FAJ
Financial Analysts Journal
Volume 66 • Number 5
©2010 CFA Institute

Skulls, Financial Turbulence, and Risk Management

Mark Kritzman, CFA, and Yuanzhen Li

Based on a methodology introduced in 1927 to analyze human skulls and later applied to turbulence in financial markets, this study shows how to use a statistically derived measure of financial turbulence to measure and manage risk and to improve investment performance.

Most investors look to their domestic equity markets as the main engine of growth for their portfolios, and they search for other assets to diversify this exposure. The typical investor considers only average correlations, however, when measuring an asset's diversification benefits, and average correlations tend to be misleading. For example, when both U.S. and non-U.S. equities produce returns greater than one standard deviation above their means, their correlation equals –17 percent; when both markets produce returns more than one standard deviation below their means, their correlation rises to +76 percent.[1] These differences explain why many investors who believed their portfolios were well diversified suffered catastrophic losses during the financial crisis of 2007–2008. Rather than rely on average measures of risk to structure portfolios and manage risk, investors should use conditional measures that take into account the behavior of assets during turbulent subperiods.

Chow, Jacquier, Lowrey, and Kritzman (1999) introduced a mathematical measure of financial turbulence, which originally was developed by Mahalanobis (1927, 1936) to analyze human skulls. We extended this research by investigating the empirical features of financial turbulence and by demonstrating how this methodology can be used to stress-test portfolios, to construct turbulence-resistant portfolios, and to scale exposure to risk to improve performance.

Measuring Financial Turbulence

We define financial turbulence as a condition in which asset prices, given their historical patterns of behavior, behave in an uncharacteristic fashion, including extreme price moves, decoupling of correlated assets, and convergence of uncorrelated assets. Financial turbulence often coincides with excessive risk aversion, illiquidity, and devaluation of risky assets.

The method we used to measure turbulence first appeared as the "Mahalanobis distance." Mahalanobis (1927) used 7–15 characteristics of the human skull to analyze distances and resemblances between various castes and tribes in India. The skull characteristics used by Mahalanobis included head length, head breadth, nasal length, nasal breadth, cephalic index, nasal index, and stature. The characteristics differed by scale and variability. That is, Mahalanobis might have considered a half-inch difference in nasal length between two groups of skulls a significant difference whereas he considered the same difference in head length to be insignificant. Mahalanobis normalized differences in each characteristic by the characteristic's standard deviation and then squared and summed the normalized differences, thus generating one composite distance measure that was invariant to the variability of each dimension. He later proposed a more generalized statistical measure of distance, the Mahalanobis distance, which takes into account not only the standard deviations of individual dimensions but also the correlations between dimensions (see Mahalanobis 1936). In his system, with n characteristics for measurements, each skull can be represented as an n-dimensional vector. The Mahalanobis distance of an individual vector $\mathbf{y}$ from a sample of vectors with average $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is defined as

$$d = \sqrt{(\mathbf{y}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})'}, \qquad (1)$$

where

d = Mahalanobis distance from sample average (scalar)

$\mathbf{y}$ = vector of multivariate measurements ($1 \times n$ vector)

Mark Kritzman, CFA, and Yuanzhen Li are at Windham Capital Management, LLC, Cambridge, Massachusetts.

$\boldsymbol{\mu}$ = sample average vector ($1 \times n$ vector)

$\boldsymbol{\Sigma}$ = sample covariance matrix ($n \times n$ matrix)

The Mahalanobis distance can be used to measure the similarity of a particular skull to a sample of skulls belonging to a group of known anthropological origin. Consider the following hypothetical example. Suppose we use two features, head length and head breadth, to measure skulls. Each skull can then be represented as a point in a two-dimensional space. Two groups of skulls belonging to distinct anthropological origins form two clusters in the two-dimensional space, as shown in **Figure 1**.

Suppose we compare a skull of unknown origin, represented by the square in Figure 1, with the two groups and categorize it. In terms of Euclidean distance, it lies closer to the center of Group 2 than to the center of Group 1. The Mahalanobis distance, however, would consider this skull more similar to Group 1 because its characteristics are less unusual in light of the more inclusive scatter plot of Group 1's characteristics.

This example reveals that the Mahalanobis distance is scale independent, which is also apparent from Equation 1. The characteristic deviations are scaled by the covariance matrix.

Chow et al. (1999) independently derived a nearly identical formula to detect turbulence in financial markets.[2] By substituting asset returns for skull characteristics, Chow et al. determined the statistical unusualness of a cross section of returns on the basis of their historical multivariate distributions. With n assets, the returns for a particular period (such as a month or a day) constitute an n-dimensional vector. Suppose a historical sample of such return vectors has an average vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The statistical measure of financial turbulence, which we term the "turbulence index," is formally defined as

$$d_t = \left(\mathbf{y_t} - \boldsymbol{\mu}\right)\boldsymbol{\Sigma}^{-1}\left(\mathbf{y_t} - \boldsymbol{\mu}\right)', \tag{2}$$

where

d_t = turbulence for a particular time period t (scalar)

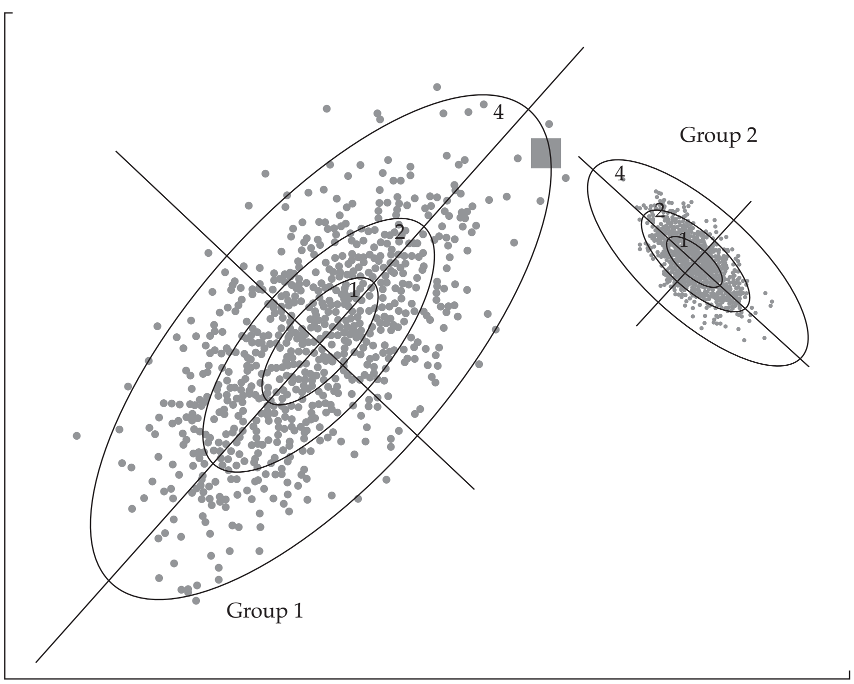
$\mathbf{y_t}$ = vector of asset returns for period t ($1 \times n$ vector)

$\boldsymbol{\mu}$ = sample average vector of historical returns ($1 \times n$ vector)

$\boldsymbol{\Sigma}$ = sample covariance matrix of historical returns ($n \times n$ matrix)

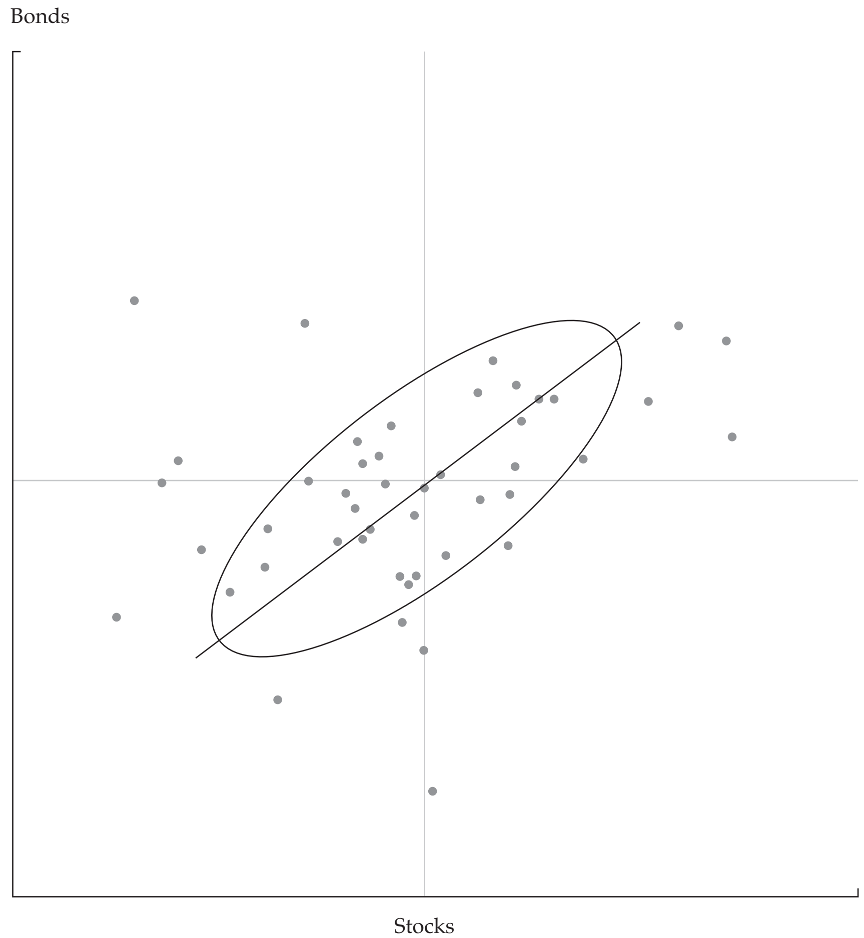
Turbulence, as we have just defined it, can be calculated for any group of n return series a user may choose. **Figure 2** illustrates this statistical measure of turbulence for a simple example with two return series—stocks and bonds. Each point represents the returns of stocks and bonds for a particular

Figure 1. Scatter Plot of Hypothetical Human Skull Characteristics



Note: Points on each ellipse have identical Mahalanobis distances from the corresponding group center.

Figure 2. Scatter Plot of Hypothetical Stock and Bond Returns



period. The center of the ellipse represents the average of the joint returns of stocks and bonds. The ellipse itself represents a tolerance boundary that encloses a certain percentage—for example, 75 percent—of the bivariate Gaussian distribution of stock and bond returns. All points on the ellipse have equal Mahalanobis distances from the center. This boundary also embodies the threshold that separates "turbulent" from "quiet" observations. Points inside the ellipse represent return combinations associated with quiet periods because the observations are within 75 percent of the distribution and are thus not particularly unusual. The observations outside the ellipse are statistically unusual and are thus likely to characterize turbulent periods. Notice that some returns just outside the middle of the ellipse are closer to the ellipse's center than some returns within the ellipse at either end. This pattern illustrates the notion that some periods qualify as turbulent not because one or more of the returns were unusually high or low but, instead, because the returns moved in the opposite direction in that period despite the fact that the assets are positively correlated, as evidenced by the positive slope of the scatter plot.

Others have addressed the problem of financial turbulence differently. Correlations conditioned on upside or downside market moves (Ang and Chen 2002); time-varying volatility models, such as GARCH (generalized autoregressive conditional heteroscedasticity) models (Bollerslev 1986); Markov regime-switching models (Ang and Bekaert 2002); mixture models, such as jump diffusion (Das and Uppal 2004); and implied volatility (Mayhew 1995) have all been proposed as measures of financial duress. The statistical measure of turbulence defined in this article has two particular advantages over the (perhaps) most commonly used indicator of financial stress—namely, implied volatility. In our measure, turbulence can be estimated for any set of assets rather than only for assets with liquid option markets. And our measure captures interactions among combinations of assets in addition to the magnitude of the assets' returns.

Analysts may be tempted to think that the volatility of an index comprising the assets used to measure turbulence captures the same information as our measure. After all, such a volatility estimate incorporates both the volatility of the individual assets and their correlations with each other. By summarizing the data in an index, however, one sacrifices the higher-dimensional information captured by the turbulence index.

This distinction is illustrated in **Figure 3**. The loosely clustered circles in the scatter plot are the returns of two assets with relatively high volatilities and a negative correlation. The tightly clustered squares are the returns of two assets with relatively low volatilities and a positive correlation. It turns out that an index comprising the circle assets has the same volatility as an index comprising the square assets, yet the turbulence estimates of each index's assets are substantially different.

Thus far, we have characterized returns as belonging to distinct turbulent and nonturbulent regimes. Depending on the data or particular application, this distinction may be somewhat arbitrary.[3] We can just as well characterize returns along a continuum ranging from quiescence to turbulence. If we follow the latter approach, we find that our mathematical measure of turbulence coincides closely with well-known turbulent events, as evidenced by **Figure 4**.

Figure 4 shows a turbulence index calculated according to Equation 2, for which we used monthly returns of six asset-class indices: U.S. stocks, non-U.S. stocks, U.S. bonds, non-U.S. bonds, commodities, and U.S. real estate. The average vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ in Equation 2 were calculated for the full sample from January 1980 to January 2009. Spikes in this index can clearly be seen to coincide with events known to have been associated with financial turbulence. Also interesting to note, but certainly not surprising, is that the 2007–08 financial crisis is by far the most turbulent episode of recent history.

Figure 3. Scatter Plot of Asset Pairs from Indices with Equal Variances

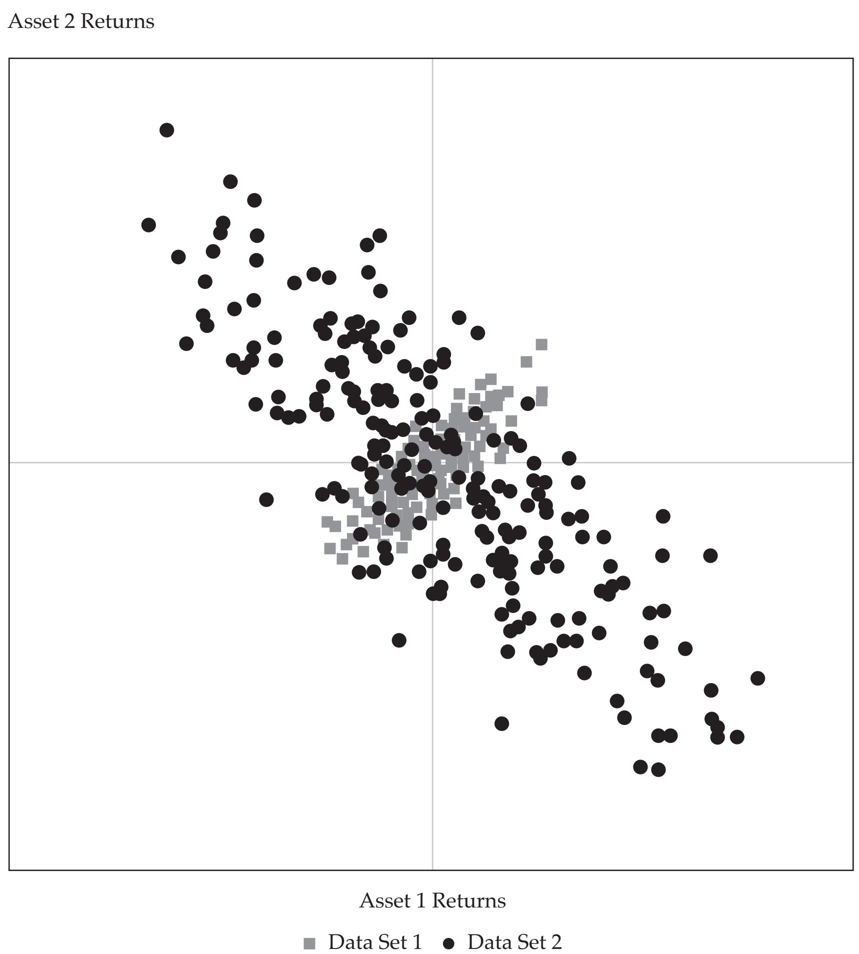
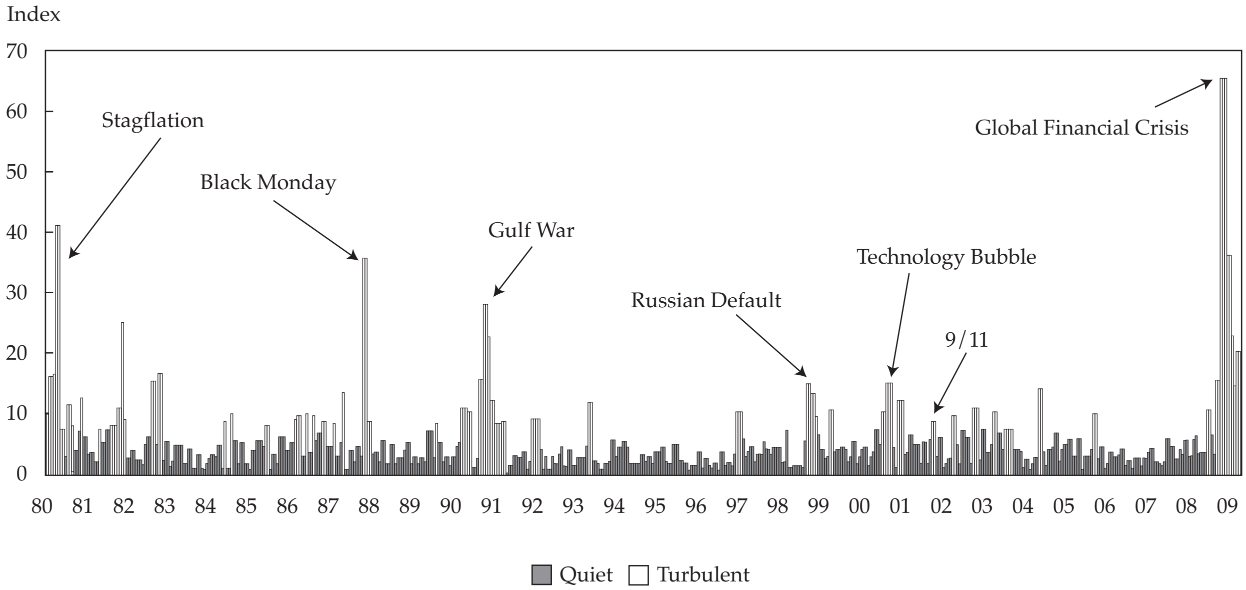


Figure 4. Historical Turbulence Index Calculated from Monthly Returns of Six Global Indices, 1980–2009



Note: Measured as of January each year.

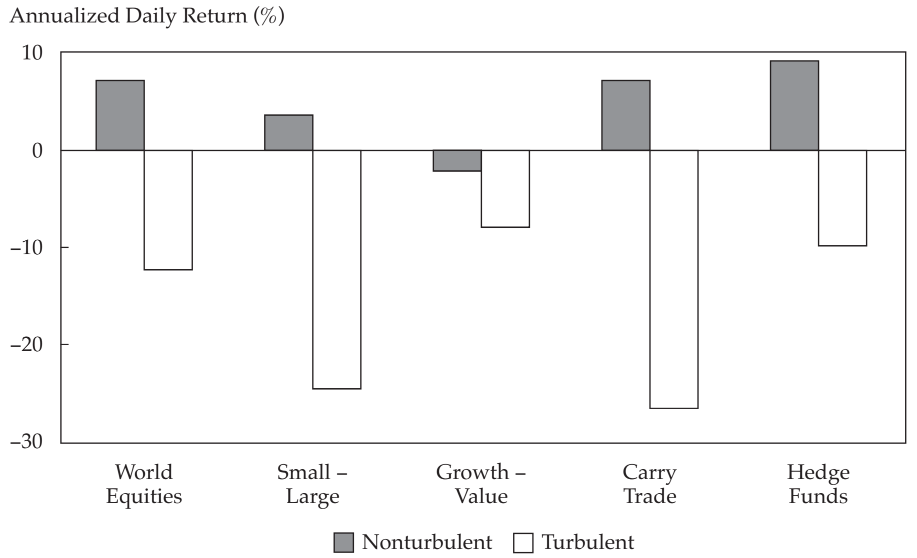
Empirical Features of Turbulence

Two empirical features of turbulence are particularly interesting. First, returns to risk are substantially lower during turbulent periods than during nonturbulent periods, irrespective of the source of turbulence. For example, the recent financial crisis began with a downturn in housing prices, which led to a sharp devaluation of mortgage derivatives. What was somewhat surprising to many investors is that this crisis in the mortgage derivatives market coincided with substantial losses in carry strategies. The carry strategy calls for long positions in currency forward contracts that sell at a discount combined with short positions in currency forward contracts that sell at a premium. Why should a crisis in the mortgage market lead to losses in a currency strategy? As mortgage derivatives fell in value, many investors, especially hedge funds, were required to raise capital. But the mortgage derivatives market is relatively illiquid. To meet margin calls from their prime brokers and redemptions from their investors, these hedge funds thus turned to the most liquid components of their portfolios—their currency positions and investments in publicly traded securities. **Figure 5** provides evidence that returns to risk are substantially lower during episodes of financial turbulence, even though the causes of turbulence may differ over time.

These differences in return suggest that knowing whether the period ahead will be turbulent would be helpful, which leads to the second empirical feature of turbulence. Financial turbulence is highly persistent. It is similar to the turbulence encountered in air travel. Weather turbulence may arrive unexpectedly, but once it begins, we know that it will take time for the airplane to pass through the weather system or for the pilot to find a smoother altitude. The same process is true of financial turbulence. Although we may not be able to anticipate the initial onset of financial turbulence, once it begins, it usually continues for a period of weeks as the markets digest it and react to the events causing the turbulence. **Table 1** provides evidence of the persistence of turbulence.

Table 1 shows the level of average daily turbulence following the initial arrival of a day that is more turbulent than 10 percent of the most turbulent days in the sample for the next 5 days, 10 days, and 20 days, not including the initial turbulent day. The turbulence values are multivariate distances calculated according to Equation 2 for full-sample averages and covariance matrices. These turbulence scores were normalized to 1 to facilitate comparisons of the different sets of assets. The rightmost column shows the normalized turbulence scores separating the 10 percent most turbulent days from

Figure 5. Returns to Risk during Turbulent and Nonturbulent Periods



Notes: Turbulent periods were identified from USD-denominated daily values of the turbulence index constructed for global asset allocation (World Equities), U.S. sectors based on the small-capitalization premium (Small – Large) and the growth premium (Growth – Value), and developed country currencies (Carry Trade) for 4 January 1993–31 December 2008. Monthly turbulence index values for global asset allocation over the period January 1993–December 2008 were used for Hedge Funds. Raw turbulence values are multivariate distances based on a full-sample covariance matrix. The market returns are daily returns of the MSCI World Index (for World Equities), the Russell 2000 Index minus the S&P 500 Index (Small – Large), the Russell 1000 Growth Index minus the Russell 1000 Value Index (Growth – Value), and a naive carry strategy over the same time period. The monthly Hedge Fund returns are from the HFRI Fund of Funds Composite Index.

Source: State Street Associates.

the rest of the sample. The 10 percent threshold for each market was identified from the full sample of returns. The percentile rankings of the average daily turbulence scores shown next to each score clearly show that markets tend to remain turbulent for up to a month or longer once turbulence begins.

Applications

This statistical measure of finance turbulence has several useful applications. Analysts can use it to stress-test portfolios more reliably than when using conventional methods. Analysts can use it to structure portfolios that are relatively resilient to

Table 1. Persistence of Turbulence

	Next 5 Days		Next 10 Days		Next 20 Days		
Market	Level	Percentile Rank	Level	Percentile Rank	Level	Percentile Rank	10th Percentile Threshold
Global assets	2.31	7	2.22	8	2.13	8	1.93
U.S. assets	2.98	5	2.9	5	2.79	6	1.95
U.S. sectors	3.12	5	3.04	6	2.87	6	2.03
Currencies	2.08	8	1.93	9	1.8	11	1.83
U.S. fixed income	4.05	4	3.85	5	3.6	5	2.12
U.S. Treasury notes	3.19	5	3.13	6	2.96	6	2.00
U.S. credit	4.17	4	4.09	4	3.69	4	1.61

Note: Time periods used for the calculations are as follows: January 1993–April 2009 for "Global assets," August 1989–April 2009 for "U.S. assets," January 1973–April 2009 for "U.S. sectors," January 1975–April 2009 for "Currencies," December 2000–April 2009 for "U.S. fixed income," September 1998–April 2009 for "U.S. Treasury notes," and August 1998–April 2009 for "U.S. credit."

turbulent episodes without significantly compromising their long-term growth prospects. And finally, analysts can use turbulence as a signal for scaling a strategy's exposure to risk.

Stress-Testing Portfolios. Investors often use value at risk (VaR) to measure a portfolio's exposure to loss. VaR provides the largest loss a portfolio might experience at a certain level of confidence. The conventional approach for measuring VaR uses the full-sample covariance matrix to compute the portfolio's standard deviation and considers the probability distribution only at the end of the investment horizon. We can measure exposure to loss more reliably by estimating covariances from the turbulent subperiods, when losses are more likely to occur, and by accounting for interim losses as well as losses that occur only at the conclusion of the investment horizon.[4]

Table 2 shows three portfolios—conservative to aggressive—together with assumptions for their expected returns and two estimates of standard deviation. One estimate of standard deviation, "Full-sample risk," is based on the full-sample covariance matrix of monthly returns beginning in January 1977 and ending in December 2006. The other estimate of standard deviation, labeled "Turbulent risk," is based on the covariance matrix from the turbulent subsample. Turbulence was calculated according to Equation 2, in which each return vector consisted of returns of the five asset-level indices for a particular month, and average vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ were calculated from monthly returns during the entire 30-year history. The threshold for identifying turbulent periods was set at 75 percent, which means that roughly 25 percent of the months fell within turbulent subperiods.[5] Note how risk increases for each portfolio when turbulence risk is used.

Table 3 shows the VaR, given a 1 percent confidence level, for each portfolio as of December 2006. The first column shows VaR estimated as of the end of a five-year horizon on the basis of standard deviations estimated from the full-sample covariance matrix. The second column reports VaR when we used the standard deviations estimated from the covariances that prevailed during the turbulent subsample and we took into account losses that could occur throughout the investment horizon. The third column shows the maximum losses actually experienced by each of these portfolios from inception during the financial crisis. The fourth column reports the maximum drawdowns realized by each of these portfolios.

If we consider the 2007–08 financial crisis as a once-in-a-century event, Table 3 shows that the conventional approach to measuring exposure to loss badly underestimated the riskiness of these portfolios. The turbulence-based approach, in contrast, anticipated the exposure to loss of these portfolios much more accurately. To be clear, we point out that the turbulence-based approach does not offer a more reliable estimate of when an extreme event will occur; rather, it gives a more reliable estimate of the consequences of such an event. Also note that turbulence is a relative measure. If the world becomes more turbulent, for example, the threshold for separating turbulent periods from nonturbulent periods will rise.

Table 2. Efficient Portfolios, Expected Returns, and Two Estimates of Risk

Asset Class	Conservative Portfolio	Moderate Portfolio	Aggressive Portfolio
U.S. stocks	22.86%	35.23%	48.15%
Non-U.S. stocks	16.59	24.22	32.19
U.S. bonds	49.95	32.81	14.89
Real estate	3.85	2.59	1.28
Commodities	6.75	5.16	3.49
Expected return	7.60%	8.37%	9.17%
Full-sample risk	7.77	10.12	12.86
Turbulent risk	10.68	13.68	17.33

Note: "Full-sample risk" was estimated from the full-sample covariance matrix; "Turbulent risk" was estimated according to the covariance matrix of the turbulent subsample.

Table 3. VaR and Realized Returns, End of 2006

Portfolio	VaR for Full Sample, End of Horizon	VaR for Turbulent Sample, within Horizon	Maximum Loss from Inception (Jan/07–Sep/09)	Maximum Drawdown (Jan/07–Sep/09)
Conservative	2.10%	26.20%	19.60%	25.80%
Moderate	9.90	35.10	29.42	35.50
Aggressive	18.70	45.00	38.96	45.30

Note: The horizon is five years.

Source: Windham Capital Management, LLC.

Building Turbulence-Resistant Portfolios. We have shown how analysts—by focusing on the behavior of assets during periods of turbulence, especially how they interact with each other—can construct portfolios that are conditioned to better withstand turbulent events and, at the same time, perform relatively well in all market conditions. Analysts can also modify two methods of optimization, mean–variance optimization and full-scale optimization, to derive turbulence-resistant portfolios.

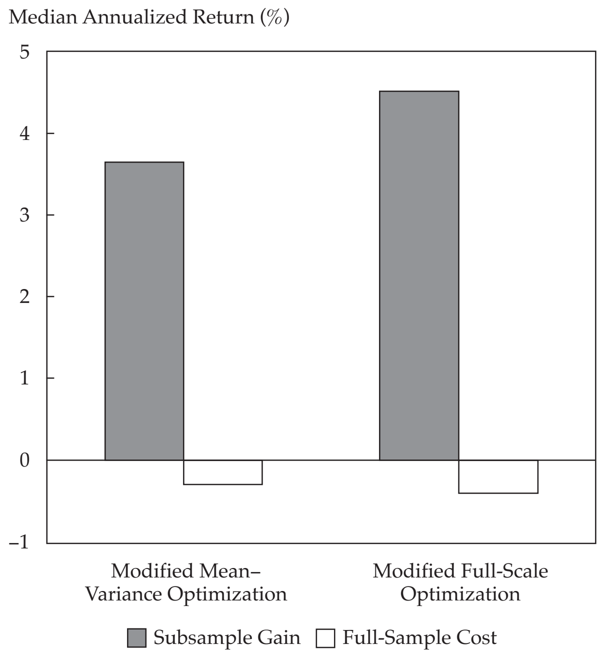
We modified mean–variance optimization by blending the differences between the realized turbulent returns and full-sample returns with equilibrium returns to estimate expected returns. We also blended the turbulent subsample covariances with the full-sample covariances in proportion to their sample sizes.

We applied a modified version of full-scale optimization. Full-scale optimization uses a search algorithm to maximize expected utility for a given sample of returns.[6] We modified this optimization method by increasing the representation of the turbulent subsample returns beyond their empirical frequency. The details of these modified optimization methods are described in Appendix A.

To evaluate the two methods, we performed 1,000 random trials of training and out-of-sample testing. For each trial, we drew a random half from the historical sample to use as training data. The other half was used as testing data. From the training data, we identified a turbulent subsample by calculating the turbulence index according to Equation 2 and, subsequently, selecting the periods with the highest turbulence index values (the highest quartile). Using the full training sample, we built an unconditioned optimal portfolio that did not take turbulence into account. Using the turbulent subsample combined with some information from the full training sample, we built a conditioned optimal portfolio that was expected to be more resistant to turbulence than an unconditioned portfolio. We then used the testing data to test the unconditioned and the conditioned optimal portfolios out of sample. We performed two types of testing: one on the full testing sample and the other on a turbulent subsample within the testing sample.

Figure 6 compares the performance of the conditioned, turbulence-resistant, optimal portfolio with that of the unconditioned optimal portfolio. The conditioned portfolios substantially outperformed the unconditioned portfolios in the out-of-sample turbulent periods and only marginally underperformed the unconditioned portfolios, on average, in all market conditions.

Figure 6. Differences between Median Annualized Returns of Conditioned and Unconditioned Portfolios



Notes: The median was taken over 1,000 out-of-sample tests. "Subsample Gain" stands for outperformance (gain in median return) of the conditioned portfolio during turbulent periods. "Full-Sample Cost" stands for underperformance (loss in median return) of the conditioned portfolio across quiet and turbulent conditions.

The conditioned portfolios also outperformed the unconditioned portfolios much more frequently in the out-of-sample turbulent periods and outperformed almost as often in all market conditions, on average, as shown in **Table 4**.

This evidence strongly suggests that by understanding the conditional behavior of assets, portfolio managers can construct turbulence-resistant portfolios without substantially compromising average performance.

Table 4. Frequency of Conditioned Portfolio Outperforming Unconditioned Portfolio

Out-of-Sample Period	Modified Mean–Variance Optimization	Modified Full-Scale Optimization
Turbulent periods	85%	86%
All market conditions	45	41

Scaling Exposure to Risk. The differential performance of risky strategies during turbulent and nonturbulent periods, together with the persistence of turbulence, raises the tantalizing prospect that portfolio managers might be able to improve performance by conditioning exposure to risk on the degree of turbulence. Indeed, we show here that a simple scaling rule applied to the carry strategy does significantly improve performance.[7] We measured the 30-day moving average of turbulence each day from the returns of the G–10 currencies and recorded whether the level of turbulence that day fell into the first, second, third, fourth, or fifth quintile of turbulence on the basis of a trailing three-year window.[8] We then weighted exposure to the carry strategy in inverse proportion to turbulence as shown in **Table 5**. We assumed a one-day lag for implementation.

Table 5. Exposure to Carry Strategy Weighted in Inverse Proportion to Turbulence

Turbulence Quintile	Exposure
1st	20%
2nd	40
3rd	60
4th	80
5th	100

We applied the same scaling rule in using other signals of market stress, including VIX (index of volatility on the S&P 500), swap spreads, and yield spreads. **Table 6** shows the performance of the unfiltered carry strategy as well as its filtered performance. The evidence shows that reducing exposure to the carry strategy in proportion to turbulence substantially improves performance and by a wider margin than using any other signal of market stress. Moreover, filtering on turbulence eliminates most of the negative skewness of the carry strategy.

Conclusion

We showed that a mathematically derived measure of financial turbulence, which measures the statistical unusualness of a set of returns given their historical pattern of behavior, coincides with well-known episodes of market turbulence.

We also showed that this measure of financial turbulence is mathematically equivalent to the Mahalanobis distance, first motivated by the study of similarity in human skulls. The following analogy provided intuition of this equivalence. The Mahalanobis distance captures several dimensions of human skulls that set them apart from a perfect sphere. If one constructed a scatter plot of three return series emanating from a single multivariate normal distribution, it would form a perfect ellipsoid. This measure of financial turbulence captures deviations from this idealized three-dimensional scatter plot.

Financial turbulence has two intriguing features. First, returns to risk are substantially lower during turbulent periods than during nonturbulent periods, and second, turbulence is persistent. It may arrive unexpectedly, but it does not immediately subside. It typically continues for weeks as market participants digest it and react to its cause.

This measure of financial turbulence can be applied in a variety of useful ways. Portfolio managers can use it to stress-test portfolios by estimating VaR from the covariances that prevailed during the turbulent subsample. They can also construct portfolios that are relatively resistant to turbulence by conditioning inputs to the portfolio construction on the performance of assets during periods of turbulence. Finally, they can enhance the performance of certain risky strategies by using turbulence as a filter for scaling exposure to risk.

We thank Erin Abouzaid, Timothy Adler, Jordan Alexiev, Robin Greenwood, Sébastien Page, and David Turkington for helpful comments and assistance.

This article qualifies for 1 CE credit.

Table 6. Filtered Carry Trade Performance

Measure	Unfiltered Carry	Turbulence Index	VIX	Five-Year Swap Spread	TED Spread	Yield Spread
Return (%)	2.81	3.02	1.98	2.67	2.21	2.10
Standard deviation (%)	5.55	3.29	3.95	4.04	3.51	3.07
Information ratio	0.51	0.92	0.50	0.66	0.63	0.68
Skewness	–0.48	–0.15	–0.26	–0.25	–0.31	–0.56

Notes: The swap spread is based on U.S. dollars and is measured against the comparable U.S. Treasury security. The bond yield spread represents the difference between 10-year and 2-year Treasury yields. "TED Spread" stands for Treasury–Eurodollar spread.

Appendix A. Two Methods for Building Turbulence-Resistant Portfolios

The investment universe we used for building turbulence-resistant portfolios is as follows: U.S. equities, non-U.S. equities (the MSCI EAFE Index), U.S. T-bonds, U.S. corporate bonds, U.S. Treasury Inflation-Protected Securities (TIPS), commodities, REITs, and cash. Our historical sample includes data between 1973 and 2009. (We back-filled TIPS data for the period between 1973 and 1996 on the basis of U.S. Consumer Price Index returns and returns of bond indices with similar durations.)

The conditioned and unconditioned portfolios were constructed 1,000 times from a randomly selected half of the training data and tested out of sample on the remaining unused data. From the training sample, we identified a turbulent subsample by comparing each period's turbulence index with a threshold. The turbulence index was calculated according to Equation 2, and the threshold was the 70th percentile of all the training-sample turbulence index values.[9]

Using the full training sample, we derived an unconditioned optimal portfolio that was expected to be optimal in all market conditions, quiet or turbulent. Using the turbulent subsample combined with information from the full training sample, we derived a conditioned optimal portfolio for withstanding market turbulence.

To evaluate the performance of the unconditioned and the conditioned optimal portfolios, we used both the full testing sample and a turbulent subsample, which was identified by placing a threshold on the testing-period turbulence index. The threshold was set equal to the one used for dissecting the training sample.

We performed this process 1,000 times. Figure 6 and Table 4 in the body of this article demonstrate that the conditioned optimal portfolios derived by incorporating the turbulence index are indeed more resistant to market turbulence yet perform reasonably well, on average, in all market conditions.

Modified Mean–Variance Optimization

1. We estimated the unconditioned expected returns as equilibrium returns on the basis of the full training sample. To estimate equilibrium returns, we used a portfolio consisting of 60 percent U.S. equities, 30 percent T-bonds, and 10 percent U.S. corporate bonds as the market portfolio.
2. We estimated the conditioned expected returns as follows:
   - We computed the average returns of the turbulent subsample and compressed them toward the return of the minimum-variance portfolio. The subsample, conditioned on its turbulence index being high, was often a relatively small sample. Therefore, expected returns estimated by taking the subsample average were subject to possibly severe estimation errors, leading to inferior out-of-sample performance. The small-sample problem can be mitigated by compression, also known as shrinkage, where the sample average is blended with another estimator that enforces more structure. Any common return constant can function as a shrinkage target and improve out-of-sample risk-adjusted performance of the resulting portfolio. In our experiments, we chose to use the return of the minimum-variance portfolio, proposed by Jorion (1986), as the shrinkage target:

     $$\boldsymbol{\mu}_{\mathbf{m}} = \frac{\mathbf{e}'\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\mu}}}{\mathbf{e}'\boldsymbol{\Sigma}^{-1}\mathbf{e}}, \tag{A1}$$

     where $\mathbf{e} = [1, 1, \ldots, 1]'$ is the unit vector, $\boldsymbol{\Sigma}$ is the sample covariance matrix, and $\hat{\boldsymbol{\mu}}$ is the sample mean.

     The shrinkage strength (i.e., the weight applied to the shrinkage target for linearly blending it with the sample mean) was determined by the following formula:

     $$w = \frac{n+2}{n+2+T\left(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\mathbf{m}}\mathbf{e}\right)'\boldsymbol{\Sigma}^{-1}\left(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\mathbf{m}}\mathbf{e}\right)}, \tag{A2}$$

     where n is the number of dimensions (i.e., number of assets in the portfolio) and T is the sample size (i.e., number of observations in the sample). The return estimate after shrinkage was found by

     $$\boldsymbol{\mu}_{\mathbf{s}} = w\boldsymbol{\mu}_{\mathbf{m}} + (1-w)\hat{\boldsymbol{\mu}}. \tag{A3}$$

   - We computed the average returns of the full training sample and compressed them toward the return of the minimum-variance portfolio. The shrinkage target and shrinkage strength were calculated in the same fashion as in the previous step with the use of Equations A1 and A2, but the calculations were performed on the full training sample instead of the turbulent subsample.

- We used ratios proportional to sample sizes to blend the differences between the compressed subsample and full-sample returns with equilibrium returns.

3. We used the full training sample to estimate the unconditioned covariances by shrinking the sample covariance matrix toward a one-factor market model covariance matrix (Ledoit and Wolf 2003).
4. Using ratios proportional to sample sizes, we estimated the conditioned covariances by blending the turbulent subsample covariance matrix (which was also shrunk toward a one-factor market model covariance matrix) with the full-sample covariance matrix.
5. We identified the unconditioned and conditioned optimal portfolios in a mean–variance framework by using expected returns and covariances estimated in Steps 1–4. We chose risk aversions in such a way that the unconditioned and the conditioned optimal portfolios would have similar risk profiles in the training data, both with standard deviations around 10 percent.
6. We computed performances of unconditioned and conditioned optimal portfolios for the full testing sample and the turbulent subsample. The results are provided in the first portions of Table 4 and Figure 6.

Modified Full-Scale Optimization

1. We estimated the unconditioned expected returns as equilibrium returns on the basis of the full training sample. To estimate equilibrium returns, we used a portfolio consisting of 60 percent U.S. equities, 30 percent T-bonds, and 10 percent U.S. corporate bonds as the market portfolio.
2. We scaled the training sample by the unconditioned expected returns estimated in Step 1 so that the average asset returns would be equal to equilibrium returns.
3. We modified the training data by increasing the representation of the turbulent subsample. The increased representation was achieved via boot strapping: The subsample was sampled at three times the frequency at which the full training sample was sampled.
4. We identified the unconditioned optimal portfolio in a mean–variance framework by using unconditioned expected returns estimated as equilibrium returns (Step 1) and unconditioned covariances estimated for the unmodified full training sample.
5. We used the modified training data to identify the conditioned full-scale optimal portfolio. Full-scale optimization uses a search algorithm to identify a set of asset weights that maximize expected utility given a sample of returns and a utility function (see Cremers et al. 2005 for a detailed description). We used a kinked utility function in which utility changes abruptly at a threshold return level. Utility was defined as a log-wealth function above the threshold return and a linear function with a steep slope below the threshold return:

$$U = \begin{cases} \ln(1+x), & \textit{for } x \geq \theta \\ \gamma(x-\theta) + \ln(1+\theta), & \textit{for } x < \theta \end{cases}, \tag{A4}$$

   where x stands for return, θ stands for the threshold, and γ stands for the slope of utility below the threshold. In our experiments, x was annually compounded portfolio return, θ was –15 percent, and γ was 2.25. For a particular set of weights, we calculated a portfolio's expected utility as the sum of its utility for each period. We considered many sets of candidate portfolio weights and identified the set that yielded the highest expected utility as the set of optimal weights.
6. We computed performances of unconditioned and conditioned optimal portfolios on the full testing sample and its turbulent subsample. The results are provided in the right-hand portions of Table 4 and Figure 6.

Notes

1. These correlations are based on monthly returns of the S&P 500 Index and MSCI World ex US Index from the period starting January 1970 and ending February 2008.
2. The only difference between the two measures is that the Mahalanobis distance is the square root of the turbulence measure as defined in Equation 2. In many of the applications in Chow et al. (1999), the authors used the ordinal relationships of the turbulence measure, which are not altered by squaring, to identify turbulent periods.
3. One approach for separating turbulent from nonturbulent regimes is to compute the absolute differences in correlations, cell by cell, between the 5 percent, 10 percent, 15 percent, 20 percent, and so forth, most turbulent regimes and the nonturbulent remaining observations and then plot the average of these differences as a function of sample size and look for obvious correlation shifts. Alternatively, one can determine the threshold as a function of returns to risk.

4. VaR that takes into account within-horizon losses is called "continuous VaR." It is estimated numerically from first passage probabilities. See Kritzman and Rich (2002) for a thorough description of continuous VaR.
5. The results are not particularly sensitive to the 75th percentile threshold.
6. See Cremers, Kritzman, and Page (2005) for a detailed description of full-scale optimization.
7. We implemented the carry strategy by first estimating each currency forward contract's expected return as its annualized monthly discount or premium. Then, we constructed an equally weighted portfolio of long and short exposures (long discount currencies and short premium currencies). We rebalanced the portfolio monthly.
8. The G–10 countries are Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Sweden, Switzerland, the United Kingdom, and the United States.
9. The results are not particularly sensitive to the 70th percentile threshold.

References

Ang, A., and G. Bekaert. 2002. "International Asset Allocation with Regime Shifts." *Review of Financial Studies*, vol. 15, no. 4 (Fall):1137–1187.

Ang, A., and J. Chen. 2002. "Asymmetric Correlations of Equity Portfolios." *Journal of Financial Economics*, vol. 63, no. 3 (March):443–494.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, vol. 31, no. 3 (April):307–327.

Chow, G., E. Jacquier, K. Lowrey, and M. Kritzman. 1999. "Optimal Portfolios in Good Times and Bad." *Financial Analysts Journal*, vol. 55, no. 3 (May/June):65–73.

Cremers, J.H., M. Kritzman, and S. Page. 2005. "Optimal Hedge Fund Allocations: Do Higher Moments Matter?" *Journal of Portfolio Management*, vol. 31, no. 3 (Spring):70–81.

Das, S.R., and R. Uppal. 2004. "Systemic Risk and International Portfolio Choice." *Journal of Finance*, vol. 59, no. 6 (December):2809–2834.

Jorion, P. 1986. "Bayes–Stein Estimation for Portfolio Analysis." *Journal of Financial and Quantitative Analysis*, vol. 21, no. 3 (September):279–292.

Kritzman, M., and D. Rich. 2002. "The Mismeasurement of Risk." *Financial Analysts Journal*, vol. 58, no. 3 (May/June):91–99.

Ledoit, O., and M. Wolf. 2003. "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection." *Journal of Empirical Finance*, vol. 10, no. 5 (December):603–621.

Mahalanobis, P.C. 1927. "Analysis of Race-Mixture in Bengal." *Journal of the Asiatic Society of Bengal*, vol. 23:301–333.

———. 1936. "On the Generalised Distance in Statistics." *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1:49–55.

Mayhew, S. 1995. "Implied Volatility." *Financial Analysts Journal*, vol. 51, no. 4 (July/August):8–20.