OFFICE OF FINANCIAL RESEARCH
U.S. DEPARTMENT OF THE TREASURY

Office of Financial Research
Working Paper #0007
April 9, 2013

Stress Scenario Selection by Empirical Likelihood

Paul Glasserman,[1] Chulmin Kang,[2] and Wanmo Kang[3]

[1] Columbia University*, pg20@columbia.org
[2] Korea Advanced Institute of Science and Technology, cm0222@kaist.ac.kr
[3] Korea Advanced Institute of Science and Technology, wanmo.kang@kaist.edu

The Office of Financial Research (OFR) Working Paper Series allows staff and their co-authors to disseminate preliminary research findings in a format intended to generate discussion and critical comments. Papers in the OFR Working Paper Series are works in progress and subject to revision.

Views and opinions expressed are those of the authors and do not necessarily represent official OFR or Treasury positions or policy. Comments are welcome as are suggestions for improvements, and should be directed to the authors. OFR Working Papers may be quoted without additional permission.

www.treasury.gov/ofr

* This paper was produced while Paul Glasserman was under contract with the Office of Financial Research.

Stress Scenario Selection by Empirical Likelihood

Paul Glasserman[*], Chulmin Kang[†], and Wanmo Kang[‡]

December 2012; this version, April 2013

Abstract

This paper develops a method for selecting and analyzing stress scenarios for financial risk assessment, with particular emphasis on identifying sensible combinations of stresses to multiple factors. We begin by focusing on reverse stress testing — finding the most likely scenarios leading to losses exceeding a given threshold. We approach this problem using a nonparametric empirical likelihood estimator (in the sense of Owen (2001)) of the conditional mean of the underlying market factors given large losses. We then scale confidence regions for the conditional mean by a coefficient that depends on the tails of the market factors to estimate the most likely loss scenarios. We provide rigorous justification for the confidence regions and the scaling procedure in three models of the joint distribution of the market factors and portfolio loss with qualitatively different tail behavior: multivariate normal (light-tailed), multivariate Laplace (exponentially tailed), and multivariate-t (regularly varying). The key to this analysis (and the differences across the three cases) lies in the asymptotics of the conditional variances and covariances in extremes. These results also lead to asymptotics for marginal expected shortfall and the corresponding variance, conditional on extreme losses; we combine these results with empirical likelihood significance tests of systemic risk rankings based on marginal expected shortfall. For the problem of selecting macro stress scenarios, we apply our results to estimate the most likely outcome for other variables given a stress in one variable, and thus to gauge the plausibility of particular combinations of stresses to financial and economic factors. Finally, we develop a scenario sampling method, suggested by the empirical likelihood contours, for exploring regions of large losses in generating stress scenarios.

1 Introduction

Stress testing has long been part of the risk management toolkit, but it has gained new prominence through the recent financial crisis. This is reflected, for example, in the impact of the Supervisory Capital Assessment Program conducted by U.S. financial regulators in 2009 (Hirtle et al. [16]), the subsequent Comprehensive Capital Assessment Reviews in 2010 and 2011 (see [5]), the corresponding stress tests undertaken by the European Banking Authority, the stress testing requirements in the Dodd-Frank Act, and greater use of stress testing for internal risk management reported in industry surveys ([15, 19]).

[*]Columbia University, pg20@columbia.edu
[†]Korea Advanced Institute of Science and Technology, cm0222@kaist.ac.kr
[‡]Korea Advanced Institute of Science and Technology, wanmo.kang@kaist.edu

Stress testing seeks to evaluate losses in extreme yet plausible scenarios that may be underweighted in a probabilistic model of market movements and absent from a historical backtest. An important challenge in designing effective stress tests lies in selecting scenarios that are indeed both sufficiently extreme and sufficiently plausible to improve risk management. Recent research and recommendations on stress testing include Alfaro and Drehmann [2], Breuer et al. [7], Financial Services Authority [12], Flood and Korenko [13], Pritsker [22], Quagliarello [23], Rebonato [24], and Schuermann [26]. Borio et al. [6] provide a critical review of current practice.

Our objective in this article is to develop a data-driven procedure to inform the selection of scenarios that are both extreme and plausible. Our primary focus is on *reverse* stress testing, which seeks to identify scenarios that result in losses exceeding a given magnitude for a particular portfolio or firm. Because many different combinations of movements of market factors can produce losses of similar magnitude, we formulate the goal of reverse stress testing more precisely as one of identifying the most likely scenario or scenarios among all such combinations. These scenarios are, by definition, of primary importance to a particular portfolio, whereas purely hypothetical scenarios often seem arbitrary and their consequences therefore difficult to interpret. With a single risk factor, it may be relatively clear in which direction and even by how much to stress the factor to get a plausible adverse outcome, but identifying a sensible *combination* of stresses to multiple factors requires further analysis. This is one of the main challenges in defining stress scenarios.

We view the selection of stress scenarios as an exploratory process. Reliance on a single scenario — even the most likely one — is potentially misleading, so our objective is to identify important regions of stress scenarios, where importance reflects both the likelihood of the outcome and the severity of the resulting loss. These regions should be anchored in the available data, though data on extreme outcomes is necessarily limited. We also want to be able to draw scenarios from the important regions in a way that is consistent with the available data.

We approach the problem of identifying reverse stress testing regions in two steps. First, we estimate the conditional mean of the underlying market factors given a portfolio loss exceeding a specified level. Then we scale the conditional mean by a multiplier that depends on the tails of the market factors to correct for the ratio of the conditional mean to the most likely loss scenario.

For the first step, the estimation of the conditional mean, we use an empirical likelihood estimator, in the sense of Owen [21]. Empirical likelihood (EL) is a nonparametric estimation procedure through which we get confidence regions for the conditional mean. Importantly, the EL estimator does not rely on significant assumptions about the conditional distribution of the market factors in extremes. The shape of the resulting confidence regions is able to capture skewness and other features present in extreme outcomes.

For the second step in our procedure — scaling the conditional mean — we derive asymptotically exact scaling multipliers in three special but important models of the joint distribution of market factors and portfolio value: multivariate normal, multivariate Laplace, and multivariate t. These represent three fundamental cases in the sense that the normal distribution has very light tails, the Laplace distribution has tails that decay exponentially, and the t distribution has regularly varying tails; the second and third cases are particularly relevant to market data. In all three cases, we provide rigorous justification for the scaling factor we derive and for the combination of the scaling factor and the EL estimator; this combination yields asymptotically valid confidence regions for the most likely scenario leading to losses exceeding a given magnitude. We also apply this procedure to examine how conditioning on a stressed value for one variable affects the levels of other variables to gauge the consistency and relative severity of a set of stresses.

As part of this analysis, we derive results for the conditional variances and covariances of the underlying market factors given an extreme move by one factor. These results illustrate qualitative differences between the normal, Laplace, and t cases and thus connect the tail behavior of market factors with conditional moments in extreme scenarios.

As a further application of these ideas, we analyze marginal expected shortfall (MES) and a corresponding marginal variance of shortfall (MVS). These are conditional moments in a stress scenario. In particular, MES measures the expected loss in part of a portfolio conditional on a stress to the full portfolio. It has also been proposed as a measure of systemic risk when applied across firms rather than across parts of a single portfolio. We show how MES and MVS change under the alternative multivariate models we consider — the normal, Laplace and t distributions.

In particular, our analysis shows how the variability in MES estimates depends on the heaviness of the tails of market risk factors. Large variance values suggest the potential for a high degree of variability in MES estimates. With this in mind, we apply EL confidence regions to test the significance of systemic risk rankings in Acharya et al. [1]. The tests suggest that the top 50 companies rank roughly equally, as measured by MES, and that the difference between this group and the 100th ranked company is highly significant. Our approach takes into account the limited data available in extreme stress scenarios.

After providing a precise estimation procedure for reverse stress testing and analyzing conditional means and covariances in extremes, we use similar ideas to draw scenarios from regions of the factor space leading to large losses. In principle, one would want to draw scenarios from the conditional distribution given a large loss, but characterizing extreme conditional distributions remains notoriously difficult (as in Balkema and Embrechts [3]). We use a mechanism suggested by the EL confidence regions. Our approach generates random weight vectors and uses these to take weighted

combinations of past extreme scenarios to generate new scenarios in important large-loss regions. This produces new scenarios that are guided by past extreme combinations of factor moves; the procedure is able to capture the shape of the empirical distribution of extreme scenarios.

We illustrate our procedures on two types of data. First, we use returns on a set of world equity indices and currencies as examples in which data is plentiful, and we use this data for our reverse stress testing procedure. We then apply our procedures to a mix of economic and financial variables drawn from the 2012 Comprehensive Capital Analysis and Review [5]; this is the Federal Reserve's stress test for large bank holding companies.

The rest of this paper is organized as follows. Section 2 introduces reverse stress testing through the simple case of a normal distribution. Section 3 introduces empirical likelihood estimation, and Section 4 converts estimates of conditional means to estimates of most likely loss scenarios. Section 5 presents the application to equity and currency portfolios. Section 6 presents our results on conditional extreme moments and applies these to analyze marginal expected shortfall and a corresponding variance. Section 7 considers some of the CCAR variables and scenarios. Section 8 presents the sampling algorithm. Proofs are deferred to an appendix.

2 The Reverse Stress Testing Problem

Let $\mathbf{Z}$ be a random d-dimensional vector representing the changes in relevant market factors — rates, prices, and economic variables. Suppose $\mathbf{Z}$ has a probability density f on $\mathbb{R}^d$. For a given portfolio exposed to these market factors, let $(\mathbf{Z}, L)$ have the joint distribution of the factors and the portfolio loss L. Write $f(\mathbf{z}|L \geq \ell)$ for the conditional density of $\mathbf{Z}$ given $L \geq \ell$, assuming it exists. The generic problem of reverse stress testing, for a loss threshold ℓ, is to find the most likely scenario (or scenarios) given a loss greater than or equal to ℓ; in other words,

$$\text{(RST)} \qquad \max_{\mathbf{z}\in\mathbb{R}^d} f(\mathbf{z}|L \geq \ell).$$

We refer to a solution of this problem as a most likely loss scenario or as a solution to the reverse stress test. This is called the "design point" in De and Tamarchenko [9] and Koyluoglu [17], based on an analogy with structural reliability.

To help fix ideas, we consider the simple setting of normally distributed market factors, $\mathbf{Z} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$. Suppose that the portfolio loss is given by a linear function of the factors, $L = \mathbf{c}^\top \mathbf{Z}$ for some $\mathbf{c} \in \mathbb{R}^d$, in which case $(\mathbf{Z}, L)$ are jointly normal. Maximizing f is equivalent to minimizing $-\log f$, so problem (RST) reduces to

$$\min_{\mathbf{z}\in\mathbb{R}^d} (\mathbf{z} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{z} - \boldsymbol{\mu}) \quad \text{subject to} \quad \mathbf{c}^\top \mathbf{z} \geq \ell.$$

This problem is easily solved and has the explicit solution

$$\mathbf{z}^*(\ell) = \boldsymbol{\mu} + \left(\frac{\ell - \mathbf{c}^\top \boldsymbol{\mu}}{\mathbf{c}^\top \Sigma \mathbf{c}} \right) \boldsymbol{\Sigma} \mathbf{c}. \tag{1}$$

We can also write this as

$$\mathbf{z}^*(\ell) = \boldsymbol{\mu} + \mathbf{b}^*(\ell - \mathbf{c}^\top \boldsymbol{\mu}), \tag{2}$$

with $\mathbf{b}^* = \boldsymbol{\Sigma}\mathbf{c}/\mathbf{c}^\top \boldsymbol{\Sigma}\mathbf{c}$. Thus, in the case of normally distributed factors and a linear loss function, the most likely loss scenario is the conditional mean $\mathbf{z}^*(\ell) = \mathbb{E}[\mathbf{Z}|\mathbf{c}^\top \mathbf{Z} = \ell]$.

3 Empirical Likelihood Estimation of the Conditional Mean

To move away from reliance on specific distributional assumptions, we adopt a nonparametric approach. Our objective remains to find the solution to (RST), but as an intermediate step we first focus on estimating $\mathbb{E}[\mathbf{Z}|L \geq \ell]$, the conditional mean of the factors given a large loss. In the setting of (2), the conditional mean overestimates the most likely loss scenario, and it is easy to see that this will typically be the case under modest restrictions on the density f.To offset this effect, we will derive a scaling correction based on the tail decay of $\mathbf{Z}$.

But first we need to estimate the conditional mean. We assume we have observations $(\mathbf{z}_i, L_i)$, $i = 1, 2, \ldots$, of past scenarios $\mathbf{z}_i$ and corresponding losses L_i. From these, we discard all observations except those for which the loss is at least ℓ. Through appropriate re-indexing, we are left with n observations $(\mathbf{z}_1, L_1), \ldots, (\mathbf{z}_n, L_n)$, all of which have $L_i \geq \ell$. Clearly, this requires that ℓ not be so large that losses of at least ℓ have never been observed. To get to yet more extreme values of ℓ, one might estimate the conditional means $\mathbb{E}[\mathbf{Z}|L \geq \ell_j]$ at a sequence of increasing values of ℓ_j and then try to extrapolate.

Once we have culled those observations for which $L_i \geq \ell$, the original problem of estimating a conditional mean reduces to once of estimating an unconditional mean. For this problem, we apply Owen's [21] empirical likelihood (EL) method. This method considers convex combinations of the observations as candidate estimates of the mean:

$$w_1\mathbf{z}_1 + w_2\mathbf{z}_2 + \cdots + w_n\mathbf{z}_n, \quad \sum_{i=1}^n w_i = 1, \quad w_i \geq 0,\, i = 1, \ldots, n,$$

The profile empirical likelihood associated with a candidate value $\mathbf{x}$ is

$$\mathcal{R}(\mathbf{x}) = \max \left\{ \prod_{i=1}^n n w_i : \sum_{i=1}^n w_i \mathbf{z}_i = \mathbf{x}, \sum_{i=1}^n w_i = 1, w_i \geq 0,\, i = 1, \ldots, n \right\}. \tag{3}$$

The product inside the braces is the likelihood ratio of the probability vector $(w_1, \ldots, w_n)$ to the uniform distribution $(1/n, \ldots, 1/n)$; $\mathcal{R}(\mathbf{x})$ is larger when $\mathbf{x}$ is a more uniform convex combination of the weights on the observations.

Suppose that the observations are i.i.d. with mean $\boldsymbol{\mu}_0$, and suppose that the convex hull of the observations contains $\boldsymbol{\mu}_0$ with probability approaching 1 as the number of observations increases. Then Owen's [21] Theorem 3.2 states that $-2\log\mathcal{R}(\boldsymbol{\mu}_0)$ has an asymptotic χ^2_d distribution for large n. This provides the basis for EL confidence regions: Fix a confidence level $1-\alpha$ and find the quantile x_α for which $\mathbb{P}(\chi^2_d \geq x_\alpha) = \alpha$; the corresponding $1-\alpha$ confidence region for $\boldsymbol{\mu}_0$ is the set

$$\mathcal{C}_{1-\alpha,n} = \left\{ \sum_{i=1}^n w_i \mathbf{z}_i : \prod_{i=1}^n n w_i \geq \exp(-x_\alpha/2), \sum_{i=1}^n w_i = 1,\, w_i \geq 0, i = 1, \ldots, n \right\}. \tag{4}$$

As discussed in Owen [21], the maximization problem defining the profile empirical likelihood is easy to solve by first reformulating it as

$$\max_{w_1,\ldots,w_n} \sum_{i=1}^n \log w_i \quad \text{subject to} \quad \sum_{i=1}^n w_i = 1,\ \sum_{i=1}^n w_i \mathbf{z}_i = \mathbf{x}.$$

The resulting confidence regions are appealing, if n is not too small, because they make minimal assumptions about the distribution of the underlying data and are able to capture skewness and other notable shape characteristics in the data. Before presenting examples, we examine the connection between the conditional mean estimated here and the most likely loss scenario.

4 From Conditional Mean to Most Likely Loss Scenario

4.1 Multivariate Models

Recall that our objective is to estimate the solution $\mathbf{z}^*(\ell)$ to the reverse stress testing problem (RST), and in the previous section we have estimated a conditional mean $\mathbb{E}[\mathbf{Z}|L \geq \ell]$, which we denote by $\bar{\mathbf{z}}(\ell)$. The next step is therefore to relate these quantities. We will do so under the assumption that the loss level ℓ is large and that the joint distribution of the market factors and the portfolio loss is multivariate normal, Laplace, or t.

The multivariate distributions we consider admit the representation

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{X}, \quad \mathbf{X} = \sqrt{W} N(0, \boldsymbol{\Sigma}), \tag{5}$$

with $N(0, \boldsymbol{\Sigma})$ representing a normal random vector with mean zero and covariance $\boldsymbol{\Sigma}$, W a mixing random variable independent of the normal vector, and $\boldsymbol{\mu}$ a constant mean vector. In other words, these are translated scale mixtures of normals, and they belong to the family of elliptically contoured distributions. The Laplace distribution has W exponentially distributed with mean $1/\lambda$, and the t distribution with ν degrees of freedom has $1/W = \chi^2_\nu/\nu$, a chi-square random variable normalized by its degrees of freedom parameter. The normal case itself corresponds to $W \equiv 1$.

These three cases represent three distinct classes of tail behavior for $\mathbb{P}(X_i \geq x)$ and $\mathbb{P}(X_i \leq -x)$, for large x. The normal tail is very light, with order $O(e^{-ax^2}/x)$, for some $a > 0$; the Laplace tail is exponential, with order $O(e^{-ax})$, for some $a > 0$; and the t_ν tail is regularly varying, with order $O(x^{-\nu})$. Thus, these three distributions capture a key distinction in models of the extreme behavior of market factors, with the last case the most relevant in most applications.[1] These cases also have qualitatively different tail dependence, with the multivariate normal having no tail dependence (except when perfectly correlated), the multivariate t exhibiting positive tail dependence even with negative correlation (except when perfectly negatively correlated); see Schmidt [25]. The mixture representation in (5) and the associated tail behavior can be interpreted as the result of heteroskedasticity or stochastic volatility in a dynamic model.

For our theoretical results, we will assume that the joint distribution of $(\mathbf{Z}, L)$, the market factors and the portfolio loss, belongs to one of these three families of distributions. This assumes, in particular, that the tail decay of the portfolio loss coincides with that that of the underlying factors. We do not assume that L is a deterministic function of $\mathbf{Z}$. We may think of $\mathbf{Z}$ as recording the most important factors influencing the portfolio, and then the model assumes that the tail behavior of the portfolio loss matches that of the most important factors. We will always assume that the restriction of $\mathbf{\Sigma}$ to the $d \times d$ covariance matrix of $\mathbf{Z}$ is positive definite, so that none of factors is redundant. This is sufficient to ensure that the most likely loss scenario $\mathbf{z}^*(\ell)$ is well-defined.

4.2 Estimation

We now turn to the problem of estimating the most likely loss scenario through the conditional mean, beginning with the following result.

Proposition 1 *Suppose the joint distribution of $(\mathbf{Z}, L)$ is multivariate normal, Laplace or t_ν, $\nu > 1$. Let $\mathbf{z}^*(\ell) \in \mathbb{R}^d$ be the most likely loss scenario and let $\bar{\mathbf{z}}(\ell) \in \mathbb{R}^d$ denote the conditional mean $\mathbb{E}[\mathbf{Z}|L \geq \ell]$. Then there exists a positive scalar sequence κ_ℓ such that*

$$\mathbf{z}^*(\ell) = \kappa_\ell \bar{\mathbf{z}}(\ell), \quad \textit{and} \quad \kappa_\ell \to \kappa \quad \textit{as } \ell \to \infty, \tag{6}$$

where

- $\kappa = 1$ *for a multivariate normal or Laplace distribution;*
- $\kappa = (\nu - 1)/\nu$ *for a multivariate t_ν distribution.*

[1] However, Heyde and Kou [14] show that it is quite difficult to distinguish t tails from Laplace tails empirically.

Based on this result, we can estimate the most likely loss scenario $\mathbf{z}^*(\ell)$ by estimating the conditional mean $\bar{\mathbf{z}}(\ell)$ and then scaling the result as needed. In the normal and Laplace cases, no scaling is needed; in the t_ν case, we multiply the estimate of the conditional mean by $(\nu-1)/\nu$ asymptotically to estimate the most likely loss scenario. Market data is often well approximated with ν in the range of 5–7, corresponding to scale factors in the range of 0.80–0.86. In addition to scaling the point estimate, we would like, more importantly, to scale the confidence regions for $\bar{\mathbf{z}}(\ell)$ to get confidence regions for $\mathbf{z}^*(\ell)$. Such a procedure involves two limits, because Proposition 1 applies as $\ell \to \infty$ whereas the chi-square limit that underpins the EL method holds as the number of observations grows. For a combined result, we therefore need an array version of the EL limit theorem, building on Owen's [21] Theorem 4.1.

In the following, we let $\mathbf{Z}_1(\ell), \mathbf{Z}_2(\ell), \ldots, \mathbf{Z}_{n_\ell}(\ell)$ denote i.i.d. observations from the conditional distribution of $\mathbf{Z}$ given $L \geq \ell$, with $n_\ell \to \infty$. As before, let x_α be the quantile defined by $\mathbb{P}(\chi^2_d \geq x_\alpha) = \alpha$. Write $\mathcal{R}_\ell(\mathbf{x})$ for the profile empirical likelihood in (3) with $n = n_\ell$. For a set $\mathcal{C} \subseteq \mathbb{R}^d$ and a constant κ, $\kappa\mathcal{C}$ denotes the set of points of the form $\kappa\mathbf{x}$ with $\mathbf{x} \in \mathcal{C}$.

Theorem 1 *Under the multivariate normal, Laplace or t_ν distribution, $\nu > 4$,*

$$-2\log\mathcal{R}_\ell(\bar{\mathbf{z}}(\ell)) = -2\log\mathcal{R}_\ell(\kappa_\ell^{-1}\mathbf{z}^*(\ell)) \to \chi^2_d$$

in distribution, and $\kappa_\ell\mathcal{C}_{1-\alpha,n_\ell}$ is an asymptotic $100(1-\alpha)\%$ confidence region for the most likely loss scenario $\mathbf{z}^*(\ell)$; i.e.,

$$\mathbb{P}(\mathbf{z}^*(\ell) \in \kappa_\ell\mathcal{C}_{1-\alpha,n_\ell}) \to 1-\alpha,$$

as $\ell \to \infty$, where $\kappa_\ell \to \kappa$, with κ as in Proposition 1.

This result leads to the following procedure. As in Section 3, we extract the large loss scenarios from the available data. Using these observations, we construct EL confidence regions (4) for the conditional mean $\bar{\mathbf{z}}(\ell)$. We then scale the confidence regions by the factor κ_ℓ to get confidence regions for the most likely loss scenario $\mathbf{z}^*(\ell)$. As a simplifying approximation, one could use the limiting value κ in place of κ_ℓ.

We have described this procedure through its application to historical data. The same approach could be used with simulated data. In some settings — stress testing an entire bank portfolio, for example — fully evaluating each scenario is extremely time-consuming. A simplified model could then be applied to simulated scenarios from which one would then estimate the most likely loss scenarios for a more extensive evaluation.

The EL procedure is nonparametric. The asymptotic scaling factor κ is "lightly" parametric in the sense that it depends on the tail decay of the factors. For both the normal and Laplace cases, we

have seen that $\kappa = 1$, so no scaling is necessary asymptotically, and we consider this representative of what one should expect in any light-tailed (meaning exponential or lighter) setting. Regularly varying tails are more typical for financial data. For data with an exponent of regular variation of ν, we expect that using a scale correction of $\kappa = (\nu - 1)/\nu$ should be reasonably robust even beyond the specific case of a t distribution. It should be noted that the procedure provided by Theorem 1 does not involve estimation of $\mathbf{\Sigma}$, which can be particularly difficult in high dimensions. Indeed, we suggest that the method provided by Theorem 1 is potentially applicable in practice even beyond the set of models for which the result provides rigorous support.

Baysal and Staum [4] compare several methods for estimating confidence intervals for value-at-risk and expected shortfall, including empirical likelihood, which they find to have the highest coverage among the methods they compare. Their setting considers confidence regions for the outputs of risk measurement — value-at-risk and expected shortfall — whereas our concern is with the inputs in the form of most likely scenarios.

4.3 Coverage

Theorem 1 provides asymptotic support for confidence regions as the sample size and loss level increase. To test the performance of the confidence regions at finite sample sizes and loss levels, we use simulation. We generate points from a multivariate t distribution with uncorrelated marginals. The loss is given by the linear function $\mathbf{c}^\top \mathbf{Z}$ with $\mathbf{c} = (1, 0, \ldots, 0)^\top$, so the most likely scenario producing a loss of ℓ is $\mathbf{z}^*(\ell) = (\ell, 0, \ldots, 0)^\top$. To test performance at sample size n, we generate enough points to get n observations for which the loss is at least ℓ; we then construct the confidence region, scaled by κ_ℓ, and check if it contains $\mathbf{z}^*(\ell)$. We repeat this 1000 times and record the percentage of times the confidence region contains $\mathbf{z}^*(\ell)$ as the estimated coverage.

Table 1 shows the results at degrees of freedom $\nu = 5$, 6, and 7; dimensions $d = 2$, 5, and 10; sample sizes $n = 10$, 50, and 500; and loss levels at the 95th, 99th, and 99.9th percentile of the t_ν distribution. The top half of the table uses a confidence level of 95%, and the bottom half uses 50%. We need at least $d + 1$ points in dimension d to get a confidence region with nonzero volume, so the entries with $n = d = 10$ are blank. As expected, the observed coverage approaches the nominal coverage as the sample size increases. The most significant shortfalls in coverage occur in high dimensions with few points. The coverage is not very sensitive to the loss level ℓ.

		$\nu = 5$			$\nu = 6$			$\nu = 7$		
	95% confidence	$n = 10$	50	500	10	50	500	10	50	500
$d = 2$,	$\ell = F_\nu^{-1}(.95)$	73.4	90.0	94.8	75.7	93.1	93.6	74.8	91.4	95.0
	$F_\nu^{-1}(.99)$	71.6	90.2	95.7	74.6	92.3	95.7	75.6	91.8	94.0
	$F_\nu^{-1}(.999)$	72.4	91.2	95.2	72.3	92.4	96.3	77.6	93.4	94.2
$d = 5$,	$\ell = F_\nu^{-1}(.95)$	30.1	84.4	94.6	29.2	86.8	94.3	30.4	86.6	95.2
	$F_\nu^{-1}(.99)$	26.3	85.9	94.2	28.6	89.1	93.7	28.5	87.2	93.9
	$F_\nu^{-1}(.999)$	25.2	86.6	93.8	28.5	89.1	95.1	31.0	86.3	94.4
$d = 10$,	$\ell = F_\nu^{-1}(.95)$		69.6	93.8		71.8	94.2		73.6	93.6
	$F_\nu^{-1}(.99)$		68.0	92.3		69.5	92.9		71.0	94.4
	$F_\nu^{-1}(.999)$		68.2	94.1		73.5	93.8		74.3	94.0
	50% confidence	$n = 10$	50	500	10	50	500	10	50	500
$d = 2$,	$\ell = F_\nu^{-1}(.95)$	35.4	45.0	48.0	35.8	47.8	48.8	36.8	47.0	46.4
	$F_\nu^{-1}(.99)$	30.8	43.2	48.4	32.5	44.4	50.2	35.4	47.0	49.1
	$F_\nu^{-1}(.999)$	33.6	45.4	51.5	35.9	46.2	50.4	35.4	50.6	52.0
$d = 5$	$\ell = F_\nu^{-1}(.95)$	10.7	39.2	50.6	12.6	40.6	48.4	11.3	39.7	50.1
	$F_\nu^{-1}(.99)$	9.4	37.8	46.5	10.7	40.8	46.6	11.5	41.8	51.2
	$F_\nu^{-1}(.999)$	8.8	37.3	45.2	11.9	39.8	51.1	12.0	39.0	46.0
$d = 10$,	$\ell = F_\nu^{-1}(.95)$		22.7	46.5		26.3	51.4		26.6	47.6
	$F_\nu^{-1}(.99)$		23.1	44.9		24.3	48.0		25.7	49.6
	$F_\nu^{-1}(.999)$		23.6	46.6		28.0	47.0		26.7	49.7

Table 1: Estimated coverage of the most likely loss scenario for dimension d, sample size $n \geq d + 1$, loss level ℓ, and degrees of freedom ν at confidence levels of 95% and 50%.

5 Application to Equity and Currency Scenarios

5.1 An Equity Portfolio

For our first application, we consider a portfolio of world equity indices: the S&P 500, FTSE, DAX, Nikkei 225, Hang Seng, and Bovespa. We consider weekly returns from May 3, 1993, to December 26, 2011, and monthly returns from June 1, 1993, to December 1, 2011. We select weights based on the market capitalization traded on each exchange, as listed in Table 2. This gives us a linear loss function proportional to $\mathbf{c} = [-0.5050; -0.1362; -0.0539; -0.1443; -0.1022; -0.0583]$; we scale this $\mathbf{c}$ to get to a 1% loss level with weekly data and a 5% loss level with monthly data.

We model the returns on the equity indices using a multivariate t distribution. The density with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, ν is given by

$$f(\mathbf{x}|\mu, \Sigma, \nu) = \frac{\Gamma(\frac{1}{2}(\nu + d))}{\Gamma(\frac{1}{2}\nu)(\pi\nu)^{d/2}|\Sigma|^{1/2}} \left(1 + \frac{(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})}{\nu}\right)^{-(\nu+d)/2}, \quad \text{for} \quad \mathbf{x} \in \mathbb{R}^d.$$

The mean and variance of the distribution are given by

$$\mathbb{E}[X] = \boldsymbol{\mu}, \quad \mathbb{V}(X) = \frac{\nu}{\nu - 2}\boldsymbol{\Sigma},$$

Exchange	Market Cap	Proportion(%)
NYSE Euronext	13,394,081.8	50.50
London SE Group	3,613,064.0	13.62
Deutsche Börse	1,429,719.1	5.39
Tokyo SE Group	3,827,774.2	14.43
Hong Kong Exchanges	2,711,316.2	10.22
BM&FBOVESPA	1,545,565.7	5.83

Table 2: Market caps of exchanges at 2010, in USD millions, from www.world-exchanges.org/statistics

assuming $\nu > 2$. To estimate ν, we first estimate the sample mean and covariance and then maximize the likelihood over ν. We get $\hat{\nu} = 5.0$ with weekly data and $\hat{\nu} = 5.8$ with monthly data.

For purposes of illustration, we show confidence regions for pairs of indices at a time, though having an automated method is particularly valuable in multiple dimensions where visualization is difficult. Figures 1 and 2 show results for weekly data. The circles show the observations, and the crosses show the extreme observations — those beyond the loss threshold. The contours show 99% and 50% confidence regions for the conditional mean and (after scaling) for the most likely loss scenario. The squares indicate the point estimates for the conditional and the most likely loss scenario. The confidence regions are clearly shaped by the data — compare the two panels in Figure 1, for example — yet tempered compared to the most extreme points. Figures 3 and 4 show corresponding results for monthly data.

5.2 A Currency Portfolio

Next we consider a basket of currencies, half held in British pounds (GBP), the rest divided evenly among the Australian dollar (AUD), the euro (EUR), the Japanese yen (JPY), and the Swiss franc (CHF). We use monthly returns against the US dollar from February 2000 through December 2011. A maximum likelihood fit of the data to a multivariate t distribution yields an estimate of $\hat{\nu} = 5.2$ to the degrees-of-freedom parameter. For the loss severity ℓ, we choose the loss threshold ℓ at the level of the worst 5% of losses in the sample period. Our estimated most likely loss scenario is

$$(\text{AUD, EUR, JPY, CHF, GBP}) = (-5.5907\%, -4.1142\%, 1.0402\%, -4.0246\%, -4.4338\%),$$

the values on the right indicating one-month returns against the US dollar.

Figure 5 illustrates the results. The circles show the observations, and the crosses show the 5% most extreme observations — those beyond the loss threshold. The contours show 99% and 50% confidence regions for the conditional mean and (after scaling) for the most likely loss scenario. The squares indicate the point estimates.

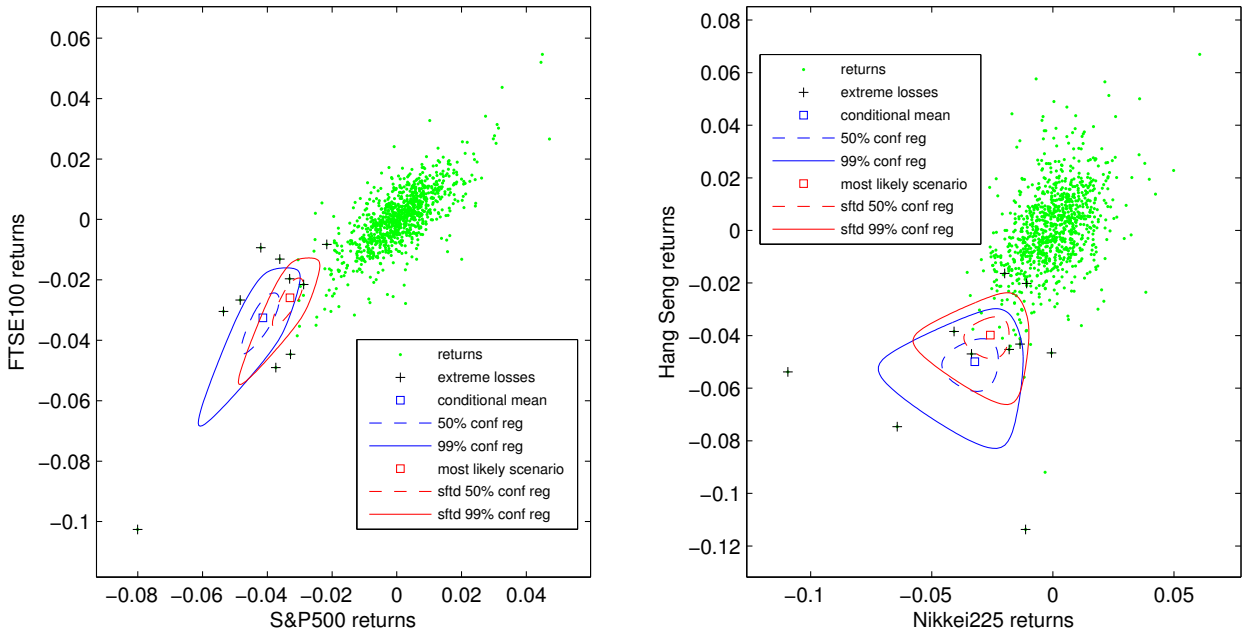


Figure 1: Equity indices, weekly data

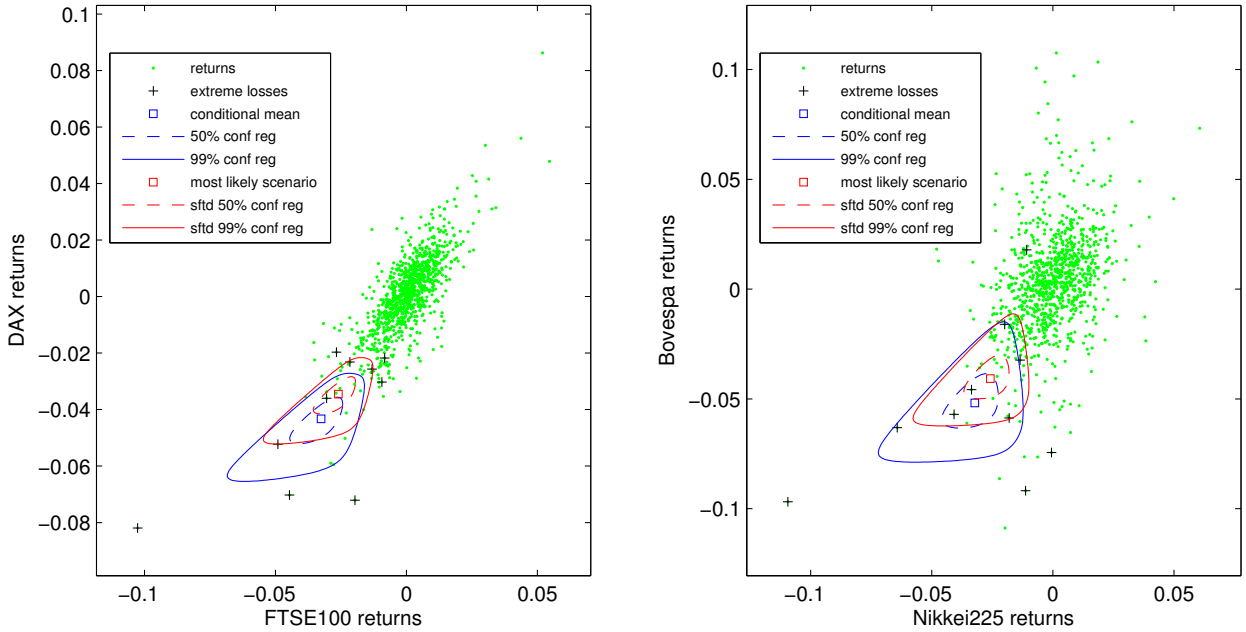


Figure 2: Equity indices, weekly data

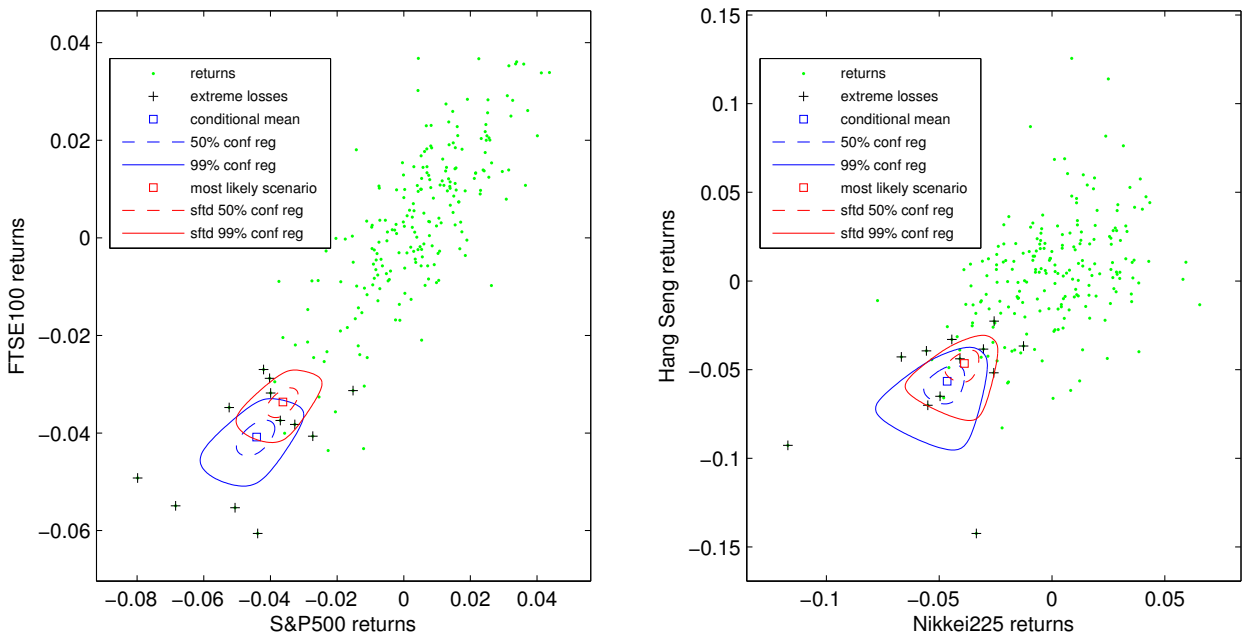


Figure 3: Equity indices, monthly data

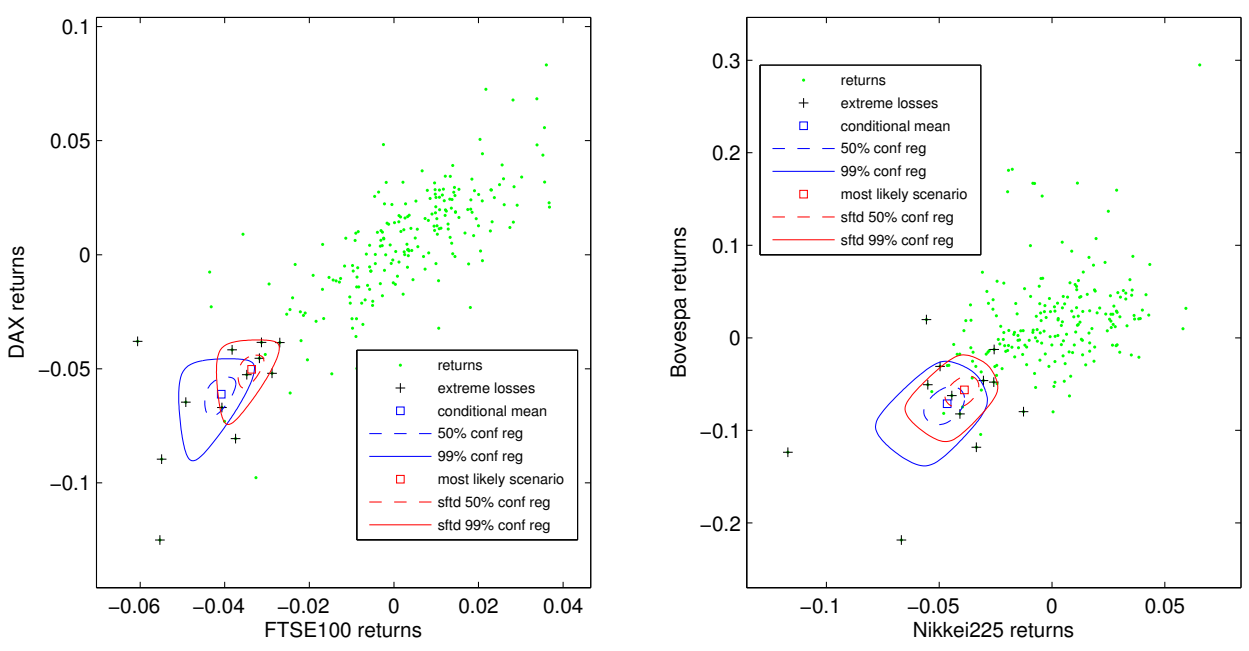


Figure 4: Equity indices, monthly data

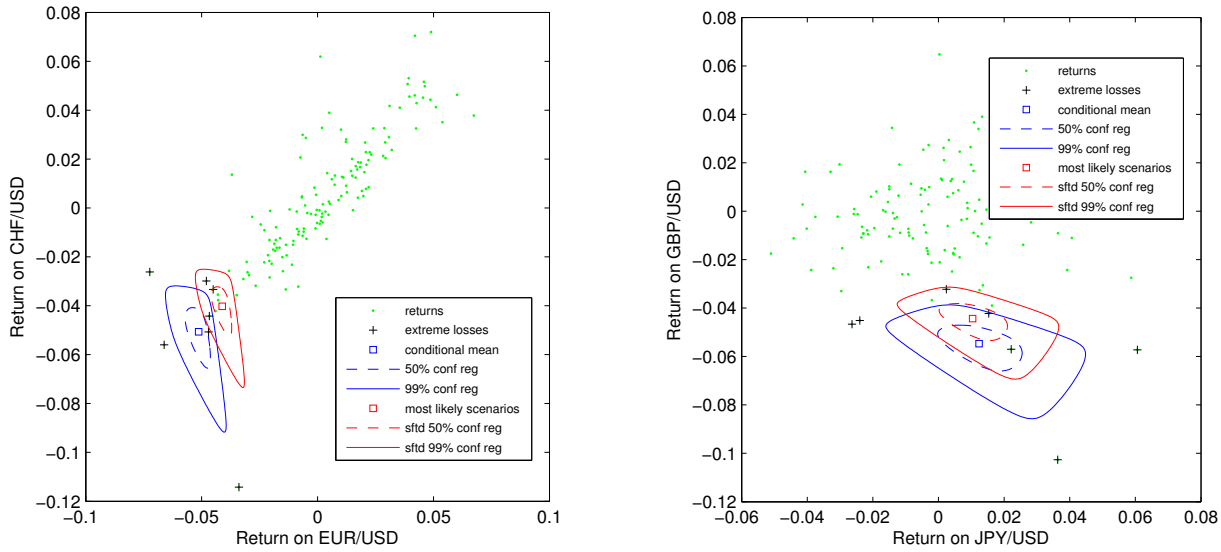


Figure 5: Foreign currencies, monthly data

In the left panel, we see that the confidence regions for the most likely loss reflect the skewness in the joint distribution of the EUR/USD and CHF/USD returns. The most likely loss scenario involves an increase in the JPY/USD rate, even though this increase would, by itself, generate a gain, not a loss, for the portfolio. This outcome is a reflection of the joint distribution of the returns: the largest drops in the GBP (which makes up 50% of the portfolio) coincide with increases in the JPY/USD rate. However, the confidence regions in the right panel of Figure 5 also indicate a wide range of outcomes for the JPY/USD rate when the GBP drops, suggesting that one should explore other scenarios in the large-loss region, a topic to which we return in Section 8.

6 Conditional Moments and Marginal Shortfall

Marginal expected shortfall (MES) provides a mechanism for attributing a portfolio's overall loss in a stress test to parts of the portfolio or to individual factors. The MES of a subportfolio is its expected loss conditional on the loss in the full portfolio exceeding some threshold. The analysis underlying Proposition 1 and Theorem 1 allows us to characterize MES and a corresponding conditional variance for the distributions we consider. Moreover, the EL procedure provides a way to measure the precision of MES estimates.

6.1 Conditional Moments in Extremes

The key to this analysis (and to the proof of Theorem 1) is the calculation of conditional moments in extremes for the multivariate distributions we consider. In fact, it suffices (see the appendix) to

consider a pair (Z_1, Z_2) having the distribution of $\sqrt{W}(N_1, N_2)$, where (N_1, N_2) are independent standard normal random variables. Table 3 summarizes the conditional means and variances of the factors, given an extreme outcome of one of the factors.

Moving from left to right in the table, we have heavier tails. As one would expect, the conditional variance of Z_1 itself, conditional on $Z_1 \geq \ell$, increases with the heaviness of the tail, increasing from $O(1/\ell^2)$ to $O(1)$ to $O(\ell^2)$. The conditional variance of Z_2 given $Z_1 \geq \ell$ also increases from left to right but by a factor of $O(\ell)$ each time, starting from a value of 1. The net effect is that under the multivariate t distribution, the conditional variances of Z_1 and Z_2, given $Z_1 \geq \ell$, have the same order of magnitude, and this underpins the fact that $\kappa \neq 1$ in this case.

	Normal	Laplace	t_ν, $\nu > 2$
$\mathbb{E}[Z_1 \vert Z_1 \geq \ell]$	$\ell + o(\ell)$	$\ell + \frac{1}{\sqrt{2\lambda}}$	$\frac{\nu}{\nu-1}\ell + o(\ell)$
$\mathbb{E}[Z_1^2 \vert Z_1 \geq \ell]$	$\ell^2 + o(\ell^2)$	$(\ell + \frac{1}{\sqrt{2\lambda}})^2 + \frac{1}{2\lambda}$	$\frac{\nu}{\nu-2}\ell^2 + o(\ell^2)$
$\mathbb{V}(Z_1 \vert Z_1 \geq \ell)$	$\frac{1}{\ell^2} + o(\frac{1}{\ell^2})$	$\frac{1}{2\lambda}$	$\frac{\nu}{(\nu-2)(\nu-1)^2}\ell^2 + o(\ell^2)$
$\mathbb{E}[Z_2 \vert Z_1 \geq \ell]$	0	0	0
$\mathbb{V}(Z_2 \vert Z_1 \geq \ell) = \mathbb{E}[Z_2^2 \vert Z_1 \geq \ell]$	1	$\frac{1}{\sqrt{2\lambda}}\ell + \frac{1}{\lambda}$	$\frac{\nu}{(\nu-2)(\nu-1)}\ell^2 + o(\ell^2)$

Table 3: Conditional means and variances of factors, given an extreme outcome of Z_1.

6.2 Marginal Shortfall

Using the asymptotic moments in Table 3, we can analyze the marginal expected shortfall (MES) of part of a portfolio conditional on a large loss in the portfolio, provided the joint distribution of returns falls within one of the cases in the table. Suppose, now, that $\mathbf{Y} = (Y_1, \ldots, Y_d)^\top$ in (5) is a vector of asset returns, with mean $\boldsymbol{\mu}$ and covariance matrix $w\boldsymbol{\Sigma}$; we have $w = 1$ in the normal case, $w = 1/\lambda$ for the Laplace distribution, and $\nu/(\nu - 2)$ for the multivariate t_ν. The loss for a portfolio holding these assets is given by $\mathbf{c}^\top \mathbf{Y}$, for some $\mathbf{c} \in \mathbb{R}^d$. For each asset i and loss level ℓ, define the marginal expected shortfall and the corresponding variance by

$$\begin{aligned} MES_i &= \mathbb{E}[Y_i \vert \mathbf{c}^\top \mathbf{Y} \geq \ell] \\ MVS_i &= \mathbb{V}[Y_i \vert \mathbf{c}^\top \mathbf{Y} \geq \ell]. \end{aligned}$$

We analyze these quantities for large loss levels ℓ. The marginal shortfall contribution for the ith subportfolio or factor is c_i times the expression given here for MES_i.

To lighten notation, let $(\beta_1, \ldots, \beta_d) = \mathbf{c}^\top \boldsymbol{\Sigma}/(\mathbf{c}^\top \boldsymbol{\Sigma} \mathbf{c})$ and write, for each $i = 1, \ldots, d$,

$$Y_i = \mu_i + \beta_i \mathbf{c}^\top (\mathbf{Y} - \boldsymbol{\mu}) + \epsilon_i; \tag{7}$$

this representation defines ϵ_i and makes it uncorrelated with $\mathbf{c}^\top \mathbf{Y}$. Letting $\sigma^2_{\epsilon_i}$ denote the variance of ϵ_i, we get $\sigma^2_{\epsilon_i}/w = \sigma^2_i - \beta^2_i \mathbf{c}^\top \mathbf{\Sigma} \mathbf{c}$, with $\sigma^2_i = \mathbf{\Sigma}_{ii}$. Denote by $\mu_c = \mathbf{c}^\top \boldsymbol{\mu}$ the expected loss and by $\sigma^2_c = w\mathbf{c}^\top \mathbf{\Sigma} \mathbf{c}$ its variance.

Proposition 2 *As $\ell \to \infty$, MES_i and MVS_i behave as follows.*

(i) If $\mathbf{Y} - \boldsymbol{\mu}$ is multivariate normal, then

$$\begin{aligned} MES_i &= \mu_i + \beta_i(\ell - \mu_c) + o(\ell) \\ MVS_i &= \sigma^2_{\epsilon_i} + \beta^2_i \sigma^4_c/(\ell - \mu_c)^2 + o(1/\ell^2). \end{aligned}$$

(ii) If $\mathbf{Y} - \boldsymbol{\mu}$ is multivariate Laplace, then

$$\begin{aligned} MES_i &= \mu_i + \beta_i(\ell - \mu_c) + \frac{\beta_i \sigma_c}{\sqrt{2}} \\ MVS_i &= \sigma^2_{\epsilon_i} + \frac{\beta^2_i \sigma^2_c}{2} + \frac{\sigma^2_{\epsilon_i}}{\sigma_c \sqrt{2}}(\ell - \mu_c). \end{aligned}$$

(iii) If $\mathbf{Y} - \boldsymbol{\mu}$ is multivariate t_ν, $\nu > 2$, then

$$\begin{aligned} MES_i &= \mu_i + \frac{\nu}{\nu - 1}\beta_i(\ell - \mu_c) + o(\ell) \\ MVS_i &= \left(\frac{\beta^2_i \nu}{(\nu - 2)(\nu - 1)^2} + \frac{\nu \sigma^2_{\epsilon_i}}{(\nu - 2)(\nu - 1)\sigma^2_c} \right)(\ell - \mu_c)^2 + o(\ell^2). \end{aligned}$$

Proof. The results follow from substituting (7) into the definitions of MES and MVS and then using Table 3 to evaluate the conditional mean and the conditional variance, taking $Z_1 = \sqrt{w}(\mathbf{c}^\top \mathbf{Y} - \mu_c)/\sigma_c$ and $Z_2 = \sqrt{w}\epsilon_i/\sigma_{\epsilon_i}$. The formulas in the table apply because this (Z_1, Z_2) is a linear transformation $\mathbf{Y} - \boldsymbol{\mu}$ and thus has a bivariate normal, Laplace or t distribution accordingly. □

Proposition 2 shows, as one might expect, that the MES is larger under the Laplace and t distributions than under the normal distribution; but the result shows an interesting distinction between the two cases: the increase is additive under the Laplace distribution and multiplicative under the t distribution. The result also highlights important differences in how the MVS depends on the loss level ℓ: the increase is linear with the Laplace distribution and quadratic with t distribution. A large MVS suggests that estimates of MES are likely to be imprecise, an issue we examine next. Notice the following difference between Theorem 1 and Proposition 2. Even though, according to Proposition 2, a heavier-tailed distribution has a larger marginal variance, Theorem 1 tells us the confidence regions are not much different from each other once the data is fixed.

6.3 EL Significance of MES Rankings

The EL method in Section 3 can be used to estimate confidence regions for a full vector (MES_1, $\ldots$, MES_d). Here we extend these ideas to measure the significance of MES rankings.

Acharya et al. [1] use an MES measure as part of their analysis of systemic risk. Their MES for a company is the expected decline in the company's stock price conditional on a large decline in the whole market, as measured by a broad market index. This quantity measures the expected amount of capital the company would lose in a crisis and is thus a measure of the company's contribution to a crisis. This is a type of stress test in which the implicit stress scenario is a decline in the market index. Acharya et al. [1] rank firms by their MES as an indication of their systemic importance.

A ranking $MES_i > MES_j$ of firm i higher than firm j is equivalent to the point (MES_i, MES_j) lying below the 45° line in the plane. In practice, we estimate MES values from historical data and check if the point estimate falls in this halfspace. We can supplement the point estimate with an EL confidence region using the procedure in Section 3. If a 95% confidence region is fully contained within the halfspace but a 99% confidence region is not, then the significance of the ranking is between 1% and 5%. Indeed, we can measure the significance of an estimated ranking by the smallest p for which the $(1-p)$ confidence region is contained within the halfspace. The same idea can be applied to test the simultaneous significance of an ordering of three or more firms.

Following Appendix B of Acharaya et al. [1], we estimate MES values using daily stock returns for the 13 months from June 2006 through June 2007. We find the 5% of days with the largest declines in the CRSP value-weighted index and estimate the MES of each firm in Appendix B by averaging the firm's stock return over those days.[2] The resulting top 50 values and rankings, displayed in Table 4, match those in Acharya et al. [1]. See Brownlees and Engle [8] for a dynamic approach to MES estimation.

In Table 4, we also report EL confidence levels (i.e., 1 minus significance levels) for pairwise comparisons between firms ranked consecutively, firms ranked ten apart, and between each firm and AIG, which is ranked 100th. None of the comparisons between consecutive firms or firms ranked ten apart approaches conventional thresholds for statistical significance, corresponding to a confidence level of 90% or higher. There is too much conditional variability in the tails to draw reliable conclusions about the MES comparisons, as one might suspect from the MVS asymptotics in Proposition 2. Nearly all the comparisons with AIG are highly significant. The overall picture that emerges from the table is that the top 50 firms can be confidently ranked higher than the 100th, but all of the top 50 should be viewed as of roughly equal importance, as measured by MES.

[2] The worst days are, in decreasing order of severity, Feb 27, Mar 13, Jun 5, 2006, Jun 7, 2007, Jun 12, 2006, Nov 27, Jul 13, May 10, Jun 13, 2006, Jun 20, 2007, Mar 2, Mar 5, Sep 6, and May 24.

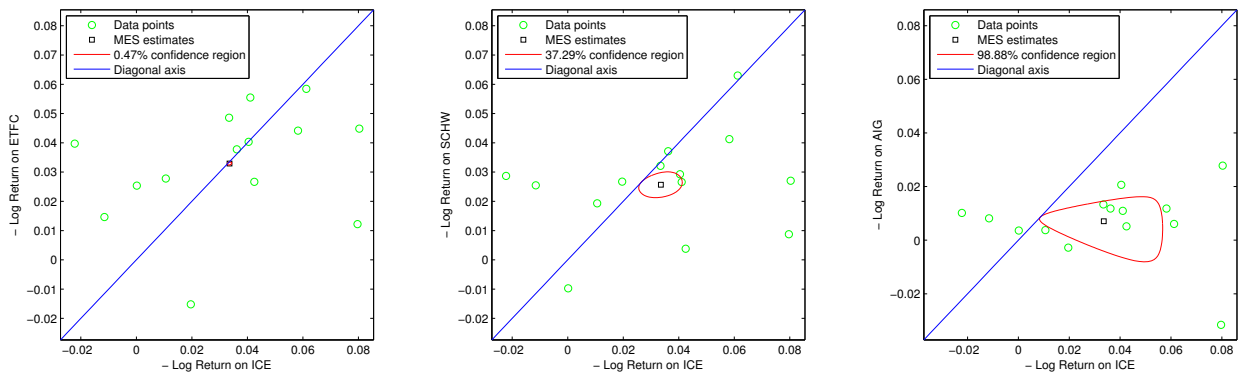


Figure 6: The figures show the largest EL confidence regions for (MES_i, MES_j) below the 45° degree line for the cases $(i, j) = (1, 2)$, $(1, 11)$, and $(1, 100)$.

The fact that AIG is ranked 100th during this period, despite its subsequent role in the financial crisis, reflects the limitations of trying to measure systemic risk purely through stock market data.

The pairwise comparison for ICE with ETFC, SCHW, and AIG are illustrated in Figure 6. Each panel plots the negative log returns for the indicated stocks on the worst 5% of days for the index. Each panel also shows the largest EL confidence region contained below the 45° line. The corresponding confidence level (or, more precisely, the amount by which the confidence level falls short of 100%) measures the significance of each pairwise ordering. These are the values reported in Table 4. The rightmost panel of Figure 6 indicates substantial skewness in the extreme outcomes and how this skewness is reflected in the confidence region.

7 Application to Macro Stress Scenario Generation

The goal of macro stress testing of banks or the full banking system is to gauge the ability of banks to sustain losses under adverse economic and financial conditions. The stress scenarios are usually defined in terms of broad economic and financial variables, rather than directly through the values of assets held by banks. Defining consistent stresses across multiple variables is important for stress tests used to set capital levels; Schuermann [26] refers to this as the problem of defining coherent scenarios. An inconsistent set of stresses implicitly lowers the capital charge for some activities relative to others, much as misspecification of risk weights does under traditional capital calculations. This creates the potential for regulatory arbitrage if banks can anticipate which variables are over-stressed or under-stressed. See Breuer et al. [7] and Kupiec [18] for related considerations.

In this section, we apply the reverse stress testing methodology to examine the consistency of stresses across variables. We condition on a large move in one variable to see the effect on other

MES ranking, i	Name of Company	Ticker	MES	Confidence ($MES_i > MES_j$) $j = i+1$	$j = i+10$	$j = 100$
1	INTERCONTINENTALEXCHANGE INC	ICE	3.36%	0.47%	37.29%	98.88%
2	E TRADE FINANCIAL CORP	ETFC	3.29%	4.57%	**	100.00%
3	BEAR STEARNS COMPANIES INC	BSC	3.15%	1.33%	48.85%	100.00%
4	N Y S E EURONEXT	NYX	3.05%	4.82%	39.18%	99.99%
5	C B RICHARD ELLIS GROUP INC	CBG	2.84%	0.16%	82.10%	100.00%
6	LEHMAN BROTHERS HOLDINGS INC	LEH	2.83%	5.10%	90.70%	100.00%
7	MORGAN STANLEY DEAN WITTER & CO	MS	2.72%	1.57%	44.53%	100.00%
8	AMERIPRISE FINANCIAL INC	AMP	2.68%	0.27%	53.85%	100.00%
9	GOLDMAN SACHS GROUP INC	GS	2.64%	0.05%	67.56%	100.00%
10	MERRILL LYNCH & CO INC	MER	2.64%	2.82%	40.74%	100.00%
11	SCHWAB CHARLES CORP NEW	SCHW	2.57%	**	22.66%	99.97%
12	NYMEX HOLDINGS INC	NMX	2.47%	**	**	**
13	C I T GROUP INC NEW	CIT	2.45%	0.06%	**	100.00%
14	T D AMERITRADE HOLDING CORP	AMTD	2.43%	7.68%	11.21%	99.99%
15	T ROWE PRICE GROUP INC	TROW	2.27%	0.76%	12.06%	100.00%
16	EDWARDS A G INC	AGE	2.26%	0.08%	19.52%	100.00%
17	FEDERAL NATIONAL MORTGAGE ASSN	FNM	2.25%	0.43%	22.70%	100.00%
18	JANUS CAP GROUP INC	JNS	2.23%	0.56%	20.89%	99.89%
19	FRANKLIN RESOURCES INC	BEN	2.20%	0.96%	19.79%	99.94%
20	LEGG MASON INC	LM	2.19%	0.34%	25.82%	99.88%
21	AMERICAN CAPITAL STRATEGIES LTD	ACAS	2.15%	0.28%	38.14%	100.00%
22	STATE STREET CORP	STT	2.12%	**	39.60%	100.00%
23	WESTERN UNION CO	WU	2.10%	**	**	**
24	COUNTRYWIDE FINANCIAL CORP	CFC	2.09%	0.11%	29.19%	96.14%
25	EATON VANCE CORP	EV	2.09%	2.08%	51.50%	99.72%
26	S E I INVESTMENTS COMPANY	SEIC	2.00%	0.72%	12.18%	99.86%
27	BERKLEY W R CORP	BER	1.95%	0.08%	16.95%	99.85%
28	SOVEREIGN BANCORP INC	SOV	1.95%	0.61%	37.21%	100.00%
29	JPMORGAN CHASE & CO	JPM	1.93%	2.20%	45.10%	100.00%
30	BANK NEW YORK INC	BK	1.90%	1.79%	63.02%	100.00%
31	M B I A INC	MBI	1.84%	0.09%	17.72%	99.54%
32	BLACKROCK INC	BLK	1.83%	0.53%	13.39%	99.58%
33	LEUCADIA NATIONAL CORP	LUK	1.80%	0.00%	13.79%	99.89%
34	WASHINGTON MUTUAL INC	WM	1.80%	2.20%	44.33%	99.95%
35	NORTHERN TRUST CORP	NTRS	1.75%	0.39%	12.06%	100.00%
36	C B O T HOLDINGS INC	BOT	1.71%	0.01%	5.91%	82.07%
37	PRINCIPAL FINANCIAL GROUP INC	PFG	1.71%	4.21%	38.22%	99.97%
38	CITIGROUP INC	C	1.66%	0.69%	18.75%	99.92%
39	LOEWS CORP	LTR	1.63%	0.99%	11.10%	98.09%
40	GENWORTH FINANCIAL INC	GNW	1.59%	0.23%	11.51%	99.77%
41	LINCOLN NATIONAL CORP IN	LNC	1.59%	0.05%		100.00%
42	UNION PACIFIC CORP	UNP	1.58%	0.86%		92.23%
43	AMERICAN EXPRESS CO	AXP	1.56%	0.57%		99.94%
44	COMERICA INC	CMA	1.55%	0.75%		99.92%
45	C I G N A CORP	CI	1.54%	0.09%		76.81%
46	FIDELITY NATIONAL INFO SVCS INC	FIS	1.54%	0.36%		99.65%
47	METLIFE INC	MET	1.52%	0.61%		96.81%
48	PROGRESSIVE CORP OH	PGR	1.51%	1.05%		99.50%
49	M & T BANK CORP	MTB	1.49%	0.21%		98.13%
50	NATIONAL CITY CORP	NCC	1.48%			91.21%
100	AMERICAN INTERNATIONAL GROUP INC	AIG	0.71%			

Table 4: Estimates of MES and top 50 rankings based on daily returns from June 2006 through June 2007, as in Acharya et al.[1]. The last three columns show EL confidence levels for the significance of pairwise comparisons of rankings. NMX and WU traded for only part of the time period and are omitted from the comparison.

variables. Conditioning on a decline in GDP, for example, is analogous to taking GDP as the loss function.

The scenarios we consider are based on several of the variables selected by the Federal Reserve for the 2012 Comprehensive Capital Analysis and Review (CCAR). The full details of the CCAR scenarios are included in [5]. Here, we confine ourselves to the following variables: US real GDP growth rate, US unemployment growth rate, a seasonally adjusted house price index (HPI), returns of the Dow Jones Total Market Index, changes in the level of the VIX, and changes in the level of the EUR/USD exchange rate.[3]

Rather than treat the data series as i.i.d., we fit a vector autoregressive model to the data and treat the residuals as i.i.d. and apply EL to the residuals. We estimate the model from quarterly values from Q1-1990 through Q3-2011, a longer series than that tabulated in [5]. After comparing alternative specifications, we fit a first-order model in which the off-diagonal entries involving HPI are zero and all other entries are unconstrained. From the residuals, we get a maximum likelihood estimate of $\nu = 8.43$. We stress each variable — more precisely, its residual — and then estimate confidence regions for the most likely values of the residuals of the other variables, which we then convert to the variables themselves. (The confidence regions do not reflect uncertainty in the estimated VAR parameters that produce the residuals.) Doing so provides a reference point for evaluating the moves in the other variables in a specific scenario, given the move in one variable.

Table 5 provides a first look at the scenarios from this perspective. For each variable, the table first reports the first-quarter CCAR scenario in the units of each variable. We then convert each variable into a residual and a standardized residual. Several of these stresses go well beyond the available data, so in the last two columns we introduce reduced stresses, chosen so that we have at least five scenarios more extreme than the indicated stress within the window of historical data. We use these reduced stress levels when we estimate the impact on other variables.

In Figures 7–10, we show EL confidence regions for the most likely values of pairs of variables when one variable (which may be one of the pair or a third variable) exceeds its stress level. The circles and crosses show past observations, the crosses indicating those that are beyond the indicated stress. In each case, we show confidence regions for the conditional mean and shifted confidence regions for the most likely outcome. For illustrative purposes, we also include the CCAR scenario in each figure, marked by an asterisk. The CCAR includes values for 13 consecutive quarters, of which we consider only the first. The procedure could be extended to subsequent quarters by taking

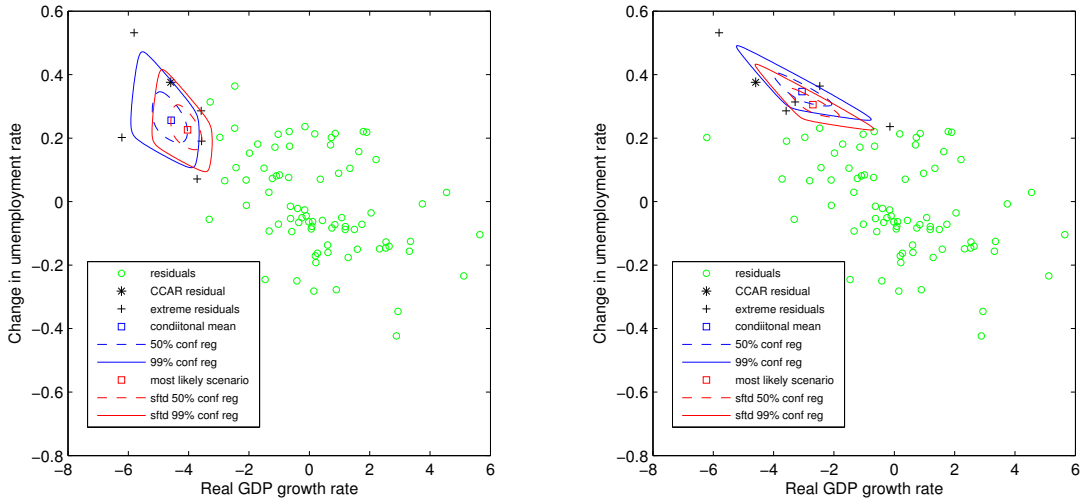
[3]We convert the CCAR scenarios for unemployment, the VIX, and EUR/USD exchange rate from levels to differences, and we convert the stock index and HPI scenarios from levels to returns. For the HPI, we use seasonally adjusted values from CoreLogic; these are very close to those reported in [5] and allow us to use a longer time series. For the VIX, we use the maximum value in each quarter, as appears to have been used in [5], correcting the value for Q2-2008 from 31.01, as reported in [5], to 24.12.

	CCAR Scenarios	Residuals of Scenarios	Standardized Residuals	Stress Level	Standardized Stress Level
Real GDP Growth Rate	-4.84	-4.5926	-2.0867	-3.5703	-1.6222
Change in Unemployment Rate	0.59	0.3756	2.0505	0.2363	1.2901
Return on DJ Total Market Index	-0.1929	-0.1714	-1.9978	-0.1714	-1.9978
Change in VIX Index	27.86	28.1367	3.2212	16.4281	1.8807
Change in EUR/USD Exchange Rate	-0.03	-0.0395	-0.5772	-0.0395	-0.5772
Return on HPI	-0.0126	-0.0105	-1.2094	-0.0105	-1.2094

Table 5: From left to right, the table shows the first quarter CCAR scenario for each variable in the units of that variable; the corresponding residual and standardized residual; a reduced stress in the units of each variable; and the standardized residual for the reduced stress.

a path of residuals as a single outcome.

In Figure 7, we stress GDP and unemployment and look at the most likely outcome for the combination of the two. The CCAR scenario falls just at the edge or just beyond the 99% confidence region, consistent with the fact that we are applying the reduced stresses from Table 5.



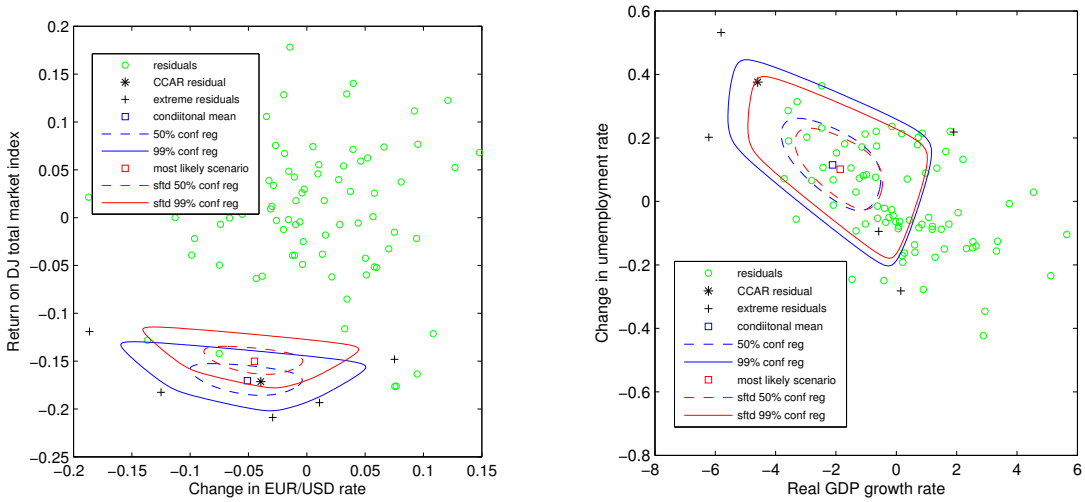
(a) Stress: Real GDP growth rate

(b) Stress: Changes in unemployment rates

Figure 7: Confidence regions conditional on the indicated stressed variable

In Figure 8, we stress the VIX. The left panel shows the impact on financial variables, and the right panel shows the impact on macro variables. Interestingly, the CCAR scenario falls on the boundary of the 99% confidence region in each case.

Figure 9 presents a rather different picture. Here we stress the EUR/USD exchange rate and again compare the impact on financial variables (left) and macro variables (right). In both cases we find that the most likely outcomes for the other variables are near the center of the historical distribution, and the CCAR scenario is very extreme by comparison. This is not surprising — it

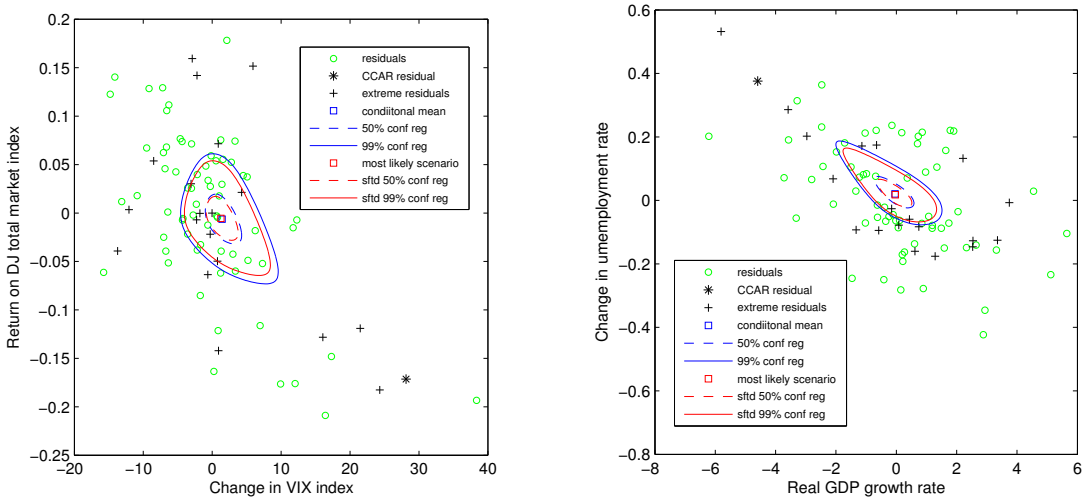


(a) Stress: Changes in VIX index

(b) Stress: Changes in VIX index

Figure 8: Confidence regions conditional on the indicated stressed variable

reflects the fact that a stressed value of the exchange rate is not, by itself, associated with extreme values of the other variables.



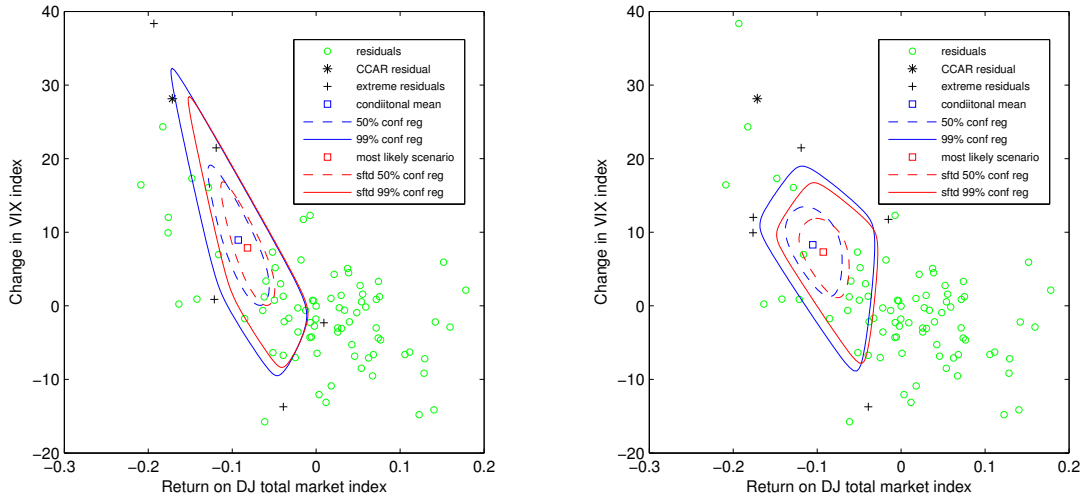
(a) Stress: Changes in EUR/USD rates

(b) Stress: Changes in EUR/USD rates

Figure 9: Confidence regions conditional on the indicated stressed variable

In our final example, Figure 10, we again stress GDP and unemployment, but we now look at the impact on financial variables. Here the CCAR scenario looks comparatively extreme, particularly in the right panel. This is partly due to the fact that we are applying reduced stresses from Table 5. However, it also suggests that the change in financial variables associated with stresses to the macro variables posited by the CCAR are relatively extreme when compared with historical

co-movements. All of these examples also point to the value of exploring important regions through multiple scenarios, a problem we take up in the next section.



(a) Stress: Real GDP growth rate

(b) Stress: Change in unemployment rate

Figure 10: Confidence regions conditional on the indicated stressed variable

8 Generating Extreme Scenarios

The examples of Sections 5 and 7 point to the need to consider multiple combinations of extreme movements, given uncertainty around the most likely conditional scenario. In this section, we propose a method for drawing extreme scenarios by simulation, building on historical scenarios but not restricting ourselves to past observations. The method is based on the idea of using the EL confidence regions (or a simplified construction) as contours of relative probability for a hypothetical distribution. The merits of this approach are that it uses available extreme observations to guide sampling while, unlike bootstrapping, generating scenarios outside the historical record. We see this sampling algorithm as a sensible way to explore the region of large-loss scenarios and to generate candidates in the vicinity of the most likely loss scenario, rather than as a way to undertake precise calculations for the conditional distribution of the factors. As we noted in Section 1, rather little is known about multivariate distributions conditioned on extremes.

A direct approach to using the EL contours to guide scenario generation would be to normalize $\mathcal{R}$ to a probability density and then sample from this density. Because $\mathcal{R}$ is bounded between 0 and 1, this can be implemented without explicit normalization through rejection sampling, as follows. Sample U uniformly from a rectangle in $\mathbb{R}^d$ that contains the convex hull of the data points $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$; accept U with probability $\mathcal{R}(U)$. The accepted values then have a probability density

proportional to $\mathcal{R}$. A shortcoming of this approach is that the acceptance probability can be quite small, particularly in high dimensions, making the procedure quite slow.

A faster alternative is to generate random weight vectors in the unit simplex and then map these to scenarios using the data points. We implement this using a Dirichlet distribution for the weights. A random vector $(W_1, \ldots, W_n)$ has a Dirichlet distribution with parameter $\alpha = (\alpha_1, \ldots, \alpha_n)$, all $\alpha_i > 0$, if $(W_1, \ldots, W_{n-1})$ has probability density

$$f(w_1, \ldots, w_{n-1}; \alpha_1, \ldots, \alpha_n) = \frac{1}{B(\alpha)} \prod_{i-1}^{n} w_i^{\alpha_i - 1},$$

for $w_1, \ldots, w_n \in (0, 1)$ with $w_n = 1 - w_2 - \cdots - w_{n-1}$. Here, $B(\alpha)$ is the normalization constant

$$B(\alpha) = \frac{\prod_{i=1}^{n} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{n} \alpha_i)}.$$

Dirichlet distributions provide a natural and flexible family of distributions for random weight vectors. Given a sample $(w_1, \ldots, w_{n-1})$ from a Dirichlet distribution and loss scenarios $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$ in $\mathbb{R}^d$, we generate a new scenario by setting

$$\mathbf{Z} = w_1 \mathbf{Z}_1 + \cdots + w_{n-1} \mathbf{Z}_{n-1} + (1 - w_1 - \cdots - w_{n-1}) \mathbf{Z}_n. \tag{8}$$

Sampling from Dirichlet distributions is easy using the algorithm in Devroye [10]. Generate $Y_1, \ldots, Y_n$ independently, each Y_i having a gamma distribution with shape parameter α_i and scale parameter 1. Then set $W_i = Y_i/(Y_1 + \cdots + Y_n)$, $i = 1, \ldots, n$.

With $\alpha = (1, \ldots, 1)$, the Dirichlet distribution is the uniform distribution on the unit simplex. With $\alpha = (2, \ldots, 2)$, the Dirichlet density is proportional to the product used to define the empirical likelihood in (3). However, sampling from this weight distribution and then mapping weights to scenarios through (8) is not equivalent to sampling scenarios from the density proportional to $\mathcal{R}$ because $\mathcal{R}(\mathbf{Z})$ is defined by the *maximum* product over all weight vectors satisfying (8). When all α_i are equal, smaller values of the parameters generate more scenarios near the extreme points $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$, and larger values generate more scenarios near the mean. In simulation experiments, we have found that the uniform case $\alpha = (1, \ldots, 1)$ often produces a scenario distribution closer to that obtained by rejection sampling from $\mathcal{R}$ unless n is very small.

As n grows, the EL confidence regions tighten around the true (conditional) mean, and samples generated in proportion to $\mathcal{R}$ or using random weights as described above concentrate around a single point. To keep the dispersion of the scenarios generated consistent with the dispersion of the observed scenarios, we need to scale the scenarios by a factor that grows with n.

The central limit theorem suggests that we should scale the distance from the conditional mean by the square root of the number of historical scenarios. In more detail, let $\bar{\mathbf{z}}$ be the sample mean

$(\mathbf{Z}_1 + \cdots + \mathbf{Z}_n)/n$, let $\mathbf{Z}$ be as in (8), and set

$$\tilde{\mathbf{Z}} = \bar{\mathbf{z}} + \sqrt{n}(\mathbf{Z} - \bar{\mathbf{z}}). \tag{9}$$

This transformation dilates the contours of the distribution defined through (8), offsetting the concentration of the confidence regions. To see the analogy with the central limit theorem, suppose $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \Sigma)$ random vectors with sample mean $\bar{X}$ and sample covariance matrix S. The distribution of $\bar{X}$ is $N(\mu, \Sigma/n)$, which is approximated by $N(\bar{X}, S/n)$; scaling the distance from the mean by $\sqrt{n}$ transforms this distribution to $N(\bar{X}, S)$, which approximates the underlying distribution $N(\mu, \Sigma)$.

To illustrate, we consider the joint distribution of weekly returns of the Dow Jones Total Market Index and weekly changes in level of the VIX from the beginning of 1990 (the earliest date for which the VIX is available) through the end of February 2012, for a total of 1,156 observations. We select the worst 11 weeks (1%) as measured by the index returns; these are the weeks in which the index declines by more than 6.27%. We record both the change in the VIX and index returns in these weeks. These outcomes are indicated by the large circles in Figures 11-12. We interpret these as samples from the joint distribution of the index return and the change in the VIX conditional on a decline in the index greater than 6.27%.

In Figure 11, we generate 500 points using (8) and random weights from symmetric Dirichlet distributions with $\alpha_i = 1$ (left) and $\alpha_i = 2$ (right). The point marked by an asterisk is the conditional mean of the historical observations. The simulated points reflect uncertainty in this conditional mean, and all lie within the convex hull of the historical data. The figure on the right shows greater concentration near the mean, consistent with the larger value of α. In both cases, if we added observations to the 11 historical points, we would see greater clustering around the mean.

In Figure 12, we use random weights with $\alpha_i = 1$, as in the left panel of Figure 11, but then scale as in (9). The objective here is to generate extreme scenarios for the joint movement of the Dow Jones index and the VIX — not to draw from the distribution of the sample mean. We want the extreme scenarios we generate to extrapolate beyond the 11 historical scenarios in a sensible way. The figure shows 50,000 simulated points to illustrate the distribution; in practice, one would typically generate far fewer scenarios.

For comparison, the figure includes an ellipse defined by the mapping

$$\mathbf{z} \mapsto (\mathbf{z} - \bar{\mathbf{z}})^\top S^{-1} (\mathbf{z} - \bar{\mathbf{z}}),$$

with S the sample covariance matrix of the 11 historical observations. This is a contour of the normal distribution with matching mean and covariance — for illustration, we have chosen the contour that passes through the most extreme of the 11 historical scenarios. The key point is that

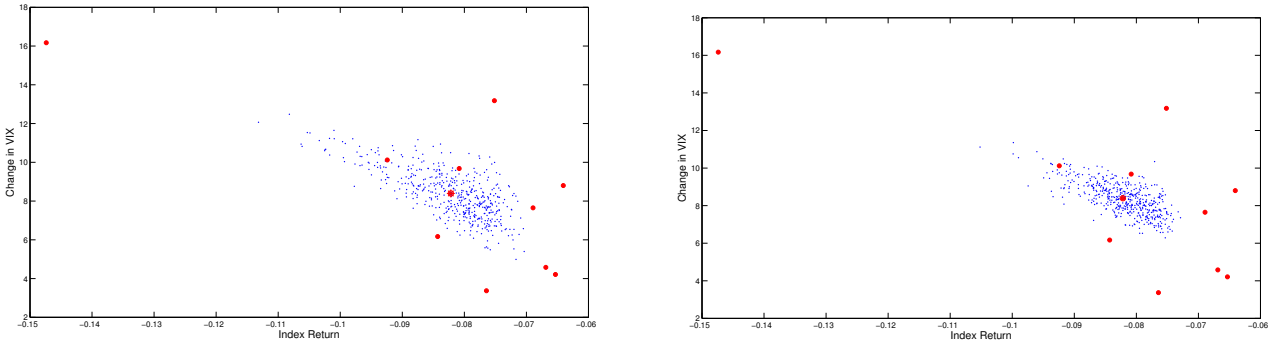


Figure 11: The large circles are historical observations for the VIX and Dow Jones Total Market Index — the worst 1% of weekly observations, as determined by the return on the Dow Jones index from January 1990 through February 2012. The asterisk indicates their conditional mean. The small points are simulated using Dirichlet weights with $\alpha_i = 1$ (left) and $\alpha_i = 2$ (right).

the simulated distribution driven by the random weights captures the skewness and general shape suggested by the historical scenarios. The ellipse indicates that sampling from a matching normal distribution instead would produce far more observations in the lower right corner and far fewer in the upper left. Skewness is to be expected in this examples (and others like it) because we have conditioned on falling below a linear combination of the variables — in this case, falling below an extreme value of the first coordinate.

9 Concluding Remarks

Stress testing is of growing importance in both industry and regulatory practice, yet there is rather little theory underpinning the selection of stress scenarios. In this paper, we have introduced an approach for selecting scenarios that combines historical data with qualitative information about the tail behavior of risk factors. Rather than simply posit a hypothetical extreme outcome, our method estimates the most likely scenarios leading to a loss of a given magnitude, which ensures the relevance of the scenarios selected for the portfolios to be stress tested.

Historical data on extreme events is, by definition, limited. Our method acknowledges uncertainty in extremes by generating confidence regions rather than just individual scenarios. These confidence regions combine a nonparametric empirical likelihood estimator, which captures skewness and other features of extreme observations, with an adjustment based on the heaviness of the tails of market risk factors.

As a further application of these ideas, we analyze marginal expected shortfall. MES measures

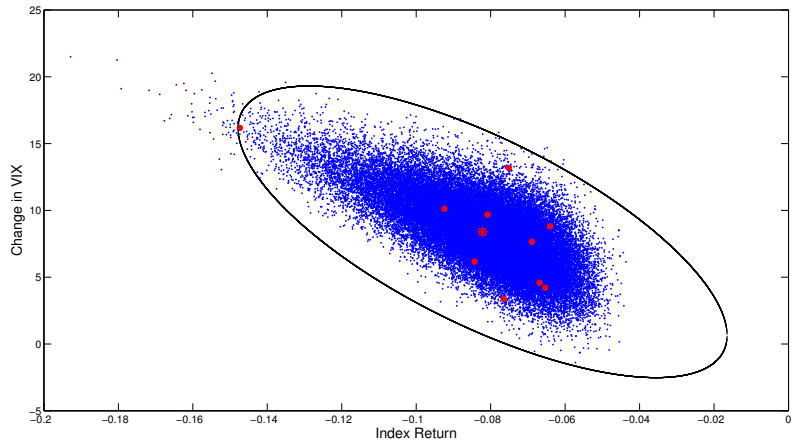


Figure 12: The small points indicate simulated scenarios using Dirichlet weights with $\alpha_i = 1$ and scaling as in (9). The ellipse is a contour of the normal distribution with the same mean and covariance as the 11 historical observations, which cannot capture the skewness in the data.

the expected loss in part of a portfolio conditional on a stress to the full portfolio. It has also been proposed as a measure of systemic risk when applied across firms rather than across parts of a single portfolio. Our analysis shows how the variability in MES estimates depends on the heaviness of the tails of market risk factors. We also show how empirical likelihood confidence regions can be used to assess the statistical significance of MES rankings, again taking into account the limited data available in extreme stress scenarios.

Acknowledgements. The authors thank Mark Flood, Matthew Pritsker, Til Schuermann for their comments and suggestions.

A Appendix: Proofs

A.1 Main Results

Most of the work in proving Proposition 1 and Theorem 1 lies in establishing the limits in Table 3. Before detailing these limits, we show how they lead to the stated results. For Proposition 1 we need finite means, hence the condition $\nu > 1$; the variance limits in Table 3 and Proposition 2 require finite variance, thus $\nu > 2$; and Theorem 1 requires finite fourth moments, hence $\nu > 4$.

Recall our standing assumption that the covariance matrix $\mathbf{\Sigma}$ of $\mathbf{Z}$ is positive definite. By relabeling L as the $(d+1)$st coordinate of $\mathbf{Z}$, we may rewrite $(\mathbf{Z}, L)$ simply as $\tilde{\mathbf{Z}} \in \mathbb{R}^{d+1}$. Conditioning on $L \geq \ell$ then reduces to conditioning on $\mathbf{c}^\top \tilde{\mathbf{Z}} \geq \ell$ with $\mathbf{c}^\top = (0, \ldots, 0, 1) \in \mathbb{R}^{d+1}$, provided the

covariance matrix $\tilde{\mathbf{\Sigma}}$ of $(\mathbf{Z}, L)$ remains positive definite. If $\tilde{\mathbf{\Sigma}}$ fails to be positive definite, then there exists a vector $\mathbf{c} \in \mathbb{R}^d$ such that $L = \mathbf{c}^\top \mathbf{Z} + h$, a.s., where h is a constant. Thus, for both cases, it suffices to consider conditioning on $\mathbf{c}^\top \tilde{\mathbf{Z}} \geq \tilde{\ell}$, for some $\tilde{\ell}$ and some $\mathbf{c} \neq \mathbf{0}$, with $\tilde{\mathbf{Z}}$ having a positive definite covariance matrix. We may therefore drop the tildes.

Proof of Proposition 1: By replacing ℓ with $\ell - \mathbf{c}^\top \boldsymbol{\mu}$, we may take $\boldsymbol{\mu} = 0$ in the representation (5). More generically, if a vector $\mathbf{Z}$ has a distribution of the form (5) with $\boldsymbol{\mu} = 0$, we can represent it as $\mathbf{Z} = \sqrt{W} A\mathbf{N}$, where the matrix A satisfies $AA^\top = \mathbf{\Sigma}$, and $\mathbf{N} \sim N(\mathbf{0}, I)$. Suppose we condition on $\mathbf{c}^\top \mathbf{Z} = \sqrt{W}\mathbf{c}^\top A\mathbf{N}$. Since $\mathbf{N}$ has a spherical distribution, we can find an orthogonal matrix R such that $\mathbf{c}^\top AR = \|A^\top \mathbf{c}\|(1, 0, \ldots, 0)$ and $R^{-1}\mathbf{N}$ is also $N(\mathbf{0}, I)$. That is,

$$(\mathbf{c}^\top AR)(\sqrt{W} R^{-1}\mathbf{N}) = \sqrt{W}\|A^\top \mathbf{c}\| Z_1'$$

where $\mathbf{Z}' = R^{-1}\mathbf{N} \sim N(\mathbf{0}, I)$. Hence, we can reduce an arbitrary linear combination to one of the form $\mathbf{c} = (1, 0, \ldots, 0)^\top$ with $\mathbf{Z} = \sqrt{W}\mathbf{N}$, where $\mathbf{N}$ is $N(\mathbf{0}, I)$.

In the setting of the proposition, this shows that we can find a linear bijection B such that $B\mathbf{z}^*(\ell) = \ell \times (1, 0, \ldots, 0)^\top$ and $B\bar{\mathbf{z}}(\ell) = \mathbb{E}[Z_1 | Z_1 \geq \ell] \times (1, 0, \ldots, 0)^\top$. By setting $\kappa_\ell = \ell / \mathbb{E}[Z_1 | Z_1 \geq \ell]$, we obtain the equation (6). The first row of Table 3 then gives the stated limits for κ_ℓ. □

Owen [21] provides a triangular array version of his EL theorem for data of the form $\mathbf{Z}_{1,n}, \ldots, \mathbf{Z}_{n,n}$, $n = 1, 2, \ldots$, in which variables with a shared second subscript are independent of each other and have a common mean. His result requires two conditions:

(i) For some $c > 0$, $\frac{\lambda_{m,n}}{\lambda_{M,n}} \geq c$, where $\lambda_{m,n}$ and $\lambda_{M,n}$ are the minimum and maximum eigenvalues of the covariance matrix associated with the n-th row of the array, respectively.

(ii) $\frac{1}{n^2} \sum_{i=1}^n \mathbb{E}\big[\|\mathbf{Z}_{i,n} - \boldsymbol{\mu}_n\|^4 \lambda_{M,n}^{-2}\big] \to 0$.

An additional convex hull condition required for the theorem is automatically satisfied by the normal, Laplace, and t distributions.

Define

$$D_n = \begin{cases} diag\{\ell_n, 1, \ldots, 1\} & \text{for normal case;} \\ diag\{1, \frac{1}{\sqrt{\ell_n}}, \ldots, \frac{1}{\sqrt{\ell_n}}\} & \text{for Lapalce case;} \\ diag\{\frac{1}{\ell_n}, \ldots, \frac{1}{\ell_n}\} & \text{for } t_\nu \text{ case} \end{cases}$$

where *diag* represents a diagonal matrix with specified elements. Let $\mathbf{Z}_1, \mathbf{Z}_2, \ldots$ are i.i.d. factor changes. Choose those satisfying $\mathbf{c}^\top \mathbf{Z}_j \geq \ell_n$, so that $\mathbf{Z}_{1,n}', \ldots, \mathbf{Z}_{n,n}'$ are i.i.d. samples from the distribution of $\mathbf{Z} | \{\mathbf{c}^\top \mathbf{Z} \geq \ell_n\}$. Then, we can apply the EL theorem for triangular arrays to the scaled factors $\mathbf{Z}_{k,n} = D_n B \mathbf{Z}_{k,n}'$ where B is the transform introduced in the proof of Proposition 1. Theorem 1 follows once we verify conditions (i) and (ii) for this array and apply Proposition 1. We

will show that the off-diagonal conditional covariances all vanish in Section A.2, so (i) will follow from limits of the conditional variances in Table 3. For (ii), we need to extend these limits to higher moments and these computations are shown in Section A.3, A.4, and A.5.

We apply the Theorem 1 to $D_n B \mathbf{Z}'_{k,n}$ for the inference of $D_n B \mathbb{E}[\mathbf{Z}|\mathbf{c}^\top \mathbf{Z} \geq \ell_n]$. To infer $\mathbb{E}[\mathbf{Z}|\mathbf{c}^\top \mathbf{Z} \geq \ell_n]$, we can apply the procedure to the original data $\mathbf{Z}'_{k,n}$ since the confidence region $C_{1-\alpha,n}$ consists of convex (so linear) combination of observed data.

A.2 Vanishing Conditional Covariances

For all $1 \leq i < j$,

$$\begin{aligned}
\mathbb{COV}(Z_i, Z_j | Z_1 \geq \ell) &= \mathbb{E}[Z_i Z_j | Z_1 \geq \ell] - \mathbb{E}[Z_i | Z_1 \geq \ell] \times \mathbb{E}[Z_j | Z_1 \geq \ell] \\
&= \mathbb{E}[Z_i Z_j | Z_1 \geq \ell] \\
&= \mathbb{E}\Big[a^2 W \ \mathbb{E}[N_i N_j | N_1, W] \Big| Z_1 \geq \ell\Big] \\
&= \mathbb{E}\Big[a^2 W \ \mathbb{E}[N_i | N_1, W] \mathbb{E}[N_j | N_1, W] \Big| Z_1 \geq \ell\Big] \\
&= 0 \qquad (10)
\end{aligned}$$

since $\mathbb{E}[Z_j | Z_1 \geq \ell] = 0$ and $\mathbb{E}[N_j | N_1, W] = 0$. Hence it is enough to consider the conditional variances to check the eigenvalue conditions (i) of conditional covariance matrices.

A.3 Normal Distribution

A.3.1 Asymptotic Conditional Moments for the Normal Distribution

Consider $\mathbf{Z} = a\sqrt{W}\mathbf{N} = \mathbf{N}$ where $\mathbf{N}$ is $N(\mathbf{0}, \mathbf{I})$. Define $A(\ell) = \{Z_1 \geq \ell\} = \{N_1 \geq \ell\}$. ϕ is the pdf of the standard normal random variable. Observe that

$$\mathbb{E}[Z_1^k | Z_1 \geq \ell] \cdot \mathbb{P}(A(\ell)) = \mathbb{E}\Big[Z_1^k \mathbf{1}_{\{Z_1 \geq \ell\}}\Big] = \int_\ell^\infty x^k \phi(x) dx$$

and

$$\mathbb{E}[Z_2^k | Z_1 \geq \ell] = \mathbb{E}[Z_2^k] = \begin{cases} 0 & \text{for } k = 1 \\ 1 & \text{for } k = 2 \\ 3 & \text{for } k = 4 \end{cases}.$$

Hence we have

$$\lim_{\ell \to \infty} \mathbb{V}(Z_2 | Z_1 \geq \ell) = 1 \qquad (11)$$

and

$$\lim_{\ell \to \infty} \mathbb{E}\Big[(Z_2 - \mathbb{E}[Z_2 | Z_1 \geq \ell])^4 \Big| Z_1 \geq \ell\Big] = 3. \qquad (12)$$

It is also easy to see

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathbb{E}[Z_1|Z_1\geq\ell]=\lim_{\ell\to\infty}\frac{\int_\ell^\infty x\phi(x)dx}{\ell\mathbb{P}(A(\ell))}=1. \tag{13}$$

A.3.2 Asymptotic Conditional Centered Moments for the Normal Distribution

Define

$$Z_1|\{Z_1\geq\ell\}=Z_\ell=\ell+\frac{1}{\ell}X_\ell.$$

Then X_ℓ has non-negative values and a pdf of

$$f_\ell(x)=\frac{1}{C_\ell}e^{-x}e^{-\frac{x^2}{2\ell^2}}\leq\frac{1}{C_\ell}e^{-x}$$

where $\lim_{\ell\to\infty}C_\ell=1$. (See Section 3.4 in Embrechts, Kluppelberg, and Mikosch [11], Set $\mu_\ell=\mathbb{E}[X_\ell]$ and observe that, by the dominated convergence theorem,

$$\lim_{\ell\to\infty}\mu_\ell=\lim_{\ell\to\infty}\int_0^\infty xf_\ell(x)dx=\int_0^\infty xe^{-x}dx=1.$$

Then $\ell(Z_\ell-\mathbb{E}[Z_\ell])=X_\ell-\mu_\ell$ and

$$\begin{aligned}\lim_{\ell\to\infty}\mathbb{E}\Big[(\ell Z_1-\mathbb{E}[\ell Z_1|Z_1\geq\ell])^k\Big|Z_1\geq\ell\Big] &= \lim_{\ell\to\infty}\mathbb{E}\Big[(\ell Z_\ell-\ell\mathbb{E}[Z_\ell])^k\Big]\\ &= \lim_{\ell\to\infty}\mathbb{E}\Big[(X_\ell-\mu_\ell)^k\Big]\\ &= \lim_{\ell\to\infty}\frac{1}{C_\ell}\int_0^\infty(x-\mu_\ell)^kf_\ell(x)dx\\ &= \int_0^\infty(x-1)^ke^{-x}dx,\end{aligned}$$

again by the dominated convergence theorem. It is also known that

$$\int_0^\infty(x-1)^2e^{-x}dx=1 \text{ and } \int_0^\infty(x-1)^4e^{-x}dx=9.$$

Hence we have

$$\lim_{\ell\to\infty}\mathbb{V}(\ell Z_1|Z_1\geq\ell)=1 \tag{14}$$

and

$$\lim_{\ell\to\infty}\mathbb{E}\Big[(\ell Z_1-\mathbb{E}[\ell Z_1|Z_1\geq\ell])^4\Big|Z_1\geq\ell\Big]=9. \tag{15}$$

From (13) we see that the conditional mean and the most likely loss scenario coincide asymptotically, as asserted in Proposition 1. Combining (11) and (14) confirms the eigenvalue condition for the covariance matrix and (12) and (15) provide the 4-th moment condition if we multiply by ℓ along the stressed direction.

A.4 Laplace Distribution

Consider $\mathbf{Z} = \sqrt{W}\mathbf{N}$ where $\mathbf{N}$ is $N(\mathbf{0}, \mathbf{I})$ and W follows an exponential distribution with mean $1/\lambda$. As before, $A(\ell) = \{Z_1 \geq \ell\} = \{\sqrt{W}N_1 \geq \ell\}$, and ϕ is the pdf of the standard normal random variable. It is well-known that $\mathbb{P}(Z_i \geq \ell) = \frac{1}{2}e^{-\sqrt{2\lambda}\ell}$ for all i. Set $\beta = \sqrt{2\lambda}$. From the memoryless property of the exponential distribution,

$$\begin{aligned}
\mathbb{E}[Z_1|Z_1 \geq \ell] &= \frac{\mathbb{E}[Z_1\mathbf{1}_{\{Z_1\geq\ell\}}]}{\mathbb{P}(Z_1 \geq \ell)} \\
&= \int_\ell^\infty x\frac{1}{2}\beta e^{-\beta x}dx \times 2e^{\beta\ell} \\
&= \ell + \frac{1}{\beta} = \ell + \frac{1}{\sqrt{2\lambda}} \qquad (16)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[Z_1^2|Z_1 \geq \ell] &= \frac{\mathbb{E}[Z_1^2\mathbf{1}_{\{Z_1\geq\ell\}}]}{\mathbb{P}(Z_1 \geq \ell)} \\
&= \int_\ell^\infty x^2\beta e^{-\beta x}dx \times e^{\beta\ell} \\
&= \Big(\ell + \frac{1}{\beta}\Big)^2 + \frac{1}{\beta^2}.
\end{aligned}$$

Hence

$$\mathbb{V}(Z_1|Z_1 \geq \ell) = 1/\beta^2 = 1/(2\lambda). \qquad (17)$$

The kurtosis satisfies

$$\mathbb{E}\Big[(Z_1 - \mathbb{E}[Z_1|Z_1 \geq \ell])^4\Big|Z_1 \geq \ell\Big] = 9. \qquad (18)$$

It is easy to see

$$\mathbb{E}[Z_2|Z_1 \geq \ell] = 0. \qquad (19)$$

The density of the generalized inverse Gaussian (GIG) distribution $N^-(\alpha, \beta, \gamma)$ is given by

$$\frac{\beta^{-\alpha/2}\gamma^{\alpha/2}}{2K_\alpha(\sqrt{\beta\gamma})}w^{\alpha-1}e^{-\frac{1}{2}(\beta/w+\gamma w)}$$

where K_α denotes a modified Bessel function of the third kind with index α. It is known that $K_{1/2}(x) = K_{-1/2}(x) = \sqrt{\frac{\pi}{2x}}e^{-x}$. Hence for $\alpha = 1/2, -1/2$,

$$\begin{aligned}
gig(\alpha, \beta, \gamma) &= \int_0^\infty w^{\alpha-1}e^{-\frac{1}{2}(\beta/w+\gamma w)}dw \\
&= 2\Big(\frac{\beta}{\gamma}\Big)^{\alpha/2}K_\alpha(\sqrt{\beta\gamma}) \\
&= \sqrt{2\pi}\beta^{\alpha/2-1/4}\gamma^{-\alpha/2-1/4}e^{-\sqrt{\beta\gamma}}.
\end{aligned}$$

For $x > 0$,

$$\mathbb{E}[W\mathbf{1}_{\{Z_1\geq\ell\}}|N_1 = x] \quad = \quad \int_{\ell^2/x^2}^{\infty} w\lambda e^{-\lambda w}dw = \Big(\frac{\ell^2}{x^2} + \frac{1}{\lambda}\Big)e^{-\lambda\ell^2/x^2}.$$

Note that

$$\begin{aligned}
&\mathbb{E}[Z_2^2|Z_1 \geq \ell] \cdot \mathbb{P}(Z_1 \geq \ell) = \mathbb{E}\Big[Z_2^2\mathbf{1}_{\{Z_1\geq\ell\}}\Big] \\
&= \mathbb{E}\Big[W\mathbb{E}[N_2^2|W]\mathbb{E}[\mathbf{1}_{\{Z_1\geq\ell\}}|W]\Big] \\
&= \mathbb{E}\Big[W\mathbb{E}[N_2^2]\mathbb{E}[\mathbf{1}_{\{Z_1\geq\ell\}}|W]\Big] \\
&= \mathbb{E}\Big[W\mathbb{E}[\mathbf{1}_{\{Z_1\geq\ell\}}|W]\Big] \\
&= \mathbb{E}\Big[W\mathbf{1}_{\{Z_1\geq\ell\}}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}[W\mathbf{1}_{\{Z_1\geq\ell\}}|N_1]\Big] \\
&= \int_0^{\infty}\Big(\frac{\ell^2}{x^2} + \frac{1}{\lambda}\Big)e^{-\lambda\ell^2/x^2}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx \\
&= \frac{1}{2\sqrt{2\pi}}\int_0^{\infty}\Big(\ell^2 y^{-1/2-1} + \frac{1}{\lambda}y^{1/2-1}\Big)e^{-\frac{1}{2}(2\lambda\ell^2/y+y)}dy \\
&= \frac{1}{2\sqrt{2\pi}}(\ell^2 \cdot gig(-1/2, 2\lambda\ell^2, 1) + \frac{1}{\lambda}gig(1/2, 2\lambda\ell^2, 1)) \\
&= \Big(\ell^2\frac{1}{\sqrt{2\lambda}\ell} + \frac{1}{\lambda}\Big)\frac{1}{2}e^{-\sqrt{2\lambda}\ell} \\
&= (\ell/\sqrt{2\lambda} + 1/\lambda)\mathbb{P}(Z_1 \geq \ell).
\end{aligned}$$

Hence we finally get

$$\mathbb{V}(Z_2|Z_1 \geq \ell]) = \mathbb{E}[Z_2^2|Z_1 \geq \ell] = (\ell/\sqrt{2\lambda} + 1/\lambda), \tag{20}$$

which implies the $\frac{1}{\sqrt{\ell}}Z_j$ for $j \neq 1$ has asymptotic variance $1/\sqrt{2\lambda}$ under the condition of $\{Z_1 \geq \ell\}$.

It can be also shown that

$$\begin{aligned}
&\mathbb{E}[Z_2^4|Z_1 \geq \ell] \cdot \mathbb{P}(Z_1 \geq \ell) = \mathbb{E}\Big[Z_2^4\mathbf{1}_{\{Z_1\geq\ell\}}\Big] \\
&= \mathbb{E}\Big[W^2\mathbf{1}_{\{Z_1\geq\ell\}}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}[W^2\mathbf{1}_{\{Z_1\geq\ell\}}|N_1]\Big] \\
&= \int_0^{\infty}\Big(\frac{\ell^4}{x^4} + \frac{2\ell^2}{\lambda x^2} + \frac{2}{\lambda^2}\Big)e^{-\lambda\ell^2/x^2}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx \\
&= \frac{1}{2\sqrt{2\pi}}\int_0^{\infty}\Big(\ell^4 y^{-3/2-1}\frac{2\ell^2}{\lambda}y^{-1/2-1} + \frac{2}{\lambda^2}y^{1/2-1}\Big)e^{-\frac{1}{2}(2\lambda\ell^2/y+y)}dy \\
&= \frac{1}{2\sqrt{2\pi}}(\ell^4 \cdot gig(-3/2, 2\lambda\ell^2, 1) + \frac{2\ell^2}{\lambda} \cdot gig(-1/2, 2\lambda\ell^2, 1) + \frac{2}{\lambda^2} \cdot gig(1/2, 2\lambda\ell^2, 1))
\end{aligned}$$

$$
\begin{aligned}
&= \Big(\ell^4 \frac{1}{2\lambda\ell^2}(1+\frac{1}{\sqrt{2\lambda}\ell}) + \frac{2\ell^2}{\lambda}\frac{1}{\sqrt{2\lambda}\ell} + \frac{2}{\lambda^2}\Big)\frac{1}{2}e^{-\sqrt{2\lambda}\ell} \\
&= \Big(\ell^2\frac{1}{2\lambda} + \ell\frac{2\lambda+1}{2\lambda\sqrt{2\lambda}} + \frac{2}{\lambda^2}\Big)\mathbb{P}(Z_1 \geq \ell)
\end{aligned}
$$

where we used $K_{-3/2}(x) = K_{3/2}(x) = \sqrt{\frac{\pi}{2x}}e^{-x}(1+1/x)$. Hence we finally get

$$
\mathbb{E}\Big[(Z_2 - \mathbb{E}[Z_2|Z_1 \geq \ell])^4 \Big| Z_1 \geq \ell\Big] = \frac{\ell^2}{2\lambda} + o(\ell^2). \tag{21}
$$

From (16) and (19), we conclude that the conditional mean and the most likely loss scenario coincide asymptotically for the Laplace distribution, as claimed in Proposition 1. Together, (20) and (17) confirm the eigenvalue condition for the covariance matrix, and (21) and (18) provide the 4-th moment condition if we divide the vectors orthogonal to the stressed direction by $\sqrt{\ell}$.

A.5 t Distribution

Consider $\mathbf{Z} = a\sqrt{W}\mathbf{N}$ where $\mathbf{N}$ is $N(\mathbf{0}, \mathbf{I})$ and W is a positive mixing random variable. As before, $A(\ell) = \{Z_1 \geq \ell\} = \{a\sqrt{W}N_1 \geq \ell\}$, ϕ is a pdf of the standard normal random variable, and a is any positive constant.

$$
\begin{aligned}
\mathbb{E}\Big[Z_1^k \mathbf{1}_{\{Z_1 \geq \ell\}}\Big] &= \mathbb{E}[Z_1^k | Z_1 \geq \ell] \cdot \mathbb{P}(A(\ell)) \\
&= a^k \mathbb{E}[W^{k/2} N_1^k \mathbf{1}_{A(\ell)}] \\
&= a^k \mathbb{E}\Big[\mathbb{E}[W^{k/2} N_1^k \mathbf{1}_{A(\ell)} | W]\Big] \\
&= a^k \mathbb{E}\Big[W^{k/2}\mathbb{E}[N_1^k \mathbf{1}_{A(\ell)} | W]\Big] \\
&= a^k \mathbb{E}\Big[W^{k/2}\int_{\ell/(a\sqrt{W})}^{\infty} x^k \phi(x)dx\Big] \\
&= a^k h_{k,k}(\ell/a).
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\Big[Z_2^k \mathbf{1}_{\{Z_1 \geq \ell\}}\Big] &= \mathbb{E}[Z_2^k | Z_1 \geq \ell] \cdot \mathbb{P}(A(\ell)) \\
&= a^k \mathbb{E}[W^{k/2} N_2^k \mathbf{1}_{A(\ell)}] \\
&= a^k \mathbb{E}\Big[\mathbb{E}[W^{k/2} N_2^k \mathbf{1}_{A(\ell)} | W]\Big] \\
&= a^k \mathbb{E}\Big[W^{k/2}\mathbb{E}[N_2^k | W]\mathbb{E}[\mathbf{1}_{A(\ell)} | W]\Big] \\
&= a^k \mathbb{E}[N_2^k]\mathbb{E}\Big[W^{k/2}\int_{\ell/(a\sqrt{W})}^{\infty} \phi(x)dx\Big] \\
&= \begin{cases} 0 & \text{for } k=1 \\ a^k \mathbb{E}\Big[W^{k/2}\int_{\ell/(a\sqrt{W})}^{\infty} \phi(x)dx\Big] = a^k h_{k,0}(\ell/a) & \text{for } k=2 \end{cases}
\end{aligned}
$$

where

$$h_{k,j}(\ell) := \mathbb{E}\Big[W^{k/2}\int_{\ell/\sqrt{W}}^{\infty} x^j\phi(x)dx\Big].$$

Define

$$g_r(\ell) := \mathbb{E}\Big[W^{r/2}\phi\Big(\frac{\ell}{\sqrt{W}}\Big)\Big].$$

If we assume $\mathbb{E}[W^k] < \infty$ for relevant k's, then we easily have the following differentiation rule:

$$\frac{d}{d\ell}g_r(\ell) \quad = \quad -\ell\cdot g_{r-2}(\ell).$$

Then we also have

$$\begin{aligned}
\frac{d}{d\ell}h_{k,j}(\ell/a) &= \frac{d}{d\ell}\mathbb{E}\Big[W^{k/2}\int_{\ell/(a\sqrt{W})}^{\infty} x^j\phi(x)dx\Big] \\
&= -\mathbb{E}\Big[W^{k/2}\frac{\ell^j}{a^jW^{j/2}}\phi\Big(\frac{\ell}{a\sqrt{W}}\Big)\frac{1}{a\sqrt{W}}\Big] \\
&= -\ell^j a^{-j-1}\mathbb{E}\Big[W^{(k-j-1)/2}\phi\Big(\frac{\ell}{a\sqrt{W}}\Big)\Big] \\
&= -\ell^j a^{-j-1}g_{k-j-1}(\ell/a).
\end{aligned}$$

A.5.1 Asymptotic Conditional Moments for the t-Distribution

Now take $\frac{1}{W}\sim\chi^2_\nu$ (that is, $W\sim IG(\frac{1}{2}\nu,\frac{1}{2})$), $a=\sqrt{\nu}$, and $f_W(w)=\frac{1}{2^{\nu/2}\Gamma(\nu/2)}w^{-\nu/2-1}e^{-\frac{1}{2w}}$. Then

$$\begin{aligned}
g_r(\ell) &= \mathbb{E}\Big[W^{r/2}\phi\Big(\frac{\ell}{\sqrt{W}}\Big)\Big] \\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty w^{r/2}e^{-\ell^2/(2w)}\frac{1}{2^{\nu/2}\Gamma(\nu/2)}w^{-\nu/2-1}e^{-\frac{1}{2w}}dw \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty w^{-(\nu-r)/2-1}e^{-\frac{\ell^2/2+1/2}{w}}dw \\
&= \frac{1}{\sqrt{2\pi}}\frac{(\frac{\ell^2+1}{2})^{-(\nu-r)/2}\Gamma((\nu-r)/2)}{2^{\nu/2}\Gamma(\nu/2)} \\
&\quad\times\int_0^\infty\frac{(\frac{\ell^2+1}{2})^{(\nu-r)/2}}{\Gamma((\nu-r)/2)}w^{-(\nu-r)/2-1}e^{-\frac{\ell^2/2+1/2}{w}}dw \\
&= \frac{\Gamma(\nu/2-r/2)}{\sqrt{2\pi}2^{r/2}\Gamma(\nu/2)}(\ell^2+1)^{-(\nu-r)/2}.
\end{aligned}$$

Since $Z_1\sim t_1(\nu,0,1)$,

$$\frac{d}{d\ell}\mathbb{P}(A(\ell)) = -\frac{\Gamma(\nu/2+1/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}(\ell^2/\nu+1)^{-(\nu+1)/2}.$$

For the case of $k+r=-1$,

$$\lim_{\ell\to\infty}\frac{\ell^k g_r(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} \quad = \quad \lim_{\ell\to\infty}\frac{\ell^k\frac{\Gamma(\nu/2-r/2)}{\sqrt{2\pi}2^{r/2}\Gamma(\nu/2)}(\ell^2/\nu+1)^{-(\nu-r)/2}}{-\frac{\Gamma(\nu/2+1/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}(\ell^2/\nu+1)^{-(\nu+1)/2}}$$

$$= \quad -\frac{\nu^{(k+1)/2}\Gamma(\nu/2 - r/2)}{2^{(r+1)/2}\Gamma(\nu/2 + 1/2)}.$$

For $r = 3, 1, -1, -3, -5$,

$$R(r) := \lim_{\ell\to\infty} \frac{\ell^k g_r(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \begin{cases} -\nu^{-3/2}(\nu-1)^{-1}(\nu-3)^{-1} & \text{for } r = 3 \\ -\nu^{-1/2}(\nu-1)^{-1} & \text{for } r = 1 \\ -\nu^{1/2} & \text{for } r = -1 \\ -\nu^{3/2}(\nu+1) & \text{for } r = -3 \\ -\nu^{5/2}(\nu+1)(\nu+3) & \text{for } r = -5 \end{cases}.$$

By the dominated convergence theorem, $\lim_{\ell\to\infty} \ell^k \mathbb{P}(A(\ell)) = 0$ for $k < \nu$ and

$$\lim_{\ell\to\infty} \mathbb{E}\Big[W^{k/2} \int_{\ell/(a\sqrt{W})}^{\infty} x^j \phi(x) dx\Big] = 0$$

for $k, j = 1, 2$. Hence we can apply the L'Hospital's rule for the limit of ratios.

$$\begin{aligned}
\lim_{\ell\to\infty} \frac{1}{\ell^k}\mathbb{E}\Big[Z_1^k \Big| Z_1 \geq \ell\Big] &= \lim_{\ell\to\infty} \frac{\mathbb{E}\Big[Z_1^k \mathbf{1}_{\{Z_1\geq\ell\}}\Big]}{\ell^k \mathbb{P}(A(\ell))} \\
&= \lim_{\ell\to\infty} \frac{\frac{d}{d\ell}\mathbb{E}\Big[Z_1^k \mathbf{1}_{\{Z_1\geq\ell\}}\Big]}{k\ell^{k-1}\mathbb{P}(A(\ell)) + \ell^k \frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= \lim_{\ell\to\infty} \frac{\nu^{k/2}\frac{d}{d\ell} h_{k,k}(\ell/\sqrt{\nu})}{k\ell^{k-1}\mathbb{P}(A(\ell)) + \ell^k \frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= \lim_{\ell\to\infty} \frac{-\nu^{k/2}\ell^k \nu^{-(k+1)/2} g_{-1}(\ell/\sqrt{\nu})}{k\ell^{k-1}\mathbb{P}(A(\ell)) + \ell^k \frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= \lim_{\ell\to\infty} \frac{-\nu^{-1/2}\ell g_{-1}(\ell/\sqrt{\nu})}{k\mathbb{P}(A(\ell)) + \ell \frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= \frac{1}{\frac{k}{-\nu} + 1} \\
&= \frac{\nu}{\nu - k} \qquad (22)
\end{aligned}$$

since

$$\lim_{\ell\to\infty} \frac{-\nu^{-1/2} g_{-1}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} = -\nu^{-1/2} R(-1) = 1$$

and

$$\begin{aligned}
\lim_{\ell\to\infty} \frac{-\nu^{-1/2}\ell g_{-1}(\ell/\sqrt{\nu})}{\mathbb{P}(A(\ell))} &= -\nu^{-1/2} \lim_{\ell\to\infty} \frac{g_{-1}(\ell/\sqrt{\nu}) + \ell\frac{d}{d\ell} g_{-1}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= -\nu^{-1/2} \lim_{\ell\to\infty} \frac{g_{-1}(\ell/\sqrt{\nu}) - \frac{\ell^2}{\nu} g_{-3}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))}
\end{aligned}$$

$$
\begin{aligned}
&= -\nu^{-1/2}(R(-1) - \nu^{-1}R(-3)) \\
&= -\nu.
\end{aligned}
$$

Hence we get

$$
\begin{aligned}
\lim_{\ell\to\infty} \mathbb{V}\Big(\frac{Z_1}{\ell}\Big|Z_1 \geq \ell\Big) &= \lim_{\ell\to\infty} \frac{1}{\ell^2}\mathbb{E}\Big[Z_1^2\Big|Z_1 \geq \ell\Big] - \lim_{\ell\to\infty}\Big(\frac{1}{\ell}\mathbb{E}\Big[Z_1\Big|Z_1 \geq \ell\Big]\Big)^2 \\
&= \frac{\nu}{\nu-2} - \Big(\frac{\nu}{\nu-1}\Big)^2 = \frac{\nu}{(\nu-2)(\nu-1)^2}. \qquad (23)
\end{aligned}
$$

For computation with the other variances, we observe that

$$
\begin{aligned}
\frac{d}{d\ell}h_{2,0}(\ell/\sqrt{\nu}) &= -\nu^{-1/2}g_1(\ell/\sqrt{\nu}); \\
\frac{d^2}{d\ell^2}h_{2,0}(\ell/\sqrt{\nu}) &= \ell\nu^{-3/2}g_{-1}(\ell/\sqrt{\nu}); \\
\frac{d^3}{d\ell^3}h_{2,0}(\ell/\sqrt{\nu}) &= \nu^{-3/2}g_{-1}(\ell/\sqrt{\nu}) - \ell^2\nu^{-5/2}g_{-3}(\ell/\sqrt{\nu}).
\end{aligned}
$$

$$
\begin{aligned}
\lim_{\ell\to\infty}\frac{1}{\ell^2}\mathbb{E}\Big[Z_2^2\Big|Z_1 \geq \ell\Big] &= \lim_{\ell\to\infty}\frac{\mathbb{E}\Big[Z_2^2\mathbf{1}_{\{Z_1\geq\ell\}}\Big]}{\ell^2\mathbb{P}(A(\ell))} \\
&= \lim_{\ell\to\infty}\frac{\nu\frac{d}{d\ell}h_{2,0}(\ell/\sqrt{\nu})}{2\ell\mathbb{P}(A(\ell)) + \ell^2\frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= \frac{\nu}{2(\frac{1}{1} + \frac{1}{-\nu^{-1}}) + \frac{1}{\nu^{-1}(\nu-1)^{-1}}} \\
&= \frac{\nu}{(\nu-2)(\nu-1)}
\end{aligned}
$$

from the following values:

$$
\lim_{\ell\to\infty}\frac{\frac{d}{d\ell}h_{2,0}(\ell/\sqrt{\nu})}{\ell^2\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \lim_{\ell\to\infty}\frac{-\nu^{-1/2}g_1(\ell/\sqrt{\nu})}{\ell^2\frac{d}{d\ell}\mathbb{P}(A(\ell))} = -\nu^{-1/2}R(1) = \nu^{-1}(\nu-1)^{-1},
$$

$$
\lim_{\ell\to\infty}\frac{\frac{d^2}{d\ell^2}h_{2,0}(\ell/\sqrt{\nu})}{\ell\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \lim_{\ell\to\infty}\frac{\ell\nu^{-3/2}g_{-1}(\ell/\sqrt{\nu})}{\ell\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \nu^{-3/2}R(-1) = -\nu^{-1},
$$

$$
\begin{aligned}
\lim_{\ell\to\infty}\frac{\frac{d^3}{d\ell^3}h_{2,0}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} &= \lim_{\ell\to\infty}\frac{\nu^{-3/2}g_{-1}(\ell/\sqrt{\nu}) - \ell^2\nu^{-5/2}g_{-3}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} \\
&= \nu^{-3/2}R(-1) - \nu^{-5/2}R(-3) \\
&= 1.
\end{aligned}
$$

Hence we get

$$
\lim_{\ell\to\infty}\mathbb{V}\Big(\frac{Z_2}{\ell}\Big|Z_1 \geq \ell\Big) = \lim_{\ell\to\infty}\frac{1}{\ell^2}\mathbb{E}\Big[Z_2^2\Big|Z_1 \geq \ell\Big] - 0^2 = \frac{\nu}{(\nu-2)(\nu-1)}. \qquad (24)
$$

For the computation of the conditional 4-th moment, we observe that

$$\begin{aligned}
\frac{d}{d\ell}h_{4,0}(\ell/\sqrt{\nu}) &= -\nu^{-1/2}g_3(\ell/\sqrt{\nu});\\
\frac{d^2}{d\ell^2}h_{4,0}(\ell/\sqrt{\nu}) &= \ell\nu^{-3/2}g_1(\ell/\sqrt{\nu});\\
\frac{d^3}{d\ell^3}h_{4,0}(\ell/\sqrt{\nu}) &= \nu^{-3/2}g_1(\ell/\sqrt{\nu}) - \ell^2\nu^{-5/2}g_{-1}(\ell/\sqrt{\nu});\\
\frac{d^4}{d\ell^4}h_{4,0}(\ell/\sqrt{\nu}) &= -3\ell\nu^{-5/2}g_{-1}(\ell/\sqrt{\nu}) + \ell^3\nu^{-7/2}g_{-3}(\ell/\sqrt{\nu});\\
\frac{d^5}{d\ell^5}h_{4,0}(\ell/\sqrt{\nu}) &= -3\nu^{-5/2}g_{-1}(\ell/\sqrt{\nu}) + 6\ell^2\nu^{-7/2}g_{-3}(\ell/\sqrt{\nu}) - \ell^4\nu^{-9/2}g_{-5}(\ell/\sqrt{\nu}).
\end{aligned}$$

$$\begin{aligned}
\lim_{\ell\to\infty}\frac{1}{\ell^4}\mathbb{E}\Big[Z_2^4\Big|Z_1\geq\ell\Big] &= \lim_{\ell\to\infty}\frac{\mathbb{E}\Big[Z_2^4\mathbf{1}_{\{Z_1\geq\ell\}}\Big]}{\ell^4\mathbb{P}(A(\ell))}\\
&= \lim_{\ell\to\infty}\frac{\nu\frac{d}{d\ell}h_{4,0}(\ell/\sqrt{\nu})}{4\ell^3\mathbb{P}(A(\ell)) + \ell^4\frac{d}{d\ell}\mathbb{P}(A(\ell))}\\
&= \frac{\nu}{\nu^2(\nu-1)(\nu-3) - 4\nu^2(\nu-1) + 12\frac{\nu^2(\nu-1)}{\nu-2} - 24\frac{\nu^2}{\nu-2} + 24\frac{\nu}{\nu-2}}\\
&= \frac{1}{(\nu-4)(\nu-3)(\nu-1)}
\end{aligned}$$

from the following values:

$$\lim_{\ell\to\infty}\frac{\frac{d}{d\ell}h_{4,0}(\ell/\sqrt{\nu})}{\ell^4\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \lim_{\ell\to\infty}\frac{-\nu^{-1/2}g_3(\ell/\sqrt{\nu})}{\ell^4\frac{d}{d\ell}\mathbb{P}(A(\ell))} = -\nu^{-1/2}R(3) = \nu^{-2}(\nu-1)^{-1}(\nu-3)^{-1},$$

$$\lim_{\ell\to\infty}\frac{\frac{d^2}{d\ell^2}h_{4,0}(\ell/\sqrt{\nu})}{\ell^3\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \lim_{\ell\to\infty}\frac{\ell\nu^{-3/2}g_1(\ell/\sqrt{\nu})}{\ell^3\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \nu^{-3/2}R(1) = -\nu^{-2}(\nu-1)^{-1},$$

$$\begin{aligned}
\lim_{\ell\to\infty}\frac{\frac{d^3}{d\ell^3}h_{4,0}(\ell/\sqrt{\nu})}{\ell^2\frac{d}{d\ell}\mathbb{P}(A(\ell))} &= \lim_{\ell\to\infty}\frac{\nu^{-3/2}g_1(\ell/\sqrt{\nu}) - \ell^2\nu^{-5/2}g_{-1}(\ell/\sqrt{\nu})}{\ell^2\frac{d}{d\ell}\mathbb{P}(A(\ell))}\\
&= \nu^{-3/2}R(1) - \nu^{-5/2}R(-1)\\
&= \frac{\nu-2}{\nu^2(\nu-1)},
\end{aligned}$$

$$\begin{aligned}
\lim_{\ell\to\infty}\frac{\frac{d^4}{d\ell^4}h_{4,0}(\ell/\sqrt{\nu})}{\ell\frac{d}{d\ell}\mathbb{P}(A(\ell))} &= \lim_{\ell\to\infty}\frac{-3\ell\nu^{-5/2}g_{-1}(\ell/\sqrt{\nu}) + \ell^3\nu^{-7/2}g_{-3}(\ell/\sqrt{\nu})}{\ell\frac{d}{d\ell}\mathbb{P}(A(\ell))}\\
&= -3\nu^{-5/2}R(-1) + \nu^{-7/2}R(-3)\\
&= \frac{2-\nu}{\nu^2},
\end{aligned}$$

$$\lim_{\ell\to\infty} \frac{\frac{d^5}{d\ell^5} h_{4,0}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))} = \lim_{\ell\to\infty} \frac{-3\nu^{-5/2} g_{-1}(\ell/\sqrt{\nu}) + 6\ell^2\nu^{-7/2} g_{-3}(\ell/\sqrt{\nu}) - \ell^4\nu^{-9/2} g_{-5}(\ell/\sqrt{\nu})}{\frac{d}{d\ell}\mathbb{P}(A(\ell))}$$
$$= -3\nu^{-5/2}R(-1) + 6\nu^{-7/2}R(-3) - \nu^{-9/2}R(-5) = (\nu-2)/\nu.$$

Hence we get

$$\lim_{\ell\to\infty} \frac{1}{\ell^4}\mathbb{E}\Big[(Z_2 - \mathbb{E}[Z_2|Z_1 \geq \ell])^4 \Big| Z_1 \geq \ell\Big] = \frac{1}{(\nu-4)(\nu-3)(\nu-1)}. \tag{25}$$

Applying (22) with $k = 1$ provides the asymptotic ratio between the conditional mean and the most likely loss scenario, as in Proposition 1. Together, (23), (24) and (10) confirm the eigenvalue condition for the covariance matrix, and (22) with $k = 4$ and (25) confirm the 4-th moment condition if we divide the vectors by ℓ.

References

[1] Acharya, V., Pedersen, L.H., Philippon, T., and Richardson, M. (2010) Measuring Systemic Risk, working paper, Stern School of Business, New York University, New York.

[2] Alfaro, R. A., and Drehmann, M. (2009) Macro Stress Tests and Crises: What Can We Learn, *BIS Quarterly Review*, December, 29–41.

[3] Balkema, G., and Embrechts, P. (2007) *High Risk Scenarios and Extremes: A Geometric Approach*, Zurich Lectures in Advanced Mathematics, European Mathematical Society Publishing, Zurich, Switzerland.

[4] Baysal, R.E., and Staum, J. (2008) Empirical Likelihood for Value-at-Risk and Expected Shortfall, *Journal of Risk* 11(1), 3–32.

[5] Board of Governors of the Federal Reserve System (2011) Comprehensive Capital Analysis and Review: Summary Instructions and Guidance, November 22, 2011. Available at http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20111122d1.pdf.

[6] Borio, C., Drehmann, M., and Tsatsaronis, K. (2012) Stress-Testing Macro Stress Testing: Does it Live up to Expectations? BIS Working Paper 369, Bank for International Settlements, Basel, Switzerland.

[7] Breuer, T., Jandacka, M., Rheinberger, K., and Summer, M. (2009) How to Find Plausible, Severe, and Useful Stress Scenarios, *International Journal of Central Banking* 5, 205–224.

[8] Brownlees, C., and Engle, R. (2011) Volatility, Correlation and Tails for Systemic Risk Measurement, working paper, Stern School of Business, New York University, New York.

[9] De, R., and Tamarchenko, T. (2002) VaR You Can Rely On, *Risk*, August, 77–81.

[10] Devroye, L. (1986) *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.

[11] Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[12] Financial Services Authority (2009) Stress and Scenario Testing, *Policy Statement 09/20*, London, United Kingdom.

[13] Flood, M., and Korenko, G. (2012) Systemic Scenario Selection, forthcoming.

[14] Heyde, C.C., and Kou, S.G. (2004) On the Controversy Over Tailweight of Distributions, *Operations Research Letters* 32, 399–408.

[15] Hida, E.T. (2011) *Deloitte Global Risk Management Survey*, Seventh Edition, Deloitte, London.

[16] Hirtle, B., Schuermann, T., and Stiroh, K. (2009) Macroprudential Supervision of Financial Institutions: Lessons from the SCAP, Staff Report 409, Federal Reserve Bank of New York.

[17] Koyluoglu, H.U. (2006) Imagine Iso-Loss Surfaces, working paper, Oliver Wyman, New York.

[18] Kupiec, P. (1998) Stress Testing in a Value at Risk Framework, *Journal of Derivatives* 6, 7–24.

[19] Moody's (2011) *Banking Industry Survey on Stress Testing*, Moody's Analytics, New York.

[20] Ong, L.L., Maino, R., and Duma, N. (2010) Into the Great Unknown: Stress Testing with Weak Data, IMF Working Paper WP/10/282, International Monetary Fund, Washington, D.C.

[21] Owen, A.B. (2001) *Empirical Likelihood*, Chapman & Hall/CRC Press, Boca Raton, Florida.

[22] Pritsker, M. (2011) Enhanced Stress Testing and Financial Stability, working paper, Federal Reserve Bank of Boston.

[23] Quagliarello, M. (2009) *Stress-Testing the Banking System*, Cambridge University Press, Cambridge, U.K.

[24] Rebonato, R. (2010) *Coherent Stress Testing: A Bayesian Approach*, Wiley, New York.

[25] Schmidt, R. (2002) Tail Dependence for Elliptically Contoured Distributions, *Mathematical Methods of Operations Research* 55, 301–327.

[26] Schuermann, T. (2012) Stress Testing Banks, Wharton Financial Institutions Working Paper 12-08, University of Pennsylvania, Philadelphia, Pennsylvania.