solution should increase as a result. Instead, we find that the critical ratio N/T decreases, so for a given portfolio size N we need longer and longer time series T in order to find a solution. We will return to this question below.

To put the instability in perspective it is worth making a comparison with the case of the variance. As long as $T > N$ the covariance matrix is always positive definite, so the optimization can always be carried out. It is well known, however, that the rank of the covariance matrix is the smaller of

N and T , so when T goes to N from above the covariance matrix develops a zero eigenvalue, and the curvature of the risk measure becomes zero in a certain direction. With this the optimization problem loses its meaning. As T decreases further, more and more such neutral directions appear, but the variance as a risk measure never becomes unbounded. The behavior of ES as a risk measure is both similar to and different from this. For finite $N < T$ there are always samples for which the optimization does not have a solution, but the reason is not the appearance of neutral directions, but the fact that the risk measure ES becomes unbounded from below. For N/T small the probability of such samples is small. As N/T approaches its critical value, the probability of samples for which the optimization is unfeasible is growing and as we go above the critical N/T it fast approaches 1. Finally, when N/T goes above 1 all the samples become unfeasible. For $N, T \rightarrow \infty$ with their ratio fixed, this transition becomes sharp: below the critical value of N/T all the samples are feasible, above none of them.

Note that the critical value of N/T is always smaller for ES than for either the variance or MAD, so ES is more sensitive to sample to sample fluctuations (noise) than either the variance or MAD [15]. The phase diagram does not only separate the regions where the optimization under ES can or cannot be carried out, but it is also the locus where the estimation error diverges. All these features and several more details were thoroughly discussed in [15], but the financial content of this instability was left somewhat implicit.

As a matter of fact, the message is almost self-evident. Let us just consider the simplest case imaginable: $N = T = 2$ for the Minimax problem. Let x_{it} be the returns on securities $i = 1, 2$ at times $t = 1, 2$, and let w_i ($i = 1, 2$) be the portfolio weights. Then the task is to find the weights w_i that make the maximal loss

$$(2) \quad \max_t \left(- \sum_{i=1}^N w_i x_{it} \right)$$

minimal. As the weights sum to 1, we let $w_1 = w$, and $w_2 = 1 - w$. Then the losses of the portfolio over periods $t = 1, 2$ are the following:

$$\begin{aligned} y_1 &= -wx_{11} - (1-w)x_{21} = w(x_{21} - x_{11}) - x_{21}, \\ y_2 &= -wx_{12} - (1-w)x_{22} = w(x_{22} - x_{12}) - x_{22}. \end{aligned}$$

Whether $\max\{y_1, y_2\}$ is bounded from below or not depends on the slopes of these two straight lines, see Fig. 2. If the two slopes are of opposite sign, $\max\{y_1, y_2\}$ is bounded from below, and the minimax problem has a solution. If the two slopes are of the same sign, there is no solution, the minimum runs

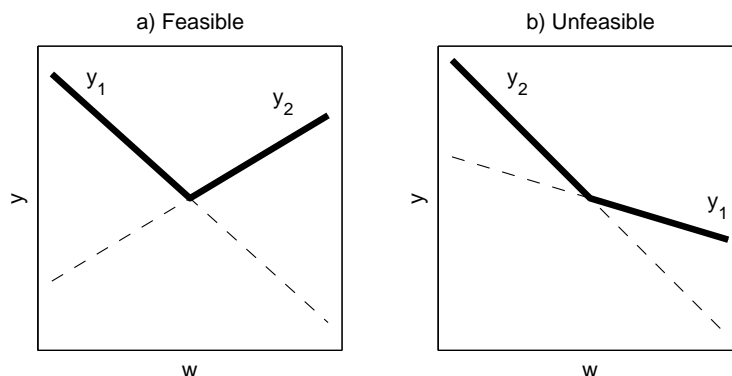


Figure 2: The losses y_1 and y_2 (dashed lines) and their maximum (heavy lines) as a function of the portfolio weight w . Panel a) shows the case where the maximal loss is bounded from below so we have a minimum, while in Panel b) the maximal loss is unbounded, hence there is no minimum.

away to infinity, and the value of the objective function in (2) goes to minus infinity. When do the two slopes have the same sign? This happens if

- (a) either $x_{21} > x_{11}$ and $x_{22} > x_{12}$,
- (b) or $x_{21} < x_{11}$ and $x_{22} < x_{12}$.

In case (a) asset $i = 2$ dominates asset $i = 1$ in the sense that the return on asset $i = 2$ is larger than the return on asset $i = 1$ for all $t = 1, 2$, and the solution runs away to infinity such that $w_2 \rightarrow +\infty$ and $w_1 \rightarrow -\infty$. In the opposite case (b) asset $i = 1$ dominates and the solution runs away to infinity with $w_1 \rightarrow +\infty$ and $w_2 \rightarrow -\infty$. But this is obvious! If the return on one of the assets is always larger than on the other, we have to go infinitely long in the dominating asset, and infinitely short in the other one, and this will produce an infinitely large negative risk, i.e. no risk at all.

The generalization to ES and beyond is immediate: if we have several assets in the portfolio and we have larger, but finite samples (i.e. finite time series segments), it can always happen that one asset, or a combination of assets, performs much better than the rest *in the given sample*, even if they do not perform better in general. In such a case the optimization of portfolios under a risk measure that penalizes large losses more severely than smaller losses will inevitably lead to the result that we have to go infinitely long in the over-performing asset(s) and infinitely short in the rest, which will make the average return unbounded. Of course, all this strongly depends on the size of

the portfolio and on the size of the sample. For N increasing it is easier and easier to find a well-performing combination, while for an increasing sample T it is less and less probable that an atypical fluctuation may lead us to a false conclusion. This explains the features displayed in Fig. 1: for a given α the feasibility of the task is lost when N/T exceeds a certain threshold, while for a given N/T the feasibility is lost with decreasing α , as we are allowing, in addition to the worst case, a larger and larger part of the worst cases to enter into competition.

Runaway solutions may, of course, arise only if we allow arbitrarily large short positions. The main message does not change much, however, even if we impose a ban on short selling, or any other constraint on the weights. For finite T there will always be a chance for one or the other security to dominate, and then the optimization will lead us to the conclusion that we should go long in the dominating item to the maximum extent allowed by the constraints. This will not result in infinite positions, nevertheless, the solution will still lack any stability: in the next sample another security may dominate, so the solution will jump around on the 'walls' of the allowed region. This 'finite volume' instability has been studied in the case of the variance and maximal loss for various sets of linear constraints in [19]. The insight gained in the above example leads us to the conjecture that coherent measures may, in general, be sensitive to this type of instability for finite samples. This will, indeed, turn out to be the case, as shown by the formal argument developed in the next Section.

3 The instability of coherent measures

In addition to the notation we have already introduced, that is N for the number of assets in the portfolio, T for the number of observation periods, x_{it} the return on asset i over the observation period t , and w_i the portfolio weight of asset i , we are also going to use the notation \mathbf{w} for the N dimensional vector of portfolio weights and \mathbf{X} for the $N \times T$ matrix of the observed returns (i.e. the matrix with entries x_{it}).

3.1 Dominant portfolios

As explained in the previous section, the minimization of ES is not feasible when there is a dominating combination of assets on the given sample. In order to spell out this proposition in a more formal and general way, let us define what we exactly mean by *domination*.

Definition 3.1. Let \mathbf{u} and \mathbf{v} be two portfolio vectors, and let us assume that both vectors are properly normalized, i.e. the sum of their components is 1. We say that \mathbf{u} dominates \mathbf{v} on the sample \mathbf{X} , if for all $t = 1, 2, \dots, T$ we have

$$(3) \quad \sum_{i=1}^N u_i x_{it} \geq \sum_{i=1}^N v_i x_{it},$$

The dominance of \mathbf{u} over \mathbf{v} is denoted by $\mathbf{u} \succeq \mathbf{v}$.

Furthermore, we say that \mathbf{u} strictly dominates \mathbf{v} on the sample \mathbf{X} if for all $t = 1, 2, \dots, T$ we have

$$(4) \quad \sum_{i=1}^N u_i x_{it} > \sum_{i=1}^N v_i x_{it}.$$

The strict dominance¹ of \mathbf{u} over \mathbf{v} is denoted by $\mathbf{u} \succ \mathbf{v}$.

It is important to emphasize that the dominance (or strict dominance) relation, as defined above, between two portfolios is specific to the underlying sample. Clearly, if two vectors form a dominant-dominated pair for (almost) all possible samples then we have an arbitrage opportunity. At the same time, it is obvious that dominance on a single sample does not imply arbitrage: even if $\mathbf{u} \succ \mathbf{v}$ on sample \mathbf{X}_1 (giving the false illusion of an arbitrage opportunity) it is quite possible that neither of them dominates the other on another sample \mathbf{X}_2 .²

3.2 Optimization of coherent risk measures

Let us assume that we have a coherent risk measure ρ and we want to optimize our portfolio under this measure, on the basis of data making up the sample \mathbf{X} . Since we do not know the true data generating process, we cannot tell exactly what the risk of a portfolio \mathbf{w} is, but we can estimate it on the basis of the given sample. Let us denote this estimate by $\hat{\rho}(\mathbf{w}|\mathbf{X})$. Since ρ is coherent, we require that its estimator $\hat{\rho}$ also be coherent on the sample \mathbf{X} . Let us recall the axioms of coherence and adapt them to the present context:

¹It would be more precise to define strict dominance so that we allow equality for some (but not all) of the observations. This would complicate our discussion, but would not affect the generality of the results.

²In fact, the notation $\mathbf{u} \succ \mathbf{v}$ is somewhat sloppy and $\mathbf{u} \succ_{\mathbf{X}} \mathbf{v}$ would be more appropriate. However, we refrain from this rather awkward notation, as it will always be clear on which sample dominance is considered.

Definition 3.2. We say that the risk measure estimator $\hat{\rho}$ is coherent on sample \mathbf{X} , if it fulfills each of the following requirements (where $\mathbf{0}$ denotes the null-vector):

1. Monotonicity: $\mathbf{u} \succeq \mathbf{0} \Rightarrow \hat{\rho}(\mathbf{u}|\mathbf{X}) \leq 0$
2. Sub-additivity: $\hat{\rho}(\mathbf{u} + \mathbf{v}|\mathbf{X}) \leq \hat{\rho}(\mathbf{u}|\mathbf{X}) + \hat{\rho}(\mathbf{v}|\mathbf{X})$
3. Positive homogeneity: $a > 0 \Rightarrow \hat{\rho}(a\mathbf{u}|\mathbf{X}) = a\hat{\rho}(\mathbf{u}|\mathbf{X})$
4. Translational invariance: $a \in \mathbb{R} \Rightarrow \hat{\rho}(\mathbf{u}|\mathbf{X} + a) = \hat{\rho}(\mathbf{u}|\mathbf{X}) - a$

where $\mathbf{X} + a$ denotes the matrix obtained by adding the number a to each element of \mathbf{X} .

It is obvious that the estimator $\hat{\rho}$ of a coherent risk measure ρ should behave coherently in the above sense on any sample \mathbf{X} .

The global risk minimization problem for the finite sample \mathbf{X} is

$$(5) \quad \min_{\mathbf{w}} \hat{\rho}(\mathbf{w}|\mathbf{X}),$$

$$\sum_{i=1}^N w_i = 1.$$

where the second equation is usually referred to as the budget constraint. The following theorem states that the existence of an optimal portfolio is not automatically guaranteed, and it is a property of the underlying sample.

Theorem 3.1. *If there exist two portfolios \mathbf{u} and \mathbf{v} such that they satisfy the budget constraint and $\mathbf{u} \succ \mathbf{v}$ then (5) has no solution.*

Proof. Let us use proof by contradiction and assume that $\mathbf{w}^{(0)}$ is the (normalized) minimal risk portfolio. Let $\mathbf{0}$ denote the null vector (empty portfolio), then clearly $\mathbf{u} - \mathbf{v} \succ \mathbf{0}$, hence, due to the monotonicity and translational invariance axioms, we get that

$$(6) \quad \hat{\rho}(\mathbf{u} - \mathbf{v}|\mathbf{X}) < 0.$$

Let us now consider the following portfolio:

$$\mathbf{w}^{(1)} = \mathbf{w}^{(0)} + a(\mathbf{u} - \mathbf{v}),$$

where a is a positive real number. Obviously, $\mathbf{w}^{(1)}$ is also normalized. Using the sub-additivity and positive homogeneity axioms:

$$\hat{\rho}(\mathbf{w}^{(1)}|\mathbf{X}) = \hat{\rho}(\mathbf{w}^{(0)} + a(\mathbf{u} - \mathbf{v})|\mathbf{X}) \leq \hat{\rho}(\mathbf{w}^{(0)}|\mathbf{X}) + a\hat{\rho}(\mathbf{u} - \mathbf{v}|\mathbf{X}).$$

Because of (6) and the positivity of a the last term on the right hand side is negative. Hence we get:

$$\hat{\rho}(\mathbf{w}^{(1)}|\mathbf{X}) < \hat{\rho}(\mathbf{w}^{(0)}|\mathbf{X}),$$

which obviously contradicts the assumption that $\mathbf{w}^{(0)}$ is optimal. \square

This theorem states that the existence of a dominant-dominated portfolio pair is a sufficient condition for the non-existence of the optimum. This condition, however, is not necessary in general: the optimization problem (5) may be unfeasible even if there is no dominating portfolio.

One important exception is the minimization of the Maximal Loss measure (Minimax problem), which is formally:

$$(7) \quad \min_{\mathbf{w}} \max_{1 \leq t \leq T} \left(- \sum_{i=1}^N w_i x_{it} \right),$$

$$\sum_{i=1}^N w_i = 1,$$

Theorem 3.2. *The optimization problem (7) has no solution, if and only if there exists a pair of portfolios so that one of them strictly dominates the other.*

Proof. In view of Theorem 3.1 we only have to prove that there must be a pair of dominating-dominated portfolios when the Minimax problem has no solution. Let us assume the contrary: there is no portfolio that strictly dominates any other portfolio. It can be shown [16] that the Minimax problem is equivalent to the following linear programming task:

$$(8) \quad \min u$$

$$u \geq - \sum_{i=1}^N w_i x_{it}$$

$$\sum_{i=1}^N w_i = 1$$

Therefore, having no solution means that for any real number u there is a portfolio \mathbf{w} that satisfies the budget constraint, and for any $t = 1, 2, \dots, T$ we have $u > - \sum_{i=1}^N w_i x_{it}$. (That, is the portfolio return is not bounded in the space of portfolios.)

More specifically, let us consider the maximum of all returns in our sample x_{it} and let the indices of this (or one of these) maximal observation(s) be i_1

and t_1 . (Since the number of assets and the number of observations are both finite, this maximum must exist, although it may not be unique.) Letting $u = -x_{i_1 t_1}$ we have a portfolio $\mathbf{w}^{(1)}$ such that $x_{i_1 t_1} < \sum_{i=1}^N w_i^{(1)} x_{it}$ for any $t = 1, 2, \dots, T$ so eventually:

$$(9) \quad x_{i_1 t_1} < \min_{1 \leq t \leq T} \left(\sum_{i=1}^N w_i^{(1)} x_{it} \right)$$

Since there is no dominating portfolio, for $\mathbf{w}^{(1)}$ there exist indices $1 \leq i_2 \leq N$ and $1 \leq t_2 \leq T$ such that $x_{i_2 t_2} > \sum_{i=1}^N w_i^{(1)} x_{it_2}$ (if that were not true $\mathbf{w}^{(1)}$ itself would dominate all of the N assets). Therefore:

$$(10) \quad x_{i_2 t_2} > \min_{1 \leq t \leq T} \left(\sum_{i=1}^N w_i^{(1)} x_{it} \right).$$

Equations (9) and (10) together mean that $x_{i_1 t_1} < x_{i_2 t_2}$. Since $x_{i_1 t_1}$ is the maximum of all sample returns, this is obviously a contradiction. \square

It is important to stress that neither Theorem 3.1 nor Theorem 3.2 assumes anything about the probability distribution of the asset returns. Further results can immediately be inferred if we now assume that the returns follow an elliptical distribution with zero mean.

Then equation (1) in Section 2 states the probability of the Minimax problem being feasible. Theorem 3.2 asserts that the Minimax problem is unfeasible, if and only if there is a dominant-dominated pair of portfolios. Thus we arrive at the following corollaries (where $p(N, T)$ is the probability given by (1)):

Corollary 3.1. *For returns drawn from an elliptical distribution the probability of the existence of a pair of portfolios such that one of them dominates the other on the sample \mathbf{X} is $1 - p(N, T)$.*

Corollary 3.2. *For elliptically distributed returns the probability of the unfeasibility of the portfolio optimization problem under any coherent measure on the sample \mathbf{X} is at least $1 - p(N, T)$.*

Moreover, in Section 2 we saw that if N, T go to infinity such that N/T is held constant then for $N/T < 1/2$ we have $p(N, T) \rightarrow 1$, while for $N/T > 1/2$ we have $p(N, T) \rightarrow 0$. By Corollary 3.2, $1 - p(N, T)$ is a lower bound of the probability of having no solution for an arbitrary coherent measure. Thus we can also state:

Corollary 3.3. *If, similarly to the case of ES, there is a sharp feasibility–non-feasibility transition for an arbitrary coherent measure in the limit where N, T go to infinity with their ratio fixed, then the critical value of N/T must be smaller than $1/2$ for portfolios of elliptically distributed returns.*

Similar statements can be made for non-elliptically distributed returns, the only difference being that the probability of finding a solution for the Minimax problem is not known explicitly in the general case.

Let us now turn to the financial interpretation of the findings of this section. Essentially, Theorem 3.1 used an arbitrage argument to show that the risk is unbounded from below. Assuming that portfolios \mathbf{u} and \mathbf{v} are such that $\mathbf{u} \succ \mathbf{v}$, we can go a unit long of \mathbf{u} and a unit short of \mathbf{v} so that the portfolio returns on the given sample will be given by the vector $a(\mathbf{u}^T \mathbf{X} - \mathbf{v}^T \mathbf{X})$, whose elements are strictly positive. Moreover, the higher the leverage, the bigger the profit, so there is no optimum: we can always do better by increasing the value of a .

There is general consensus among financial experts that on real markets such arbitrage opportunities rarely occur and quickly disappear. However, in finite length time series such a dominance may be observed with non-zero probability even between assets that do not have a dominance relation between them on the long run. Clearly, the shorter the time series, the higher the probability of the occurrence of such a finite size dominance.

4 Summary

We have shown that the axioms of [7, 4] imply that for finite T there will always be samples for which portfolio optimization cannot be carried out under a given coherent risk measure, because the measure becomes unbounded from below. We would like to stress again that the point here is not that such an instability can occur for large values of the ratio of the portfolio size N to the sample size T : that is a general feature of any reasonable risk measure. The point is that the optimization under coherent measures can become impossible even for small N/T 's. This is a purely random effect, it depends on the sample. The transition from the feasible to the unfeasible region was found, at least for the Minimax and ES measures, to be smooth for finite N, T and became sharp, with a well-defined critical value of N/T separating the two regions, only for N, T going to infinity with a fixed ratio between them [15]. This transition was identified as an algorithmic phase transition in [15]. (On algorithmic phase transitions see [20].) It may be conjectured that the instability for other coherent measures is similar, although at the present time nothing is known about their critical N/T or the behavior around it.

Paradoxically, this instability is related to a very attractive feature of coherent measures: if one of the assets dominates the rest for all times (that is for infinitely large samples), then the coherent measures signal this by going to minus infinity. (This is a simple corollary of our Theorem 3.1 above). It may happen, however, that in a given finite sample a single asset dominates the others even when there is no such dominance relationship between them for infinitely long observation periods, and the coherent measures become unbounded from below also in this case, thereby giving a false signal. The probability of such a false alarm is known for any finite N and T for the Minimax problem [15], and its behavior around the critical N/T ratio is known for the Expected Shortfall [18]. These results suggest that the probability is falling off rather fast as we go below the critical value of N/T , but in the context of other coherent measures this point would deserve further investigation.

In the somewhat extreme case of the Minimax problem we have been able to show via Theorem 3.2 that the dominance of a single portfolio is not only sufficient but also necessary for the unfeasibility of the optimization. Theorem 3.1 states that the dominance of one portfolio is sufficient for the instability for any coherent measure, but the necessary conditions are not known even for ES. It is evident that they may be more relaxed than for the Minimax: the instability may set in even if none of the portfolios is dominant, but some of them perform exceptionally well on the given sample.

Although we have made use of all four coherence axioms to prove Theorem 3.1, it can be shown that the range of risk measures that suffer from this kind of instability goes well beyond coherent measures (including Value-at-Risk, for instance). The generalization of Theorem 3.1 and exploration of the set of risk measures it applies to will be the subject of an upcoming study.

For the sake of simplicity, we have not put any restriction on short positions in this paper. Infinitely large long and short positions, or infinitely large values of the objective function can, of course, only occur if there is no constraint on short selling. As we briefly argued above, however, possible constraints on the weights can only mask, rather than remove the instability. If we introduce constraints on the weights, they will make the domain within which the optimum is sought finite. It may still happen that one of the assets dominates the rest in a given sample, and then the objective function will monotonically decrease within the allowed domain along some direction. The solution will then be found where this unstable direction runs into the "wall" constituted by the constraint on the weights. For another sample, another unstable direction develops, so ultimately the solutions will not reflect any objective, stable feature of the risk measure, only the nature of random fluctuations. In a way, this is even worse than the appearance of

runaway solutions in the unrestricted case: the fact that the solution sticks to the boundary of the allowed domain is much less conspicuous than the divergence of the risk measure, so it may well go unnoticed.

The introduction of the coherent risk axioms was a great achievement. They grasp some of the most important features we expect from any reasonable risk measure. We have seen, however, that coherent measures are sensitive to sample to sample fluctuations, and break down if one of the assets happens to dominate the others in a given sample. We argue that, in addition to mathematical consistency, noise tolerance, or robustness to sample to sample fluctuations under real life conditions where samples are always too small, is also a highly desirable feature of risk measures. It remains to be seen how much modification of the coherence framework is needed in order to accommodate the requirement of robustness.

Acknowledgements

One of us (I. K.) is grateful to E. Berlinger for a short but inspiring discussion. This work has been supported by the "Cooperative Center for Communication Networks Data Analysis", a NAP project sponsored by the National Office of Research and Technology under grant No. KCKHA005.

References

- [1] G. Szegö. On the (Non)Acceptance of Innovations. In G. Szegö, editor, *Risk Measures for the 21st Century*. John Wiley & Sons, 2004.
- [2] Ph. Jorion. *VaR: The New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, 2000.
- [3] J.P. Morgan. *Risk Metrics*. New York, 1994-1995.
- [4] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent Measures of Risk. *Math. Fin.*, 9(3):203–228, 1999.
- [5] C. Acerbi, C. Nardio, and C. Sirtori. Expected Shortfall as a Tool for Financial Risk Management. Working paper, 2001. (<http://www.gloriamundi.org/detailpopup.asp?ID=453055940>).
- [6] P. Embrechts. Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool. *Derivatives Use, Trading & Regulation*, 6:449–456, 2000.

- [7] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Thinking Coherently. *Risk*, 10(11):68–71, 1997.
- [8] S.S. Wang, V.R. Young, and H.H. Panjer. Axiomatic Characterization of Insurance Prices. *Insurance: Mathematics and Economics*, 21:173–183, 1997.
- [9] S. Hodges. A Generalization of the Sharpe Ratio and its Application to Valuation Bounds and Risk Measures. FORC Preprint 98/88, University of Warwick, 1998.
- [10] P. Carr, H. Geman, and D. Madan. Pricing and Hedging in Incomplete Markets. *Journal of Financial Economics*, 62:131–168, 2001.
- [11] C. Acerbi and D. Tasche. Expected Shortfall: a Natural Coherent Alternative to Value at Risk. *Economic Notes*, 31(2):379–388, 2002.
- [12] C. Acerbi and D. Tasche. On the Coherence of Expected Shortfall. *Journal of Banking and Finance*, 26(7):1487–1503, 2002.
- [13] C. Acerbi. Spectral Measures of Risk: A Coherent Representation of Subjective Risk Aversion. *Journal of Banking and Finance*, 26(7):1505–1518, 2002.
- [14] R. T. Rockafellar and S. Uryasev. Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2(3):21–41, 2000.
- [15] I. Kondor, S. Pafka, and G. Nagy. Noise Sensitivity of Portfolio Selection under Various Risk Measures. *Journal of Banking and Finance*, 31:1545–1573, 2007.
- [16] M.R. Young. A Minimax Portfolio Selection Rule with Linear Programming Solution. *Management Science*, 44:673–683, 1998.
- [17] H. Konno. Portfolio Optimization Using L_1 Risk Function. Technical Report IHSS 88-9, Institute of Human and Social Sciences, Tokyo Institute of Technology, 1988.
- [18] S. Ciliberti, I. Kondor, and M. Mézard. On the Feasibility of Portfolio Optimization under Expected Shortfall. *Quantitative Finance*, 4:389–396, 2007.
- [19] N. Gulyás and I. Kondor. Portfolio Instability and Linear Constraints. To be published, 2008.

- [20] M. Mézard and A. Montanari. Information, Physics and Computation. To be published in 2008 by Oxford University Press, 2008. (<http://www.lptms.u-psud.fr/membres/mezard/>).