CHAPTER 14

Stress Testing

This is one of those cases in which the imagination is baffled by the facts.

—*Winston Churchill*

The main purpose of value-at-risk (VAR) measures is to quantify potential losses under "normal" market conditions, where *normal* is defined by the confidence level, typically 99 percent. In principle, increasing the confidence level could uncover progressively larger but less likely losses. In practice, VAR measures based on recent historical data can fail to identify extreme unusual situations that could cause severe losses. This is why VAR methods should be supplemented by a regular program of stress testing. Stress testing is a *nonstatistical* risk measure because it is not associated with a probability statement like VAR.

Stress testing is indeed required by the Basel Committee as one of seven conditions to be satisfied to use internal models. It is also endorsed by the Derivatives Policy Group and by the Group of Thirty. *Stress testing* can be described as a process to identify and manage situations that could cause extraordinary losses. This can be made with a set of tools, including (1) scenario analysis; (2) stressing models, volatilities, and correlations; and (3) policy responses.

Scenario analysis consists of evaluating the portfolio under various extreme but probable states of the world. Typically, these involve large movements in key variables, which requires the application of full-valuation methods. The earliest application of stress tests consisted of sequentially moving key variables by a large amount. This is also called *sensitivity testing.*

This approach, however, ignores correlations, which are crucial to large-scale risk measurement. More generally, scenarios provide a description of the joint movements in financial variables. Scenarios can be *historical*, that is, drawn from historical events, or *prospective*, that is, drawn from plausible economic and political developments. Prospective scenarios are also called *hypothetical*. More recently, the industry has realized that the identification of scenarios should be driven by the particular portfolio at hand. Scenarios that matter are those that could generate extreme losses.

Stress tests are used primarily for understanding the risk profile of a firm. Increasingly, however, they are also used, in conjunction with VAR, for *capital allocation*. Whenever the stress tests reveal some weakness, management must take steps to manage the identified risks. One solution could be to set aside enough capital to absorb potential large losses. Too often, however, this amount will be cripplingly large, reducing the return on capital. Alternatively, positions can be altered to reduce the exposure. The goal is to ensure that the institution can ride out the turmoil.

Section 14.1 discusses why stress testing is required at all. In theory, extreme losses could be identified by increasing the confidence level of VAR measures. Section 14.2 shows how to use scenarios to generate portfolio losses. Sections 14.3 and 14.4 then examine scenario analysis in great detail. This is no easy matter owing to the large number of risk factors that global financial institutions are exposed to. Next, Section 14.5 turns to stress testing of models and parameters. Section 14.6 then discusses management actions that can be taken in response to stress-test results.

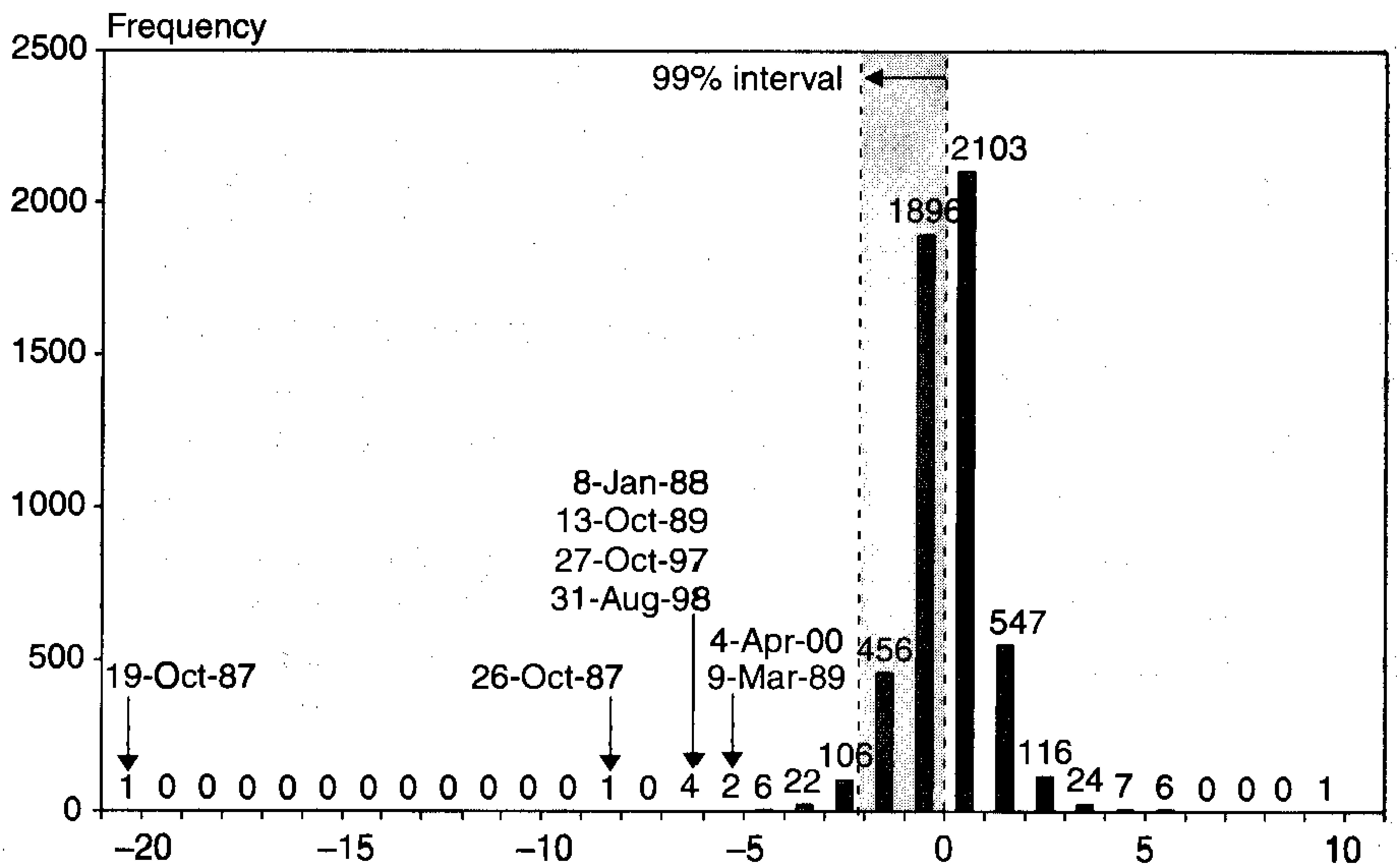
14.1 WHY STRESS TESTING?

Compared with VAR methods, stress testing appears refreshingly simple and intuitive. The first step is scenario analysis, which examines the effect of simulated large movements in key financial variables on the portfolio. Such scenarios have the advantage of linking the loss to a specific event, which is more intuitive to many managers than a draw from a statistical distribution. Owing to its simplicity, this approach actually predates VAR methods.

To understand the need for scenario analysis, consider, for instance, the stock market crash of October 19, 1987. Figure 14-1 displays the distribution of U.S. daily stock returns using data from 1984 to 2004. Over this period, the average volatility was about 1 percent per day. On Monday, October 19, the Standard & Poor's (S&P) Index lost 20 percent of its value.

FIGURE 14-1

Distribution of daily U.S. stock returns, 1984–2004.



Even if there was some time variation in volatility, this 20 standard deviation event was so far away in the tail that it never have should happened under a normal distribution. The figure also shows that a standard 99 percent VAR interval would have totally missed the magnitude of the actual loss.

More generally, Bookstaber (1997) says that there is

> . . . a general rule of thumb that every financial market experiences one or more daily price moves of 4 standard deviations or more each year. And in any year, there is usually at least one market that has a daily move that is greater than 10 standard deviations.

These observations, however, are an indictment of the distributional assumption rather than VAR itself. In theory, one could fit a better distribution to the data and vary the confidence level so as to cover more and more of the left-tail events. This can be accomplished with historical simulations or, if a smoother distribution is required, through the use of extreme-value theory (EVT). In other words, the generation of a scenario is akin to a particular point in the distribution drawn from historical data. So what is special about stress testing?

The goal of stress testing is to identify unusual scenarios that would not occur under standard VAR models. Berkowitz (2000) classifies these scenarios into the following categories:

1. Simulating shocks that have never occurred or are more likely to occur than historical observation suggests.
2. Simulating shocks that reflect permanent structural breaks or temporarily changed statistical patterns.

Thus one reason to stress test is that VAR measures typically use recent historical data. Stress testing, in contrast, considers situations that are absent from historical data or not well represented but nonetheless likely. Alternatively, stress tests are useful to identify states of the world where historical relationships break down, either temporarily or permanently.

A direct example of the need for stress testing is Niederhoffer's belief, described in Box 14-1, that the market would not drop by more

BOX 14-1

VICTOR NIEDERHOFFER: THE EDUCATION OF A SPECULATOR

Victor Niederhoffer outlined his investment philosphy in his book, *Education of a Speculator*, which quickly became a best-seller. An eccentric and brilliant investor, he was a legend of the hedge-fund business. Indeed, he had compiled an outstanding track record—a 32 percent compound annual return since 1982.

Niederhoffer's mission was to "apply science" to the market. Although he had a Ph.D. in business from the University of Chicago, he did not believe in efficient markets and traded on statistical anomalies. He believed, for instance, that the market would never drop by more than 5 percent in a single day. Putting this theory into practice, Niederhoffer sold naked out-of-the-money puts on stock index futures. When the stock market plummeted by 7 percent on October 27, 1997, he was unable to meet margin calls for some $50 million. His brokers liquidated the positions, wiping out his funds.

Apparently, his views were narrowly based on recent history. It is true that the worst loss had been 3 percent in the previous 5-year period. Larger losses do occur once in a while, however. Most notably, the market lost 20 percent on October 19, 1987.

than 5 percent in a day. Indeed, this never happened from 1990 to October 1997. This does not mean that a loss of this magnitude can never happen.

Another illustration is a breakup of a fixed exchange-rate system. In the summer of 1992, it would have been useful to assess potential vulnerabilities in the European monetary system. Indeed, in September 1992, the Italian lira and the British pound abandoned their fixed exchange rates, which led to a disastrous fall in their value. Historical volatilities based on the previous 2 years would have completely missed the possibility of a devaluation. Thus scenario analysis forces risk managers to consider events they otherwise might ignore.

14.2 PRINCIPLES OF SCENARIO ANALYSIS

We now consider the implementation of scenario analysis. Define s as a selected scenario. This is constructed as a set of changes in risk factors $\Delta f_{k,s}$ for various k. Based on the new hypothetical risk-factor values, $f_{k,0} + \Delta f_{k,s}$, all the securities in the portfolio are revalued, preferably using a full-valuation method if the portfolio has nonlinear components. The portfolio return then is derived from changes in the portfolio value V, which depends on positions and risk factors, that is,

$$R_{p,s} = V_s - V_0 = V(f_{1,0}+\Delta f_{1,s} \ldots, f_{K,0}+\Delta f_{K,s}) - V(f_{1,0} \ldots, f_{K,0}) \qquad (14.1)$$

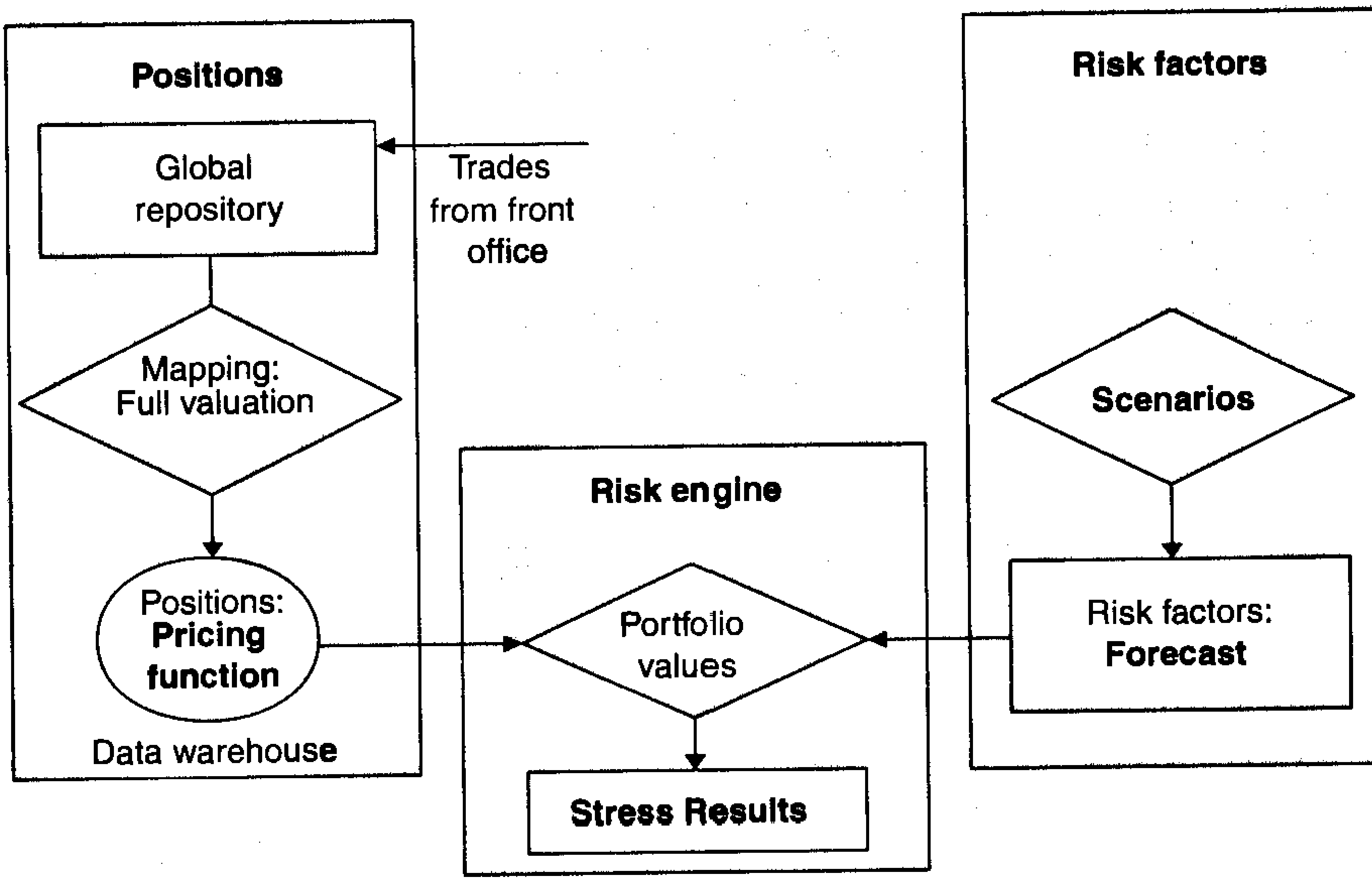
Note that this is a special case of the historical simulation method. Thus scenario analysis can be implemented easily once a VAR system is in place. Figure 14-2 details the steps involved in this approach. The question is how to generate realistic scenarios.

14.2.1 Portfolio- versus Event-Driven

The generation of scenarios can be either *event-driven* or *portfolio-driven*. In the first case, the scenario is formulated from plausible events that generate movements in the risk factors. In the second case, risk vulnerabilities in the current portfolio are identified first that translate into adverse movements in risk factors. These lead to the generation of scenarios. For instance, institutions that invest in long-term bonds funded by short-term debt are vulnerable to upward movements in the yield curve. It is therefore essential to consider scenarios that reflect such changes.

FIGURE 14-2

Scenario-analysis approach.



14.3 GENERATING UNIDIMENSIONAL SCENARIOS

14.3.1 Sensitivity Tests

The traditional approach to scenario analysis focuses on one variable at a time. For instance, the Derivatives Policy Group (DPG) provides specific guidelines for scenarios. It recommends focusing on a set of specific movements:

1. Parallel yield-curve shifting by ±100 basis points
2. Yield-curve twisting by ±25 basis points
3. Each of the four combinations or yield-curve shifts and twist
4. Implied volatilities changing by ±20 percent of current values
5. Equity index values changing by ±10 percent
6. Currencies moving by ±6 percent for major currencies and ±20 percent for others
7. Swap spreads changing by ±20 basis points

While these movements are quite large for a daily horizon, the DPG's goal was to provide comparable results across institutions in order to assess zones of vulnerabilities. By specifying consistent guidelines, it tried to ensure that all the models used by brokers "possess broadly similar performance."

These scenarios shock risk factors generally one at a time. The loss in value, scaled by the size of the factor movement, is a *sensitivity* measure. These tests can be run relatively quickly and are intuitive.

This approach is appropriate in situations where the portfolio depends primarily on one source of risk. The Office of Thrift Supervision (OTS), for instance, uses scenario analysis to assess the market risk of savings and loans associations (S&Ls).[1] The OTS requires institutions to estimate what would happen to their economic value under parallel shifts in the yield curve varying from -400 to $+400$ basis points. The OTS recently has imposed a risk-based capital requirement linked directly to the interest-rate exposure of supervised institutions.

14.3.2 An Example: The SPAN System

The standard portfolio analysis of risk (SPAN) system is a good example of a scenario-based method for measuring portfolio risk. SPAN was introduced in 1988 by the Chicago Mercantile Exchange (CME) to calculate collateral requirements on the basis of overall *portfolio risk* as opposed to position by position. Since its inception, SPAN had become widely used by futures and options exchanges as a mechanism to set margin requirements.

The objective of the SPAN system is to identify movements in portfolio values under a series of scenarios. SPAN then searches for the largest loss that a portfolio may suffer and sets the margin at that level. The SPAN system only aggregates futures and options on the same underlying instrument. It uses full-valuation methods, which is important given that portfolios usually include options.

Consider a portfolio of futures and options on futures involving the dollar/euro exchange rate. SPAN scans the portfolio value over a range of prices and volatilities. These ranges are selected so that they cover a fixed percentage of losses, for example, 99 percent. Contracts have a notional of 125,000 euros. Assume a current price of \$1.05/euro and a

[1] The OTS is a U.S. agency created in 1989 to supervise S&Ls.

12 percent annual volatility. The value range for the contract is set at the daily VAR, that is,

$$\text{Price range} = 2.33 \times 12 \text{ percent } \sqrt{252}) \times (\text{euro}125{,}000 \times 1.05\$/\text{euro}) = \$2310$$

This is indeed close to the daily margin for an outright futures position, which is around $2500 for this contract. This corresponds to a price range of $0.0176 around the current price of $1.05 per euro. Next, the volatility range is set at 1 percent.

Table 14-1 presents an example of scenario generation. We select scenarios starting from the initial rate plus and minus three equal steps that cover the price range, as well as an up-and-down move for the volatility range. In addition, to provide protection for short positions in deep

TABLE 14-1

Example of SPAN Scenario System

	Scenario			Gain/Loss	
Number	Fraction Considered for P&L	Price Scan Expressed in Range	Volatility Scan Expressed in Range	Long Call	Long Futures
1	100%	0	1	$198	$0
2	100%	0	−1	−$188	$0
3	100%	+1/3	1	$395	$767
4	100%	+1/3	−1	−$21	$767
5	100%	−1/3	1	$23	−$767
6	100%	−1/3	−1	−$332	−$767
7	100%	+2/3	1	$615	$1,533
8	100%	+2/3	−1	$170	$1,533
9	100%	−2/3	1	−$132	−$1,533
10	100%	−2/3	−1	−$455	−$1,533
11	100%	+1	1	$858	$2,300
12	100%	+1	−1	$388	$2,300
13	100%	−1	1	−$268	**−$2,300**
14	100%	−1	−1	**−$559**	**−$2,300**
15	35%	+2	0	$517	$1,610
16	35%	−2	0	−$240	−$1,610
Ranges:		$0.0176	1%		

Note: Euro-FX futures and option on futures with notional of 125,000 euros, spot of 1.05$/euro, strike of $1.10, 12 percent annual volatility, 90 days to maturity, and interest rate of 5 percent.

out-of-the-money options, two scenarios are added with extreme price movements, defined as double the maximum range. Because such price changes are rare, the margin required is 35 percent of the resulting loss.

Next, the value of each option and futures position is calculated under each scenario, using full valuation. The table presents calculations for two positions only, long one call and long one futures, under each of the 16 scenarios. The result of the computation for each risk scenario is called a *risk-array value*. The set of risk array values for the position is called the *risk array*.

The long-call position would suffer the most under scenario 14, with a large downward move in the futures accompanied by a drop in the volatility. Similarly, the worst loss for a long-futures position also occurs under a large downward move. This analysis is repeated for all options and futures in the portfolio and aggregated across all positions. Finally, the margin is set to the worst portfolio loss under all scenarios.

The SPAN system is a scenario-based approach with full valuation. Its systematic scanning approach is feasible because it considers only two risk factors. The number of combinations, however, soon would become unmanageable for a greater number of factors. This is perhaps the greatest hurdle to systematic scenario analysis.

Another drawback is that the approach essentially places the same probability on most of the scenarios, which ignores correlations between risk factors. And as we have seen, correlations are an essential component of portfolio risk.

14.4 MULTIDIMENSIONAL SCENARIO ANALYSIS

Unidimensional scenarios provide an intuitive understanding of the effect of movements in key variables. The problem is that these scenarios do not account for correlations. This is where multidimensional scenarios are so valuable. The approach consists of (1) positing a state of the world and (2) inferring movements in market variables.

14.4.1 Prospective Scenarios

Prospective scenarios represent *hypothetical* one-off surprises that are analyzed in terms of their repercussions on financial markets. One might want to examine, for instance, the effect of an earthquake in Tokyo, of

Korean reunification, of a war in an oil-producing region, or of a major sovereign default. The definition of scenarios should be done with input from top managers, who are most familiar with the firm's business and extreme events that may affect it.

Let us go back to the example of a scenario analysis for a potential breakup in the exchange-rate mechanism (ERM), evaluated as of summer 1992. The risk manager could hypothesize a 20 percent fall in the value of the Italian lira against the German mark. One could surmise further that if the Italian central bank let the lira float, short-term rates likewise could drop, and the stock market would rally. Beyond the effect on Italian interest rates and equity prices, however, it may not be obvious to come up with plausible movements for other financial variables. The problem is that the portfolio may have large exposures to these other risk factors that remain hidden. Thus this type of subjective scenario analysis is not well suited to large, complex portfolios.

Factor Push Method

Some implementations of stress testing try to account for multidimensionality using a rough two-step procedure. First, push up and down all risk-factor variables individually by, say, $\alpha = 2.33$ standard deviations, and then compute the changes to the portfolio. Second, evaluate a worst-case scenario, where all variables are pushed in the direction that creates the worst loss. For instance, variable 1 could be pushed up by $\alpha\sigma_1$, whereas variable 2 could be pushed down by $\alpha\sigma_2$, and so on.

This approach is very conservative but completely ignores correlations. If variables 1 and 2 are positively correlated, it makes little sense to consider moves in opposite directions. Further, looking at extreme movements may not be appropriate. Some positions such as combinations of long positions in options will lose the most money if the underlying variables do not move at all.

Conditional Scenario Method

There is a systematic method, however, to incorporate correlations across all variables consistently. Let us represent the "key" variables that are subject to some extreme movements as R^*. The other variables are simply represented by R. The usual approach to stress testing focuses solely on R^*, setting the other values to zero. Simplifying Equation (14.1) to a linear movement, what we call the *narrow stress loss* (NSL), is $\Sigma_i w_i^* R_i^*$.

To account for multidimensionality, we first regress the R variables on the controlled R^* variables, obtaining the conditional forecast from

$$R_j = \alpha_j + \sum_i \beta_{j,i} \mathrm{R}_i^* + \epsilon_j = E(R_j \mid R^*) + \epsilon_j \tag{14.2}$$

This allows us to predict other variables conditional on movements in key variables using information in the covariance matrix. We construct a *predicted stress loss* (PSL) as $\Sigma_i w_i^* R_i^* + \Sigma_j w_j E[R_j | R^*]$. This can be compared with the realized *actual stress loss* (ASL), which is $\Sigma_i w_i^* R_i^* + \Sigma_j w_j R_j$.

Kupiec (1998) illustrates this method with episodes of large moves from 1993 to 1998 using a $1 million portfolio invested in global equity, bond, and currency markets. Table 14-2 presents typical results. For the Philippine peso, for instance, the event was a devaluation, which was a 5.50 standard deviation move. The notional value of the position on this risk factor was $40,700, which led to a narrow stress loss (NSL) of $3070. This number, however, fails to account for other markets, such as Philippine equities, that moved in the opposite direction. Taking this correlation into account, the predicted stress loss (PSL) is much smaller than the NSL, even close to zero. In some other cases, PSL is much worse than NSL.

Interestingly, the table shows that, in all cases, the PSL produces results that are much closer to the ASL than in the simple, narrow model that zeroes out nonkey variables. The conclusion is that the covariance matrix, which is at the core of conditional normal VAR modeling, does provide useful information for stress-testing analysis.

TABLE 14-2

Comparison of Forecast Losses on a $1 Million Portfolio

				Stress Loss		
Key Variable	**Period**	**Event Size (σ)**	**Position on Key Variable**	**Narrow**	**Predicted**	**Actual**
Philippine peso	11 Jul 1997	−5.50	$40,700	−$3,070	$43	$190
Japanese equities	23 Jan 1995	−5.23	$72,120	−$2,700	−$7,730	−$11,700
U.S.equities	27 Oct 1997	−4.93	$136,480	−$6,650	−$5,330	−$5,420
U.K.bonds	29 Dec 1994	−4.84	$122,910	−$2,640	−$3,550	−$3,030
U.S.bonds	20 Feb 1996	4.86	$122,970	−$1,210	−$7,070	−$10,380

The main drawback of this conditional scenario method is that it relies on correlations derived from the entire sample period. This includes normal periods and *hectic* periods. Should correlations change systematically across these periods, however, the results will differ. For portfolios with long positions only, increases in correlations will increase the worst loss. Banks often reexamine the stress-test results using correlations estimated over hectic periods.[2]

As an example, the correlation between stocks and bonds typically is positive in normal times. In times of stress, however, this correlation often turns negative. When equity markets drop sharply, the demand for Treasury bonds typically increases, reflecting a flight to quality. At the short end of the yield curve, this effect usually is reinforced when the central bank injects liquidity into the financial system, pushing down short-term rates. Thus government bonds are good diversifiers for stocks in times of stress.

A related, useful approach relies on the output from a VAR Monte Carlo analysis or historical simulation. The risk manager could examine the worst losses from the simulation, which specifically accounts for correlations. Such analysis provides valuable insight into the vulnerabilities of a particular portfolio and could guide the construction of scenarios.

14.4.2 Historical Scenarios

Alternatively, scenario analysis can examine historical data to provide examples of joint movements in financial variables. The role of the risk manager is to identify scenarios, such as those listed in Table 14-3, that may be outside the VAR window. Each of these scenarios then will yield a set of joint movements in financial variables that automatically takes correlations into account.

Table 14-3 displays a list of scenarios, both historical and prospective, used by a large group of banks. The largest category of stress tests focuses on interest rates. Historical scenarios include the 1994 bond market crash, the 1997 Asian currency crisis, the LTCM and Russia crises, and the terrorist attack on the World Trade Center, all of which led to global interest-rate shocks. Also common are stress tests involving equities, currencies, commodities, and credit.

[2] See, for instance, Kim and Finger (2000). Using volatility measures to sort the sample, however, creates potential biases in correlation estimates, as explained in Boyer et al. (1999). Another issue is that because hectic periods cover, by definition, fewer observations than the entire sample, correlations are estimated less precisely.

TABLE 14-3

Description of Scenarios (Number of Tests Reported by 64 Institutions Surveyed)

Category	Historical	Prospective
Interest rates (173)	1994: Bond market crash (18)	U.S. economic crash (21)
	1997: Asian currency crisis (22)	Global economic crash (11)
	1998: LTCM, Russia crises (26)	Inflation pickup (8)
	2001: World Trade Center (30)	
Equities (86)	1994: Black Monday (83)	Geopolitical unrest (5)
	1997: Asian currency crisis (22)	Terrorist attack (5)
	2000: IT bubble bursting (6)	
	2001: World Trade Center (30)	
Currencies (56)	1992: EMS crisis (8)	Pegged-currencies breakdown (7)
	1997: Asian currency crisis (22)	
	1998: LTCM, Russia crises (26)	
Commodities (22)		Oil-price jump (11)
		Unrest in Middle East (5)
Credit (104)	1997: Asian currency crisis (22)	Emerging market crash (10)
	1998: LTCM, Russia crises (26)	Euro area economic crash (7)
	2001: World Trade Center (30)	Global economic crash (6)
Property (19)		Various episodes

Source: Basel Committee (2005b) survey of financial institutions.

TABLE 14-4

October 1987 Market Crash: Change in Market Variables

	Equities				Fixed Income			Currencies	
Date	U.S., S&P	Japan, Nikkei	U.K., FTSE	Germany, DAX	Fed Funds	3-Month T-Bill	30-year T-Bond	Yen/ $	DM/ $
Oct 19	−20.4%	−2.4%	−9.1%	−9.4%	−0.14 bp	−0.52 bp	0.01 bp	−0.2%	1.3%
Oct 20	5.3%	−14.9%	−11.4%	−1.4%	−0.77 bp	−0.62 bp	−0.76 bp	1.5%	1.7%

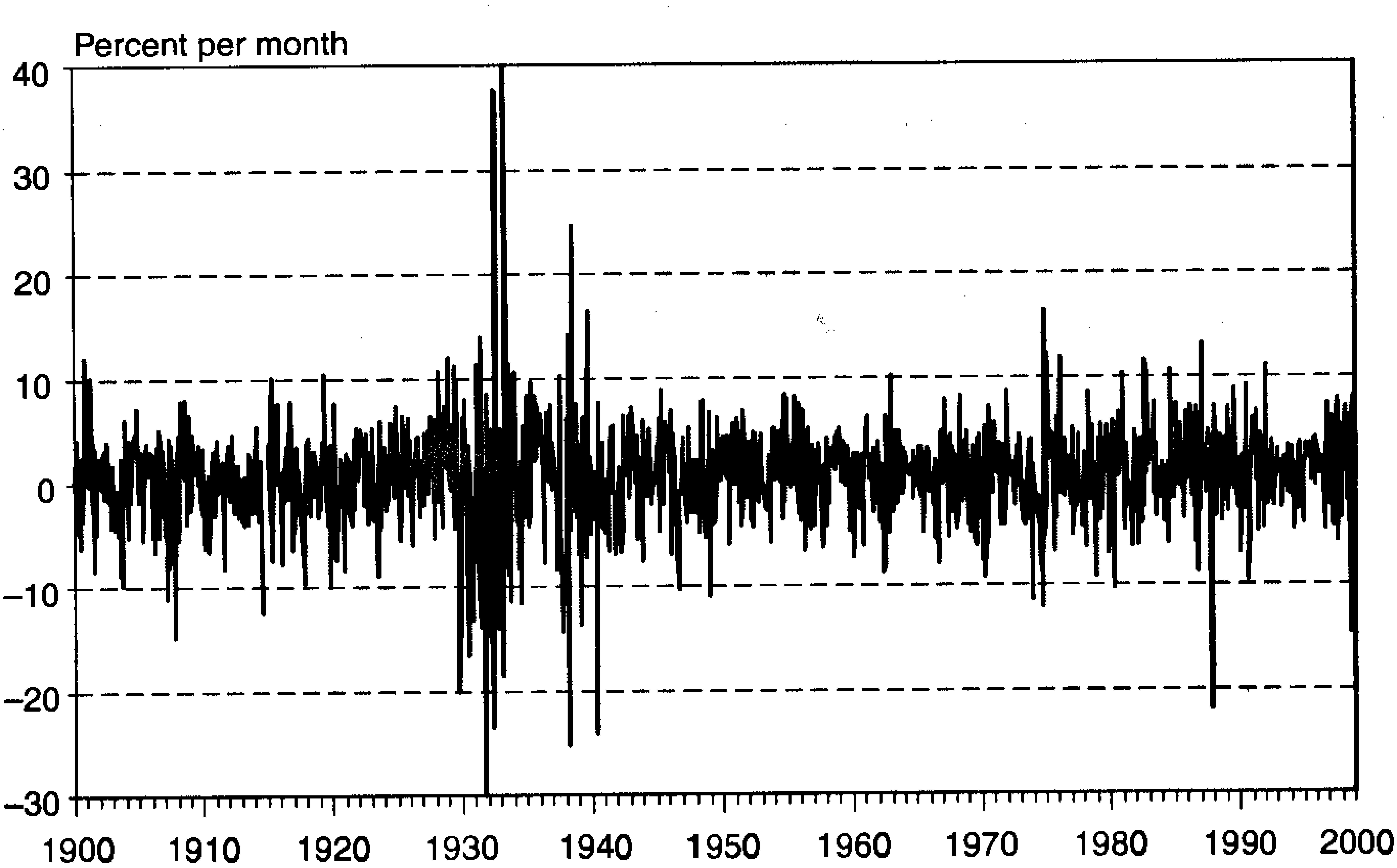
As an example, Table 14-4 details the vector of movements in risk factors for the stock market crash of October 1987. On Monday, October 19, the S&P Index lost more than 20 percent of its value. The sheer size of this movement had dramatic effects on other prices. The next day, the Nikkei fell by 15 percent. In order to forestall the failure of financial institutions, the Federal Reserve injected liquidity into the financial system, pushing short-term interest rates down by 91 basis points in 2 days. This chaos, however, apparently had little effect on currency markets. Overall, there was little diversification benefit across global stocks. As is often the case, however, long positions in bonds helped to cushion the equity losses.

Historical scenarios are quite useful to measure joint movements in financial variables. Their drawback, from the risk manager's perspective, is the limited number of extreme events. Whenever possible, very long histories should be considered. These give a long-term perspective that may be absent from recent data.

The October 1987 crash, for instance, is viewed widely as an extremely unusual event. Perhaps so if one looks only at recent data. On the other hand, a different perspective is offered by Figure 14-3, which reports monthly returns on U.S. stocks since the beginning of the century. The figure shows that if one goes back sufficiently in time, there have been many other instances of losses exceeding 20 percent over a month and that the recent past has not been unusually volatile.

FIGURE 14-3

Monthly U.S. stock returns, 1900–2000.



Likewise, events such as sovereign defaults are extremely rare. Recently, however, Russia defaulted on its domestic debt and Ecuador on its external debt. One would need to go back to the 1930s to encounter sovereign defaults on external debt. Defaults should be expected to occur, however. Otherwise, there would be no rationale for the wide yield spread observed on some sovereign debt.

As should be clear by now, the generation of relevant scenarios is a time-consuming process that requires quantitative skills as well as a good economic understanding of the factors driving financial shocks.

14.5 STRESS-TESTING MODELS AND PARAMETERS

Going back to the schematics of scenario analysis in Figure 14-2, the risk manager should critically examine all the steps in the generation of risk measures. A stress-testing process should consider not only movements in market variables but also the other components of the risk-management system, that is, the securities valuation models and the risk engine.

A distinction usually is made between *sensitivity analysis*, which examines the effect of changing the functional form of the model, and stress-testing *model parameters*, which are inputs into the model.

Let us consider sensitivity analysis first. Derivative securities can be priced using a variety of models. Interest-rate derivatives, for instance, can be valued using one- or multiple-factor models, with parameters typically estimated from historical data. Mortgage-backed securities (MBS) must in addition model prepayments. All these assumptions introduce insidious risks. Current model prices may fit the current market data but may not provide a good approximation under large movements in key variables. The Askin story (see Box 1-2) provides an example of an MBS portfolio that was thought to be hedged but led to large losses under severe interest-rate shocks. Pricing models may fail in changing environments.

Likewise, simplifications in the risk measurement system also may create hidden risks. For example, bond mapping replaces a continuous yield curve by a finite number of risk factors. If there is insufficient *granularity*, or detail, in the choice of the risk factors, the portfolio could be exposed to losses that are not measured by the risk management system.

Turning to *model parameters*, pricing and risk management systems rely on particular input data, such as a set of volatilities and correlations. Correlations, however, may deviate sharply from historical averages in times of stress. A key issue is whether a traditional variance-covariance-based VAR system provides adequate measures of risk when historical correlation patterns break down.

To some extent, this question can be answered directly by scenario analysis based on historical data. It is also informative, however, to check how sensitive a VAR number is to changes in the risk measures. As will be seen in the example of LTCM, this is especially important if the same period is used to measure risk and optimize the portfolio.

As an example, consider a covariance matrix measured with recent data that shows a high correlation between two series. The risk manager, however, cannot believe that this high correlation will remain in the future and could alter the covariance matrix toward values that are considered reasonable.[3] The stress test then compares the new VAR measure with the original one.

In all these cases, there is no simple rule to follow for stress testing. Rather, the risk manager must be aware of limitations, assumptions, and measurement errors in the system. Stress testing can be described as the art of checking whether the risk forecasts are robust to changes in the structure of the system.

14.6 POLICY RESPONSES

The main function of stress tests is to communicate and understand the risks that the institution is exposed to. The next question is what to do if the size of the stress loss appears unacceptably large. This is the crux of the issue with stress testing. Too often stress-testing results are ignored because they create large losses that are dismissed as irrelevant.

Indeed, institutions do not need to withstand every single state of the world. Central banks, in particular, are supposed to provide protection against systemic banking crises. Likewise, there is little point in trying to protect against a widespread nuclear war.

Relevant scenarios, however, require careful planning. One response is for the institution to set aside enough *economic capital* to absorb the worst losses revealed by stress tests. In many cases, however, this amount

[3] This is no easy matter because changing some entries manually could produce an inconsistent matrix, which is not positive semidefinite (see Chapter 8). A positive-semidefinite matrix ensures that for any value of the vector w, the product $w'\Sigma w$ will never be negative. Rebonato and Jäckel (1999) describe methods to ensure that this will be the case.

BOX 14-2

STRESS TESTING'S BENEFITS

A risk manager at a U.S. investment bank recalls that in December 1997 stress tests showed that the firm could be put in jeopardy should Russia default on its debt. The firm reduced its exposure to Russia in part through the purchase of credit derivatives.

The bank was able to ride the turmoil but still suffered losses owing to the fact that some counterparties defaulted on the credit protection. This illustrates that stress testing generally is useful but still is a subjective exercise that cannot possibly cover all contingencies.

may be much too large, which will make it uneconomical. A number of other actions can be considered, though. The institution could

- Purchase protection or insurance for the events in question (although this may transform market risk into counterparty risk)
- Modify the portfolio to decrease the impact of a particular event through exposure reduction or diversification across assets
- Restructure the business or product mix for better diversification
- Develop a plan for a corrective course of action should a particular scenario start to unfold
- Prepare sources of alternative funding should the portfolio liquidity suffer

Risk limits often are based on stress-test results, in addition to the usual VAR or notional-amount limits. This plan of action should help to ensure that the institution will survive this scenario, as shown in Box 14-2.

14.7 CONCLUSIONS

While VAR focuses on the dispersion of revenues, stress testing instead examines the tails. Stress testing is an essential component of a risk management system because it can help to identify crucial vulnerabilities in an institution's position. The methodology of stress testing also applies to credit and operational risks.

In some sense, stress testing can be viewed as an extension of the historical simulation method at increasingly higher confidence levels.

Stress testing, however, is a complement to standard VAR methods because it allows users to include scenarios that did not occur over the VAR window but nonetheless are likely. It also allows risk managers to assess "blind spots" in their pricing or risk management systems. Stress testing can help to ensure the survival of an institution in times of market turmoil.

The drawback of the method is that it is highly subjective. Bad or implausible scenarios will lead to irrelevant potential losses. Even worse, plausible scenarios may not be considered. The history of some firms has shown that people can be very bad at predicting extreme situations. Generally, stress-test results are presented without an attached probability, which makes them difficult to interpret. Unlike VAR, stress testing can lead to a large amount of unfiltered information. There may be a temptation for the risk manager to produce large numbers of scenarios just to be sure that any likely scenario is covered. The problem is that this makes it harder for top management to decide what to do.

Overall, stress testing should be considered an essential complement to rather than a replacement for traditional VAR measures. Stress testing is useful to evaluate the worst-case effect of large movements in key variables. This is akin to drawing a few points in the extreme tails: useful information, but only after the rest of the distribution has been specified. Still, stress testing provides a useful reminder that VAR is no guarantee of a worst-case loss.

QUESTIONS

1. Why should VAR measures be supplemented by portfolio stress testing?
2. How is stress testing different from backtesting when evaluating risk models?
3. "Stress testing is not necessary because the same results are obtained by a VAR model with an increasing confidence level." Comment.
4. Why is unidimensional scenario analysis not sufficient for stress-testing purposes?
5. If historical scenarios automatically take into account correlations, why not rely exclusively on them?
6. What is the major difficulty in generating hypothetical stress tests across large portfolios?

7. Describe the SPAN system for setting margins on a portfolio of futures and options on the same currency.
8. A pension fund has a portfolio with \$1 billion invested in U.S. stocks and \$1 billion invested in Japanese stocks. The 99 percent 1-week VAR analysis reveals a VAR of \$112 million. The risk manager, however, is concerned about extreme moves not reflected in VAR. Compute the stress loss for the following situations:
   (a) A univariate scenario where U.S. stocks fall by 20 percent.
   (b) A univariate scenario where Japanese stocks fall by 25 percent.
   (c) A prospective scenario where U.S. stocks fall by 20 percent and Japanese stocks by 15 percent.
   (d) A prospective scenario where U.S. stocks fall by 5 percent and Japanese stocks by 25 percent. As a risk manager, which stress loss do you think would be most relevant?
9. Just to be sure, the risk manager runs a regression of the Japanese stocks on U.S. stocks and finds a slope coefficient of 0.9. Assuming that a shock of 20 percent originates from the United States, compute the predicted stress loss for the portfolio. What is the danger of this approach?
10. What is the scenario most used for stress tests of firm-wide positions, (a) The 1987 equity crash, (b) money tightening by central banks, (c) a country default, or (d) widening of credit spreads?
11. The crash of 1987 was a 20 standard deviation event. Based on the normal density function, a movement of this magnitude should never happen. In addition, it was aggravated by a failure of the stock exchange to absorb the volume of trading, which has been fixed since then. Based on this information, a portfolio manager argues that this event should not be used in stress tests. Discuss.
12. If a scenario analysis reveals unacceptably large losses, what is a possible response?