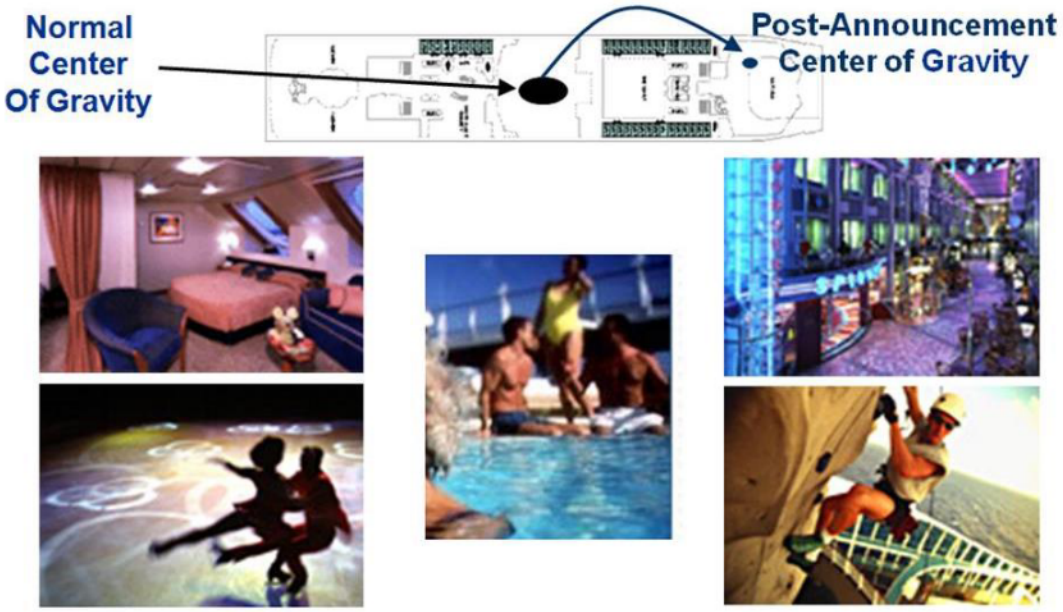
distributions that are unstable to some extent. Generally, this instability is manifested slowly as technology changes, communication becomes faster, and other structural changes evolve. The real source of many black swans is failure of the other underlined word: independent.

The Parable of the Fat Tail

It is useful to consider an example unrelated to financial markets to illustrate the importance of independence of random components. Suppose you equip the passengers of a single-deck cruise ship with a device that allows you to locate them exactly at any given moment, and then proceed to calculate, once every minute, the center of gravity of all these locations, with reference to the two-dimensional framework of the ship, and plot the resulting distribution. At most times, passengers will be in a variety of locations, based on their personal preferences, their energy levels, their mood at the moment, and available alternatives. The resulting distribution of their center of gravity over time will be a cloud of points bunched around the center of the available passenger areas. We expect it to exhibit something close to a bivariate, normal distribution.

Now assume there is an announcement over the ship's loudspeaker saying, "There is a large pod of whales breaching off the port bow." It does not take a Ph.D. in statistics or degree in finance to know what will happen next. We would see a sudden outlier in the distribution as passengers rush to find good viewing spots among the limited spaces available.



In the immediate aftermath of the announcement, all passengers know several things:

- There is an opportunity to see something unusual;
- The time to see it is limited;
- There is limited space at the ideal location for viewing the phenomenon;
- ***Everyone knows that everyone else knows these things and is likely to act accordingly.***

This final point, this mutual self-awareness, makes for the sudden mad rush to the port bow. Each passenger reacts to the knowledge that speed is of the essence if a good viewing place is to be secured. Were the ship nearly empty, or had only a few people been aware of the opportunity or likely to take advantage of it (if, say, most passengers were confined to their cabins with seasickness), the sense of urgency would have been greatly reduced.

There is a relevant scene in the movie *Rogue Trader* about Nick Leeson and the failure of Baring Brothers Bank. He is awakened by a call at his home in the early hours of the morning from one of the firm's traders in New York. The voice at the other end of the phone says urgently, "Turn on CNN!" The TV in the bedroom flickers to life, showing scenes of the Kobe earthquake. The voice at the other end of the phone says, "This is just going to *kill* the market!"

In effect, this is much like the announcement on the ship, but global. Observers worldwide suddenly focus on a common, crystallizing event, with obviously directional implications for the market. In addition, everyone knows that everyone else knows. Suddenly, the millions of decisions that drive the market are no longer randomly independent; they are subject to a common, shared perception. The core structural assumptions that underpin a normal distribution have temporarily broken down, and we see a sudden, extreme observation.

Is Risk Management Playing the Wrong Game?

We often conduct our lives based on unconscious metaphors we find compelling. Such metaphors carry with them important assumptions that are only rarely examined. Most risk managers would deny it if directly confronted, but all too often, we conduct our work as if the game we are playing is roulette. A properly serviced and maintained roulette wheel is a classic example of a process that generates a stable, random distribution of numbers. In particular, what number will come up next is unrelated to the number that came up before – each result is statistically independent of all others.

In the aftermath of September 2008, risk managers learned the hard way that roulette is a poor metaphor to use as a guide. This raises the question of what alternatives serve to condition our thinking more effectively. If we want to stick with gambling, a far superior metaphor would be poker rather than roulette. Poker involves all the volatile human emotions such as greed, fear, and revenge, in addition to odds based on the draw of the cards. It is a sufficiently fruitful metaphor that an accomplished financial risk manager named Aaron Brown wrote a book entitled *The Poker Face of Wall Street* to elucidate its implications. It undoubtedly did not hurt that Brown is also known as a successful, professional poker player.

If we want to stick with a metaphor from the physical world, two are superior to roulette: earthquakes and avalanches. Consider a seismologist who seeks to estimate the magnitude of the largest possible earthquake in a given area. Assume he proposes to collect data on the frequent small tremors detected by seismometers along the earth's fault lines. Finally, assume he creates a distribution, and argues that the biggest earthquake could not possibly be greater

than three or four standard deviations of this distribution. Would we take this analysis seriously? Certainly not.

Appropriate and Inappropriate Metaphors

Earthquakes

Avalanches

We know that an earthquake is not just another random event from a stable distribution. It is a structural shift in the system that occurs when pressure along a fault line passes a critical point. Parts of the fault give way, imposing increased pressure on adjoining sections and creating a reinforcing feedback loop. For a short time, a different random dynamic takes over until the system reaches a new, more stable configuration. Although we would not take seriously our fictional seismologist's distributional approach to estimating the largest possible earthquake, too many were ready to accept this logic when applied to financial markets.

Implications for Risk Management

Various statistical methods can be used to build periodic crises into distributions of risk factor returns. What these approaches cannot do, however, is predict when such events will occur. Thoughtful consideration of such potential scenarios, especially those that present special threats given existing exposures, is an essential component of effective risk management. Such analysis remains in the realm of experience and seasoned judgment that no amount of advanced analytical technique can replace.

In 2009, Roger Bootle, a London-based capital markets economist, made an appearance at the major risk management conference held each year in early December. With the financial crisis in full swing, he offered some harsh remarks about risk managers during his presentation. During the Q&A period, a member of the audience asked what he thought risk managers should do differently. His thought-provoking answer was, "I think they should read less mathematics and more history and literature."

This is not easy advice for many who migrated into financial risk management since the mid-1980s. Many are highly skilled in the technical mathematics of risk management but are not nearly as well-versed in history and literature. Nevertheless, if we are going to be successful in treating extreme event risk more effectively in the future, we need to broaden our horizons and/or bring people with a wider variety of backgrounds into the conversation. Synthesizing the contributions of political scientists, country risk analysts, social psychologists, behavioral finance specialists, and others is softer and less quantitative than building and operating VaR models, but it is an essential element of successful stress testing and scenario analysis. If risk managers are to cope with the full range of dangerous unknowns, especially those that pose the most serious threats to organizations, they can do no less.

Stress Testing: Static and Otherwise

The Shape of the Challenge

The global financial crisis hammered home the lesson that explicit examination of "what lurks beyond the VaR threshold" is an essential part of a comprehensive risk management program. Unfortunately, this realization can leave us feeling like we are staring into a dark abyss; the range and magnitude of things that can go wrong are so vast and their likelihoods so obscure, it is hard to know where to begin.

The challenge is compounded by the fact that humans have a difficult time incorporating highly unlikely events into our thought processes. More often, we either ignore them as too unlikely to be relevant or obsess about them beyond any justifiable degree of concern. Which of these extreme mental states prevails usually depends on whether we have recently experienced something that dramatizes a given event. Thus, most people today have an unrealistically high level of concern about the individual risk from a terrorist attack but worry little about being struck by lightning. This attitude prevails even though, for most of us in the industrial world, the

odds of injury or death from these two sources are comparably remote. Surely, this reaction is induced by recurring news stories of terrorism in action in several remote parts of the world.

Stress testing in organizational risk management is complicated by this human tendency to swing from complacency to obsession and back again. The usual reaction of many to any given stress test is to dismiss the exercise as useless because "that could never happen." And yet, whole organizations can be paralyzed by fear in the aftermath of a specific, disastrous event. The challenge for any organization is to incorporate extreme stress scenarios into its decision-making, without being whipsawed between complacency and despair.

Comparative Statistics - A Three-Pronged Approach

A useful starting point for stress testing is to apply a three-pronged approach. The three components of the approach can be characterized as:

- Historic data analysis;
- Vulnerability-based, reverse stress testing;
- What-if scenario analysis.

These are described individually below.

Historic Data: The Market's Greatest Hits

This involves defining stress scenarios that replicate relative changes in all applicable market variables for selected historical events. A far from exhaustive list of examples is shown below:

- Oct 1973 First OPEC Oil Shock;
- 1979 Iranian Revolution & Second OPEC Oil Shock;
- Aug 1982 Mexican Debt Crisis;
- Sep 1985 The Plaza Accord to weaken the USD;
- Oct 1987 Black Monday in U.S. Stocks;
- Sep 1992 Speculative Attack on the European ERM;
- Feb 1994 Dramatic Federal Reserve Tightening;
- Jul 1997 Thai Baht Collapse → Asian Currency Crisis;
- Aug 1998 Russian Default → Emerging Market Debt Crisis;
- Sep 1998 Long-term Capital Management Failure;
- 2000-2001 Dot.com Bust;
- Jan 2002 Argentine Peso Devaluation;
- Mar 2008 Bear Stearns Rescue;
- Sep 2008 Lehman, AIG, FNMA & FHLMC Collapse;
- Apr 2010 Greece Sovereign Downgrade → Euro Crisis;
- Feb 2014 Ukraine/Crimea Crisis.

One advantage of this exercise is that no one can defend the position that "this scenario could never happen" because it already has. Such simulations can be useful when addressing the anxiety of experienced managers who have lived through the trauma of these events.

Simulated market changes might appear to be inconsistent with prevailing circumstances when they vary considerably from those that surrounded the event. Historical scenarios represent a scattershot approach that is not guaranteed to highlight the sources of current potential for extreme losses. The market movements being simulated usually has little to do with the vulnerabilities of current trading positions. See the section titled *The Market's Greatest Hits - Calibrating Stress Scenarios Based on History* for detailed examples of historic stress testing.

The Achilles Heel Approach: Reverse Stress Testing

This approach systemically examines an existing portfolio to define its vulnerabilities and then constructs stress scenarios that fully exploit these vulnerabilities. Among other things, this type of exercise reveals cases in which traders are systemically "selling the wings" by writing out-of-the-money options. Often, this will not become obvious in standard VaR results, without analyzing market scenarios that generate losses beyond the 1% cutoff.

The worst 1% of losses in Monte Carlo VaR simulations (those greater than the VaR estimate itself) provide useful clues for formulating stress scenarios. It is unusual for most random scenario generators to produce changes greater than three to four standard deviations, even in 10,000 draws. In contrast, we know that market shifts of six, ten, or even larger numbers of standard deviations happen every few years rather than once in millions of years as the normal distribution implies. The worst 1% of losses in Monte Carlo simulations might be just the beginning of even larger losses were one of these major market shocks to occur. Examining market scenarios behind these worst loss results provides clues to even more extreme stress scenarios that should be simulated, and the resulting losses examined.

A regulatory demand that emerged in the aftermath of the global financial crisis was the performance of reverse stress tests. This requirement is driven by the realization that business people tend to be inherently optimistic and find it hard to contemplate failure. For banks, this tendency is particularly strong since executives worry that word of internal examination of such scenarios might undermine public confidence in the bank and create a self-fulfilling feedback loop. Such a dangerous public perception is far less likely when all banks are forced to perform reverse stress tests.

A reverse stress test differs fundamentally from the usual approach. Most stress tests begin with a scenario, and then examine how the hypothetical market conditions would affect an institution's viability. This might be augmented by some form of pessimization in which permissible exposures under approved limits are adjusted to create maximum losses given the assumed scenario.

During a reverse stress test, the exercise begins with the premise that the institution fails and proceeds to construct market conditions that bring this about. Hence, the method can also be called stress to destruction. The idea is to assess how severe conditions would have to be to cause an institution to fail. An important part of the exercise is to construct the most plausible scenario given an institution's business profile and vulnerabilities. In this sense, it employs

some of the same thinking involved in the Achilles Heel approach. See the section titled *The Achilles Heel Approach* for detailed examples of this method.

Structural Imagination: What-if Scenario Analysis

The third prong of this approach is to use structural imagination when assessing current socioeconomic and geopolitical conditions to define dangerous scenarios. This requires thinking through both the initial and potential secondary effects of a hypothetical disaster. Like the market's greatest hits, this approach fails to tailor scenarios to current vulnerabilities. In contrast, it is forward-looking and driven by current external conditions. Such scenarios tend to be taken more seriously since they involve thinking about the details of current and prospective events. Structurally thinking such scenarios through also helps identify early warning indicators, and might provide insights on whether a full-blown crisis is becoming more or less likely over time.

Any set of scenarios of interest is inevitably perishable. At the risk of looking dated in retrospect, the following list represents scenarios that appear worthy of consideration:

- A European energy disruption as a result of renewed geopolitical tensions with Russia.
- A spike in energy prices because of military conflict that damages Middle Eastern production facilities;
- A prolonged decline in energy prices, causing internal political unrest in major oil producers such as Russia, Saudi Arabia, and Venezuela;
- Political unrest in China, resulting in a significant slowdown or reversal in the country's economic growth and demand for raw materials;
- A renewed Eurozone crisis, and a necessary currency realignment and debt redenomination;
- Renewed inflation once a recovery is under way, underpinned by massive liquidity injected into the global banking system after 2008;
- Failure of a systemically important central clearing facility, disrupting derivative markets.

The approach to scenario design based on imagination involves assessing broad, external trends in such things as geopolitical events, the legal and regulatory environment, technological discoveries and their applications and social impact, environmental changes, and the consequences of real or potential natural disasters. One word of caution is in order. The frenetic global environment presents dozens of potential scenarios to the imagination. Some of these are global, and some are regional or local. It is clearly impossible to investigate all possible scenarios in detail. Coming from a tradition that grew up around the detailed statistical mechanics of distributional analysis, the temptation is often to carry a similar level of rigor into the analysis of stress scenarios. This can result in becoming bogged down in excessive analytical detail while effectively restricting the number and range of scenarios that are worthy of attention.

There is significant value in engaging a broad group of senior managers in a regular brainstorming exercise to review and assess potential scenarios based on imagination. That can involve keeping a candidate list of potential scenarios to which group members can propose additions from time to time. This can be segmented further into levels of analysis based on perceptions of rising or falling likelihood and rising or falling impact. Some might be retained in the background as reminders, others warranting commissioning of external, expert input, still others suggesting detailed impact analysis is necessary. A periodic session with active senior management engagement, say three hours once a month, could:

- Review candidate scenarios;
- Propose additions or deletions;
- Propose shifting perceived relevance of scenarios;
- Discuss results of formal analyses conducted.

A benefit of such a regular, dedicated meeting is to help senior managers prevent the urgent from crowding out the important. Big but latent threats are often in the back of people's minds, but they might be too amorphous to allow them to be articulated easily. The type of regular session described here presents an excellent opportunity to air such concerns and elicit input from colleagues with multiple perspectives on the issue.

A Successful Example

An example of how this works occurred prior to 2008. The senior managers of a large U.S. pension fund came to realize that many among them were uneasy about the state of the subprime mortgage market in late 2007. After some discussion, it was agreed to clear everyone's calendar for three hours on a Friday afternoon and have a dedicated discussion of these concerns. Many facts and figures were presented, and many points of view discussed. It is possible that one of the facts considered was that broad indices of home prices in the United States had been falling since July 2006, with limited influence on the subprime mortgage security market. At the end of three hours, someone asked a question to which no one had a reasonable answer: "Where's the upside?" The point was that spreads to risk-free, government securities had been squeezed to historically low levels. The further upside would require still further compression of spreads. This seemed much less likely than a move back to much wider historical levels that would depress the market significantly. Thus, despite sacrificing some current yield, it was decided to liquidate the entire portfolio. They did this in just two days, and the liquidity was so great that it had no perceptible influence on the market.

This fund was not in a position to make a large profit as some did by shorting the market, but it did succeed in avoiding damage from the meltdown that eventually broke in September 2008. Without the type of intense sharing of viewpoints possible in a dedicated and extended discussion, it is unlikely that such a dramatic and ultimately correct decision would have been possible. Clearly, a regular series of meetings to discuss potential stress scenarios and their influences does not ensure success in avoiding the influence of sudden crises. Such practices do, however, enhance the likelihood that important but difficult decisions of this type can be deliberated and agreed on a timely basis.

Beyond Comparative Static Analysis

A legitimate criticism of traditional comparative static stress tests is that they are unrealistic. One reason such stress tests are only part of the answer is that crises rarely occur at once. There are exceptions such as the almost 23% drop in the Dow-Jones Industrial Average on 19 October 1987, but major, systemic crises unfold over time rather than instantaneously. The global financial crisis that began in September 2008 certainly did so.

The typical pattern is that some harmful events create immediate damage and create weaknesses elsewhere. These increase the chances that further events, which might otherwise have been absorbed with limited impact, create further damage and additional weaknesses, etc. Thinking through such a loosely connected series of related events is useful in forcing assessment of secondary implications that are not immediately obvious. No crisis unfolds exactly as it might be visualized in advance. As militaries often say, the battle plan goes out the window when the fighting begins. Nevertheless, a shared vision developed when creating the plan can improve coordination and performance even in the heat of battle. The same is true when reacting to a financial crisis.

Clearly, such dynamic crisis scenarios cannot be modeled at the level of detail that is possible for comparative static analysis. Trying to do so results in the exercise bogging down in unsupportable detail. Gaining insight into the broad though sometimes loose cause and effect process that could play out over time is the important point. Among other things, it can alert us to early warning signals of circumstances that trigger such a sequence of harmful events.

Such dynamic analysis is also a valuable means of making such events plausible in a way that extreme instantaneous shocks are not. By thinking through the potential impact of initial events, we also gain insight into their potential secondary and tertiary implications over time. In some cases, this allows organizations to insure, hedge, or more likely avoid the consequences of such follow-on events. By engaging a range of staff from a variety of functions across the institution, it stimulates diverse ideas about how to respond to such a series of events. An exercise such as this is bound to be too time-consuming to be carried out during a crisis. Nevertheless, the advance thinking such an exercise stimulates about consequences and potential responses to unfolding events can facilitate faster action in the midst of a crisis when reaction speed is of the essence.

The Myth of Sisyphus

In the end, of course, stress testing is like trying to cure the common cold. Like infectious viruses, crises come in too many varieties to allow a single silver-bullet solution. The effort to anticipate a crisis and protect against its consequences is like Sisyphus and his stone that always rolls back to the bottom of the hill no matter how many time he laboriously shoves it to the top. Nevertheless, applying a three-pronged approach to comparative static analysis, combined with thoughtful consideration of dynamic scenarios, can do a great deal to limit exposure to a potential crisis and assist in responding more effectively when a crisis occurs. As even Nassim

Taleb is prepared to say, thoughtful analysis of history can potentially convert black swans into gray swans, offering at least a fuzzy outline of how a potential crisis might unfold.

During all of this analysis, however, it is imperative that we learn from past mistakes and absorb lessons from relevant research outside our own area of expertise. We now turn to these.

Lessons from History

Risk managers are sometimes accused of failing to prevent the global financial crisis. This is, of course, unfair at one level. Risk managers do not set the level of risk in an organization; this is the role of senior managers and business heads. The risk manager's role is to illuminate risks implied by how business is conducted. Many risk managers warned that the boom in subprime lending in the United States was unsustainable. Often, however, they failed to make their case with sufficient force to change management actions. In addition, many experienced risk managers were surprised by the speed and magnitude of the crisis that erupted in September 2008.

It is important to take a hard look at where our thinking failed and how we might have made our case more convincing while the boom was still in progress. What follows is a list of areas where our thinking went wrong, but it is far from a complete inventory of such areas.

Beware Statistical Entropy

Noted earlier, the assumptions of classical statistical theory lie at the core of most risk analysis techniques developed over the past 25 to 30 years. Central to these methods is the assumption of a stable random process. In such a process, the value of any single draw is unknowable in advance, but sizable samples exhibit broadly similar characteristics such as the mean, dispersion (i.e., standard deviation), degree of symmetry or lack thereof (i.e., skewness), and tendency for probability in the tails to dwindle rapidly or slowly (i.e., kurtosis). The larger the sample, the more nearly identical will such characteristics be across sets selected randomly.

Classical statistical analysis recognizes that sampling techniques can never produce fully exact values for these characteristics, and has developed measures for the uncertainty of such estimates. The most common of these is the standard error of estimate, simply the standard deviation of the implied distribution of sample estimates for the true underlying parameter.

What is vital for both risk managers and general business executives to recognize, however, is that these errors of estimate assume stability of the underlying stochastic process. This is often a realistic assumption when dealing with physical processes. It is virtually never the case, however, in a social scientific setting; structural change is the constant bane of econometric forecasters. Such changes are driven by a variety of influences, including technological advances, demographic shifts, political upheavals, natural disasters, and perhaps most importantly behavioral feedback loops.

Structural change creates a fundamental dilemma for socio-statistical analysis. Classical statistics suggests that the more data, the better since, assuming stochastic stability, this results in smaller estimation errors. For analysis based on time series, however, a longer data set implies incorporation of a greater range of structural changes that undermine the classical assumption of stochastic stability.

This makes it all the more important for risk managers to focus obsessively on what can be call statistical entropy. Like water, information can never rise higher than its source. For information drawn from statistical analysis, that source is the set of data on which analysis is based. When assessing the reliability of any risk estimate, including such things as credit ratings, always begin with a review of the volume and quality of available data. No amount of complex mathematical/statistical analysis can squeeze more information from a data set than it initially contains. In complex settings, it is nearly impossible to extract all of the information that exists; something is always lost in the process of aggregating and summarizing. This is why it is more appropriate to refer to the law of statistical entropy than the more common law of the conservation of information, drawing an analogy to the second rather than the first law of thermodynamics.

A glaring example of failure to focus on the weakness of available data was the way many banks and investors blindly accepted the AAA rating for senior tranches of subprime mortgage portfolios. Before the crisis, such holdings were often treated as equivalent to AAA corporate bonds. Of course, rating agencies have about a century of experience with rating bonds. This provides a wealth of experience and data to support the effort. Subprime mortgages were a recent phenomenon, and their default experience had been dominated by a period of benign housing markets with stable to rising prices. Determining how much subordination was necessary to bring the chances of failure of timely payment of principal or interest down to a target level required estimating behavior deep into the tail of the default distribution.

A glance at available data for conducting this analysis would have made one thing clear: any estimate of the required level of subordination would necessarily be surrounded by significant uncertainty. We know that this market was undermined by serious erosion in underwriting standards to meet the apparently insatiable appetite for these securities in 2005 and 2006. Even before consideration of this type of structural change, however, the limited volume of data supporting the original AAA rating alone should have made banks wary of building up uncontrolled volumes of such securities.

The fundamental lesson to take from this experience is always to ask how much uncertainty surrounds risk estimates given the volume and applicability of available data. When such uncertainty is clearly excessive, be especially cautious when taking on corresponding exposures.

Recognize Self-referential Feedback

The most important reason to include dynamic analysis through time during stress tests is to capture some sense of potentially dangerous, self-referential feedback. Such feedback can take either favorable or unfavorable forms. Unfavorable feedback is the type we most often

remember since it can be sudden and frightening. Favorable feedback loops, however, can serve to lull us into complacency, and allow dangerously brittle circumstances to develop largely unchecked. The rise and fall of the subprime mortgage market offers an excellent example.

Self-justifying Success

Sometimes, the success of innovations creates secondary consequences that appear to justify initial success. Early development of the subprime mortgage market is a classic example. When subprime mortgages played only a small part in the U.S. housing finance market, defaults were driven primarily by idiosyncratic events that befell individual families, including:

- A prolonged period of unemployment;
- A messy and expensive divorce proceeding;
- The death of the primary breadwinner;
- A major uninsured property loss;
- A sizable medical bill.

The important characteristic of these drivers of default is that they tended to be statistically independent across borrowing households. This meant that the risk of such events could be diversified through a geographically distributed portfolio.

By reducing the variance of default risk through diversification, subprime mortgage investors could offer rates lower than those of small local lenders and still expect to earn an attractive return relative to risk. This created an initial surge of lending to supply the raw material for such securities. This increased lending, however, stimulated greater housing demand and pushed up prices in the face of a sluggish supply response. Rising prices, in turn, seemed to vindicate the appropriateness of such loans. Defaults substantially dropped as prices rose since financially troubled households could readily find a buyer and liquidate their investment at a profit. This was a classic favorable feedback loop driving the market forward.

From Virtuous Circle to Vicious Cycle

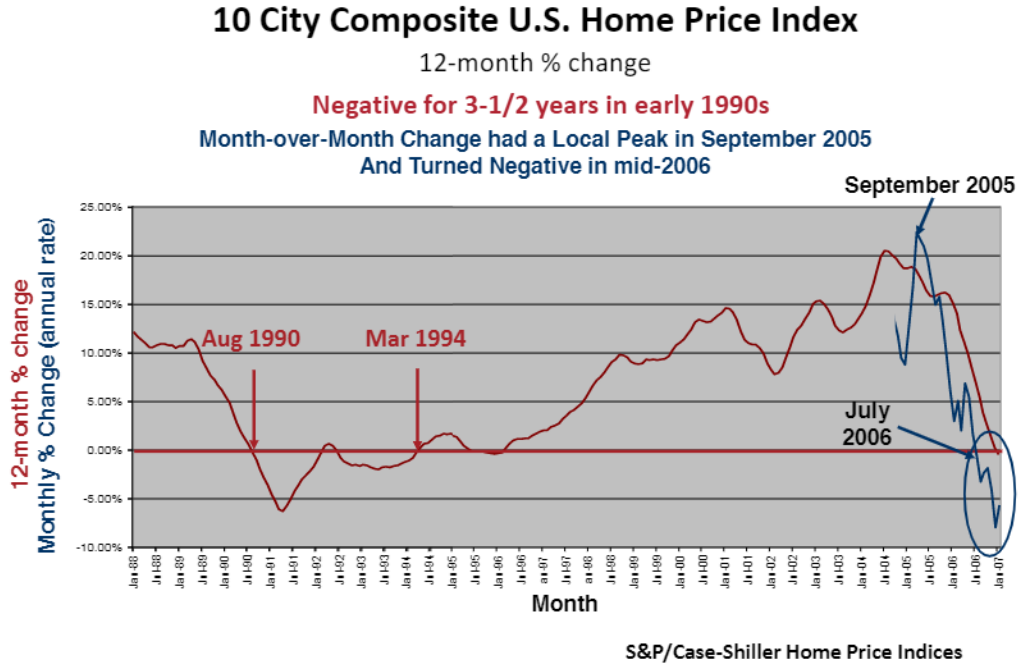
In retrospect, it is obvious that this market's heavy dependence on continued home price increases created a dangerously brittle dynamic. When price increases slowed dramatically in early 2006 and then turned negative in the second half of that year, this key support for the market was removed. By early 2007, troubled borrowers found it increasingly difficult to sell at a profit, and in many cases, to sell at a price that would cover their mortgages. The virtuous circle of rising prices and falling defaults suddenly reversed. Increasing foreclosures led to rising liquidation sales, putting further downward pressure on prices in a self-reinforcing vicious cycle.

As described in *The Big Short* by Michael Lewis, those who profited handsomely from the subprime collapse had recognized that prices did not have to fall to create a crisis; they only had to stabilize to trigger a feedback loop that could eventually drive prices down significantly. In light of recent history, seeing prices stabilize or fall slightly was far from a remote possibility.

Based on the Case-Shiller 10-City Home Price Index, this is exactly what happened for 3 1/2 years in the first half of the 1990s (see chart).

Given the possibility of a reinforcing feedback loop driving a severe downward spiral, it should have been clear that an institution with long exposure to subprime mortgages needed to be carefully monitoring home prices. In addition, it was important to raise the profile of unfolding home price behavior throughout senior management. Advance discussion of how large an adverse movement in an agreed early warning indicator should be to trigger aggressive reductions in a firm's exposure is an important means of minimizing the danger of organizational inertia.

Looking back, it is clear that some early warning signs were available. The rate of increase in the Case-Shiller 10-City Home Price Index slowed dramatically during the first six months of 2006. Even more startling is the fact that monthly, the index began to drop in absolute terms (i.e., deflation had set in) as of July 2006.



Firms paying close attention to home price inflation in mid-2006 had over six months to liquidate their exposure. Many argue that the first real "canary in the mine" was when HSBC announced in February 2007 that its bad-debt charges would be 20% higher than had been forecasted. Those who were paying close attention to the behavior of home prices, however, had a much earlier signal of problems ahead.

Trust Your Structural Imagination

In the midst of a bull market that appears to be driven by irrational exuberance (a term coined by Robert Shiller and famously used by Alan Greenspan in late 1996), it can be difficult to stand up and tell your organization "enough!" It is precisely in this situation that risk managers must

take their structural instincts seriously. As the economist Herbert Stein was fond of saying, "If something cannot continue forever, it will stop!" The only question is not whether it will stop, but when and how. Timing is always impossible to predict with any precision. Nevertheless, some thought concerning the structural mechanics on the upside often can offer valuable clues as to early warning signals of an eventual reversal. This is where applying a dynamic framework, rather than simply focusing on comparative static scenarios, is very effective.

The ultimate lesson from the rise and fall of the subprime mortgage market is to use our structural imagination to isolate variables that are crucial drivers of an unsustainable trend. Often, these variables are either not considered in our risk models or included only based on data for a favorable period of history. Then it is essential to insist that the unfolding realizations of such variables have very high visibility in senior risk management deliberations. Once a crisis hits, it is too late to begin thinking through the structural process involved and to avoid serious financial damage in any case.

Recognizing the central role of rapidly rising home prices in driving the subprime mortgage market was the first step. Then it was necessary to raise the visibility of this crucial indicator and prepare contingency plans for when and how to react if it began to signal danger. Those who did focused on the slowing increase and eventual fall in home prices over the course of 2006 and were prepared to take defensive action well before the mass stampede for the exits materialized.

The Role of Complexity

Necessary Complexity

Complaints about complexity are not new, nor are they without foundation. Since the dawn of the industrial revolution, the world has become progressively more complicated. Although such complexity causes a degree of discomfort and frustration for everyone, most of us believe these negatives are far outweighed by the associated comforts and conveniences that technological advances make possible. This familiar phenomenon can be observed in the evolution of financial derivatives. Despite daunting complexity that hampered broad public understanding, financial derivatives play a positive role in diversifying risk and shifting it, at a cost, from those less able to bear it to those more able and willing to do so. Nor are only the basic forms of derivative contracts socially useful. Average price options and those with knock-in or knock-out features are harder to price and hedge than traditional European options. Nevertheless, they are very useful at meeting the needs of some end-users.

Gratuitous Complexity

Although complexity is often a necessary byproduct of beneficial advances, it is not itself a good thing. Unnecessary complexity demands time and resources for training just to stay current with unfolding innovations. More insidiously, however, complexity can be a tool for the highly sophisticated to take advantage of those less knowledgeable, either consciously or inadvertently. This is largely what transpired in the interest rate derivative market in the early 1990s, and repeated in the subprime mortgage market in the mid-2000s.

In one of the most infamous cases, in the early 1990s, Bankers' Trust entered a swap with Gibson Greeting Cards in which Gibson received a then above-market fixed rate of 5.5% while paying LIBOR-squared divided by 6%. In this transaction, net payments remain in favor of Gibson for LIBOR up to nearly 5.75%. Beyond this point, losses mounted rapidly since increases in LIBOR soon caused the floating leg payments to rise more than twice as fast as the increase in payments on a simple LIBOR leg. The question that must be asked of such a structure is: what rationale it has, other than to hide the risk, that justifies a higher fixed rate, behind a haze of complexity?

From a risk-management standpoint, the lesson is to look with considerable skepticism at complex structures in which complexity cannot be justified as meeting structural needs of end-users. A structure for which the sole rationale is to use complexity to obscure the underlying risk creates significant reputational and compliance exposure for market makers. They should be forbidden unless managers can demonstrate that they fulfill an end-user's need that cannot be met by a more transparent product.

Not All Prices are Created Equal

Dangerous complexity is not confined to cleverly designed derivative trades; it can infect whole product markets, sometimes very large and active ones. This is largely what happened in the collateralized debt obligation (CDO) market in the first decade of the 2000s. CDOs structured with tiered loss tranches had attracted a variety of investors, with a range of risk/reward profiles in debt markets. Everyone understood that default correlation was central to the distribution of total credit losses in a CDO. Nevertheless, the casual, even simplistic, manner in which correlations were treated when quoting prices for these instruments should have been a cause for concern. Rather than building on the characteristics of the underlying instruments in a portfolio, each tranche was priced based on one pair-wise correlation across all names. Not only that, but the single, common correlation used for all names was different for different tranches, leading to what was known as the correlation smile.

Some attempts were made to address this anomaly by introducing a fat-tailed, multivariate distribution of default drivers instead of the usual assumption of a multivariate, normal distribution. This was in much the same spirit as introducing leptokurtosis into the distribution of underlying price changes to reduce the logical inconsistency of the traditional volatility smile produced by Black-Scholes option prices. Unfortunately, it proved to be a seriously inadequate extension of the Gaussian Copula model for capturing the high dimensional structure of multivariate correlations.

The assumption of a single, constant, default correlation across all pairs of underlying names was, and is, a simplification of breathtaking proportions. Despite promising proprietary research, no generally accepted market convention has yet been developed that can offer a reasonably robust means of addressing default correlations in a CDO pricing context.

Despite the obviously weak foundation for treating default correlation in the pricing of CDOs, the market began to introduce even greater complexity in the form of CDO-squared structures. These were compound structures in which the tranches of a CDO-squared were comprised of tranches of simple CDOs or mixed pools of such tranches. There was even talk at one point of CDO-cubed structures, in which the tranches would be comprised of tranches of a CDO-squared security.

This problem of complex compound structures was not confined to the corporate CDO market. By 2005, the concept of CDO-squared structures was quietly being extended in the burgeoning market for subprime mortgage securities. Even simple subprime CDOs were more complex than corporate CDOs because underlying collateral was far more varied, and relevant details were far harder to assemble. Despite this, subprime securities began to emerge in which unsold mezzanine tranches of simple subprime CDOs were rolled into new complex CDOs, effectively creating a CDO-squared structure. At this point, complexity was well and truly out of control.

Quite clearly, the ever-increasing complexity of collateralized debt obligations, driven by compound repackaging and lack of ready access to the characteristics of the underlying collateral, made fundamental analysis of these securities increasingly difficult. When the information required by investors to make sensible independent judgments becomes sufficiently complex and inaccessible, the normal efficiency of markets can fail. Simply put, markets require a minimum degree of transparency to operate effectively. When this condition is violated in the extreme, it is essential to remember that markets can fail to fulfill their central price-discovery function.

The lesson for risk managers is that complexity breeds opacity, which creates irresolvable uncertainty. Assuring sufficient transparency for markets to perform their function effectively should be an important role for public policy. Obviously, however, attempts to accomplish this are far from assured. Many interests benefit from opacity and struggle to preserve it. When transparency drops below a critical level, organizations are effectively operating in the dark. As risk managers, we must be prepared to assert that the risk cannot be estimated when the raw data to conduct a proper analysis are unavailable. Producing a risk estimate with our standard risk measurement techniques based on whatever data is available, regardless of how limited it may be, is professionally irresponsible. The sensible response is the same as when our headlights fail while driving at night—slow down!

Systemic Risk Lessons from Beyond Finance

Financial markets can be viewed as one member of a class of extremely complex, highly interconnected, dynamic, adaptive systems. This description fits many other systems that do not often enter discussions about the structure, behavior, and regulation of financial markets. Such systems include:

- Immunology;
- Epidemiology and infectious disease control;

- Pharmaceuticals;
- Fisheries management;
- Forest fire prevention and control;
- Air traffic control;
- Telecommunications networks and the Internet;
- Power grid management.

Considerable research has been conducted in these other disciplines concerning characteristics that make such systems either prone or resistant to periodic, large, systemic instability. Such analysis is typically concerned with the characteristics and interactions of entities within a system. Macroeconomics and finance too often focus inward, ignoring the potentially valuable lessons we might learn from these other disciplines. The following sections discuss a few of the lessons that financial risk managers would do well to incorporate into their thinking.

The Role of Super Nodes

An excellent example of this line of thinking is a 2009 paper by Andrew Haldane, chief economist at the Bank of England.[24] He discusses a variety of ways to characterize highly interconnected networks. One interesting concept for this purpose is the degree distribution of the network. The degree of any node in the network is the number of connections it has to other nodes. Were pair-wise connections distributed randomly, the degree distribution of the network would be normal, with a fat center and rapidly thinning tails. Many networks exhibit degree distributions that are thin in the middle and fat in the tails. Both nodes with very few links, and ones with a massive number of links, are overrepresented. Two obvious examples of such networks are the Internet (e.g., Google, MSN, Yahoo, Facebook, etc.) and derivative markets (e.g., JP Morgan, Goldman Sachs, Citibank, RBS, Barclays, Credit Suisse, Deutsche Bank, etc.)

Haldane points out that networks with "long-tailed distributions have been shown to be more robust to random disturbances, but more susceptible to targeted attacks" because most random shocks strike at the periphery of a network where the influence is distributed and can be absorbed easily. Failure of one of the massively connected nodes can have catastrophic consequences throughout the entire network by placing severe stress on a large number of other connected nodes.

This insight appears to fit well with the pattern of markets from 2004 to 2008. The years 2004 through 2006 were characterized by an almost eerie sense of calm. A major conference session in mid-2006 was entitled "Where is all the risk?" Interest rates were low, volatility was low, and the effect of occasional shocks appeared to dissipate with remarkable speed. Large institutions were able to absorb most peripheral shocks with little difficulty, causing daily volatility to be constrained.

[24] Haldane, A., *Rethinking the Financial Network*, April 2009. Available at http://www.bankofengland.co.uk/publications/speeches/2009/speech386.pdf

Such a period can lead to the anomaly Hyman Minsky describes as stability is destabilizing. Prolonged low volatility can induce a casual attitude toward higher leverage, resulting in less structural ability to absorb a genuine systemic shock. The growing leverage of U.S. investment banks from 2003 to 2008 is a classic example of such behavior. The trend certainly contributed to the ultimate failure of Lehman Brothers in September of 2008, which triggered massive reverberations through the system, affecting all of Lehman's vast multitude of business customers and counterparties. The essential lesson of the early 2000s and the role of super-nodes is that low, short-term volatility does not imply improved structural resistance to a major shock; the opposite is more likely true.

Let Small Fires Burn

Another related lesson arises from the area of forest management and forest fire prevention. For many years, the U.S. Forest Service followed a policy of quenching every small fire as quickly as possible. Exceptions to this policy were introduced in 1972. The rationale for this change was that "reduction of hazard fuels in selected areas reduces risks and costs." The idea was that small fires consume brush and other dry tinder that fuel fires. Periodic small blazes consume this fuel without doing irreparable damage to the trees themselves. Extinguishing every small fire allowed dry tinder to build up progressively. Eventually, it reached a stage at which a fire could spread out of control quickly, becoming so massive that it enveloped the crown of the trees and destroyed large sections of the forest.

Former Federal Reserve Board Chairman, William McChesney Martin, was known well for characterizing his job as "taking away the punch bowl just when the party is getting interesting." In effect, he was motivated by the essence of the lesson above. Small disturbances foster caution and greater robustness in a system. Eliminating all such disturbances fosters gradual relaxation of such caution and creates conditions that if prolonged, increase the likelihood of a major crisis.

Hide or Flee

Epidemiology is another source of systemic risk lessons. Communicable diseases have long been fertile ground for fiction writers. Epidemics represent universal and recurring situations that place individuals and societies under severe stress. Examples from literature include *The Plague* by Albert Camus and *Love in the time of Cholera* by Gabriel Garcia Marquez. A common theme in such works is the dilemma of whether the best way to avoid infection is to hide or flee. For individuals, both strategies are designed to avoid contact with others who have already succumbed to the disease. The broader, societal consequences of the two strategies, however, can be very different. Hiding effectively amounts to voluntary quarantine, supporting the public health goal of disrupting spread of the disease. Although flight can be effective for an individual, it risks contributing to the spread of the epidemic when a person who flees has unknowingly already contracted the disease and infects other previously unaffected areas.

In financial markets, the analogy to hiding is to withdraw from participation in market activities for fear that counterparties will fail to fulfill their obligations. The analog to flight is liquidating market holdings already in place because the quality of the assets has become highly uncertain. Examples include the wholesale unloading of Enron stock in the aftermath of revelations of liquidity problems and management conflicts of interest and a traditional bank run such as the one that occurred at Northern Rock in the United Kingdom in September 2007. Unlike the communicable disease situation, however, both of these reactions exacerbate a developing financial crisis.

Feedback, uncertainty, and innovation relate closely, and in combination, produce incentives to hide or flee. Innovation often leads to added complexity that increases uncertainty. Such uncertainty fosters panic when a sudden shock hits. As the quality of subprime mortgage securities suddenly became suspect, investors fled the market, creating major losses for banks that were substantial holders. Not having a clear idea of who might be a safe counterparty, the tendency was to hide by avoiding as many interactions as possible. Thus, when banks were suddenly unsure of each other's capital adequacy in the aftermath of these widespread loss announcements, they became unwilling to lend to each other for fear of who would be next to announce a big financial write-down. These are classic examples of individually logical actions that can result in reinforcing feedback loops that intensify a crisis.

Contributors to Contagion

Homogeneity

Another important relationship exists between nodal diversity and network stability. When the nodes in a network pursue diverse strategies, the network tends to be stable. When nodes are broadly similar, the network is increasingly susceptible to systemic crises. This is true of ecological systems such as fisheries, and financial systems. One example originated in the salmon industry in Chile. Salmon are not native to the cold waters of southern Chile, but they have been widely farm-raised in the area. Methods of aquaculture involved both breeding techniques to develop salmon that were resistant to disease and heavy use of antibiotics to suppress diseases that did arise. In 2009, however, a previously unknown strain of infectious salmon anemia appeared that proved impervious to both genetic resistance and antibiotic treatment. As a result, production collapsed by 87% in one season.

Homogeneity plays a similarly dangerous role in financial markets. When many institutions pursue similar strategies and measure risk similarly, the entire system becomes susceptible to a significant failure in one area. Clearly, the widespread tendency before 2007 to view AAA subprime mortgage tranches as safe and highly liquid allowed exposure to these instruments to pervade the system, creating a systemic crisis when the market collapsed.

Super Nodes and the Small-World Syndrome

Another characteristic of networks with concentrated super-nodes is the small-world phenomenon. The mechanics of six degrees of separation works through super-connected individuals. People frequently know their local legislative representative, and most elected

representatives know each other. This is a common channel by which people are connected by only three degrees of separation (i.e., citizen to representative, representative to representative, and second representative to another citizen.) This same process works globally through agencies such as the United Nations and G-20, and in the private sector, through organizations such as the World Economic Forum.

One implication of this small-world property of networks is the potential for local disturbances to make long leaps. For example, in March and April of 2009, a new strain of influenza that became known as swine flu appeared in Mexico. It soon spread globally as infected individuals moved among major world airports (i.e., transportation super nodes). This also is how the subprime collapse infected money center banks – another form of super node – worldwide. These, in turn, distributed subprime securities to many local municipal and corporate investors outside the United States, who were affected directly by the market's collapse. Had the role of super nodes as conduits of contagion been absent, the immediate consequences of the collapse would have been far more localized in the United States, where these securities originated.

Moving Beyond Value at Risk

Volatility Time and Risk Management Clock Speed

In a speech in late 2009, Myron Scholes spoke about a concept he called volatility time.[25] He meant the pace at which decisions under uncertainty need to be made to be effective. His point was that when volatility is high, it shortens the clock time available to make a decision early enough to be effective. In such an environment, delay can make an ultimately correct decision moot if the resulting action is undertaken too late to avoid losses or other damage; volatility time accelerates relative to chronological time.

A related concept is risk management clock speed,[26] the rate at which risk information can be assembled, processed, and made available to risk managers to support their decisions. Slow risk management clock speed can be fine when volatility is low and volatility time is unfolding at a leisurely pace. Unfortunately, when volatility time accelerates, it can overwhelm a risk management information system operating at slow clock speed. During a crisis, decisions are required in minutes or hours rather than days or weeks; decisions are increasingly made heuristically, and are informed by analysis of data that are often stale, incomplete, and too poorly structured to address the issue. To resolve this shortcoming, technology plays a distinctly double-edged role.

Technology as Both Enabler and Obstacle

Over the past 50 years, computing technology enabled a dramatic increase in the range and sophistication of quantitative analysis that can be applied in practice to both physical and social

[25] Remarks made in an address to the ICBI RiskMinds Conference in Geneva in December 2009.
[26] I first heard this concept enunciated in an exchange with Keith Smith of Risk Covered Limited.

systems. Despite this, the problem of slow risk management clock speed links directly to the way information systems have grown by accretion over years and even decades. A large institution's systems typically include components developed during different eras, with widely varied technological capabilities, constraints, and relative prices.

In the 1990s, global communication bandwidth was a serious constraint, and systems were designed to economize on this scarce resource. Most source systems had no message-based means of transmitting incremental, event-driven output to external systems. As a result, periodic (usually daily) production of batch files became the standard method of feeding data into central risk systems. Even in the best of circumstances, this approach meant that data would be at least 16 hours out of date. For a global trading operation with a daily close at the end of the New York business day, data for the opening of trading in Asia would often be from two days earlier.

A second problem that slows risk management clock speed is incomplete data coverage. When a counterparty encounters serious problems, having complete information on exposure to that entity in all areas of business is crucial. When Lehman Brothers failed in September 2008, many firms were unable to determine the full extent of their net exposure for days or even weeks, which hampered their actions targeted to limit losses. Even in institutions in which aggregate exposure was known at the enterprise level, it was not always easy for individual trading desks to isolate the specific trades that were now in default. This slowed the process of executing the necessary hedges to rebalance the book, resulting in unwanted market risk in an extremely volatile environment.

At the root of all of these difficulties is inconsistent organization of data across most institutions, resulting in limited ability to tabulate and analyze non-standard aggregations at the bank level. The common denominator in this situation is the nearly universal deployment of SQL relational databases. These were a dramatic advance on previous data storage methods when the concept was introduced in the mid- to late-1980s. The technology provided significantly greater flexibility in the analysis of any well-defined class of data for which customized table layouts, field definitions, and links had been defined. Within the framework of these table definitions, such databases enabled highly complex queries to access, filter, sort, group, and otherwise manipulate data with great flexibility.

However, it is important to remember that SQL technology was designed to economize on storage capacity and CPU processing cycles that used to be far more expensive than they became over the following 25 years. Relational databases were, and will continue to be, a powerful and effective tool in environments in which the structure of underlying datasets is stable. They fall short when dealing with a dynamic environment in which the underlying data structure itself and the types of questions that need to be answered are unstable and subject to frequent revisions. As long as data that need to be assembled and analyzed holistically are stored exclusively in multiple, inconsistently specified SQL databases, this problem will remain unsolved.

Technology as the Solution

In the mid-teens of the 21st Century, an excellent case can be made that multiple emerging trends in technology offer an opportunity for significant improvement in risk management clock speed. Among these trends are:

- Dramatic declines in the cost of RAM and CPU processing;
- Advances in indexing and search methodology;
- Dramatic expansion of the potential for parallel processing;
- Cloud computing, offering an elastic supply of CPUs on a cost-per-second basis;
- Radical modularity in:
    - Data storage formats;
    - Independently executable code components;
    - CPU/GPU resources.

Economists discuss putty/clay models of growth; once capital investments are made, they shift from malleable putty into largely inflexible clay. Once in place, such investments are not altered easily. Given that wholesale removal and replacement is not commercially sensible in anything short of life-and-death situations, deployment of new and improved techniques inevitably occurs gradually. Moving beyond the current dysfunctional state of enterprise data storage will take time. Eventually, core systems of record will need to shift their data storage to a more modular and readily indexed framework. In the meantime, exporting data into such a framework can make important advances, even though this step introduces opportunities for errors and delay.

Institutions need to have a clear vision of a more transparent enterprise data environment, a strategy for moving steadily toward this environment, and the discipline to stay the course. Risk management needs to be an active participant in, and a firm advocate of, such a change. Failure to progress toward a significantly more functional architecture for enterprise data hampers a firm's risk management capabilities relative to competitors, especially during a crisis.

Practical and Organizational Considerations

Clearly, stress testing is far less prescriptive and tightly defined than traditional distributional analysis based on statistical theory. As a result, variations during stress testing across institutions are inevitably greater. Nevertheless, some general, best-practice principles are now emerging.

Severity, Breadth, and Speed

Previous sections presented three broad approaches to defining stress scenarios based on:

- Historical events;
- Assessment of an institution's own vulnerabilities;
- Imagination applied to current events.

In addition to these approaches, it is useful to segregate scenarios based on three more characteristics:

- Severity;
- Breadth;
- Speed.

Severity relates to the significance of the adverse, external conditions, which are reflected in scenarios. Although we can think of severity as inherent in a given scenario, in practical terms, severity is an institution-specific concept. It relates to how big a threat to an institution's prosperity or survival would be created by a scenario. In this sense, severity can be classified broadly into four categories.

Normal Stress Scenarios

It is reasonable to expect the occurrence of these scenarios once or twice during a ten-year period. They represent the low point of normal, recurring business cycles that occur several times in a forty-year professional lifetime. Generally, this type of scenario should be manageable within the normal structure of roles and responsibilities for daily decisions. Critical guidelines might need to be tightened and credit criteria made more rigorous, but these fall within the normal scope of regular policy adjustments. Although these types of events generally lead to reduced earnings and increased loan losses, they usually do not present a serious threat to the survival of a well-run financial institution.

Severe Stress Scenarios

These are scenarios that one would expect only once or twice in a professional lifetime. The two oil shocks of the mid- and late-1970s triggered unusually severe economic consequences that were unmatched for another 25 years until the onset of the global financial crisis in 2008. These episodes represent severe stress scenarios for many institutions, but would be classified in one of the next two categories by others. Severe stress scenarios are normally included in regular stress-testing exercises. They will inevitably result in declines in earnings and some periods of losses. They might involve significant defensive decisions, including cutting back on selected business areas and suspending dividend payments to conserve cash. With proper early-warning indicators and timely action, however, institutions should be able to avoid serious risk of default in this environment.

Near-Default Stress Scenarios

The global financial crisis that began in late 2008 falls into this category for some institutions, especially many involved heavily in the creation and sale of subprime mortgage securities. This would generally be true for institutions that came close to default but were able to weather the storm without extraordinary assistance from the government. Near-default stress scenarios form the basis for the development of a detailed recovery plan. Such a plan represents an institution's response to extraordinary conditions during which comparably extraordinary actions are required. Careful thought should be given to the range of actions that might be

required. More importantly, a clear definition of lines of authority, once a recovery plan is invoked, is essential if it is to be effective.

Stress to Default Scenarios (Also Called Reverse Stress Test Scenarios)

Some institutions failed amid the global financial crisis, and several others would have failed had public support not been forthcoming. For all, this period represented a stress to default scenario. Reverse stress tests are a means of overcoming the inherent tendency to allow the past to constrain our imaginations about possible future events. The idea is to begin with the premise that the institution failed and then formulate conditions that are sufficiently severe to cause this. The presumption is that this will involve extremely unlikely events, but the exercise forces managers to think through a firm's most serious vulnerabilities and design stress to default scenarios accordingly. Again, broad organizational involvement is essential when defining appropriate events—failure of a major counterparty, rogue trading losses, internal fraud, reputational damage, or loss of funding—that might contribute to institutional failure.

Breadth relates to whether a scenario is idiosyncratic or systemic; does it involve damaging circumstances that are specific to one institution such as a major cybersecurity breach or a rogue trader loss, or ones that are damaging to the industry and economy as a whole? This can have significant implications for the effectiveness of potential response measures such as asset sales to bolster liquidity. Broad, systemic scenarios often mean that other firms might be attempting similar corrective measures simultaneously, making them less effective for all.

Speed relates to how quickly damages will be sustained once an event occurs, relating to volatility time and risk management clock speed described above.

Organizational Involvement

Scenario Development

Broad, active organizational involvement is essential if stress testing is to be more than an isolated compliance exercise. Done well, it is an excellent opportunity to encourage constructive discussion across traditional business silos. Success requires explicit and highly visible support from top managers affirming the value and importance of stress testing as a core competency of good long-term management.

Formulation of scenarios that need to be examined involves both top-down and bottom-up perspectives. Scenarios often originate from senior managers' concerns about potential macroeconomic or geopolitical events. These will generally be given specific definition by corporate staff members, including members of the economics department. In some cases, however, market circumstances might present serious risks that are obvious only to those closer to the daily activity. Stress testing should be an opportunity to encourage those with such concerns to bring them forward. This can be a useful way to bring such concerns out of the shadows and, where appropriate, to bring specialized resources to bear and assess their severity.

Broad participation is also vital in the necessary process of translating aggregate stress assumptions into detailed implications for individual business areas. Those involved in each individual business will have valuable experience to bring to the discussion. It is important, however, for a centralized stress testing team to ensure individual businesses do not lowball the estimated impact on their activities to make them look less risky.

For scenarios of continuing concern, it is particularly important to develop a consensus around early warning indicators and potential trigger points for remedial action. Tracking such metrics contributes to organizational awareness of whether scenarios are becoming more or less likely. Setting action triggers in advance makes it harder to dismiss unfolding events as not yet severe enough to warrant difficult decisions once triggers are reached.

Recovery Plans

Recovery plans should be at the core of long-term strategic management. A plan should consider a range of potential corrective actions and their likely effectiveness and timing, including:

- Potential sources of additional capital;
- Back-up sources of funding liquidity;
- Cutting back or disposing of selected businesses;
- General expense reduction options.

Such actions need to be material in their impact, feasible in a crisis environment, and plausibly effective at restoring confidence in an institution's viability. In some cases, these potential crisis measures suggest pre-crisis steps to ensure their feasibility and/or enhance their effectiveness when needed. One obvious example is setting up guaranteed backup credit lines to be available during a crisis despite the associated short-term increase in cost.

It is ultimately the responsibility of senior managers and the board to decide when to mobilize a recovery plan. Nevertheless, a capital and risk committee should be charged with monitoring events and recommending to senior managers and the board when it believes a recovery plan needs to be invoked.

Once a recovery plan is invoked, time is always of the essence. For this reason, the underlying near-default scenario, including its capital and liquidity impacts and the associated recovery plan, should be defined with a more granular time grid than is used during a standard stress-test exercise. This imposes an obligation to be explicit about the speed of the damage and the necessary speed of implementing required responses if they are to be effective.

Effective execution of a recovery plan also requires clear lines of authority. To ensure this, a plan should have a well-defined, senior-manager structure for implementation, including *ex officio* definitions of who will play what roles and with what authority as part of a recovery response committee during execution. A lower level recovery management committee should also be defined in advance, designating who will have daily responsibility for coordinating responses and assessing their effectiveness.

Resolution Plan

By definition, stress to default scenarios involve an institution's failure. For this reason, they are the basis for resolution planning rather than recovery planning. A resolution plan involves specification of how an institution's assets would be allocated to claimants at various levels of seniority, and how cross-entity guarantees and collateral agreements would function in liquidation. Regulators demand that this plan must specify how such a liquidation plan can be accomplished with minimal government support. In particular, it should reflect a proposed resolution of conflicting claims involving uncertainty around seniority status and allocation of credit mitigants.

Gap Identification

One benefit of thinking through detailed recovery and resolution plans is to identify gaps in current arrangements that might prevent successful execution. Filling these gaps enhances the robustness of an institution by increasing the likelihood that such plans can be successfully executed.

Challenges of Stress Testing

The Challenge of Two Cultures

Technology is not the only obstacle to improving bank risk management capabilities. Another problem is cultural – a failure between quants and the larger community of traditional finance managers to communicate effectively with each other. Quantitative pricing techniques and statistical risk management are little more than opaque black boxes for all too many general financial executives. What is more, those who understand the technical details often have limited insight into broader structural, behavioral, and business issues.

Running a financial institution demands a constant series of large and small decisions to be made under uncertainty. Such decisions can never be effective if they are made mechanically. Effective decisions must reflect experience and judgment conditioned by available empirical evidence. As finance became more complex and quantitative, the communications gap between finance's two cultures has become ever more consequential. Most senior bank managers have limited ability to weigh the subtle mathematical details of modern finance, and few state-of-the-art quants are well equipped to assist them, even if they are motivated to do so.

One example of the consequences of this cultural divide is the fragile nature of the Gaussian Copula Model for pricing CDOs, described earlier. The framework depended heavily on market liquidity as a basis for implied default correlations, for which no objective historical analogues were available. In the absence of liquidity, valuing such tranches is reduced to little more than guesswork. If bank senior managers had realized this critical dependence on liquidity to determine fair values, they might have hesitated to hold such massive amounts of these securities on their balance sheets.

The blind acceptance of AAA ratings assigned to senior tranches of subprime mortgage CDOs was noted earlier. A little thought to the paucity of data available to estimate behavior deep in the tails of the loss distribution of subprime mortgage portfolios should have raised serious doubts about both the reliability and stability of such ratings. Most senior bank managers simply did not have sufficient quantitative insight to recognize the need to ask the question in the first place.

Some quantitative risk managers might be able to broaden their horizons to incorporate softer cultural, sociological, and political issues. A more assured approach at an organizational level is to incorporate specialists in these areas into the continuing risk dialog. Doing so requires mutual respect across disparate disciplines, and concerted effort to enhance the quality of communication across the cultural divide.

In his essay entitled *The Two Cultures and the Scientific Revolution,* Snow argues that a society that allows what he calls two cultures to grow apart is unable to think with wisdom. Similarly, if organizations are to think with wisdom, it is essential that they strive to bridge the communications gap between quantitative specialists and general financial executives.

Dangerous Adaptation

Adaptation is one of the most powerful phenomena in nature. It is the means by which species survive changes, often major changes, in their environments. In this sense, we are conditioned to look on adaptation as a favorable characteristic. It is important to remember, however, that effective adaptation has its dark side because it also can be a source of strength and resilience for dangerous threats. The 2009 outbreak of flu caused by the H1N1 virus is a classic example of the challenges faced by epidemiologists. As in many previous cases, this was a new and unknown virus. Although similar to other viruses, it developed a mutation that made it resistant to known forms of prevention and treatment.

Those of us in financial risk management would do well to recognize that in terms of adaptation, the problems we face are similar to those confronting epidemiologists. We are not external observers of a distinct and independent system. Risk management is an integral part of the system, the risk of which we seek to control. This points out the ultimate futility of trying to control financial institutions by detailed micro-regulations. This is particularly relevant when such regulations evolve not over weeks or months, but over years. The underlying institutions and systems adapt much faster than such rules and regulations can be updated.

One classic example of financial sector adaptation is the behavioral adjustments traders made to the introduction of VaR as the standard metric for controlling market risk.[27] Once traders were constrained by a comprehensive measure of 99% risk, one way to take positions with additional risk, and the associated expected return, was to sell out-of-the-money options. Since

[27] See Rowe, D.; Corrosive feedback, *Risk*, January 2009, p. 99. Available at: http://www.dmrra.com/publications/Risk%20Magazine/200901%20Corrosive%20Feedback.pdf

the value of these positions initially did not change significantly in response to daily market moves of 2 to 2-1/2 standard deviations, they had little initial influence on the reported 99% VaR figure. In the face of larger, black-swan events, however, losses could accelerate rapidly as such options approached an at-the-money position. The corrosive feedback effect was that widespread use of VaR as a control metric encouraged exactly the type of risk-taking VaR fails to measure: exposure to extreme events. VaR not only fails to address the most extreme losses, but it also encourages behaviors that increase their magnitude. This is a perfect example of how controlling one form of risk induces increased risk in other, often more virulent, forms.

The subprime mortgage experience highlights a related issue known as Goodhart's Law.[28] One way of stating this is that when a reliable indicator is made a target of social, economic, or organizational policy, it quickly loses its effectiveness as a meaningful indicator. There are multiple grounds for criticizing the methodology of the rating agencies when determining how much subordination was required before a tranche of a subprime CDO should qualify for AAA status. An additional problem, however, was that once agencies published their methodologies, the market began to game them in every way possible. Such gaming undermined what limited reliability these ratings initially had.

Conclusion

In the aftermath of the global financial crisis, it is perfectly acceptable to recognize that the shortcomings of banks' risk technology and risk management will not be solved quickly or cheaply. What is not acceptable is simply ignoring the issue of how far current risk management information systems and other resources fall short of what is required until the next crisis drives the point home, again.

Although politicians and regulators are busy fighting the last war, it is essential for risk managers to remain alert to how markets and institutions are adapting products and strategies. Dealing with such an ever-shifting problem requires information systems and analytical tools beyond those currently in place. Even if we take responsible steps to improve risk management information resources, we must always remember that ours is a profession in which there are no final victories.

Adaptation is the fundamental reason that any claims that regulation can ensure that "this will never happen again" cannot and should not be taken seriously. Human beings are too ingenious and too much a part of the highly adaptive biological system for such claims to be sustained. Only constant vigilance, with special attention to the risk inherent in the adaptive changes taking place around us, allows organizations to avoid the worst consequences when

[28] The law was named for Charles Goodhart, a former advisor to the Bank of England and emeritus professor at the London School of Economics.

the next crisis occurs, as it inevitably will. As William Shakespeare has Cassius say, "The fault, dear Brutus, is not in our stars, but in ourselves."[29]

Appendix A - Examples of Stress Testing
Scenario Formulation - The Fundamental Challenge of Stress Testing

By its nature, stress testing differs from distributional approaches to risk assessment. Traditional distributional methods begin with a sizable sample of historical data on the phenomenon under analysis. These are then used either directly in a non-parametric approach such as historical simulation, sometimes called a walk through history, or are used to derive statistical parameters of the supposed underlying random process. These parameters are then used either to perform a variance/covariance analysis or to construct a Monte Carlo simulation. In all cases, once the relevant historical data are collected and organized, analysis is tightly prescribed by standard statistical theory. Some methodological tweaks might be implemented such as using some form of fat-tailed distribution rather than a simple normal distribution, but room for variation is highly constrained.

In contrast, stress testing is concerned with the types of events during which there is nothing like an adequate sample of historical data to justify applying classic statistical methods. One might seek guidance from a handful of extreme historical events, but one must always interpret the relevance of such input cautiously in light of the inevitable distinctions between the present and past. In the end, we are forced back on the sage insight supposedly spoken by Mark Twain when he said, "History doesn't repeat itself, but it does rhyme."

Lack of a clear, defining structure creates the fundamental challenge of stress testing, that is, deciding which hypothetical scenarios are worth the time and effort to analyze at some level of detail. It is sometimes said that "We cannot forecast the future from the past." When interpreted as saying that the past can tell us nothing about the future, this becomes much too categorical a statement. It is true that we cannot forecast the future accurately from simple or even highly complex extrapolations of the past. Nevertheless, when formulating ideas about the shape and magnitude of uncertain future events, we have nothing but the past to guide us. The crucial question is how we can realistically use our knowledge about the past. Careful study of the past never allows us to anticipate every potential crisis. In short, we will never turn black swans into white swans, but even Nassim Taleb concedes that careful attention to the rhymes of history might allow us to convert some black swans into gray swans.

The following sections offer thoughts and examples regarding approaches to formulating relevant hypothetical scenarios. Relevance in this sense depends on both the plausibility of such scenarios and their likely impact should they occur. Managers are usually perceptive

[29] Shakespeare, William, *Julies Caesar*, Act I Scene ii.

enough to avoid scenarios that are both highly likely and highly damaging. This leaves us needing to define some critical combinations of plausibility and impact to justify the cost of analyzing a particular stress scenario in detail.

The Market's Greatest Hits - Calibrating Stress Scenarios Based on History

The major organizational challenge of stress tests is reaching broad consensus on the magnitude of the shocks to analyze. The typical reaction is, "That could never happen." This points to the central reason for including the market's greatest hits among a firm's scenario development strategies. Stress scenarios rooted in historical events are shielded against this casual basis for dismissal. Critics are at least reduced to saying, "That could not happen now." That, however, is a far weaker sounding argument that has a ring of special pleading.

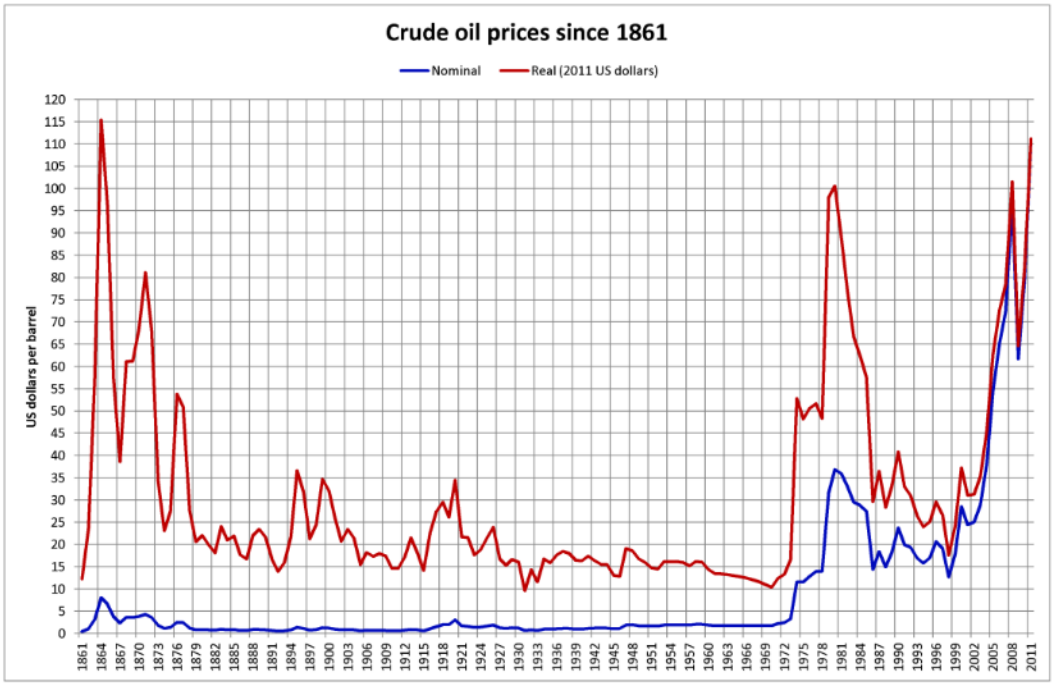
In this context, it is useful to maintain an archive of the magnitude and timing of relevant historical events. Obviously, the relevance of any given event depends on assessment of:

- Its current plausibility;
- Its expected general economic and financial impact if it occurs;
- Consequences to an institution's financial performance.

Examples of what an archive of historical events looks like follow.

Oil Price Shocks

The graph below shows a historical time series of crude oil prices in the 150-year period 1861 to 2011 and is thus the starting point for any discussion about stress testing in any portfolio for which the market price of oil is a variable.



The graph illustrates some important points. It is important to adjust nominal prices for changes in the general price level when making long-term comparisons. The blue line looks benign from 1861 to 1973, with only small fluctuations in the 1860s and 1870s, and again during World War I. Once you adjust for the general price level however, the situation looks very different. In fact, the volatility in the earliest years was comparable to what we have seen since 1973. The chart also shows that the remarkably stable period from 1948 to 1970, when real oil prices experienced an extended gradual decline with only minor fluctuations, was exceptional. We also tend to remember periods of sharp oil price increases but sometimes forget times when prices have shown comparable declines. From 1973 to 1974, crude oil prices tripled, from about $17 to $53 per barrel in real 2011 prices. From 1978 to 1979, prices doubled from $48 to $97 per barrel. As a historical guideline, it is fair to say that one-year stress increases of 300% and six-year stress increases of 600% have historical precedent.

From 1981 to 1982, prices dropped from $100 to $65 per barrel or 35%.[30] The nominal decline was smaller, but it occurred in the face of some of the highest general inflation in recent U.S. history, accentuating the size of the decline in real terms. By 1986, the real price had declined to $30 per barrel, for a cumulative, seven-year drop of 70%. As a historical guideline, it is fair to

[30] A nearly identical percent decline in the real price occurred from 2008 to 2009, followed by a sharp reversal to above the 2008 price in 2011.

say that one-year stress decreases of 35% and six-year stress decreases of 70% have historical precedent.

Clearly, these kinds of changes would be plausible only in the context of significant changes in global energy markets. For price increases of this magnitude, a major disruption in multiple sources of supply worldwide would likely be required. For such price decreases to occur requires either major expansion in energy supply (say from an expansion in hydraulic fracturing in the U.S. and elsewhere, or a major breakthrough in the efficiency of alternate energy sources such as solar) or a major decline in total energy demand due to severe economic contraction.

Exchange Rate Movements

This section provides motivating historical examples relevant to stress testing a portfolio for which FX rates are a variable.

Emerging-Market Countries

Exchange rate movements depend on a complex mix of world trade flows, capital flows, shifts in psychology, and changes to public policy. Generally, heavily managed currencies are subject to bigger and sudden shifts. These typically occur when attempts to defend a fixed peg need to be abandoned due to excessive cost. Some selected historical episodes offer useful perspectives.

The Argentine Peso/U.S. Dollar Peg

In 1991, after 15 years of high and sometimes devastating hyperinflation, Argentine policy-makers decided their only option was to establish a currency board charged with maintaining a one-to-one exchange rate between the peso and the U.S. dollar, with full convertibility of pesos to dollars. This was successful in the short-term at stabilizing the economy, but by late 2001, economic stresses made maintaining the peg increasingly difficult politically. In January 2002, rising public debt denominated in foreign currencies and pressure on the currency board's resources forced the government to abandon the peg to the U.S. dollar. The peso dropped 29% overnight, and eventually lost 75% of its value in about four months.

The East Asian Currency Crisis of 1997/1998

Thailand. From 1985 through 1996, the Thai baht was pegged at 25 to the U.S. dollar. In mid-May of 1997, the currency experienced significant pressure from speculators. At first, the prime minister said he would not devalue the currency, but by July, the government was forced to do so. In the end, the exchange rate fell to 56 to the U.S. dollar, a decline of 55%. Other currencies in East Asia came under speculative pressure partly because of the Thai baht crisis. Initial devaluations pressured the local currency profits of companies that had borrowed in foreign currencies and failed to hedge their resulting currency exposures. This led to further loss of confidence and further exchange-rate pressure.

Indonesia. In mid-1997, the Indonesian rupiah traded at 2600 to the U.S. dollar. By early January 1998, the rate had plunged to 11,000 to the U.S. dollar, and the currency traded as low

as 14,000 to the U.S. dollar over the next six months, before recovering to 8000 to the U.S. dollar by the end of the year. Thus, the currency dropped by 76% to 81% versus the pre-crisis level, before recovering to a drop of 67% from that level by the end of the year.

South Korea. In mid-1997, South Korean banks were suffering from an increase in non-performing loans since domestic conglomerates had borrowed heavily to establish themselves on the global stage. By the end of the year, poor corporate performance and a falling stock market were accompanied by credit downgrades for both private borrowers and the country's sovereign debt. In the midst of this, the South Korean won fell from 890 to the U.S. dollar in July 1997 to nearly 1700 in January 1998, before recovering to 1200 by the end of the year. Thus, it experienced a six-month decline of 48% before recovering over the following 12 months to 26% below the pre-crisis level.

Based on the experience of East Asia in 1997/1998 and Argentina in 2002, emerging market currency declines of 50% to 80% represent historically precedent events.

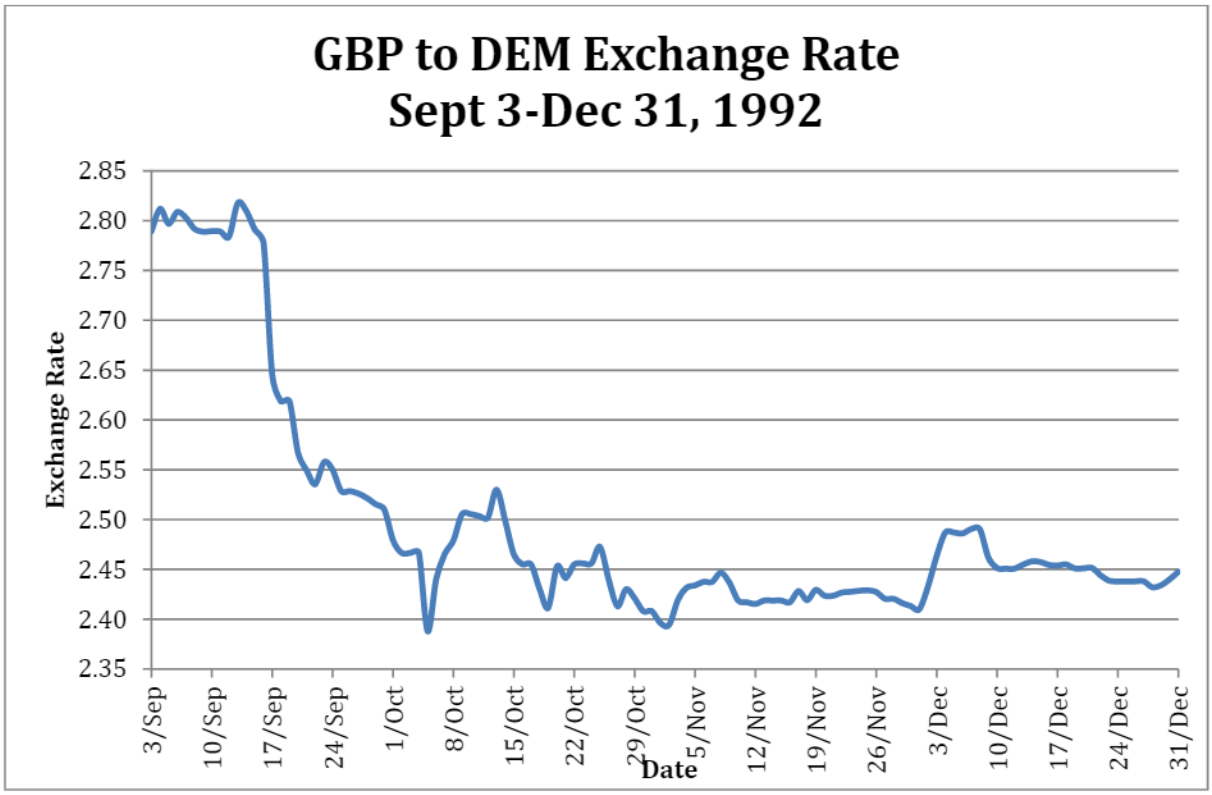
Developed Country Currencies

The range and speed of plausible exchange rate movements for developed country currencies depend heavily on whether the value of the currency has been allowed to float freely or subjected to management by political authorities. Generally, managed currencies are more subject to sudden value changes than those allowed to float freely. Despite this, significant exchange-rate movements can result from shifting market perceptions. The British pound offers examples of both types of movement.

The Pound and The Exchange Rate Mechanism (ERM) Crisis of 1992

Britain has had a checkered relationship with attempts to stabilize exchange rates between various European currencies. The country refused to join the original ERM when it was set up in 1979, but during the late 1980s, authorities began a semi-official policy of shadowing the value of the German Mark. After considerable internal clashes within the Thatcher government, agreement was reached to join the ERM formally in October 1990. Shortly thereafter, Margaret Thatcher was unseated as Prime Minister and replaced by John Major.

Unfortunately, at the time Britain entered the ERM, it was experiencing inflation rates three times higher than those in Germany. By late 1992, traders had become convinced that the U.K. government could not sustain policies needed to maintain the value of the pound within the 6% band specified in the agreement. The crisis broke on 15 September 1992. Massive market sell orders had to be covered by the U.K. Treasury buying pounds at the bottom of the agreed range. The next day, the government first raised interest rates from 10% to 12% and then to 15% to encourage purchases of pounds, but to no avail. Sell orders continued to pour in. That evening, the chancellor announced that Britain would leave the ERM and devalue its currency.



As shown in the chart, the GBP/DEM exchange rate immediately fell by 4-3/4% on 17 September. It continued to fall over the ensuing two weeks, breaking through 2.5 DEM to the pound on 1 October, for a total decline of over 10%. During the remainder of the year, the exchange rate fluctuated between 2.4 and 2.5 D-Marks to the pound. For the fourth quarter as a whole, the rate averaged 2.44 D-Marks per pound, or 12% below its close on 16 September prior to the decision to devalue.

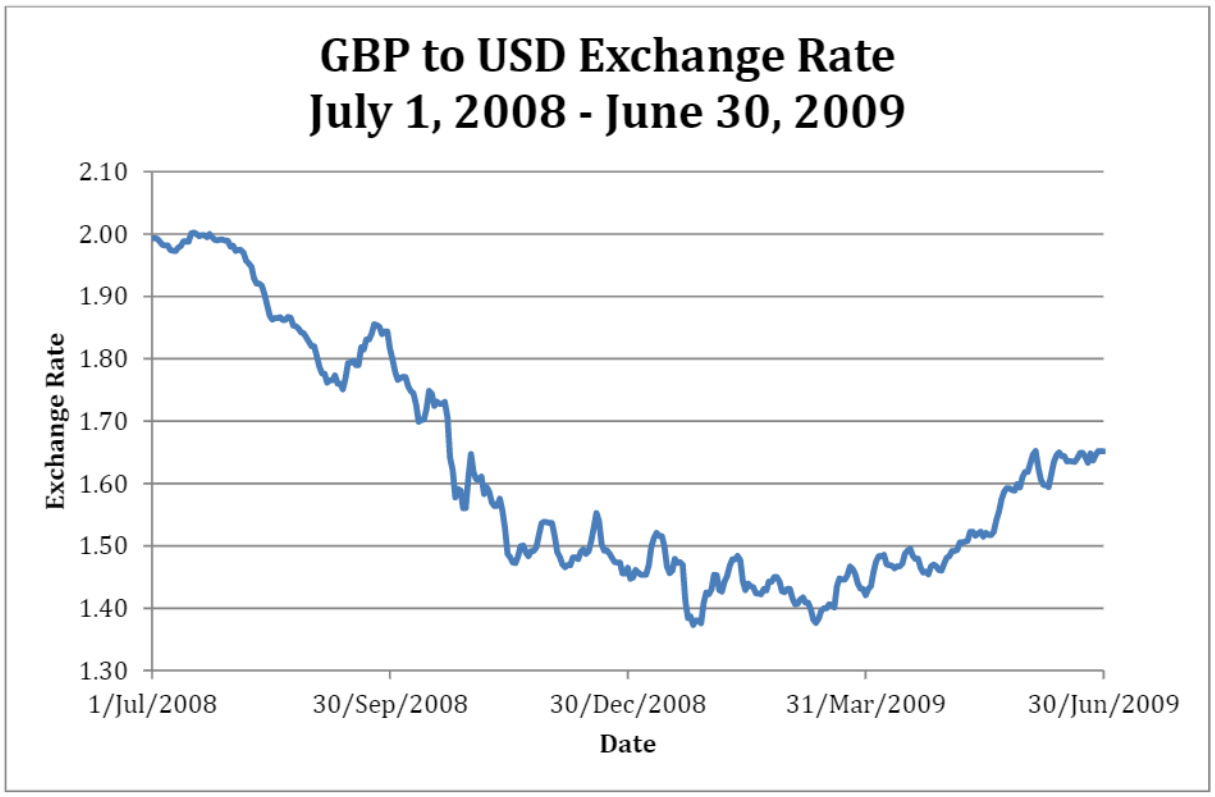
Based on this experience, a major industrial country's currency facing a crisis in a managed peg could reasonably devalue by as much as 15% over a few days if it decides to stop defending the currency. Obviously, such an estimate should be analyzed against purchasing power parity to ensure it is likely to be sufficient to reestablish its competitive position in international trade. If purchasing power parity is more out of line than this, larger potential changes need to be considered.

The Pound after the 2008 Financial Crisis

A different example is represented by performance of the British pound after the onset of the global financial crisis in September 2008. The currency had been floating freely for many years, so there was not an issue of trying to defend a fixed exchange rate in the face of adverse fundamental conditions. The last half of 2008 through the first half of 2009 was one of the most tumultuous periods in global financial markets in decades. Arguably the most important financial market in the world, London, and Britain as a whole, experienced significant economic prosperity in the boom years from 2003 to 2007. As the subprime financial crisis simmered

through 2007 and early 2008, the U.K. exchange rate continued to hold up. Starting in August as concerns continued to mount, the rate began to slide. Interestingly, it rallied after the Lehman Brothers collapse on 15 September, perhaps based on the idea that Lehman was a major U.S. firm and London might gain competitive advantage from a setback to its largest competitor as a global financial center.

Soon, however, problems began to surface in the United Kingdom, eventually leading to the rescue and partial nationalization of the Royal Bank of Scotland on 20 November. By this time, the USD to GBP exchange rate had fallen to 1.50, a drop of 25% from its mid-year value. The rate fluctuated around 1.50 until early January 2009 when it declined further to a cyclical low of 1.38 late in the month and again touched this level in mid-March (thus, having experienced a cumulative decline of over 30%). After mid-March, it began a steady rise to 1.65 at the end of June, recovering almost half of its previous decline and leaving it nearly 17% below its year-earlier level.



It is interesting to compare the magnitude of GBP rate moves against the D-Mark around the ERM crisis and those versus the USD during the onset of the global financial crisis. Surprisingly, both periods experienced peak daily declines of over 4% and peak two-week declines of over 10%. The longer-term decline in sterling after the ERM crisis was about 15%. The maximum decline from early July 2008 through the first quarter of 2009 was almost twice as large, reaching over 30%. By the end of June 2009, however, the decline from the previous July had come back to somewhat over 15%, similar to the enduring structural shift after Britain left the ERM.

It appears that economic fundamentals drive the needed adjustment and eventual shift in exchange rates more than public policy to manage a country's currency. Of course, much depends on how far public policy allowed the exchange rate to diverge from a value that is economically sustainable. Thus, when exchange rates begin to shift, keeping an eye on purchasing power parity is a useful exercise.

The Plaza Accord to Weaken the U.S. Dollar in 1985

Since the end of World War II, the U.S. dollar has been the dominant unit of account for most foreign-trade transactions. This provided a source of demand that makes the market for dollars the largest and deepest in the world. Its value behaves uniquely, and the depth of the market limits the ability of individual traders to move the market significantly. A useful benchmark for the potential impact of a coordinated official intervention occurred in 1985.

From 1980 through the end of 1984, the trade-weighted value of the U.S. dollar rose dramatically from 95 (based on 1973 = 100) at the beginning of 1980 to 137.3 at the end of 1984, an increase of nearly 45%. This created a significant balance of payments deficit in the United States and placed intense pressure on the profits of U.S. manufacturing firms. In September 1985, the finance ministers of West Germany, France, Britain, Japan, and the United States met at the Plaza Hotel in New York City. They agreed to a coordinated currency market intervention designed to weaken the U.S. dollar relative to the other four national currencies. The accord was signed on 22 September 1985. On the previous Thursday, the trade-weighted value of the dollar closed at 132.8. Two weeks later, on 4 October, the index stood at 123.9, a decline of roughly 7.5%. Many people note that the value of the dollar continued to fall over the next 27 months, reaching 87.8 at the end of 1987, a total drop of 34% from the pre-accord level. It is unclear, however, how much of this was attributable to the coordinated intervention agreed on during the Accord. In fact, the value of the dollar peaked at 146.8 on 6 March 1985, six months before the Plaza Accord was signed and was already trending rapidly lower by September. At a minimum, it is clear that some portion of the 7.5% drop in the two weeks following the accord was due to the policy intervention. During a period of stress on financial and economic fundamentals, it appears reasonable to use a short-term shift of 10% in the trade-weighted value of the dollar as a plausible stress scenario.

The Achilles Heel Approach

Diversification is arguably the most dependable risk-reduction tool we have. This applies not only to narrowly defined risk in the Frank Knight sense of recurring results from a fairly stable random process, but also to uncertainty surrounding rare events that are not amenable to analysis by distributional methods. In many cases, diversification is the only reasonably effective tool we have when managing tail risk.

Developing scenarios to evaluate as part of the Achilles Heel Approach is obviously an institution-specific exercise. Relevant assessments depend on each bank's business strategy and exposure profile. For this reason, no prescriptive set of stresses is relevant to all institutions. Knowing the institution's own exposures, and being coldly objective about potential negative events and their impacts, are essential. It is also useful to follow events in troubled banks with business profiles similar to one's own. Such events might inspire managers to examine circumstances that have not previously occurred to them. In addition, such analysis can be extremely useful when regulators or the media raise questions. Having answers ready based on institution-specific facts and careful analysis can do a great deal to maintain public confidence that every bank needs so badly.

Liquidity and Asset/Liability Management (ALM)

How a bank is funded is a critical issue since banks inherently rely on the faith and confidence of depositors and borrowers. Even solvent institutions can fail due to illiquidity or instability of funding sources. A bank that is heavily reliant on capital markets' willingness to buy its debt needs to be particularly careful. Given enough time to work out its troubled loan assets, a bank might be able to pay off all of its obligations to depositors and bond holders. Nevertheless, if the market is refusing to roll over a bank's debt paper and/or its depositors have lost confidence and are demanding withdrawals to which they are legally entitled, the institution can fail. Thus, asset/liability management (ALM) is a crucial area in which stress testing is important for any bank.

It is ironic that ALM was not a part of the Basel Capital Accord for much of its life. The assumption seemed to be that banks clearly understood this longstanding form of risk and ensuring it was under control was a supervisory oversight task. In the aftermath of the global financial crisis, liquidity risk was introduced as an explicit element in the Capital Accord. Rules were set for assessing potential net cash outflows, reflecting not only the contractual runoffs of both assets and liabilities, but ratios to apply to elective withdrawals and drawdowns under legally committed credit lines plus increases in non-performing loans. These are a useful starting point for liquidity stress simulations rooted in a bank's own balance sheet configuration. See the chapter on liquidity risk for details.

Credit Risk

A Portfolio Perspective

As with all other forms of risk, it is essential to recognize that credit risk is a portfolio concept. In contrast, traditional credit risk analysis involves detailed micro-underwriting. Such careful analysis of individual borrowers' conditions is essential, and an important function of good bank management is continuous audits to ensure that such analysis is being conducted accurately in accordance with bank guidelines and is being documented properly. In the 1990s, banks started to realize that portfolio analysis is also an important aspect of daily credit decision-making. In essence, a loan with given default characteristics is riskier if the borrower is in an already heavily represented class of borrowers. Some banks have a centralized system that analyzes individual loans in the context of the evolving portfolio and adjust the base terms and rates

based on whether a proposed loan increases or decreases portfolio diversification. Others simply update guidelines periodically, showing more favorable rates for potential borrowers in underrepresented segments of the portfolio.

Clearly, stress testing must be conducted at the portfolio level. Bank-specific historical data, tracking defaults and recoveries against pre-default credit assessments, provide a vital, empirical foundation for simulating how deterioration in credit ratings translates to losses.

Retail Credit Risk

The large number of small credits in a retail portfolio usually makes customized individual analysis impossible. Such borrowers are usually rated based on their FICO scores or some similar rating index of relevant characteristics. When possible, it is useful to track a bank's own distribution of its borrowers' FICO scores against broader industry norms. Stress tests can then take the form of simulating a benchmark shift in industry figures potentially combined with a deterioration of a bank's own borrowers' FICO scores relative to industry norms. This is especially relevant when assessing the influence of a national downturn triggered by an industry that plays a disproportionate role in a bank's own portfolio. Examples include the impact of an oil price collapse on banks in oil-producing regions or a decline in defense spending for a regional bank in an area involved heavily in military contracts. It is also useful to stress the translation of deteriorating FICO scores on default experiences on both a national and bank-specific basis. Finally, attention should be given to recovery rates and timing. An important aspect of this is assessment of how the stress scenario being analyzed would affect the bank's ability to liquidate collateral it holds. See a later section entitled The Second Means of Repayment.

Wholesale Credit Risk

The practice of wholesale credit risk management is different from that for retail credit risk. Generally, expected wholesale credit losses relative to exposure are lower, but the volatility of such losses is higher than for retail credit risk. The typical approach is to estimate a probability of default (PD) and a loss given default (LGD) based on a combination of objective data such as:

- Sales;
- Sales growth;
- Cash flow;
- Fixed interest coverage;
- Balance sheet ratios;
- The size and growth of the firm's market;

and more qualitative factors such as:

- The size and nature of a firm's competition;
- Legal or structural barriers to entry such as patents, trademarks, or network effects;
- The quality of a company's management;
- Potential legal or environmental risks it might face.

Once established, simulated PDs and LGDs provide quantitative handles to drive stress scenarios. Historical data on the cyclical behavior of these variables is helpful as a basis for determining extreme but plausible adverse movements. Once again, however, special attention must be paid to local factors that intensify the impact of a given scenario based on an individual bank's portfolio concentrations.

Derivative Credit Risk

During the 1980s, a few money-center banks developed derivative trading operations to help customers manage their interest rate and foreign exchange risk. For such intangible sources of risk, this provided management tools long available for hedging exposure to prices of agricultural and industrial commodities such as wheat, pork bellies, crude oil, base metals such as aluminum, copper, nickel and zinc, and precious metals such as gold and silver. As these markets grew in size and importance, they also represented in increasingly important source of credit risk to banks involved.

Derivative credit risk has long presented a cultural challenge to bank staff accustomed to a historical cost-accounting world. If a bank makes a long-term, fixed-rate loan and interest rates fall, the market value of that loan rises. In a historical cost-accounting world, however, the loan continues to appear on the books at its original face amount. As a result, credit officers became accustomed to viewing the principal amount of a loan as constant except for explicit repayments or additional credit extensions.

Derivative market-making is accounted for on a mark-to-market basis; credit exposure of a transaction such as a term interest rate swap (equal to the net difference in the market value of the pay and receive legs of the trade) fluctuates over time. Since most such trades are initially executed at prevailing market prices, credit exposure at origination is very small (representing only the net value of the bid-offer spread charged by the bank to the client requesting the transaction). What needs to be considered, however, is an estimate of the profile of potential future exposure of the transaction over the full life of the contract. The problem becomes more complex when parties agree to net multiple contracts made at different times, and often involve different underlying variables in a single, consolidated transaction for legal purposes.

This is not the place for a fuller discussion of counterparty credit risk, as it is commonly called. There is one important point to recognize, however, in the context of credit risk stress testing. The same variables that drive the value of a derivative transaction also influence the credit quality of a counterparty. Rising oil prices, for example, have a negative influence on the credit quality of an airline but a positive impact on that of an oil-exploration company. Falling prices have the reverse effect. The widely discussed issue of wrong-way risk arises when significant moves in one or more underlying rates or prices weaken the credit quality of a counterparty while simultaneously increasing the value of trades with that entity. A classic example of this is a currency swap with a foreign commercial bank during which the trade value increases if the counterparty's home currency declines in value. Major currency depreciations usually imply a domestic economic crisis that threatens a country's banks.

The upshot for stress testing within the Achilles Heel approach is to be aware of instances in which obvious wrong-way risk occurs in the derivative book, including stress changes in the relevant variables in the direction that increase both exposure and default probabilities of the counterparties in question.

The Second Means of Repayment

It was once common when discussing a proposed credit extension for bank credit officers to ask, "What is the second means of repayment?" It seems that this question is not asked as much today as it once was, but it should be. Normally, a loan to an individual or business is made on the basis of the borrower's expected future earning capacity. For individuals, this depends importantly on education, experience, and past earnings, and current net worth. For a business, it depends on competitive position, quality of management, strength of the balance sheet, and net cash flow experience, among others. Responsible lending should involve a realistic expectation that the borrower will be able to service the loan to maturity. This, however, should not be the only consideration. Obviously, assessments of borrower quality are never perfect, and unexpected things can go wrong. The first line of defense when this happens is a second means of repayment. When funding a home or commercial building, the cash flow of the borrower is backed up by acquiring a line on the property being financed. This serves as the second means of repayment. If a borrower fails to repay, the lender can seize the property and sell it to recoup some, or hopefully, all of the unpaid balance of the loan.

During a credit stress scenario, these second means of repayment are bound to come into the analysis. How realistic is it that the collateral can be sold for the full outstanding balance? How long would this take? Answers to these questions often depend on whether the scenario being analyzed is a bank-specific situation or a broader economic crisis. A general crisis might affect the timing and potential proceeds of required collateral liquidations. Assessing the cyclical behavior of asset prices for the type of collateral being held is an important component of assessing the full impact of weakening credit conditions.

The Interaction of Credit and Liquidity Risk

The timing of collateral liquidations also affects potential liquidity risk. Close coordination between analysis of credit risk and liquidity risk is essential. During the global financial crisis, many banks felt forced to bring subprime securities they had sold, and that had fallen in value due to credit problems, back onto their balance sheets. This manifested itself as a liquidity crisis since these banks had to fund assets they did not expect to be holding, but its root cause was the rising delinquency and default rates on the underlying mortgages in the securities.

The lesson is that credit-related stress tests need to look not only at the direct losses due to deteriorating credit quality and higher defaults but also at potential adverse liquidity events that might result from declining faith in a bank's management.

Market Risk

Short-Term Volatility Versus Structural Shifts

The most widely discussed, and frequently deplored, measure of market risk is VaR. This is not the place to discuss its merits and limitations. VaR is an important technique to measure, compare, and aggregate the potential impact of short-term market volatility. For this important but limited purpose, it serves well if implemented properly. VaR is not, however, an effective measure of exposure to sudden and significant structural shifts in a firm's surrounding economic and financial environments. This is the role of stress testing.

Within the Achilles Heel Approach, the task is to identify market scenarios that present significant threats of major loss. Greek sensitivities are of limited value here because they were designed to measure sensitivity to small changes, usually standardized on one basis point or a move of 1/100th of 1 percent in the variable in question. In many cases, these sensitivities are non-linear; they might increase of decrease as markets move away from current levels.

One exercise is to shock individual variables by large amounts such as four, eight, and twelve standard deviations and record their influence on P&L. This provides useful hints to hidden risk in the tails of the distribution, especially if traders booked large volumes of far out-of-the-money options on the underlying variables they trade. The value of such trades does not move much in response to two or three standard deviation daily moves, and as a result, they do not influence daily VaR estimate much. They might, however, have a huge impact if such options move close to the money.

Clearly, single-variable stress scenarios are not exhaustive. Some large losses might arise only from logically consistent combined moves in several variables. This opens the search to a vast number of variable combinations, some of which are logically inconsistent. If a bank's market VaR system uses Monte Carlo simulation, this can provide additional hints to scenarios worth investigating. The idea is to evaluate extreme scenarios that produce losses larger than the (assume 99% confidence) VaR estimate. If a Monte Carlo simulation produces 5000 draws, there are 50 such scenarios. Rather than just throwing these away, it can be fruitful to preserve and display them in rank order of the losses they produce. Since the simulation imposes historical correlations, there should be no implausibly inconsistent combinations of market moves.

Some presentation designed to summarize the market configuration of each is important. Shifts from current market values need to be normalized by presenting them as standard deviations. This allows relevant comparisons of what constitutes a statistically large or small move across variables of different absolute numerical values. When a scenario is driven by one or more similar market moves larger than two standard deviations, it is useful to construct scenarios with even larger moves, say four and eight standard deviations, and assess what losses are produced. The particularly worrisome case is when increases in the losses are super-linear, that is, doubling the size of the market shift results in more than a doubling in the loss.

There also might be cases in which large loss scenarios do not appear to exhibit obviously outsized market moves. These might be worth investigating to see which books, and even which trades, contribute most to losses greater than VaR. This might reveal complex trades, perhaps involving large digital options, creating unacceptable risk.

Pins in the Option Book

One source of lurking risk in an options book is what is called pins, notional volumes of options on a given underlying variable that mature on a given date and have the same strike price. If the market value of the underlying is near the strike price of an option close to its maturity date, the option cannot be hedged effectively.[31] When these pins are diversified, so no single combination of strike values and maturity dates represents a major concentration, this is just part of the normal risk of running an option book. A sufficiently large concentration, however, might represent a lurking threat. The threat might not materialize, but if it does, and the market takes a large jump just before the maturity date, the losses could be unacceptable.

This is also a case in which standard VaR methods might not reveal a problem. As long as the position is delta neutral on the date of the VaR simulation, its value is largely insensitive to current rate movements. Gamma balloons only when and if the market approaches the maturity date with a spot price close to the strike. Hence, the best approach is to produce a tabulation of notional amounts by combinations of option strikes and maturity dates ranked by the total notional amount for each. Since the problem is likely to arise only occasionally, it makes sense to scan these total notional amounts and produce only a warning to risk managers if some notional is above a threshold.

[31] This is sometimes called *getting gamma-ed*. Gamma becomes infinite as delta swings wildly when the market price fluctuates around the strike price. The best one can do is hedge half the notional amount and hope the market does not take a big jump in either direction.