Coherent Asset Allocation and Diversification in the Presence of Stress Events

Rebonato, Riccardo and Denev, Alexander

27 April 2011

Coherent Asset Allocation and Diversification in the Presence of Stress Events

R Rebonato[*+], Alexander Denev[+ˆ]
[*]RBS - Client Analytics, GBM
[+]Oxford University - OCIAM
[ˆ]EIB

April 27, 2011

Abstract

We propose a method to integrate frequentist and subjective probabilities in order to obtain a coherent asset allocation in the presence of stress events. Our working assumption is that in normal market asset returns are sufficiently regular for frequentist statistical techniques to identify their joint distribution, once the outliers have been removed from the data set. We also argue, however, that the exceptional events facing the portfolio manager at any point in time are specific to the each individual crisis, and that past regularities cannot be relied upon. We therefore deal with exceptional returns by eliciting subjective probabilities, and by employing the Bayesian net technology to ensure logical consistency. The portfolio allocation is then obtained by utility maximization over the combined (normal plus exceptional) distribution of returns. We show the procedure in detail in a stylized case.

1 Introduction and Motivation

Diversification has been at the core of modern portfolio construction at least since the seminal work by Markowitz in the late 1950s (Markowitz, 1958, 1987). One of the most appealing features of the Markowitz approach is the ability it affords to turn a complex problem of utility maximization under constraints into a simple exercise of optimization of variance-return trade-off. The advantages this approach offers are not just computational (this could have been a major issue when the approach was first introduced, but hardly a consideration in the modern age of fast computers). The main reason for the enduring appeal of the method is the intuitive translation that it affords of the concept of risk into the simple variance statistic. One can discuss at great length the realism of the assumptions about either the utility function (quadratic) or the nature of the returns (Gaussian) required for this result to hold exactly.[1] The fact remains that the intuition it offers to the portfolio manager is powerful, and so are the graphical and conceptual tools it offers (efficient frontiers, minimum-variance portfolios, two-fund theorem, etc). Unfortunately, this intuitive appeal carries a high cost: within the Markowitz framework, 'risk' becomes synonymous with variance; and codependence becomes correlation.

The market dislocations that have punctuated the last decades (the Peso crisis, the South-East Asian currency crises, Russia's default, LTCM, the dot.com bubble, the September 11 events, and, of course, the credit crisis of 2007 onwards) unfortunately keep on reminding asset managers that the paradigm of 'normal markets' presents shortcomings that cannot be ignored: return distributions with the same variance display very different degrees of 'extreme risk' (tail behaviour); and just when most needed all the correlations estimated using the plentiful 'normal' data break down and diversification fails.

This sentiment is well exemplified by in the following quote by Mohamed A. El-Erian, PIMCO's CEO:

> '...[Recent] developments will serve to further highlight the danger of [...] being overly committed to an historical policy portfolio whose rigid backward-looking characterization no longer corresponds to the realities of today and tomorrow....'
>
> '...[W]ith its conventional (or, to be more precise, reduced-form) analytical foundation now subject to some motion, it will become even more difficult to rely just on a traditional portfolio diversification as both necessary and sufficient to deliver high returns and mitigate risks. Diversification will remain necessary, *but a lot more attention will be devoted to the appropriate specification of tail scenarios and tail hedges...*'[2]

This paper starts from the observation that, occasionally but violently, the regularities that prevail during normal times break down: supposedly 'six-standard-deviations events' occur with much higher frequency than the data collected during the normal times would suggest to be possible; and, most important, when this happens correlations tend to become 'polarized' towards 1 or -1. This view is gaining currency not only among practitioners and many economists (see the discussion below), but even among regulators, who had been wedded until recently to a strictly frequentist view of risk. See, for instance the views recently expressed by the Basel Committee:

> '...The financial crisis has shown that estimating *ex ante* the probabilities of stress events is problematic. The statistical relationships used to derive the probability tend to break down in stressed

[1] We note in passing that the quadratic-utility assumption need not be taken literally. Markowitz (1959, 1991) presents quadratic utlity functions as approximation to richer and more realistic utility functions. See pages 286 and *passim*.

[2] Mohamed A. El-Erian, CEO and co-CIO of PIMCO, (2010), *Journal of Portfolio Management*, Vol 36. No 2, Winter 2010, Page 4, my emphasis.

conditions. In this respect, the crisis has underscored the importance of giving appropriate weight to expert judgment in defining relevant scenarios with a forward looking perspective...' (BIS, 2009)

When it comes to integrating these 'impossible' events into a coherent asset-allocation and risk-management framework, some researchers (notably, many in the so-called Econophysics school[3]) believe that these exceptional events present recognizable 'signatures', ie, different, but nonetheless persistent, regularities that transcend the specific cause of each market turmoil.

According to this school of thought, during periods of stress codependencies (both across asset classes and, for each asset class, across time[4]) change, but do so in a predictable and regular way, irrespective of whether the cause of the market turmoil is, say, the bursting of the dot.com bubble, the September 11 attacks or the subprime crisis. This is a fascinating claim, but one for which empirical evidence is by definition very difficult to provide (as we are dealing with very rare events).

Given the rather extraordinary nature of the econophysicists' claim, the burden of proof, or the provision of a reasonable generating mechanism for such 'universal' regularities, should be squarely on the shoulders of the econophysicists. I must stress that the 'proof' I refer to here is *not* that the exceptional events that populate the extreme tails of the joint return distribution have a much higher probability than implied by fast-decaying distributions fitted to the body of the data – this is now a well-established fact. The statement that would be far more interesting, and of practical use for asset allocation and diversification, is that the tail behaviour and tail codependence is not just different, but *stable*[5].

Absent a convincing proof of the econphysicists' claim, many economists have recently taken a rather skeptical view of the ability to provide a frequentist statistical description of future economic events,[6] and a more agnostic approach seems justifiable. The approach presented here indeed takes such an agnostic stance. With the econophysicists we also believe that extreme events belong to a class of their own. But we do not assume that market behaviour in situations

[3] See, eg, Mantegna and Stanley (2000)

[4] See, eg, Sornette and Johanssen (2001)

[5] Indeed, Rebonato and Chen (2009) and Rebonato and Gaspari (2006) analyize the draw-down properites of interest rates and suggest that exceptional events do belong to a different statistical class, as Sornette (2004) and Sornette and Johanssen (2001) suggest. However, they also point out that *several*, and not one, 'signatures' exist, each corresponding to a different way for the market to be 'distressed'. These studies suggest that during exceptional periods market do behave differently, but not always in the same way.

[6] For instance, Rochet (2010) points out, 'financial risks are not exogenous, but arise from the behavior of economic agents. Consequently, the statistical distribution of [...] financial returns is not stationary but depends critically on the economic and regulatory environment, as well as the individual incentives confronting the many economic agents who participate in different aspects of financial intemediation'. See also Davidson (2010), page 18 and passim, for a discussion of limits of the ergodic theorem, or the body of work that goes under the rubric of Imperfect Knowledge Economics for a more genral criticism of the predictability of economic change (see, eg, Frymanand Goldberg (2007, 2009)).

of distress displays a universal 'signature'. Instead we attempt to specify the exceptional codependencies and stand-alone properties of the returns distribution on a case-by-case basis by making use of our (imperfect but useful) understanding of 'how the world works'. As we shall see, the exceptional co-depedencies will be a direct output of the specific causal links between the drivers of exceptional returns that the asset manager will have identified.[7]

The premise of our paper is therefore that, in extreme market conditions, a frequentist approach cannot be relied upon, as there is no *a priori* reason to believe that exceptional events should display stable and predictable universal 'signatures'.[8] The approach we propose, however, does not disregard frequentist information. Rather, it makes the assumption that, during normal market conditions, the co-dependence between the returns of different asset classes is much more stable than that observed from the whole-data sample, and can therefore be identified using traditional statistical techniques.[9]

Once the two components of the joint returns ('normal' and approximately time invariant on the one hand, and 'exceptional' and current-state-of-the-world-dependent on the other) have been identified, a coherent utility maximization brings about the conditionally optimal portfolio.

Of course, the idea of complementing mean-variance-based approaches with *ad hoc* stress testing and scenario analysis is not new. Asset managers have been aware for a long time of the exceptional and unpredictable market behaviour in conditions of distress, and have traditionally tried to 'defend' their portfolios by adding, *after the optimization*, various types of 'insurance trades' (eg, out-of-the-money puts). However, these 'defensive' positions have typically been added to optimized portfolios as an afterthought, and outside a coherent probabilistic and utility-maximization framework. Our work does adopt a probabilistic approach, but we do not equate 'probabilistic' with 'purely frequentist'. We propose that between the fully quantifiable risk of neoclassical economic analysis and Knightian uncertainty there exists a grey area where imperfect, but still useful, probabilistic statements can profitably be made.

This paper therefore attempts to bridge this gap, ie, to provide an intuitively appealing, yet logically justifiable, framework to integrate the statistical information about the diversification that applies in normal times with the altered co-dependencies that prevail in exceptional times. It does so by combining objective (frequentist) probabilities with subjective probabilities: more precisely,

[7] For an interesting contrast of opinions about the merits (or otherwsie) of the econophysics view of financial phenomena, see the positions by Davidson (page 18 and *passim*) and Potters in Skidelski and Wigstroem (2010) (page 3 and *passim*).

[8] We note in passing that our position does not mean that we believe that every new exceptional event must necessarily bring about a completely different codependence structure among the various assets. Perhaps a realtievly small number of crisis patterns may be present. However, which paricular set of codependencies will prevail in a given crisis will in general depend on the specifics of the then-prevailing state of the world. For instance, the correlation between moves in the Fed-administred rates and changes in equity prices generated by the 'Greenspan put' may have been strong and stable in the mid 2000s, but may have disapparead in the post 2007-crisis monetary conditions.

[9] See, eg, Engle (2009) for a state-of-the-art review of, and for innovative suggestions about, techniques for data-based estimation of conditional correlations.

by complementing the traditional association-based approach (based on correlations, covariances and copulae), with a causality-based framework that makes use of our imperfect, but nonetheless useful, understanding of 'how the world works'. Given these two inputs, it provides a coherent portfolio optimization framework that is conditional on the current state of the world. Its subjective inputs are transparent, auditable and challengeable by non-specialists. In this sense the approach here proposed translates and expands to an asset management context many of the ideas introduced for stress testing in Rebonato (2010a), Rebonato (2010b), Rebonato (2010c).

Concretely, we try to offer a solution the problem of static, one-period portfolio optimization under budget constraints. As it is well known, (see, eg, Samuelson (1969)), if the returns are independent and identically distributed (iid) through time and the investor has a constant relative risk aversion, then the one-period (myopic) solution iteratively solves the dynamic problem. Given the emphasis given in the present approach to the state-specific nature of the exceptional returns, the iid hypothesis is clearly not applicable. However, Merton (1971) showed that, if the logarithmic utility function is used, then the iid assumption can be relaxed, and therefore the myopic solution also offers the optimal solution. In our study we shall use both the logarithmic and the power utility function. When the former (logarithmic) choice is applicable, we shall obtain a solution that is also valid in the dynamic setting. For the power-law utility function we shall simply deal with a static, one-period problem. With Milton Friedman (1953), we believe that 'models are to be used, not to be believed.'[10]

Finally, the technique we propose is suited to determining the optimal exposure of a portfolio to a relatively small number of risk factors (say, equity, fixed income, corporate spreads, commodities, etc). Once the allocation weights to various asset classes in the top portfolio has been determined, the same procedure can be cascaded down to the component sub-portfolios (ie, the equity, fixed-income, commodities sub-portfolios).

2 Outline of the Approach

The best way to illustrate the approach is to analyze step by step a stylized example of the proposed procedure.

In the stylized example below we consider the problem of how to choose, given a fixed investment amount (the budget constraint), the optimal composition of a top portfolio, Π, (ie, the weights, $\boldsymbol{w}$, to be given to several sub-portfolios). For simplicity of exposition we assume that each sub-portfolio is invested in a different asset class and that it can be associated with a single dominant risk factor. This assumption makes the exposition clearer, but can be easily relaxed. So, in our example one sub-portfolio could be a long-only US equity portfolio (and the risk factor would be an appropriate US equity index); another sub-portfolio could be a portfolio of US investment grade corporate

[10] quoted in Connor, Goldberg and Korajczyk (2010), page 14.

bonds (and the risk factor would be an investment grade bond index, or, say, a AA spread); etc. Given the stylized nature of the example, *mutatis mutdandis* we can therefore speak interchangeably of a portfolio, of a risk factor or of an asset class.

2.1 Identification of Normal Data

We assume that, for each sub-portfolio, a distribution of (arithmetic returns) is available. The time-t_k arithmetic return for sub-portfolio i, r_i^k, is defined as

$$r_i^k = \frac{x_i^{k+1}}{p_i^k} - 1$$

where p_i^k denotes the price level of sub-portfolio i at time t_k, x_i^{k+1} denotes the value at time t_{k+1} of the same sub-portfolio plus the (reinvested) dividends accrued from time t_k to time t_{k+1}.

The distribution of arithmetic returns (simply 'returns' in the following for brevity) will in general display fatter tails then the Gaussian one. However, 'mundane' and 'regular' phenomena such as, say, stochastic volatility or small or medium-size jumps can give rise to fat tails, without necessarily implying that the market is in a state of distress.

The idea at this point is to distinguish the fat-tailedness that comes from the 'mundane' features from the fat-tailedness that comes from the existence of the truly exceptional events. To do so we must identify the 'body' of a distribution, without necessarily equating the body with its Gaussian part.

Several techniques can be employed to identify the 'body' of a distribution. The simplest (an crudest) one is to truncate all the data points over a threshold (eg, a given number of standard deviations) in the empirical margins, and define as 'body' the rest of the data. This procedure is very simple but rather arbitrary. More sophisticated and 'objective' methods have been developed which aim at eliminating outliers in the data. The earliest papers on the topic (see, eg, Grubbs (1950, 1969), Ferguson (1961)) focused on outliers of the Gaussian distribution. More recent studies broaden the idea to fat-tailed distributions such as t-Student or other power law decaying distributions (see Olmo (2009), Schluter and Trede (2002)). The central idea is that, although heavy-tailed distribution are used by definition to accommodate 'large' returns, there are still values which are so extreme that they cannot be explained by the hypothesized distribution. The outcome of these studies is a statistically formalized test procedure to identify these 'outliers'.

The focus of these studies has been on the unidimensional case. This narrow focus could hide some interesting dynamics in the joint behaviour of asset classes. In order to capture these effects, a number of multi-dimensional techniques have been proposed, that are based on the estimation of the volume of the minimum volume ellipsoid (MVE) that contains the data[11] and of the minimum covariance determinant. The central idea behind these techniques is that the 'body' of the

[11] A detailed description can be found in Meucci (2007).

data is tightly packed while the outliers are much more distant and scattered around it. Given the complexity of the multidimensional problem a numerical approach must be used. However, the output is transparent and lends itself readily to inspection, and for this reason the MVE approach has been used in the present study. The method works as follows.

Given n assets, we call MVE the ellipsoid with the smallest volume in the region of space that contains the observations. The volume of the ellipsoid and the determinant of the covariance matrix are first calculated using the entire data set. The farthest outlier is then removed and the same quantities are recalculated[12]. The procedure is repeated a large number of times. The volume of the ellipsoid and the determinant of the covariance matrix are plotted against the number of observations removed from the data sample. When the first outliers are removed abrupt changes in the volume and in the determinant are observed. As the more and more outliers are removed, both the volume of ellipsoid and the determinant of the covariance matrix are observed to stabilize. Monitoring the changes in these two quantities as a function of the residual number of points in the sample can suggest how many points should be culled.

An example of this procedure is shown in Figures 1 to 8, obtained from the data described later in the paper. Figure 1 shows the volume of the ellipsoid as a function of numbers of observations removed from the data set. Figures 2 and 3 display the determinant of the covariance and correlation matrices, respectively, also as function of numbers of observations. Figures 4 to 6 then show changes in the same quantities.

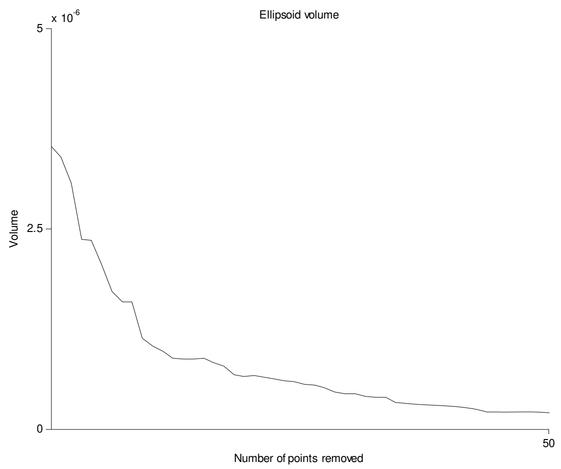

Figure 1: Volume of the ellipsoid as the number of points removed.

For the sake of brevity we do not go into details here. It is however apparent

[12] See again Meucci (2007) for an explanation of how this can be accomplished.

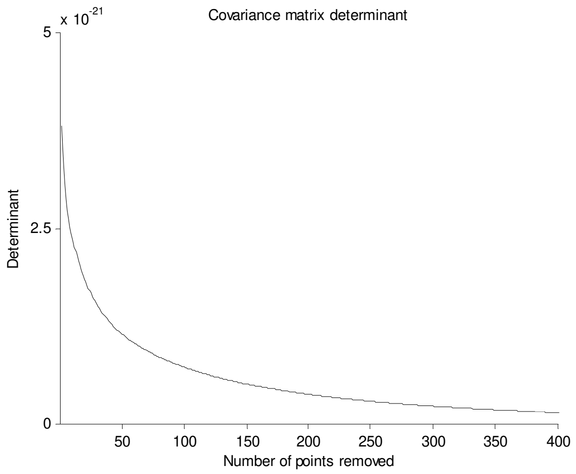

Figure 2: Determinant of the covariance matrix as a function of the number of points removed.

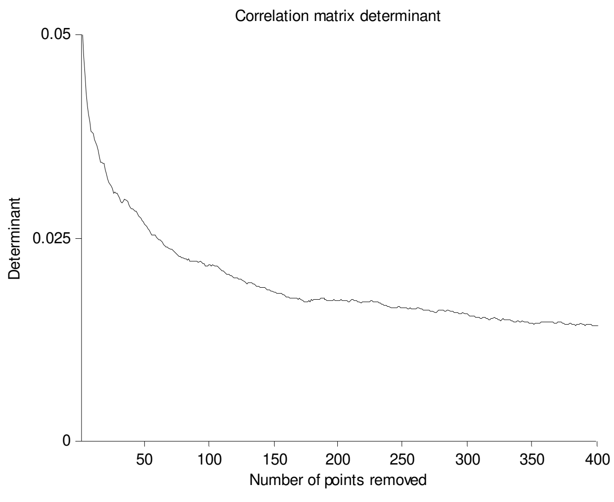

Figure 3: Determinant of the correlation matrix as function of the number of points removed.

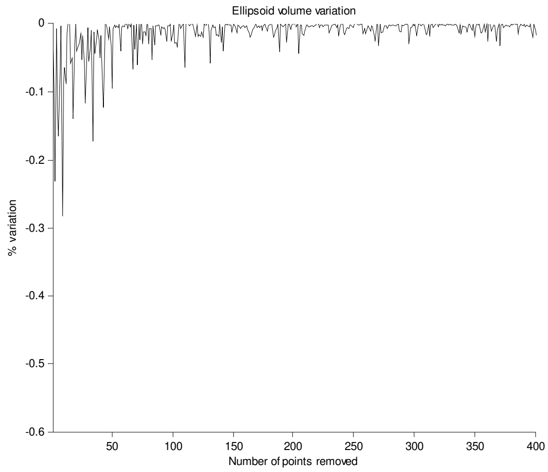

Figure 4: Percentage changes for Fig 1

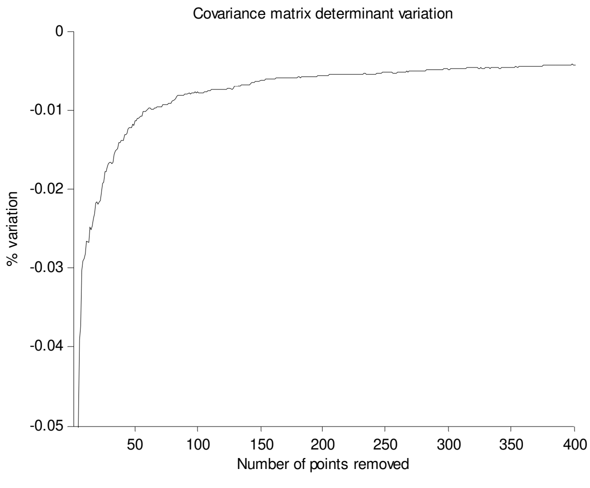

Figure 5: Percentage changes for Fig 2

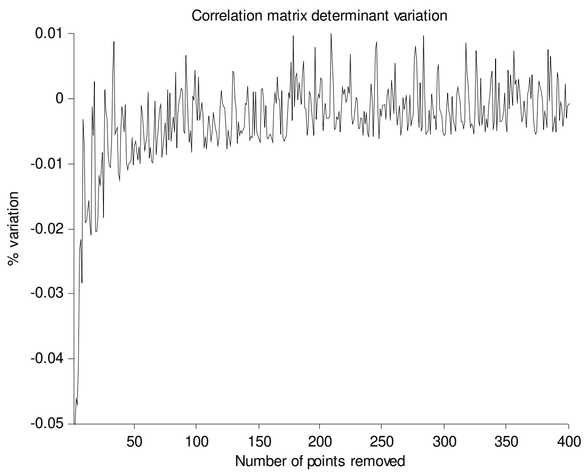

Figure 6: Changes for Fig 3

that excluding 100 to 200 points (3% to 5% of the whole data set described in detail in Section 3, footnote 27) renders the covariance and correlation structure much more stable. This is what we need for our purposes: we want to ensure that, after the exceptional events have been culled, the marginal distributions of and, above all, *the correlation among*, the variables should be reasonably stable under changes of estimation methodology. Figure 7 shows that this is indeed the case by displaying the individual elements of the correlation matrix among the *Bond*, *Credit*, *Mortgage* and *Equity* time series described in the following as a function of the number of points removed. Figure 8 displays the stabilization of the four eigenvalues of the correlation matrix. Finally, Figure 9 shows the scatterplot of the body of the data and the outliers for the Equity, Bond and Credit returns.

As we shall see, as long as these requirements are met, the exact identification of which data points belong to the body and the tail of the distribution is in itself less important.

We note in closing this section that Bouchaud and Potters (2000, 2003) find that, by excluding the highest eigenvalues from the orthogonalization of an empirical correlation matrix obtained from S&P returns, the resulting efficient frontier becomes much more stable as a function of the temporal subsection of data used (eg, first half versus second half). They also find that the 'purged' correlation matrices become much more stable over time. They point out that, even if an investor availed herself of the knowledge of the realized returns, investing according to a Markowitz prescription using all the eigenvectors of the correlation matrix would result in much riskier *ex post* portfolios than if the investment had been made using a small number of stable eigenvalues. Bouchaud and Potters tend to explain this by positing that the highest eigenvalues are

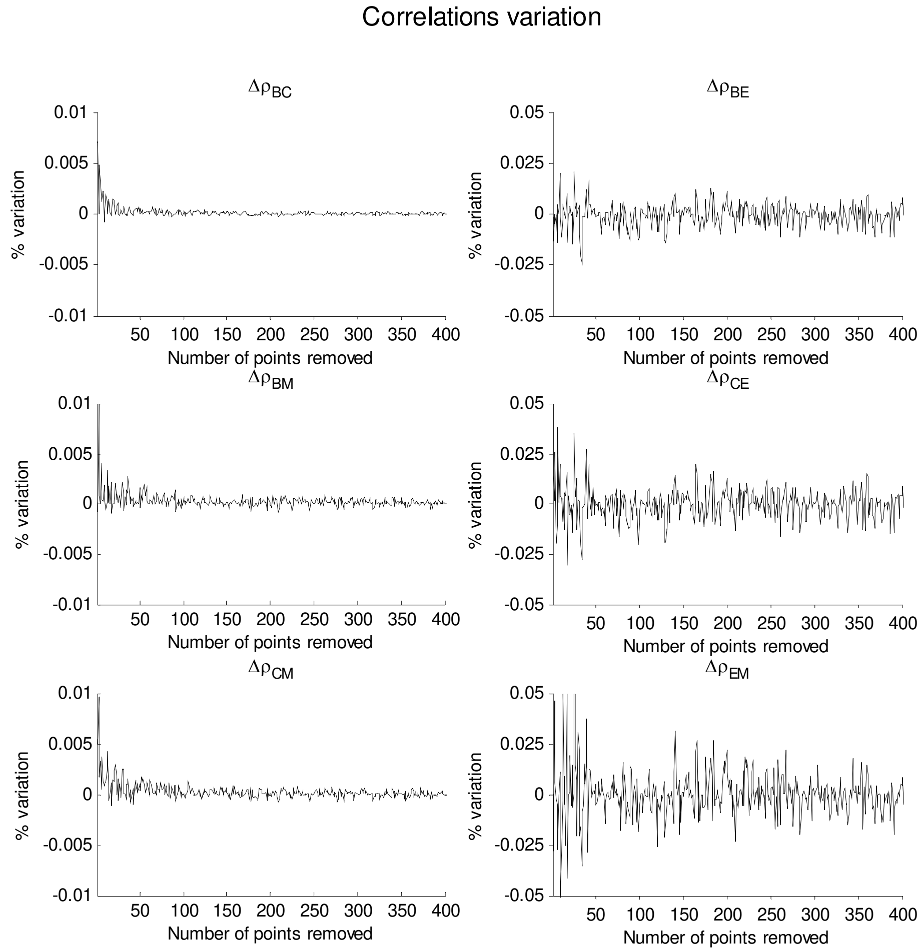

Figure 7: Changes in specific correlation elements as a function of the number of points removed. The symbol ρ signifies correlation and the subscripts C, B, M and E denote Credit, Bond, Mortgages and Equities, respectively, as described in Section 3.

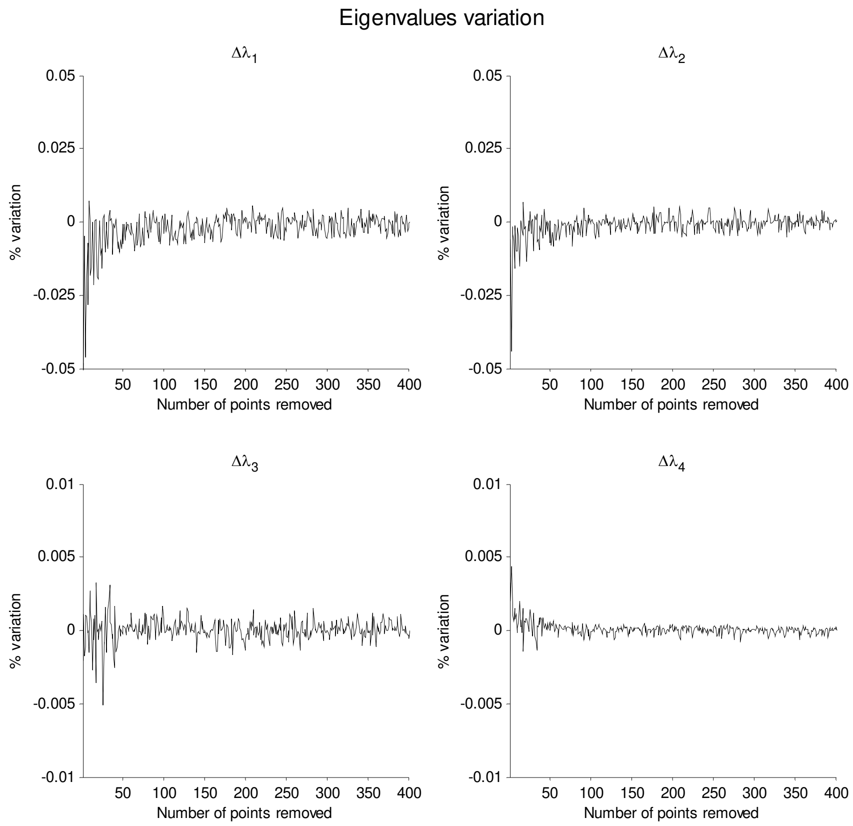

Figure 8: Changes in the eigenvalues as a function of the number of points removed.

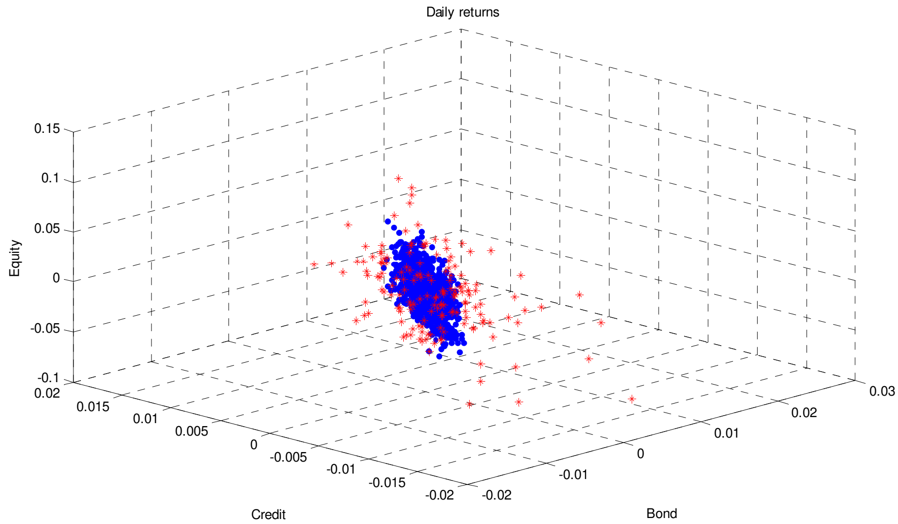

Figure 9: The ellipsoid with its body (circles) and 'outliers' (stars).

simply 'picking up noise'. They concede, however, that the effect may be due to 'genuine time dependence in the structure of meaningful correlations'.[13] The latter, of course, is the explanation we implicitly predominantly subscribe to in our work.

2.2 Constructing the Marginal Distribution of the 'Normal' Risk Factors

Once the 'normal'[14] portion of the data has been identified as suggested in the previous section, a parametric distribution can be fitted to the empirical distribution of 'normal' returns obtained for each individual risk factor. Depending on the risk factor and the sampling interval, a Gaussian distribution may or may not be adequate.[15] If for each risk factors the normal portion of the data can be satisfactorily described by a Gaussian distribution, we shall see that important computational savings can be achieved (because a simple closed-form expression can be obtained linking the distribution of portfolio returns and the

[13] Bouchaud and Potters (2000), page 120.

[14] In this article the adjective 'normal' is often used referred to data to mean 'non-exceptional'. To avoid confusion, the word 'normal' is never used in the statistical sense, for which sense the word 'Gaussian' is invariably used. To further avoid confusion, the adjective 'normal' is often placed among quotes (as in 'normal').

[15] For most asset classes, daily return display strong deviations from normality. For quarterly returns the Gaussian approximation is much better. See, eg, Connor, Goldberg and Korajczyk (2010), Chapter 1, Figure 1.3a to 1.3f, and the discussion in the text.

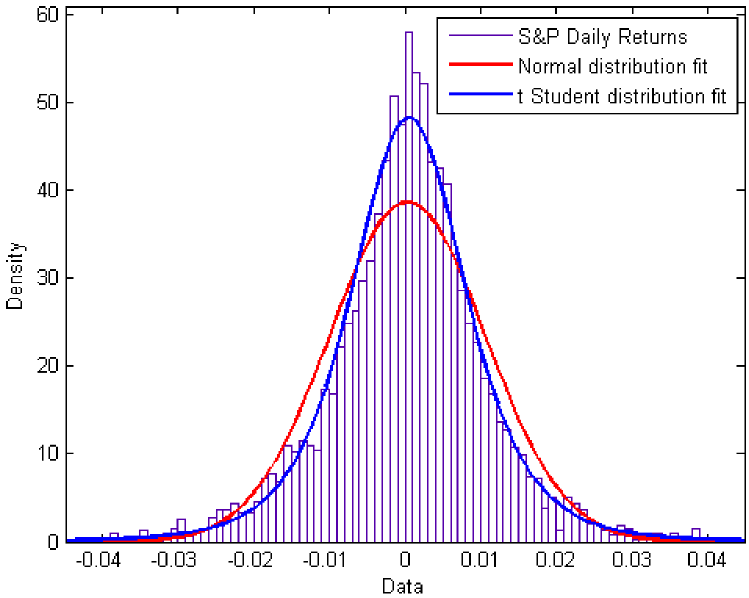

Figure 10: The fit to the S&P dialy returns obtained using a Gaussian and a t-Student distribution.

weights of the individual sub-portfolios). If this is not the case, the procedure described in the following can still be applied, but the final numerical search becomes somewhat more burdensome. Conceptually, however, nothing changes.

For daily returns, even the body of most time series is not well described by a Gaussian distribution. We find that often a Student-t distribution does an acceptable job. We also find that the results (eg, the variances of the marginals and, as we shall see, the correlation matrix) are not sensitive to the details of the chosen parametric distribution. However, for the procedure to remain computationally feasible, it is important that the chosen distribution should be extendable in the multi-variate case to a member of the elliptical class.

As an example, Figures 10 and 11 present the unidimensional fit in the form of a histogram and a Q-Q plot for the asset class *Equity* using the Gaussian and the t-Student distributions.

2.3 Constructing a Joint Distribution of 'Normal' Risk Factors

The procedure described in the previous subsection will produce univariate marginal distributions for each risk factor. These marginals may display fatter tails than Gaussians, but nonetheless refer to non-extreme market conditions. The

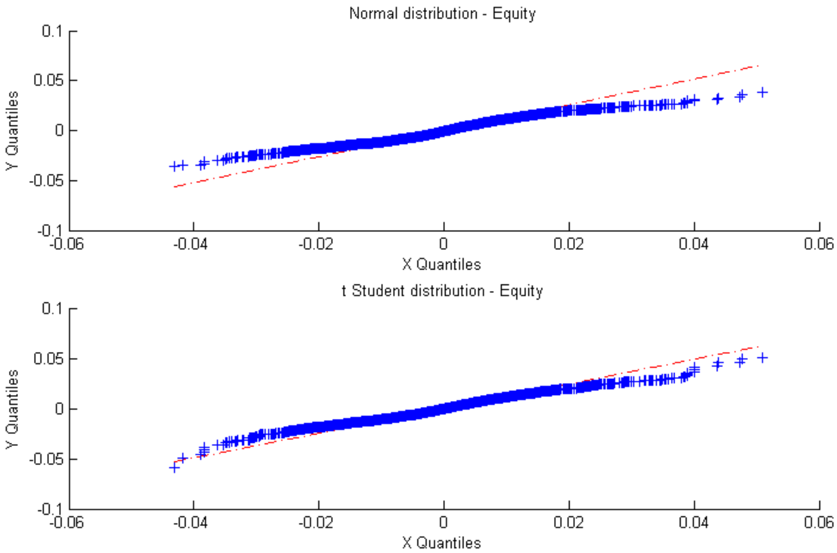

Figure 11: The Q-Q plot for the two fits in Figure 10.

assumption is now made that they can be conjoined using a pairwise copula. A t-Student copula is used in this example. If this choice is made the marginal distributions are mapped on a percentile-by-percentile basis to a t-Student distribution, and the correlation matrix between each pair can be determined.

An example of this procedure is shown in Figures 12 and 13, which show a percentile-by-percentile scatter plot of the returns for assets *Equity* and *Bond* after fitting the marginals to Student-t distributions, and after joining them by means of a Gaussian copula.

Broadly speaking, there are three methods to estimate the parameters of a copula from data (see, eg, Cherubini, Luciano and Vecchiato, 2004):

1. Exact Likelihood Method: a one-step method that estimates jointly the parameters of the marginals and those of the copula. Statistically, it is the most efficient method, but it is computationally very onerous.

2. Inference from Marginals Method: a two-step procedure whereby the marginals are fitted first, and the copula then. This is the method that has been used in this study.

3. Canonical Maximum Likelihood Method: it estimates directly the parameters of the copula without fitting the marginals, as the empirical marginals are used instead.

Assessing the goodness of fit for copulae is not straightforward, and the methods suggested in the literature (which by and large require the calculation

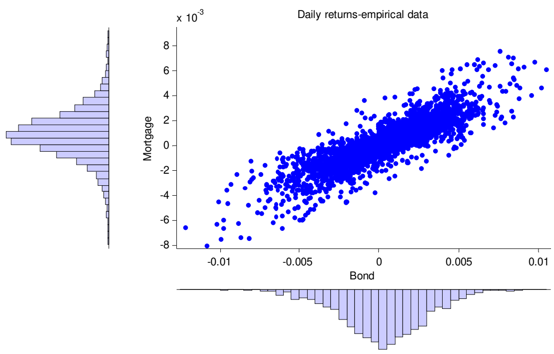

Figure 12: Scatterplot of the empirical results for the assets Mortages and Bonds with the fitted Student-t Copula

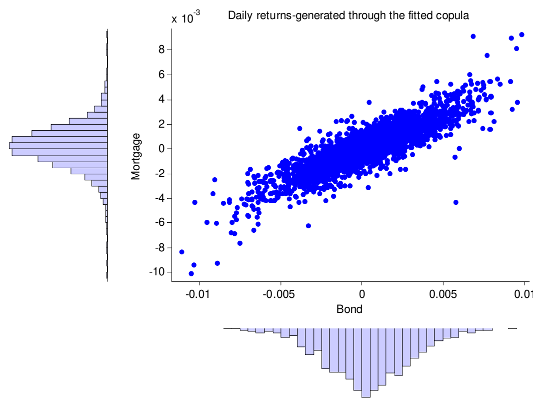

Figure 13: Scatterplot of the generated random numbers with the same fitted coupla for the same asset classes.

of the multidimensional distance between the empirical copula and the fitted one) tend to be numerically demanding.

Goodness-of-fit (GOF) measures for copulas are quite well described in Genest et al (2009). Appendix I presents a brief discussion of the merits of competing methods.

Once the marginal distributions have been conjoined using the copula of choice, the modelling of the 'normal' portion of the data is complete.

2.4 Treatment of Exceptional Events

The next step in the proposed procedure is the specification of those tail events that, given the particular macroeconomic and financial conditions of the moment, may be expected to have a large effect on a given portfolio. These significant events could be identified on the basis of macro analysis, or starting from the vulnerabilities of a portfolio, as suggested in Rebonato (2010a) in a stress-testing context[16]. The analysis should be limited to a handful of 'extreme but plausible'[17] occurrences that have the potential to inflict very serious damage (or, indeed, to generate exceptional gains) in a given portfolio.

In approach we propose these extreme events are modelled as Boolean variables, as variables, that is, that can assume 'true' or 'false' logical values. For instance, at the time of this writing, the possibility of the break-up of the EMU could be one such event.

2.5 Constructing a Bayesian Net

The Boolean variables defined above can be organized in a Bayesian net. Bayesian nets are directed, acyclical graphs that can provide both an intuitive description of the causal links among variables, and a synthetic representation of (and an algorithm to construct) the joint probabilities among them. (See Pearl (2009), Williamson (2005) for a treatment of Bayesian nets in general, and Rebonato (2010a, 2010b, 2010c) for an application to stress testing.)

A simple example of Bayesian net is shown in Fig 14, that shows a situation where variable A exerts a causal influence on variables D and C, variable B affects variable C, and variables A and B are independent of each other.

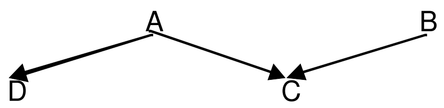

Figure 14: A simple Bayesian net displaying causal links amomg four varibales, as discussed in the text.

[16] See the discussion in Chapter 13 in particular.

[17] This is the expression used by Bank of International Settlements (2009) in a stress-testing context.

The specification of the causal structure (the topology) of a Bayesian net is one of the most important (and most delicate) steps in the procedure proposed here. This is where the asset manager makes use of her understanding of the world today. This information, imperfect as it may be, is invaluable: we know, say, that a fall in equities will cause equity implied volatilities to rise, and not the other way around. Yet, when we use associative measure (such as correlations and copulae) we discard this information. We do so at a great informational loss.

In order to move from the structure of the net (its topology) to the joint probabilities of the various elementary events (that, as shown, will be required for a consistent asset allocation) one has to assign the so-called conditional probability tables. One such table is associated with each variable. Each table in turn contains the marginal probability of the variable associated with its node, and the conditional probability of occurrence of that variable, given the occurrence of its direct parents[18]. For instance, in Fig 14 this means that the conditional probability table for node C will contain the marginal probability of occurrence of C, $P(C)$, and the conditional probabilities of C given the occurrence of A and given the occurrence of B, $P(C|A)$, $P(C|B)$.

In general, the highest order of conditioning (single, double, etc) depends on the maximum number of parents any given node is allowed to have. Once the conditional probability tables are given, the joint probabilities of all possible combinations of the chosen rare events can be obtained. As discussed in Pearl (2009), Moore (2001) and Rebonato (2010a), independence and conditional independence are the tools by which one can move from marginal and lowly-conditioned probabilities (the 'easy' probabilities) to joint probabilities (the 'difficult' probabilities).

Since, as discussed below, the inputs to the tables are arrived at using a mixture of frequentist, implied and subjective probabilities, this imposes a 'cognitive' limit on the complexity of the underlying net[19]. This is, admittedly, a limitation of the approach. However, assigning a simplified picture of 'how the world works' is certainly better than providing no such information at all – at least as long as the inputs are transparent and auditable and sensitivity analysis is easy to carry out.

The next section shows how the conditional probability tables can be filled in. It must be stressed at this stage, however, that assigning conditional probabilities can often be simpler than specifying the marginal probabilities for the same events: it may be very difficult to assign a probability for, say, a 1987-like market crash or for an overnight doubling of equity implied volatilities. However, one can safely venture that the conditional probability of the increase in equity implied volatility *given the equity market crash* should be well above 50%. For the purposes of our approach, this degree of precision is more than adequate.

[18] If a node has no parents, then just the marginal probability appears in the conditional probability table.

[19] It must be stressed that, for reasonable application, the limit of the proposed approach is not at all comptational, as the whole procedure is not CPU-time-intensive.

2.6 Building the Conditional Probability Tables

Conditional probability tables are made up of marginal and conditional probabilities. Since subjective probabilities are often to be elicited from the asset manager, only reasonably 'simple' probabilistic statements can be asked for. As far as marginal probabilities are concerned, market-implied information (such as spreads of credit default swaps) can also be profitably used. When this information is not available (or deemed not reliable[20]) the subjective input of the asset manager is in general required. Frequentist information is, of course, always available as a fall-back, complement or 'sanity check'. For conditional probabilities rarely do market-implied quantities provide useful information, and greater reliance must therefore be placed on subjective input.

As far as the actual assignments are concerned, spurious precision should be avoided. For the marginal probabilities only an order-of-magnitude assessments is required (probably nothing more precise than once-a-year, once-every-few-years, once-a-business-cycle, once-a-century statements can be made). As for the conditional probabilities, once the marginal probabilities have been (approximately) elicited, the singly-conditioned probabilities can be assigned to 'buckets' determined by asking the asset manager whether the occurrence of event A decreases (increases) the occurrence of event B by a little, a lot, makes no difference or makes the occurrence of B almost a dead certainty.

It is well known that elicitation of conditional probabilities is a delicate task, and that many cognitive biases make it harder than it seems[21]. Rebonato (2010) discusses at length elicitation techniques that can make the task easier and less error-prone. Given these cognitive difficulties in assigning conditional probabilities, it is not uncommon that the subjectively-assigned conditional probabilities may end up being incoherent. This means that there exists no set of (non-negative, smaller-than-one) joint probabilities from which the conditional probabilities could have been obtained. When this is the case, well-established techniques (typically based on Linear programming) exist to 'cleanse' the incoherent conditional probability matrix, and return the closest (in some sense) coherent conditional probability table. See, eg, Kwiatkowski and Rebonato (2010), de Kluyvert and Moskowitz (1984), Moskowitz and Safrin (1983) and Gilio (1995).

2.7 From the Conditional Probability Tables to the Joint Probabilities of Rare Events

After completing the previous steps the asset manager has at her disposal a Bayesian net, which describes the assumed causal relationships among the rare events, and the associated conditional probability tables. In building this

[20] for instance because the objective probability may be 'drowned' by liquidity and/or risk premia, or by imbalance of suply and demand.

[21] The most common bias occurs when the conditional probability is interpeted in a diagnostyic rather than causal direction. See, eg, Tversky and Kahneman (1979). Overestimation of conditonal probability when a causal link is assumed to exist is another common bias.

net, she has only been asked to provide 'easy' (ie, cognitively resonant and/or market-implied) probabilities. However, Nature does not generate elementary stress events in isolation. What the asset manager therefore really needs are the 'difficult' joint probabilities of occurrence of the various possible combinations of the rare events. For n events there are $2^n - 1$ such probabilities to assign (the -1, of course, comes from the normalization condition). Even for a relatively small number of variables, this number can be rather large (for as few as, say, 6 elementary rare events, 63 joint probabilities have to be assigned). But it is not the sheer number of probabilities that creates a problem; it is the cognitive difficulty in assigning, say, the probability of events A, E and F occurring, and B, C and D not occurring[22].

This is where the Bayesian-net technology provides the required bridge between the 'easy' probabilities the asset manager can feel confident to assign and the 'difficult' probabilities that she requires. For the example in Fig 14, for instance, only four marginal, three singly-conditioned and one doubly-conditioned probability have to be assigned in order to specify fully the 16 joint probabilities.

Once the topology of the net has been chosen, and all the conditional probability tables have been filled in, the construction of the joint probability is a mechanical task, that can always be accomplished by the relationship[23]

$$P(E_1 = e_1, E_2 = e_2, ..., E_{n-1} = e_{n-1}, E_n = e_n) = \prod_{i=1}^{n} P(E_i = e_i | \text{Parents of } E_i) \quad (1)$$

2.8 Mapping from the 'Normal' Joint Distribution to the 'Normal' Return Distribution

At this point in the procedure we have established the joint distribution for the risk factors, both in the 'normal' and in the exceptional regime. We now have to map these joint distributions onto the univariate distribution of portfolio returns. We start from the joint distribution of the 'normal' risk factors. We assume that an exact or approximate evaluation engine is available to translate from changes in the risk factors to changes in the portfolio returns. This 'evaluation engine' can be as simple as multiplication of a basis point move by a PVBP sensitivity, to a full revaluation of a complex option valuation model.

If the 'normal' body of the marginal distributions of sub-portfolio returns has been satisfactorily modelled by univariate Gaussian distributions conjoined by a Gaussian copula, and all the products in the portfolio display a linear dependence on the risk factors in the sub-portfolios, then the univariate distribution

[22] Alternatively and equivalently, absent any additinal information, five-times-, four-times, three-times and doubly conditioned probabilities would have to be assigned.

[23] I leave the precise definition of an expression such as $P(E_i = e_i | \text{Parents of } E_i)$ somewhat to the intuition of the reader. In reality it it possible to introduce a partial ordering among nodes in a directed net, and to write the conditioning , in a way that a piece computer code can understand. For the sake of brevity, I simply refer to the reader to Moore (2001) or Pearl (2009).

of portfolio returns is normal, with mean and variance given by

$$\boldsymbol{\mu}_{\Pi(\mathbf{w})} = \boldsymbol{w}^T \boldsymbol{\mu}$$

$$var\left[\Pi(\mathbf{w})\right] = \boldsymbol{w}^T \boldsymbol{C} \boldsymbol{w}$$

where $\boldsymbol{\mu}$ denotes the vector of risk-factor returns and $\boldsymbol{C}$ the covariance matrix of the returns. The notation $\boldsymbol{\mu}_{\Pi(\mathbf{w})}$ and $var\left[\Pi(\mathbf{w})\right]$ emphasizes the dependence of the portfolio on the vector of weights, $\mathbf{w}$. In the portfolio optimization over weights described below this allows a straightforward analytical link between the weights and the portfolio return distribution.

When the asset manager feels that the marginal distributions of sub-portfolio returns cannot be satisfactorily modelled by a Gaussian distribution, or when the portfolio contains strongly non-linear products, then the mapping from the joint distribution of risk-factors to the univariate portfolio return distribution must be carried out by Monte Carlo simulation. This is, of course, more computationally demanding, but, given the relatively small number of sub-portfolios, does not pose a prohibitive burden on the whole procedure. Moreover, resampling techniques can be used, and they are known to give rise to reduced sensitivity to the estimated parameters. This is a well-known problem with the Markowitz optimization technique, as the results can be highly variable even for small changes in input parameters (see, eg, Ceria and Stubbs (2004)). This was one of the drivers behind the well-known Black-Litterman approach (Black and Litterman (1992)). Unlike the Black-Litterman model, where the sensitivity is diminished by smoothing the inputs, in the resampling procedure the same result is reached by averaging the outputs of several scenarios.

2.9 Mapping from the Exceptional Joint Distribution to the Exceptional Return Distribution

If the evaluation engine mentioned above is available, the mapping from the joint distribution of the extreme occurrences of the rare events to the corresponding portfolio returns is straightforward. Each joint event will give rise to a unique return, to which a Dirac-δ distribution is associated, with 'area' equal to its joint probability.

2.10 Splicing the 'Normal' and 'Exceptional' Portions of the P&L Distribution

One of the probabilities of the joint events obtained from the Bayesian net (the probability $P(\widetilde{E}_1, \widetilde{E}_2, ..., \widetilde{E}_N)$) is the probability that none of the N exceptional events, $E_1, E_2, ..., E_N$, that the asset manager has identified will occur. If the asset manager is sufficiently confident in the Bayesian net part of the exercise, the joint probability $P(\widetilde{E}_1, \widetilde{E}_2, ..., \widetilde{E}_N)$ directly provides the required 'normalization' of the 'normal' (continuous) and exceptional (Dirac-δ) contributions. Call $P(0)$ the probability that none of the extreme events materializes. Then the area of the discrete portion of the distribution (that describes the non-exceptional events) is simply equal to $1 - P(0)$. All the other combination of elementary extreme events would then have exactly the probabilities implied by the Bayesian net.

Alternatively, the asset manager can take a less 'pure' but more pragmatic approach, and treat the quantity $P(0)$ as a 'free parameter' of the procedure – call it k. In this 'pragmatic' interpretation, the quantity k can be profitably compared with the 'degree of confidence' (also known as the 'shrinking parameter') in the Black-Litterman (1992) approach. We prefer to call it the 'normalization factor'.

If this pragmatic choice is made, the 'normal' portion of the distribution is given mass $1 - k$, and the sum of the masses of Dirac-δ distributions are renormalized (uniformly scaled) so as to have total mass k. As shown in the following, this allows the determination of the sensitivity of the outputs to an important part of the procedure.

2.11 Maximization of the Utility

By this stage of the procedure we have obtained, for any set of sub-portfolio weights $\boldsymbol{w}$, a univariate composite distribution of returns, which contains information both about the 'normal' market relationship, and about the specific links among the asset classes that we have posited to prevail if the particular market dislocations we have identified were to materialize. For a given set of sub-portfolio weights $\widehat{\boldsymbol{w}}$, the univariate distribution would therefore contain information both about the diversification that can be expected in normal market conditions, and about the particular co-movement of asset classes that are expected to materialize if, say, the Euro were to break up. This information, of course, is not present in any data base – yet any attempt at diversification that did not take into account, albeit in an approximate manner, this possibility would be both misguided and dangerous. The special danger would stem from the fact that real prices and spreads observed in the market would embed the market's conditional knowledge that the Euro may break up. However, an optimization based on a data set that neglected this eventuality would see 'trading opportunities' and 'anomalous risk premia' where the market actually points to clear signs of danger.

The next step of the procedure is the choice of a utility function. The pros and cons of various choices have been discussed in literally thousands of papers. For the application at hand, it is useful to choose a utility function that allows a simple and continuous 'tuning' of the degree of risk aversion.

In the application described below, we therefore make use of a power utility function, parametrized by the coefficient β:

$$U_{power}(c) = \frac{1}{1-\beta}\left(c^{1-\beta} - 1\right)$$

so as to give degrees of risk aversion greater and smaller than the logarithmic function (that obtains in the limit as $\beta \Longrightarrow 1$):

$$U_{\log}(c) = \lim_{\beta \to 1} \frac{1}{1-\beta}\left(c^{1-\beta} - 1\right) = \ln(c)$$

For the logarithmic function, as mentioned in the introduction, a one-period optimization ceases to be myopic, in the sense that a static optimization coincides with (the first step of) a dynamic optimization. We take the end-of-period wealth generated by the portfolio as a proxy for consumption.

The elements of the vector of weights, $\boldsymbol{w}$, are varied under the budget constraint

$$\boldsymbol{w}^T \mathbf{1}^{\boldsymbol{n}} = 1$$

until the object function (the end-of-period utility) is maximized. The numerical search over the simplex of the non-negative weights can be carried out very efficiently in an unconstrained manner using the trigonometric technique suggested in Rebonato and Jaeckel (1999).

The result provides a coherent allocation to the various sub-portfolios, where the adjective 'coherent' stresses that the allocation has been arrived at be taking into account in a consistent manner the investor's preferences over the outcomes associated both with normal and exceptional market conditions. 'Protection trades' are not attached as an incoherent afterthought to an optimization carried out assuming a stable investing universe.

The sensitivity of the outputs to the subjective inputs can be readily explored, as is shown in the worked-out example below.

3 A Worked-Out Example

In order to illustrate the strengths and weaknesses of the procedure, we present in this section a simple worked-out example in some detail.

The stylized problem we examine in this section is that of the asset allocation among four asset classes, Government Bonds, Investment-Grade Credit Bonds, Equities and High-Grade RMBS Securities (called asset class *Bond*, *Credit*, *Equity* and *Mortgage*, respectively, in the following.)[24]

The analysis to identify the body of the distribution was carried out as described above. The cut-off point was determined using a combination of the ellipsoid and the minimum covariance determinant methods. 160 data points were excluded by the algorithm, corresponding to approximately 5% of the

[24] More precisely, the following indices were used:

- for *Bond* the BarCap US Treasury Index;
- for *Credit* the BarCap US Credit Index;
- for *Equity* the S&P 500;
- for *Mortgage* the BarCap US MBS Index.

The data set consisted of 3360 x 4 data points, covering the period February 1997 to June 2010.

full data set. We then observed that, when the body of the distribution was truncated in this manner, its correlation structure remained much more stable over the whole observation period than if the whole data set had been used. This check is important because, in a way, the assumption of a relatively stable correlation structure during 'normal' periods underpins our whole approach.

The marginals of the truncated data set were modelled using a t-Student distribution, and were conjoined through a t-Student copula. Figs 12 and 13 present illustrations of this data analysis for *Equity* (marginal results), and for *Bond* and *Credit* for the copula.

The traditional inputs to a Markowitz optimization (ie, the correlation matrix and the stand-alone expected returns and volatilities) are given below[25].

Correlation	*Bond*	*Credit*	*Equity*	*Mortgage*
Bond	1	0.96	−0.22	0.87
Credit	0.96	1	−0.16	0.87
Equity	−0.22	−0.16	1	−0.11
Mortgage	0.87	0.87	−0.11	1

	Expected Return
Bond	6.90%
Credit	7.90%
Equity	7.80%
Mortgage	6.40%

	Volatility
Bond	4.4%
Credit	4.6%
Equity	18.4%
Mortgage	2.8%

The Bayesian net in Fig 14 (repeated below with new labelling for ease of reference) was assumed, with the following interpretation for the four events:

- Event A: Large sell-off in the Government Bond market due to unexpected change in monetary stance, with 1994-like increase by 75 bp of the Fed target rate.
- Event B: Large widening of Mortgage spreads due to the forced selling of inventories prompted by new capital rules.
- Event C: Large widening of Credit spreads.
- Event D: Sell of in the Equity market.

In this simplified example it is assumed that both the sell-off in Government Bonds and the widening of mortgage spreads would increase the probability of

[25] In real-life applications, the expected returns are likely to be assigned using a subjective Black-Litterman (1992) approach. This extension has not been incorporated in the present analysis for the sake of brevity, but it is conceptually straightforward.

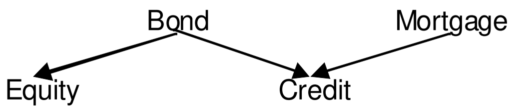

Figure 15: Fig 6: The Bayesian net connecting the four stress events associated to the four asset classes.

investment-grade spreads widening significantly. It is also assumed that the sell-off in Government Bonds and the widening in mortgage spreads do not directly affect each other (given the specific reason for the mortgage weakening). Finally the sell-off in Government Bonds is assumed to have a strong negative effect on the equity market. In constructing this example it was assumed that the event were identified on the basis of the vulnerabilities of each portfolio. This being the case, each stress event is simply associated with a single portfolio[26]. This simplifies the presentation, and the example could be straightforwardly generalized.

Given the known composition of each subportfolio, x, the stress loss, $L(x)$, that would occur if the associated stress event occurred with certainty can be readily calculated. These subportfolio losses are given in the vector below:

$$\left[\begin{array}{cc} & \text{Stress Losses} \\ L(Bond) & -5\% \\ L\left(Credit\right) & -20\% \\ L\left(Equity\right) & -10\% \\ L\left(Mortgage\right) & -2\% \end{array}\right]$$

For clarity, each entry gives the loss if the stress event associated with each asset class materialized, and the whole portfolio had been invested in that asset class.

The marginal probabilities of occurrence, $P(x)$, of the four stress events were chosen as follows:

$$\left[\begin{array}{cc} & \text{Marginal probability} \\ P(Bond) & 0.04 \\ P\left(Credit\right) & 0.07 \\ P\left(Equity\right) & 0.07 \\ P\left(Mortgage\right) & 0.03 \end{array}\right]$$

Given the causal structure embedded in the Bayesian net above, only three singly-conditioned probabilities, $P(x|y)$, are required in order to specify the full joint probability matrix. For this example these were assumed to be as follows:

$$\begin{array}{cc} & \text{Conditional probability} \\ P(Credit|Bond) & 0.40 \\ P(Credit|Bond) & 0.25 \\ P(Equity|Bond) & 0.50 \end{array}$$

[26] Needless to say, this example is overly stylized. It is only intended to give a flavour for the type of analysis and the results.

The doubly-conditioned probability, $P(Credit|Bond, Mortgage)$, was assumed to be 50%.[27] The resulting joint probabilities are shown in the table below.

$$\begin{bmatrix} Bond & Credit & Equity & Mortgage & \text{Joint Probability} \\ 0 & 0 & 0 & 0 & 0.8381 \\ 0 & 0 & 0 & 1 & 0.0208 \\ 0 & 0 & 1 & 0 & 0.0460 \\ 0 & 1 & 0 & 0 & 0.0446 \\ 1 & 0 & 0 & 0 & 0.0117 \\ 0 & 0 & 1 & 1 & 0.0011 \\ 0 & 1 & 0 & 1 & 0.0065 \\ 0 & 1 & 1 & 0 & 0.0025 \\ 1 & 0 & 0 & 1 & 0.0003 \\ 1 & 0 & 1 & 0 & 0.0117 \\ 1 & 1 & 0 & 0 & 0.0077 \\ 0 & 1 & 1 & 1 & 0.0004 \\ 1 & 0 & 1 & 1 & 0.0003 \\ 1 & 0 & 1 & 1 & 0.0003 \\ 1 & 1 & 1 & 0 & 0.0077 \\ 1 & 1 & 1 & 1 & 0.0003 \end{bmatrix} \tag{2}$$

Note that, given the assumed causal dependence among the four events, the most unlikely event is not the simultaneous occurrence of all the four stress events. For instance, the joint events:

- only *Bond* and *Mortgage* happen,
- *Equity*, *Bond* and *Mortgage* events happen with no widening of spreads credit spreads
- *Credit*, *Mortgage* and *Bond* happen without a sell-off in *Equity*

have the same probability of occurrence as 'all events happen simultaneously'. A glance at the posited causal structure confirms that this makes intuitive sense.

The asset manager can either 'believe' that the probability of no stress event provided by the Bayesian net procedure is fundamentally correct (in which case the mass of the 'normal' distribution would be 0.8381); or, as suggested above, can treat the normalization factor as a free parameter, akin to the Black-Litterman 'degree of confidence'.

Given this information, and for any vector of weights $\boldsymbol{w}$, the portfolio return can be calculated. A logarithmic and a power utility function (with exponents

[27] As the subjective conditional probabilities are obtained by elicitation, the asset manager may find it too difficult to venture a gues for $P(Credit|Bond, Mortgage)$. Moskowitz and Safrin (1983) show that very tight bounds for the joint probabilities can be obtained even with a small number of marginals and singly-conditioned probabilities. See also the discussion and the examples in Rebonato (2010a).

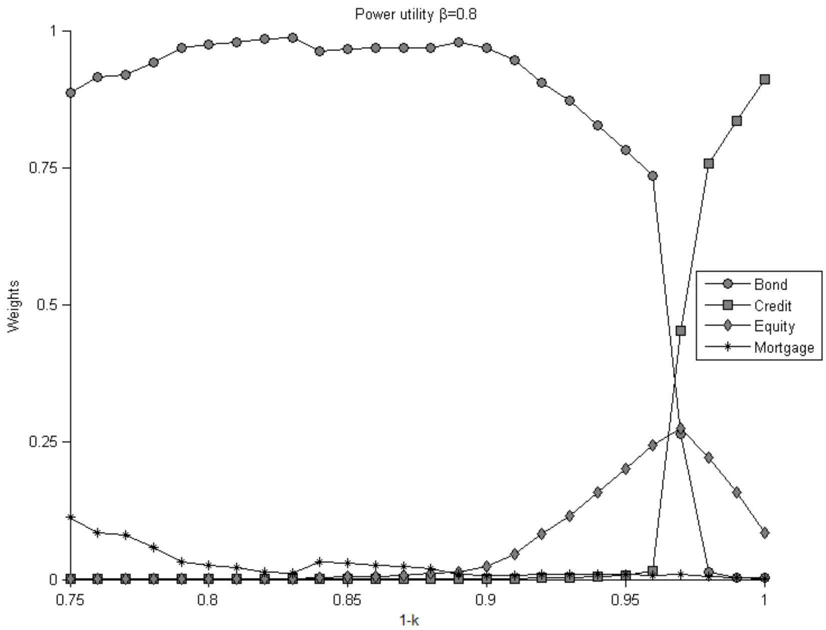

Figure 16: Asset alloaction for a power utility function with $\beta = 0.8$. The asset alloaction at the far right (corresponding to $k = 0$) represents the allocation that would be obtained by making use purely of the statistical information about the body of the distribution, and ignoring the Bayesian net.

β of 1.2 and 0.8 to straddle in terms of risk aversion the logarithmic utility) were chosen for the optimization.

The results are shown below for three different degrees of risk aversion. The asset allocation is shown as a function of the mass in the 'normal' body of the distribution. (This quantity is referred to as the 'normalization factor' in the captions below.) Therefore the asset allocation that can be read on the right vertical axis is (apart from the choice of a non-quadratic utility function) a 'Markowitz-like' allocation: the allocation that would be chosen if the exceptional events never materialized, and correlations and variances were therefore homoskedastic. The smooth curves that evolve to the left of these four points show how these allocations change as the probability of the world entering a distressed state increases. The results are discussed in detail in the following section.

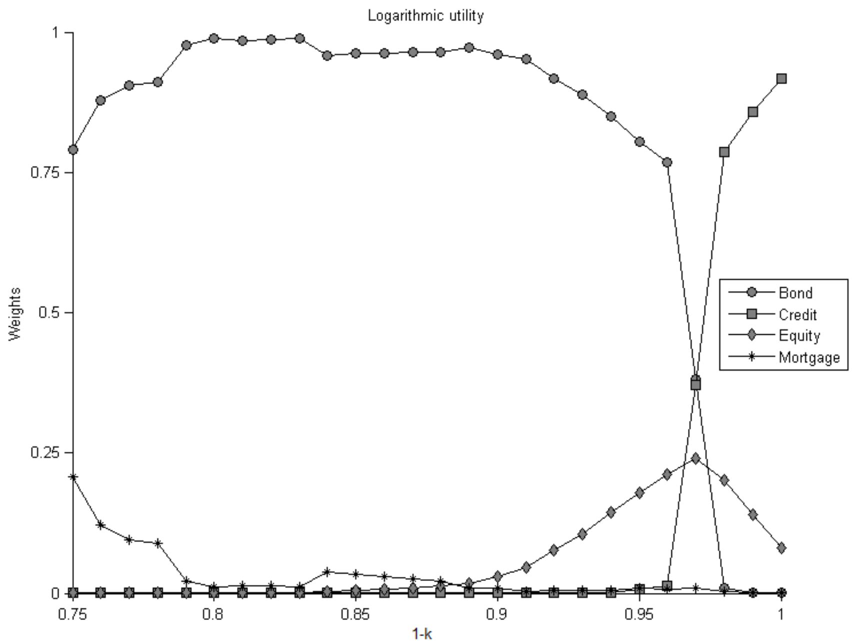

Figure 17: Same as the figure above, for the logarithmic utility function.

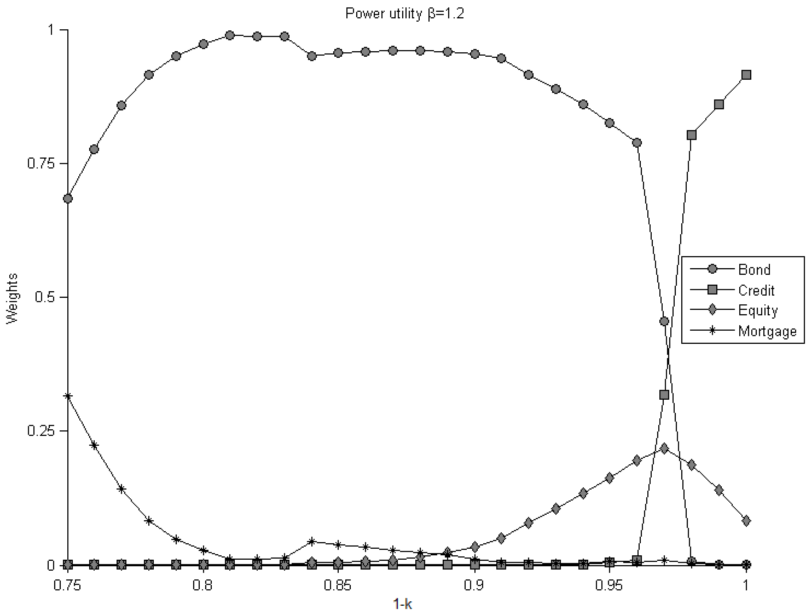

Figure 18: Same as the figure above, for a power utilty function with $\beta = 1.2$.

4 Discussion of the Results

The first observation is that the qualitative features of the results display a reasonably mild dependence on the degree of risk aversion. See Figs 16 to 18. Given the well-known difficulties in estimating this quantity in a reliable manner, this is a nice robustness feature of the procedure.

One can then profitably begin the analysis from the central case of the logarithmic utility function (see Fig 17). In the 'Markowitz-like' case (no exceptional events) the allocations are divided among two of the four asset classes: *Credit* and *Equity*. In this benchmark case, the highest allocation is for the *Credit* subportfolio (83%). This is not surprising, given the high return expected from this asset class. The *Credit* subportfolio, however, has been assumed to be the most vulnerable to the associated stress event ($L(Credit) = 20\%$). Therefore, as the probability mass in the 'normal' state decreases, the allocation to *Credit* quickly decreases. Indeed, for all degrees of risk aversion, the allocation to the *Credit* rapidly goes to zero as soon as the probability of the world not remaining in a normal state over the investment horizon is non-negligible. This result is both intuitive and, to some extent obvious. It pays stressing, however, the value of such an 'obvious' recommendation: in some situations the expected returns from some assets may appear very attractive (when their risk is assessed by looking at the historical record) exactly because the market is pricing in the vulnerability to events that have not materialized yet. In the run-up to the 2007 crisis, for instance, exotic instruments such as Constant-Proportionality Debt Obligations (CPDOs) commanded a AAA rating, but a yield of 200 basis points above other government AAA debt. Similarly, the attractive yields for peripheral European government debt in the spring and summer of 2010 may have appeared inexplicably attractive if gauged on their historical record. A frequentist-based asset allocation technique would suggest very high allocations to these 'inexplicably cheap' assets.

The non-monotonic behaviour of the allocation to the *Equity* and to the *Bond* subportfolios is less intuitively obvious and deserves some discussion. The allocation to *Equity* is low close to the 'normal' state but as the probability of stress events increases, ie, moving towards the left in the three graphs, its allocation begins to increase. This is easy to understand, because the stress loss associated with the *Equity* subportfolio is lower than that for *Credit* ($L(Equity) = 10\%$ vs $L(Credit) = 20\%$), but their marginal probability of occurrence is the same: 7%. In order to understand the non-monotonic behaviour of the *Equity* allocation, we have to look at the event correlation implied by the Bayesian net assumed above[28]:

$$\begin{bmatrix} \text{Event Correlation} & Bond & Credit & Equity & Mortgage \\ Bond & 1 & 0.26 & 0.34 & 0.00 \\ Credit & 0.26 & 1 & 0.09 & 0.12 \\ Equity & 0.34 & 0.09 & 1 & 0.0 \\ Mortgage & 0.00 & 0.12 & 0.00 & 1 \end{bmatrix}$$

From the table we can see that the probability of joint occurrence of the loss associated with $Equity$ and $Credit$ is low: 0:09. A simple calculation of the total expected return (ie, the expected return including stress events) then shows that this quantity decreases less rapidly as we move towards the left in the graphs above for $Equity$ than for $Credit$, but the two assets continue to retain a negative *total* correlation (ie, a correlation that takes into account both the normal and the excited states). Therefore, as long as the probability of being in the stressed state is low, the diversification benefit for allocating between $Equity$ and $Credit$ arising from the negative correlation in the "normal" correlation matrix is not completely lost. However, as we move further to the left (ie, as the probability of entering the excited state increases) the associated losses both to $Equity$ and to $Credit$ start to be non-negligible compared to the other two assets and their total expected returns becomes comparable to them.

Let's now examine closely the allocation to the *Bond* asset class. To understand how it changes with respect to the allocation to $Credit$, we have to consider again the total expected return. Note first that the 'Markowitz-like' (no-stress) allocation to *Bond* is 0 for all degrees of risk aversion. This is because of its low expected return compared to $Credit$ with which it is highly correlated (96%). The allocation to *Bond* replaces completely the allocation to $Credit$ after $1-k$ approaches 4% since its total expected return starts to be comparable to that of $Credit$ (due to a lower loss $L(Bond) = 5\%$ and lower marginal probability of occurrence, $P(Bond) = 4\%$). Indeed, a calculation of the total expected return for *Bond* and $Credit$ shows that, for $k = 96\%$, they become very similar (1.69% versus 1.64%), but the total standard deviation for *Bond* is significantly lower (2.34% versus 3.70%). Furthermore, *Bond* is more negatively correlated (and therefore can better diversify) with $Equity$ than with $Credit$ (-18% versus -6% for that level of the normalization factor, k) For these reasons *Bond* starts replacing $Credit$ as a 'partner' of $Equity$.

The subportfolio *Bond* replaces $Credit$ but its allocation weight does not keep on increasing: instead it starts to shift gradually to $Mortgage$. This happens because its stress marginal probability of occurrence is higher than that of $Mortgage$, ($P(Mortgage) = 3\%$ vs $P(Bond) = 4\%$) but it has the same stress loss ($L(Mortgage) = L(Bond) = 5\%$). The total return from the two assets becomes comparable for $1-k = 0.75$. However, $Mortgage$ starts rising before that point because it consistently enjoys lower volatility than *Bond*..

[28] For a discussion of how to obtain an event correlation from the joint probabilities, see Rebonato (2010a) and Tzani and Polychronakos (2008). The event correlation matrix clearly has nothing to do with the correlation matrix among the returns of the assets.

5 Conclusions and Extensions

A simple method has been presented to carry out a coherent programme of asset allocation based on utility maximization in the presence of stress events with subjectively assigned probabilities. The method relies on the ability of the asset manager to specify in an approximate way the causal links, if any, among the stress events that can most affect a given portfolio.

The approach has several advantages. First of all, it is very transparent. Its crucial 'ingredients' (ie, the subjective probabilities, the assumed causal links, the fraction of the probability in the distresses state, etc) are clearly visible, and affect the result in an intuitively understandable manner. This helps the asset manager's intuition, as the discussion in the previous section shows.

The approach also lends itself to thorough and 'thoughtful' sensitivity analysis, as literally every input can be 'understood', changed 'by hand' and the results inspected. Thanks to its intuitional appeal, the approach therefore lends itself to analysis and scrutiny by non-specialists, trustees and other senior investment officers.

It is easy to show that no other constraints are required of the input conditional probabilities in Equation (1) than being non-negative and smaller or equal to one. This has two important and positive consequences. First, for the size of Bayesian nets required for these applications it is very easy to carry out sensitivity analysis by varying the inputs over a confidence range around the central input value – this can be done without having to worry about the admissibility of the resulting joint probabilities. Second, if the expert felt unable to provide some of the required conditional probabilities, these could be obtained using the principle of Maximum Entropy for the joint distribution (and again the search can be carried out in an unconstrained manner). There is a vast literature on the topic (see, eg, Markham and Rhodes (1999), Garside, Rhodes and Holmes (1999), Schramm and Fronhoeffer (2005)), but for the application at hand most of the computational complexities (that arise from having to deal with hundreds and sometimes thousand of variables) can be avoided. For a real-life application with 10 to 15 nodes and up to three parents per node the entropy maximization is almost instantaneous.

The framework relies heavily on subjective inputs, but, in the light of the transparency and auditability of the approach, this should be viewed as a strength, not a drawback.

6 Appendix I

The computational complexity of the methods use to asses the goodness of fit (GOF) of copulae is much higher than those used to assess the GOF of one-dimensional distributions. The usual starting point is the construction of the empirical copula:

$$C_n(\mathbf{u}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(U_{i1} \leq u_1, ..., U_{id} \leq u_d) \tag{3}$$

$$\mathbf{u} = (u_1, u_2, ..., u_d) \in [0,1]^d$$

The second step is to use a kind of Kolmogorov-Smirnov or Anderson-Darling test for the distance between this copula and a hypothesized copula. However, the empirical copula may become very hard to estimate especially in the case of many assets and observations. The storage of information is another problem.

Malevergne and Sornette (2001) propose another method, much faster, to tackle this problem for a Gaussian copula. They show that under the hypothesis H_0 that the dependence between random variables can be described by a Gaussian copula, the variable:

$$z^2 = \sum_{j,i=1}^{d} \Phi^{-1}(F_i(x_i))(\rho^{-1})_{ij}\Phi^{-1}(F_j(x_j)) \tag{4}$$

follows a χ^2 distribution with d degrees of freedom. $\Phi^{-1}(.)$ is the inverse cumulative standardized normal distribution, the F_i are the cumulative marginal distribution functions and d is the number of assets. The matrix ρ is defined as:

$$\rho_{ij} = Cov[\Phi^{-1}(F_i(x_i)), \Phi^{-1}(F_j(x_j))]$$

They propose the following four distance measures:

$$KS : d_1 = max_z|F_{z^2}(z^2) - F_{\chi^2}(z^2)|$$

$$AverageKS : d_2 = \int |F_{z^2}(z^2) - F_{\chi^2}(z^2)|dF_{\chi^2}(z^2)$$

$$Anderson - Darling : d_3 = max_z \frac{|F_{z^2}(z^2) - F_{\chi^2}(z^2)|}{\sqrt{F_{\chi^2}(z^2)[1 - F_{\chi^2}(z^2)]}}$$

$$AverageAnderson - Darling : d_4 = \int \frac{|F_{z^2}(z^2) - F_{\chi^2}(z^2)|}{\sqrt{F_{\chi^2}(z^2)[1 - F_{\chi^2}(z^2)]}}dF_{\chi^2}(z^2)$$

The KS distances are more sensitive to deviations in the bulk of the distributions. On the contrary, the AD statistics are more sensitive to the tails. A deviation from statistics of common use is the presence of moduli in the numerator instead of squares. The advantage is that such distances are less sensitive to outliers. The disadvantage is that standard statistical tests (ω-test and Ω-test) cannot be used. However this disadvantage disappears, since the covariance matrix is not known but is estimated from data, and in any case the exact parameters needed in the derivation of such statistics are not known. To overcome this. a bootstrap method can be used (see, eg, Efron and Tibshirani (1986)).

The procedure can be extended also to a t-Student copula as done in Kole et al. (2007). They show that if H_0 is the hypothesis that the dependence structure comes from a t-Student copula with correlation matrix ρ and ν degrees of freedom then the variable:

$$z^2 = \sum_{j,i=1}^{d} \Psi^{-1}(F_i(x_i), \nu)(\rho^{-1})_{ij}\Psi^{-1}(F_j(x_j), \nu)/d$$

where $\Psi^{-1}(F_i(x_i), \nu)$ is the inverse of the cumulative standard Student's t distribution with degrees of freedom parameter ν, is distributed according to an F-distribution with degrees of freedom d and ν. The same steps as in the case of Gaussian copula with obvious modifications can be applied to this variable and a distribution of the four GOF distances estimated.

7 Bibliography

Bank of International Settlements (BIS), (2009), *Principles for Sound Stress Testing and Supervision* – Consultative Paper, January 2009, latest version May 2009

Black, F and Litterman, R, (1992), *Global Portfolio Optimization*, Financial Analysts Journal, September 1992, pp. 28-43

Bouchaud, J-P, Potter, M, (2000, 2004), *Theory of Financial Risk – From Statistical Physics to Risk Management*, Cambridge University Press, Cambridge, UK

Ceria, S., Stubbs R.A. (2004), *Incorporating Estimation Errors into Portfolio Selection: Robust Efficient Frontiers*, Axioma Inc. Technical Report

Cherubini, U, Luciano, E, Vecchiato, W, (2004), *Copula Methods in Finance*, John Wiley & sons, Chicester

Cochrane, J, (2004), *Asset Pricing*, Princeton University Press, Princeton, NJ

Connor, G, Goldberg L R, Korajczyk R A, (2010), *Portfolio Risk Analysis*, Princenton University Press, Princeton, NJ

de Kluyver, C, A, Moskowitz, H, (1984), *Assessing Scenario Probabilities Via Goal Programming*, Management Science, Vol 30, (March 1984), No 3, 273-278

Efron B, Tibshirani R, (1986), *Bootstrap Methods for Standard Errors, Confidnce Intervals and Other Measures of Statistical Accuracy*, Statistical Sceince, Vol 1, number 1, 54-75

El-Erian M A, (2010)*, Looking Ahead,* Journal of Portfolio Management*,* Vol 36. No 2, Winter 2010, Page 4

Frydman R, Goldberg M D, (2007), *Imperfect Knowledge Economics: Exchange Rates and Risk*, Princeton University Press, Princeton, NJ

Frydman R, Goldberg M D, (2009), *Financial Markets and the State: Long Swings, Risk and the Scope of Regulation, Capitalism and Society*, Vol 4, Issue 2, Article 2, available at: http://www.bepress.com/cas/vol4/iss2/art2, last accessed 20th November, 2009, Berkeley Electronic Press

Davidson P, (2010), *Risk and Uncertainty*, in *The Economic Crisis and the State of Economics*, Skidelski and Wigstroem, Editors, Palgrave, Macmillan, New York

Engle, R, (2009), *Anticipating Correlations: A New Paradigm for Risk Management*, Princeton Univeristy Press, Princeton, NJ

Ferguson, T S, (1961), *On the Rejection of Outliers*, in Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability,pp. 253—287.

Friedman, M, (1953), *Essays in Positive Economics*, University of Chicago press, Chicago, IL

Garside, G R, Rhodes, P C, Holmes, D E, (1999), *The Efficient Estimation of Missing Information in Causal Inverted Multiway Trees*, Knowledge-Based Systems, 12, 101–111

Genest, C, Remillard, B, Beaudoin D, (2009), Goodness-of-Fit Tests for Copulas: *A Review and Power Study*, Insurance: Mathematics and Economics, Volume 44, issue 2, April 2009, 199-213

Gilio, A, (1995), *Probabilistic Consistency of Conditional Probability Bounds*, Lecture Notes in Computer Science, Vol 945, 200-209

Grubbs, F E, (1969), *Procedures for Detecting Outlying Observations in Samples*, Techno-
metrics, 11, 1, 1-21.

Kole, E, Koedijk K, Verbeek M, (2007), *Selecting Copulas for Risk Management*, Journal of Banking & Finance, Vol. 31, number 8, pag. 2405-2423

Kwiatkowski, J, Rebonato, R, (2010), *A Coherent Aggregation Framework for Stress Testing and Scenario Analysis*, accepted for publication in *Applied Mathematical Finance*

Malevergne, Y, Sornette, D, (2001), *Testing the Gaussian Copula Hypothesis for Financial Assets Dependences*, Finance 0111003, EconWPA

Mantegna, R N, Stanley, H E, (2000), *An Introduction to Econophysics – Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, UK

Markham, M J, Rhodes, P C, (1999), *Maximixing Entropy to Deduce an Initial Probability Distribution for a Causal Network*, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol 7, No 1, 63-80

Markowitz, H, (1959, 1991), *Portfolio Selection*, Blackwell, Oxford

Markowitz, H, (1987), *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford

Merton R, (1971), *Optimum Consumption and Portfolio Rules in a Continuous Time Model*, Journal of Economic Theory, 3, 373-413

Meucci, A, (2007), *Risk and Asset Allocation*, Springer Finance, Berlin

Moore, A, (2001), *Bayes Nets for Representing and Reasoning About Uncertainty*, Carnegie Mellon University, http://www.autonlab.org/tutorials/bayesnet09.pdf, last accessed 18th August, 2009

Moskowitz, H, Sarin, R K, (1983), *Improving the Consistency of Conditional Probability Assessment for Forecasting and Decision Making*, Management Science, Vol 29, No 6, (June 1983), 735-749

Olmo J.(2009), *Extreme Value Theory Filtering Techniques for Outlier Detection*, City University Economics Discussion Papers No 09/09

Potters, M, (2010), *Lessons from Statistical Finance*, in *The Economic Crisis and the State of Economics*, Skidelski and Wigstroem, Editors, Palgrave, Macmillan, New York

Pearl, J , (2009), *Causality*, 2^{nd} edition, Cambridge University Press, Cambridge, UK

Rebonato, R, (2010a), *Coherent Stress Testing – A Bayesian Approach to Financial Stress*, John Wiley, Chichester

Rebonato, R, (2010b), *A Bayesian Approach to Stress Testing and Scenario Analysis*, Journal of Investment Management, Vol. 8, No. 3, (2010), pp. 1–13

Rebonato, R (2010c), *A Bayesian Approach to Coherent Stress Testing*, in *Rethinking Risk Measurement*, Klaus Bloeker Editor, Incisive Media, London

Rebonato, R, Chen J, (2009), *Evidence for State Transition and Altered Serial Co-dependence in US$ Interest Rates*, Quantitative Finance, vol 9, No 3, April 2009, 259-278

Rebonato, R, Gaspari V, (2006), *Analysis of Drawdowns and Drawups in the US Interest-Rate Market*, Quantitative Finance, 6(4), 297-326

Rebonato, R, Jaeckel, P, (1999), *The Most General Methodology to Create a Valid Correlation Matrix for Risk Management and Option Pricing Purposes*, Journal of Risk, Vol 2, Number 2, Winter 1999/2000, 17-28

Rochet, J-C, (2010), *The Future of Banking Regulation*, Chapter 3 in *Balancing the Banks - Global Lessons from the Financial Crisis*, M Dewatripont, J-C Rochet and J Tirole, Princeton University Press, Princeton, NJ

Samuleson P, (1969), *Lifetime Portfolio Selection by Dynamic Stochastic Programming*, Review of Economics and Statistics, 51, 239-246

Schluter, C, Trede M, (2002), *Identifying Multiple Outliers in Heavy-Tailed Distributions with an Application to Market Crashes*, Journal of Empirical Finance Vol. 15, issue 4, 700-713

Schramm, M, Fronhoefer, B, (2005), *Completing Incomplete Bayesian Networks*, in G. Kern-Isberner, W. Roedder, and F. Kulmann (Eds.): WCII 2002, LNAI 3301, pp. 200–218, Springer-Verlag, Berlin Heidelberg.

Sornette, D, (2004), *Critical Phenomena in Physical Sciences*, Springer Verlag, Berlin

Sornette, D and Johansen, A, (2001), *Significance of Log-Periodic Precursors to Financial Crashes*, Quantitative Finance, Vol 1, 452-471

Tzani, R, and Polychronakos A P, (2008), *Correlation Breakdown, Copula Credit Models and Arbitrage*, GARP Risk Review, December 2008, 27-37

Tverski, A, Kahneman, D, (1979), *Causal Schemata in Judgments under Uncertainty*, in Progress in Social Psychology, M. Fishbein (Ed.), Lawrence Erlbaum Associates, Hillsdale, 1979

Williamson, J, (2005), *Bayesian Nets and Causality – Philosohpical and Computational Foundations*, Oxford University Press, Oxford