Axioma In-Practice Series

Toward Dynamic Stress Tests

Iulian Cotoi and Robert Stamicar

November 2017

In an earlier paper about best practices for stress testing market risk factors, we discussed different types of stress tests.[1] One of these types, called *transitive* or *predictive*, is a correlated stress test. Here a set of core (explicit) factors is selected and shocked while the remaining shifts of peripheral (implied) factors are inferred from the covariance of factor returns. For example, in testing a portfolio's sensitivity to oil prices, oil prices are the core factors. With their embedded factor return correlations, transitive stress tests provide more realistic P&L scenarios than user-defined stress tests.

But how should we select the appropriate look-back period for transitive stress tests? Typically, risk managers subjectively select periods of elevated correlation and volatilities, which are then used as inputs for transitive stress tests. But what if, instead, the stress scenario determined the period when the shift was most likely to occur? More precisely, given a stress test scenario, what is the probability that a given regime contributed to this shift?

[1] See Cotoi and Stamicar [2].

Flexible is better.™

We still select the core factor shifts, but in the real world, a shift always occurs in a market context. For instance, a large shift is more likely to occur under adverse market conditions while a mild shift is more likely to occur in a calm period. Thus, transitive stress tests should be designed in which a large shock in core factors uses periods of elevated volatility and correlations for covariance matrix calibration, while a small shift in core factors incorporates normal periods of volatilities and correlations.

Rather than explicitly selecting specific periods,[2] the transitive stress test itself implicitly selects the applicable period.[3] Accordingly, we then use this period for the transitive stress test. Moreover, a blended or probability-weighted combination of transitive stress tests can be implemented to incorporate intermediate shocks (along with extreme shocks). Thus, the transitive stress test dynamically adjusts as the shock size varies.

It still remains to identify different regimes for our analysis. Statistical tools used in machine learning (ML), such as cluster analysis, can identify specific regimes for volatility and correlation estimates for transitive stress tests. In fact, there has been a surge in utilizing ML techniques in different areas of finance, such as validation of bank's models built for regulatory stress tests, such as CCAR, despite regulatory concern over black-box approaches.

Our goal in this note is to enhance traditional transitive stress tests in two ways:

- By using statistical techniques to identify different regimes
- When appropriate, by dynamically adjusting each regime's impact on the stress test. The core factor shifts will determine the probability that each regime generated the observation.

In the second point above, a link is introduced between risk factor shocks and different regimes. Although a link between shock size and regime is useful, this is relevant only if we believe that migrating to different regimes (as opposed to an absorbing state resulting from a structural shift) is possible.

Ultimately, to achieve our goal, we will blend multiple transitive stress tests via a mixture model. In each regime, separate transitive stress tests are applied, and then probability weighted for the final result.

Transitive Stress Test Background

Under a transitive stress test, the movement of peripheral factors is inferred from the core factor shifts. We can represent this relationship as:

$$r^{(\mathrm{p})} = B r^{(\mathrm{c})} \,, \tag{1}$$

where $r^{(\mathrm{c})}$ is a vector of core factor returns, $r^{(\mathrm{p})}$ is a vector of peripheral factor returns, and B is a matrix of betas. Once B is known, we shift $r^{(\mathrm{c})}$, compute $r^{(\mathrm{p})}$ via (1), and then reprice the portfolio.[4] The computation of B is straightforward if we assume that factor returns are normally distributed. Consider the following covariance ma-

[2] Other than the overall look-back period.

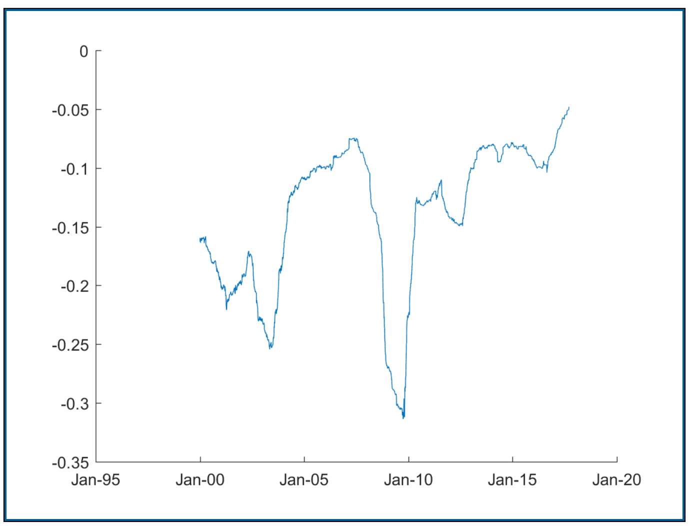
[3] The applicable periods will be determined by weighting historical returns with probabilities.

[4] Asset prices can be modeled using nonlinear models.

Table 1: Transitive Stress Test with Different Look-Back Periods on 20-Sept-2017.

Reporting Levels	**VIX +10% Current Period (%)**	**VIX +10% 2011-2012 (%)**
US Equity Sept 20 2017	−0.45	−1.38
▼ Manager 1		
► Stock (121)	−0.24	−0.74
▼ Manager 2		
► Stock (77)	−0.21	−0.64

Figure 1: Rolling Stock Beta to VIX.



trix, partitioned under core and peripheral factor returns:

$$\Sigma = \begin{bmatrix} \Sigma_{\text{pp}} & \Sigma_{\text{pc}} \\ \Sigma_{\text{cp}} & \Sigma_{\text{cc}} \end{bmatrix} \tag{2}$$

Taking conditional expectations with respect to core factor returns gives

$$E[r^{(\text{p})} \mid r^{(\text{c})}] = \Sigma_{\text{pc}}\Sigma_{\text{cc}}^{-1} r^{(\text{c})}\,, \tag{3}$$

Thus, $B = \Sigma_{\text{pc}}\Sigma_{\text{cc}}^{-1}$.

The computation of B depends on the look-back period and can vary significantly under different periods. As a motivating example, we consider a transitive stress test where the VIX is the core factor. In Table 1, we apply a shift of 10% and measure its P&L impact on a US-based equity portfolio. The analysis date is 20-Sept-2017, and two different look-back periods are chosen. In the first period, we simply utilize a look-back of one year from the analysis date. The resulting loss under current correlations is only 0.45%. The second period is from July 2011 to July 2012, resulting in a much greater loss of 1.38%.

Clearly, the look-back is important. In the first case, we were using current correlations under a *quiet* period while the second period was *noisy* with elevated volatility. We can generally represent our results in Table 1 as

$$r^{(\text{p})} = \overbrace{w\beta^{(1)} r^{(\text{c})}}^{\text{Noisy Period}} + \overbrace{(1-w)\beta^{(2)} r^{(\text{c})}}^{\text{Quiet Period}}\,, \tag{4}$$

where $w = 1$ represents the *noisy* period from July 2011 to July 2012 while $w = 0$ represents the recent *quiet* period.

We might argue the stress test under the noisy period ($w = 1$) in Equation (4) is more appropriate since the goal of stress tests is to apply adverse conditions to our portfolio. Nevertheless, measuring the difference in stress test P&Ls between current and volatile periods is helpful.

It is common to identify different periods of volatilities and correlations by examining rolling plots of volatilities, correlations, or betas. In Figure 1, which plots the rolling beta of the S&P 500 to the VIX, we see that the beta varied significantly over the historical periods that were used in the risk report in Table 1. Another way to identify regimes is to set a bound on VIX returns as illustrated in Figure 2. This is one way to create a library of relevant look-back periods for covariance matrix calibration.

Regime-Switching Approach

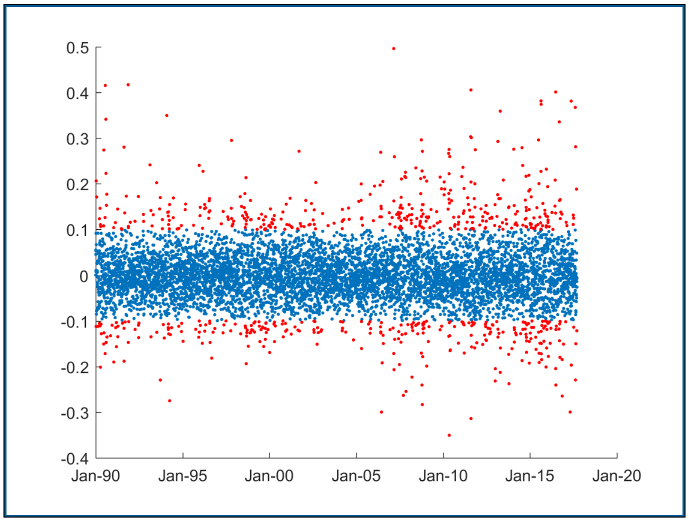
As mentioned earlier, traditional transitive stress tests rely on selecting look-back periods, usually based on subjective choices for historical volatilities and correlations.

Identifying regimes, such as low-volatility regimes, or identifying a structural change in an asset class, can help us design better stress tests (or even a better pricing model).

Logistic Mixture of Linear Components

One standard way to model different regimes is to use a mixture model such as a Gaussian mixture model (GMM). In this setting, returns are modeled as a linear combination of Gaussian distributions. GMM is classified as unsupervised learning since we do not know the regimes beforehand. The mixture probabilities are key parameters that need to be es-

Figure 2: Daily Returns of the VIX.



timated. However, the mixture probabilities under a traditional GMM are not specified in a parametric form. Instead, we examine the Logistic Mixture of Linear Components (LMLC) model introduced by Tashman [4] and Tashman and Frey [5]. Under the LMLC model, prior probabilities are modeled using a logistic function, which can now depend on the sensitivity of stress factor shifts. Thus, the LMLC model is a linear combination of linear regression models, like Equation (4), but in the LMLC model, the weights in Equation (4) are probabilities of the different regimes, which are in turn functions of stress factor returns.

In fact, the LMLC model setting specified by Tashman [4] is quite general. The core factors are grouped into (i) *mixing* components that determine the probability that each regime is responsible for producing an observation and (ii) *regression* factors that are used for each regression. These sets of core factors can overlap or even be disjoint.

For simplicity, we consider only two regimes. (Extending to multiple regimes is straightforward.) The core factors, at time t_i, are grouped as follows:

s_i : d-dimensional vector of mixing components

$x_i^{(1)}$: d_1-dimensional vector of stress factors for Component 1

$x_i^{(2)}$: d_2-dimensional vector of stress factors for Component 2

Let $y = (y_1, y_2, \ldots, y_n)^T$ represent the response variable that is generated from different regimes with n observations. We assume each y_i is generated from a different regime, which is not known beforehand.

Dynamically adjusting transitive stress tests to shocks

The stress factors in the mixing component s can be distinct from the stress factors appearing in each of the linear components. These risk factors deter-

Expectation-Maximization (EM) Algorithm

EM is a powerful statistical technique that estimates parameters of statistical models via likelihood functions.

Under EM, the model depends on latent or unobserved random variables–in our case, an unknown function that identifies the regimes. The EM algorithm starts with an initial guess of parameters, and then iterates between the following two steps:

- **E-step:** Given the current value of the parameters, estimate the distribution of the hidden variable.

  For each observation, estimate the (posterior) probability that each mixture distribution generated it.

- **M-step:** Update the parameters in order to maximize the joint distribution of the data and the hidden variable.

  Compute parameters maximizing the expected log-likelihood found in the E step.

mine the (prior) probability that each regime generated the observation. The prior probabilities for Regimes 1 and 2 are modeled using a logistic function:

$$p_1(s_i \mid \alpha) = \frac{e^{\alpha^T s_i}}{1 + e^{\alpha^T s_i}} \tag{5}$$

$$p_2(s_i \mid \alpha) = 1 - p_1(s_i \mid \alpha) \tag{6}$$

The vector of parameters α is calibrated to data and measures how sensitive the probabilities are to stress factors. These parameters thus provide a link between prior probabilities and stress factors s. For example, a large shock in s might correspond to a probability close to one for a more volatile regime, while a moderate shock might have a significant probability for another regime.

The main result of the LMLC model can be represented by the following equation, which gives the expected return of the response variable conditioned on the core factor shifts $(s_j, x_j^{(1)}, x_j^{(2)})$:

$$\begin{aligned}\mathbf{E}(y_j \mid s_j, x_j^{(1)}, x_j^{(2)}) = p_1(s_j|\alpha)\mathbf{E}_1(y_j \mid x_j^{(1)}) \\ + p_2(s_j|\alpha)\mathbf{E}_2(y_j \mid x_j^{(2)}),\end{aligned} \tag{7}$$

where the expected return conditional on regime k is

$$\mathbf{E}_k[y_j \mid x_j^{(k)}] = \left(\beta^{(k)}\right)^T x_j^{(k)} \tag{8}$$

Equation (7) extends transitive stress tests to multiple regimes. For each regime, the regression functions are referred to as *linear components*.

The parameters in Equations (7)–(8), $(\alpha, \beta^{(1)}, \beta^{(2)})$, are estimated using the Expectation Maximization (EM) algorithm.[5] The EM algorithm computes the maximum likelihood parameters for

[5] See Dempster, Laird, and Rubin [3].

How do we select regimes? How many?

We mentioned that regimes can be selected subjectively or by expert knowledge.

Under the logistic mixture of linear components (LMLC), we first select the number of regimes. The fitting of the LMLC model is based on the EM algorithm, which provides estimates for beta for each regime.

The number of regimes can be determined from an iterative procedure. We can sequentially increase the number of regimes, apply the EM algorithm, and determine if the parameters from the regressions are still statistically significant or not. In addition, we can use the Bayesian information criterion (BIC) to determine the appropriate number of regimes:

$$\mathit{BIC} = -2l + p \log(n) \,,$$

where l is the log-likelihood and p is the number of parameters in the model. Under BIC, there is a trade-off by increasing the likelihood by fitting the model with more parameters (resulting in overfitting). The model with the more negative BIC value is desirable.

statistical models, which have unknown or latent random variables. In our case, the unknown random variable itself indicates which component generated an observation. (See Appendix A for more details.)

We can interpret (7), loosely, to be a combination of two transitive stress tests. Under the general setting, the core risk factors in each regime can be different, and the probabilities themselves can be driven from yet another set of core factors. Although the LMLC setting is fairly general, our examples in the next section will be simple, as we start by considering only one core factor.

Examples

Our first example involves a dividend future. Dividends are usually described as an asset class because they are a dedicated trading instrument and have specific dynamics (see [1]). In June 2008, Eurex launched listed futures benchmarked on the EURO STOXX 50 Dividend Points index (SX5ED) to replace or complement dividend swap offerings. The SX5ED futures have annual expiry dates set to the third week of December. The EURO SX5E is the most active market for dividend futures as the European stocks pay a superior dividend yield.

We consider a specific futures contract, FEXDZ7, which expires in December 2017. As an explanatory variable, we choose the SX5E's relative return. We apply the LMLC model, as specified by Equation (7). Figure 3 provides a plot of FEXDZ7 returns along with the regimes. In this plot, we observe two regimes, where the gray blocks indicates one of the regimes. Figure 4 shows the histogram of estimated posterior probabilities of the first component. (The second component has a similar plot.) A peak at the probability of one indicates that the mixture component is well separated from the other components while significant mass in the middle of the histogram indicates overlap with other components. From this histogram, we observe a significant count for probability equal to one, suggesting that two regimes exist.

The results of the regressions for each component are provided in Table 2. The p-values for the betas provided in Table 2 are extremely low, indicating that they are statistically significant. The dividend future has a beta of 0.35 in the first regime, and a low beta of 0.04 in the second regime.

Stress test the dividend future

Now suppose the analysis date is 23-Aug-2017 (the last observation in Figure 3). At this date, the dividend future is in the low beta regime. We apply a stress test of -10% to SX5E returns. Adjusting the probabilities, we obtain a relative return of -2.7% for the future contract. This is a weighted average of -3.48% and -0.36% with probabilities of 0.76 and 0.24, respectively. Since this is a fairly large shock, we have a greater contribution from the first regime.

But is this correct? Identifying regimes is a useful first step. But adjusting probabilities makes sense if reverting to the first regime seems possible. In this example, one should not adjust the probability since the dividend future is in an *absorbing* state, which we will describe below. Here expert knowl-

Avoid pitfalls in regime-switching models

Regime-switching models, such as LMLC, are powerful, as are algorithms like EM for parameter estimation.

Knowing the current regime is important. But be careful of absorbing states. These are obvious in cases involving, say, firm default, but less obvious in the example involving the dividend future. From a risk perspective, we need to know when to turn on/off probabilities.

edge should override our temptation to alter probabilities.

Pull-to-realized effect and regime-switching

The pull-to-realized effect common in dividend futures makes them an instrument with very specific dynamics. Around 1.5 years to expiry, dividend futures undergo a big change in their risk dynamics and start trading at a discount to a fixed bullet payment corresponding to the expected dividends (see [1]). Consequently, volatility decreases as early as Q4 of the previous year. The pull-to-realized effect is believed to come from three sources: long-term dividends trade more like equities; short-term dividends trade like zero-coupon bond; and the dividend uncertainty disappears early in the process (after a few months).

This example demonstrates that identifying regimes is useful, but that one should not blindly adjust probabilities. In this case, there is a structural change in the dividend future's dynamics. So a more appropriate stress test for a -10% shift in SX5E is to use the current regime, which results in a P&L of -0.36%.

Although adjusting the component probabilities for dividend futures did not make sense, adjusting component probabilities for other risk factors such as FX rates, implied volatilities, or asset class returns such as hedge fund returns is worthwhile. For instance, we have empirical evidence of interest rate regimes in which yield curves steepen or flatten, corresponding to economic cycles. In this case, a mixing component based on a macroeconomic factor, such as GDP or manufacturing output, would be appropriate.

Another application of regime-switching is the modeling of hedge fund strategies that exhibit asymmetric payoff profiles. Tashman [4] investigates the returns of merger arbitrage strategies and shows that stress tests under LMLC produce more severe expected losses than a linear model would predict.

We finish this section with our second example involving the VIX. The VIX is a key gauge of fear in the market and is sometimes referred to as the *fear index*. We treat VIX returns as the response variable and select the S&P 500 return as the core factor. We fit two regimes under the LMLC model. Table 3 shows the results. The p-values for the

Figure 3: Evidence of regimes for FEXDZ7

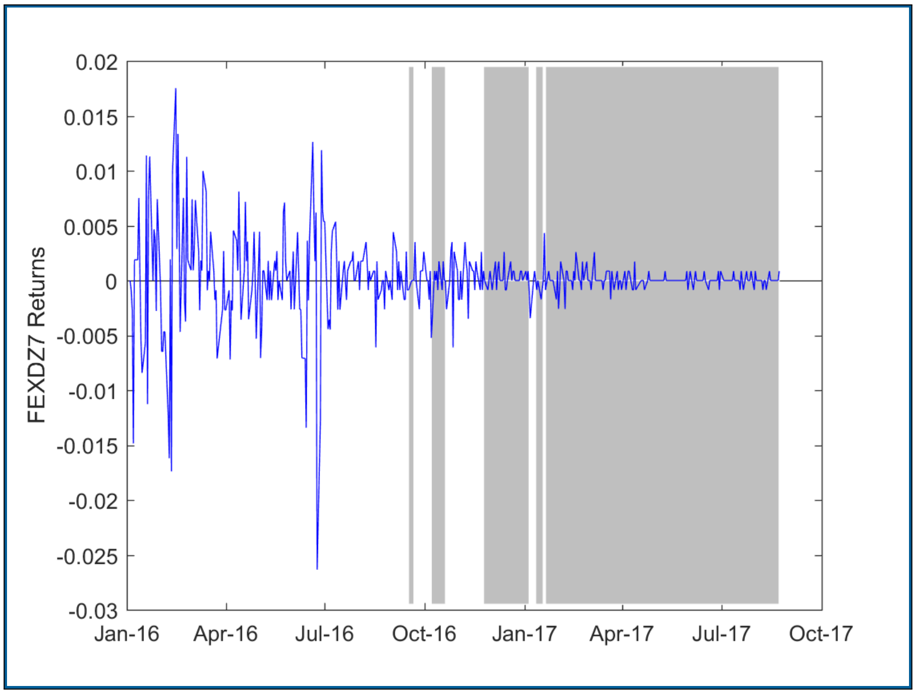


Figure 4: Regime Probabilities for Component 1

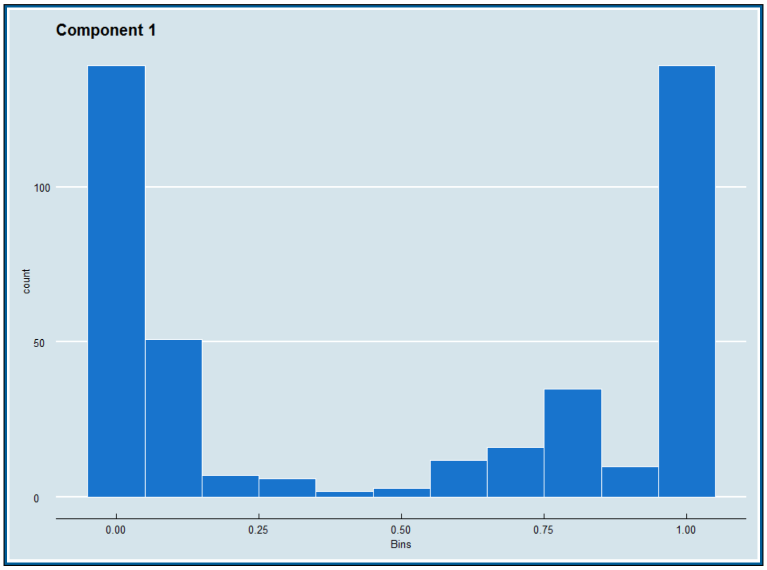


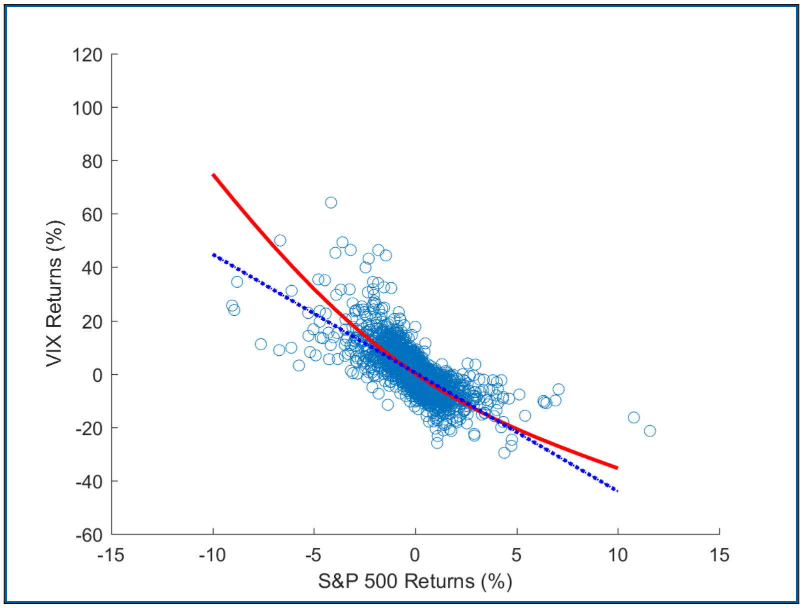
Table 2: Logistic Mixture Model Fit for Dividend Future FEXDZ7

	Coefficient	**Estimate**	**Std. Error**	**P-value**
Regime 1	Intercept	0.045	0.022	0.04
	β_{S5XE}	0.35	0.017	$< 2 \times 10^{-16}$
Regime 2	Intercept	0.012	0.007	0.08
	β_{S5XE}	0.037	0.013	0.004

Table 3: Logistic Mixture Model Fit for VIX for 2007-2017

	Coefficient	**Estimate**	**Std. Error**	**P-value**
Regime 1	Intercept	-0.33	0.12	0.002
	β_{SP500}	-2.98	0.086	$< 2.2 \times 10^{-16}$
Regime 2	Intercept	0.80	0.20	8.5×10^{-05}
	β_{SP500}	-9.04	0.25	$< 2.2 \times 10^{-16}$

Figure 5: Nonlinear Stress Test for VIX



beta coefficients are small, and thus significant. The second regime includes elevated returns of the VIX during adverse market periods.

One key benefit to applying mixture models is that a nonlinear profile can arise from core factor shifts. A linear combination of Gaussian distributions can produce a fat-tailed distribution. Figure 5 displays this nonlinear profile. Applying a linear model (the dashed line in Figure 5) would underestimate the movement of the VIX for negative shifts in the S&P 500.

Concluding Remarks

Transitive stress tests incorporate correlations and volatility estimates for risk factors. They allow a parsimonious set of core factors to be shifted, while the remaining shifts in peripheral factors are inferred from a calibrated covariance matrix. Transitive stress tests are useful since they enable risk and portfolio managers to construct forward-looking scenarios and examine relationships among different asset classes.

The required volatility and correlation estimates for transitive stress tests are commonly estimated from specific historical periods. However, as we have shown, a logistic mixture model can instead allow the stress tests themselves to determine the probability that each regime is responsible for producing the stress observation. This approach should be used alongside the traditional approach in which specific periods are explicitly specified.

References

[1] Bourgois, S. and Tom, E. (2013). Extracting the dividend risk premium, *Equity Derivatives Strategy*, Market Commentary, September 25, 2013, Credit Suisse.

[2] Cotoi, I. and Stamicar, R. (2017). Stress testing best practices, *Axioma In-Practice Series*, June 2017.

[3] Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society*, Series B, 39(1), 1-38.

[4] Tashman A. (2009). A regime-switching approach to model-based stress testing, *Journal of Risk Model Validation* 3(4), Winter 2009/10, 1–13.

[5] Tashman A. and Frey, R.J. (2009). Modeling risk in arbitrage strategies using finite mixtures, *Quantitative Finance* 9(5), 495–503.

Appendix A: Expectation-Maximization (EM) Algorithm

We outline the maximum likelihood parameter estimation for LMLC using EM (which is outlined in Tashnam and Frey [5]). First, the *complete* maximum likelikehood function is

$$L(\alpha, \beta^{(1)}, \beta^{(2)}, \sigma_1, \sigma_2) = \prod_{i=1}^{n} \left[p_1(s_i|\alpha) f_1(y_i \mid x^{(1)}, \beta^{(1)}, \sigma_1^2) \right]^{z_{1i}} \left[p_2(s_i|\alpha) f_2(y_i \mid x^{(2)}, \beta^{(2)}, \sigma_2^2) \right]^{z_{2i}} \quad \text{(A1)}$$

where z is the random variable indicating which component is responsible for the observation. For example, z_{1i} is one if the i-th observation is generated from the first component, otherwise it is equal to zero. The probability of observing the response in regime k conditional on stress factor settings at t_i is $f_k(y_i|x^{(k)}, \beta^{(k)}, \sigma_k^2)$. If the errors are normally distributed then

$$f_k(y_i \mid x^{(k)}, \beta^{(k)}, \sigma_k^2) = \frac{1}{2\pi\sigma_k^2} e^{-(y_i - \beta^{(k)T} x^{(k)})^2/(2\sigma_k^2)} \quad \text{(A2)}$$

The log-likelihood equation is

$$l(\alpha, \beta^{(1)}, \beta^{(2)}, \sigma_1, \sigma_2) = \sum_{k=1}^{2} \sum_{i=1}^{n} z_{ki} \left[\log(p_k(s_i|\alpha) + \log(f_k(y_i \mid x^{(k)}, \beta^{(k)}, \sigma_k^2)) \right] \quad \text{(A3)}$$

Substituting prior probabilities and the probability density f gives

$$\begin{aligned} l = &\sum_{i=1}^{n} z_{1i} \left[\log\left(\frac{e^{\alpha^T s}}{1 + e^{\alpha^T s}} \right) - \frac{1}{2} \log(2\pi\sigma_1^2) - \frac{(y_i - \beta^{(1)T} x_i^{(1)})^2}{2\sigma_1^2} \right] \\ + &\sum_{i=1}^{n} z_{2i} \left[\log\left(\frac{1}{1 + e^{\alpha^T s}} \right) - \frac{1}{2} \log(2\pi\sigma_2^2) - \frac{(y_i - \beta^{(2)T} x_i^{(2)})^2}{2\sigma_2^2} \right] \end{aligned} \quad \text{(A4)}$$

E-Step

For each i, compute posterior probabilities for regime k:

$$\tau_{ki} = \frac{p_k(s_i|\alpha) f_k(y_i \mid x_i^{(k)}, \beta^{(k)}, \sigma_k^2)}{\sum_{j=1}^{2} p_j(s_i|\alpha) f_j(y_i \mid x_i^{(j)}, \beta^{(j)}, \sigma_j^2)} \quad \text{(A5)}$$

M-Step

First, we update α (using Newton-Raphson):

$$\alpha^* = \arg\max_{\alpha} \left[\sum_{i=1}^{n} z_{1i} \log\left(\frac{e^{\alpha^T s_i}}{1 + e^{\alpha^T s_i}} \right) + z_{2i} \log\left(\frac{1}{1 + e^{\alpha^T s_i}} \right) \right] \quad \text{(A6)}$$

We update Equation (A5) using α^* to obtain τ^*_{ki}. We then update $\beta^{(k)}$ by running a weighted regression for each regime:

$$\left(\beta^{(k)}\right)^* = \arg\min_{\beta^{(k)}} \sum_{i=1}^{n} \tau^*_{ki} \frac{(y_i - \beta^{(k)T} x_i^{(k)})^2}{\sigma_k^2} \tag{A7}$$

The residual volatility is updated using posterior probabilities:

$$(\sigma_k^*)^2 = \frac{\sum_{i=1}^{n} \tau^*_{ki} (y_i - \beta^{(k)T} x_i^{(k)})^2}{\sum_{i=1}^{n} \tau^*_{ki}} \tag{A8}$$

Now we repeat the E-step and M-step until convergence.

Multiple initial points

It is possible that convergence can occur at a local maximum of l instead of the global maximum. One way to overcome this problem is to implement EM with multiple initial estimates. See [5] for more details.

Contact us to discuss how Axioma can bring more information and insight to your investment process.

United States and Canada: +1-212-991-4500
Europe: +44-207-856-2424
Asia: +852-8203-2770

Sales: sales@axioma.com
Customer Support: support@axioma.com