Model Error

Modern finance would not have been possible without models. Increasingly complex quantitative models drive financial innovation and the growth of derivatives markets. Models are necessary to value financial instruments and to measure the risks of individual positions and portfolios. Yet when used inappropriately, the models themselves can become an important source of risk. Recently, several well-publicized instances occurred of institutions suffering significant losses attributed to model error. This has sharpened the interest in model risk among financial institutions and their regulators.

In March of 1997, NatWest Markets, an investment banking subsidiary of National Westminster Bank, announced a loss of £90 million due to mispriced sterling interest rate options. Shortly thereafter, BZW, an investment banking subsidiary of Barclays, sustained a £15 million loss on mispriced currency options and Bank of Tokyo-Mitsubishi announced a loss of $83 million. In April of 1997, Deutsche Morgan Grenfell lost an undisclosed amount. Model errors have been blamed for all these losses.[1]

This article will describe various models and discuss model errors characteristic of two types—valuation models for individual securities, and models of market risk. The article will discuss the statistical issues that complicate the use of such models, namely the probability distributions of asset returns and estimates of their volatility. It will also discuss a number of practical issues related to model development and describe the approach taken by bank regulators to model risk.

Katerina Simons

Economist, Federal Reserve Bank of Boston. *The author wishes to thank Richard Kopcke and Peter Fortune for helpful comments. Kevin Daly provided valuable research assistance.*

I. Types of Models

Models can be roughly divided into four categories. The first type is the macroeconomic forecasting model, which seeks to predict real output, employment, inflation, the unemployment rate, and the level of interest rates. Macroeconomic models range from naive one-equation models that extrapolate the values of these variables from their recent past levels to elaborate ones containing hundreds of equations and thousands of variables.

A second, related type of model is a microeconomic model that seeks to explain relationships in particular markets. Such models also use regression analysis, but to explain microeconomic variables, for example, the effect of changes in interest rates on mortgage prepayment rates, or the effect of winter temperatures on the demand for heating oil. Macroeconomic models often use the results of these models as inputs.

Increasingly complex quantitative models drive financial innovation and the growth of derivatives markets.

The third type is the valuation model, used by traders to price derivatives instruments. Derivatives are instruments whose value depends on the price of underlying assets such as stocks, bonds, commodities, or foreign currencies. A simple example of a derivative is a call option on a stock. It gives the buyer the right, but not the obligation, to purchase the stock at some point in the future at a price agreed upon today (known as the strike price). A call option can be priced with the Black-Scholes model, which determines the option's value as a function of five factors: the strike price, the time to expiration, the current stock price, its volatility, and the risk-free interest rate. This model, developed by Fischer Black and Myron Scholes (1973) is often cited as the foundation of modern derivatives markets. For the first time, options could be priced accurately, and this gave rise to an explosion of options trading in the 1980s.

In fact, many newer derivative pricing models are modifications or extensions of Black-Scholes. Note that the Black-Scholes model is different from a statistical model. It is an analytical formula—theoretically, given the five specified factors, one can determine the value of the option *exactly*. In practice, the results of statistical models often serve as inputs into valuation models. For example, to value a mortgage-backed security, one needs an estimate of the mortgage prepayment rate, which has to come from a statistical model. To value a call option with the Black-Scholes formula, one needs to know the volatility of the stock price, which can also be statistically estimated. Statistical estimates are subject to errors and thus can lead to error in the valuation model for which they serve as an input.

The fourth type of model is a model of market risk, used to estimate how the value of an individual position or the whole portfolio changes with a change in the prices of the underlying instruments. Currently, the standard for market risk measurement is a family of models known as "Value at Risk" or VAR. VAR is defined as the maximum amount the portfolio can lose with a given probability in a given period of time. For example, if the daily VAR is estimated to be $1 million at the 95th percent confidence level, one would expect to lose no more than $1 million in one day 95 percent of the time.

To estimate VAR, one first needs to know the probability distributions of the prices of the underlying assets. Second, one must use these prices to arrive at the values of the securities in one's portfolio (that is, use valuation models); and third, one must aggregate these values into one summary estimate of the value at risk. Thus, it is clear that errors in valuation models feed into errors in VAR. In many financial institutions, the valuation methods used in VAR are different and independent of the valuation models used on the trading desks. However, an accurate VAR is not a substitute for accurate valuation models. Clearly, it is not very useful to know with a great degree of confidence and precision how much money one's portfolio can lose in a given period of time if one is seriously mistaken about what the portfolio is worth in the first place.

Statistical issues related to macroeconomic and microeconomic models are outside the scope of this article. We will be concerned only with the last two types of models mentioned—individual valuation and Value at Risk.

II. Types of Errors

It can be useful to think of valuation models and Value at Risk as subject to two distinct types of error—error in the inputs or data, and error in the structure of the model itself. To return to the example of pricing a call option, one is unlikely to make a mistake about the strike price and the expiration date, as these are specific to the contract and do not change

[1] These events are discussed in more detail in Paul-Choudhury (1997).

during the life of the option. Similarly, while there can be some disagreement about the exact number that should be used for the risk-free interest rate, the option price is not very sensitive to small variations in it. A more likely and serious error in the option value can result from an incorrect stock price. Price errors are especially likely if the stock is illiquid, infrequently traded, or subject to abrupt price swings. Then the price at the last transaction may be outdated and give the wrong price for the option.

It is always a matter of judgment which simplifying assumptions can be made and whether the resulting model is sufficiently accurate for the purpose for which it is being used.

Of the five factors in the Black-Scholes model affecting the value of the option, only the volatility is not directly observable. It must be estimated from the historical data; alternatively, it can be "backed out" from prices of other options on the same underlying security. Different methods of volatility estimation can yield widely different results, and it is probably safe to say that incorrect volatility estimates are the single largest cause of error in both valuation models and risk management models.

The second type of error is specification error, that is, error in the structure of the model itself. The model is misspecified if, for example, an additional factor affecting the price of the derivative is not part of the model. This kind of error often occurs when a model intended for one product is used for another without sufficient modification. A more complex issue arises with respect to the simplifying assumptions used in the model. Such assumptions are necessarily unrealistic, but this may not be a problem if they preserve the important features of the market while maintaining analytical tractability. However, it is always a matter of judgment which simplifying assumptions can be made and whether the resulting model is sufficiently accurate for the purpose for which it is being used. Simplifying assumptions, particularly those made in modeling the stochastic process governing asset returns, will be discussed below.

III. The Perils of VAR

Perhaps no model is more fraught with controversy than Value at Risk, or VAR. First popularized by the Group of Thirty report, *Derivatives: Practices and Principles*, VAR has become a de facto industry standard for measuring market risk, that is, the risk of losses from the prices of financial instruments. VAR forms the backbone of the "Internal Models Approach" to allocating capital for market risk in bank trading accounts. This approach, endorsed by Basle Committee on Banking Supervision of the Bank for International Settlements (BIS) and adopted by U.S. bank regulators, allows banks to use their own models to set aside capital for the risks in their trading accounts.

To be acceptable to regulators, banks' models must be based on VAR and fulfill certain other conditions. Specifically, the confidence level for the VAR calculation must be 99 percent and the holding period must be two weeks. For instance, if VAR were found to be $1 million under these conditions, the bank would expect to lose more than $1 million 1 times out of 100 over a two-week period. Figure 1 shows the probability of possible profits and losses based on the normal distribution and the VAR at the 99 percent

Figure 1

Value at Risk Based on Normal Distribution

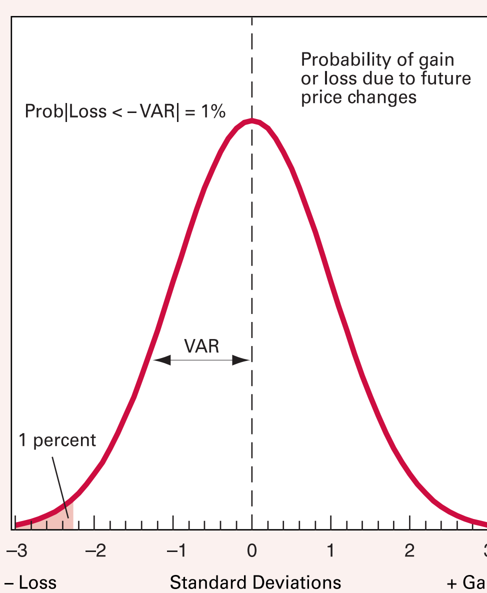
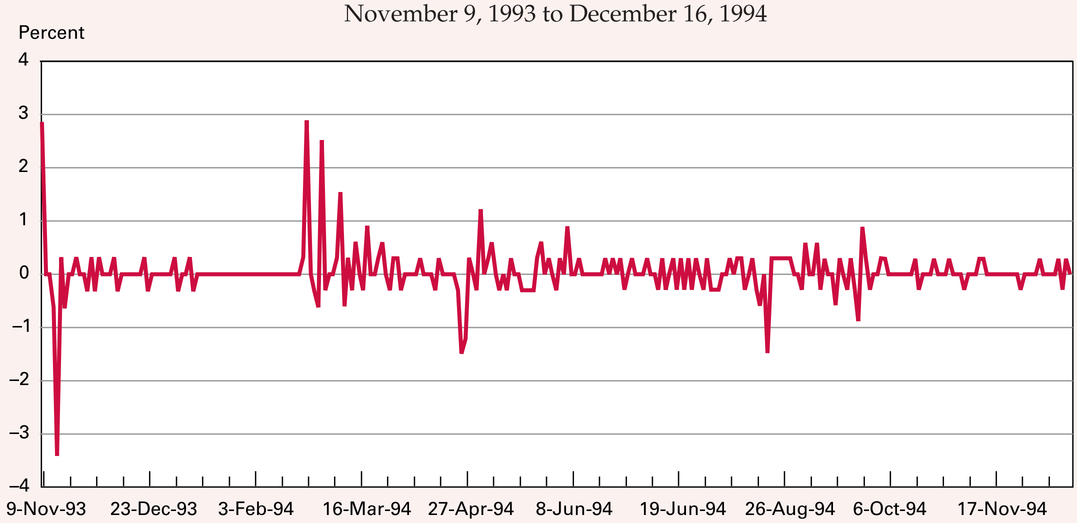


Figure 2

Daily Percent Change in the Peso/Dollar Exchange Rate

November 9, 1993 to December 16, 1994



confidence level, expressed as a number of standard deviations from the mean.

Many market participants consider the 99 percent confidence level and the two-week interval to be too conservative; the industry practice is to limit the confidence level to 95 percent and the holding period to one day. However, since the purpose of the regulatory VAR calculation is to set aside enough capital to protect the bank against failure due to market risk, the regulators consider that a conservative VAR is justified.

Recently, there has been something of a backlash against VAR. Its detractors argue that VAR is misleading and too often wrong to be useful and that, by providing an illusion of scientific precision, it creates a false sense of security, thereby leading traders to carry larger positions and take more risk.[2] The following example is a demonstration of how one can go wrong with VAR.

The Mexican Peso Crisis

Suppose a bank trading desk held a position of $10 million in Mexican pesos on December 19, 1994 and wanted to calculate the VAR of this position based on the historical behavior of the peso/dollar exchange rate. Following the BIS guidelines for internal models, the bank uses a 99 percent confidence interval and a two-week holding period. Figure 2 shows a plot of the percentage changes in the exchange rate from November 1993 through the first half of December 1994. The bank calculates the standard deviation, or volatility, of the daily return on its peso holdings during the previous year starting on November 9, 1993. (Before November 9 the peso was pegged to the dollar, so there was no variation in the exchange rate.) The daily return on the peso position can be calculated as the daily percentage change in the exchange rate, namely:

$$R_t = \frac{E_t - E_{t-1}}{E_{t-1}} \cdot 100\%.$$

R_t is the daily return and E_t is the exchange rate on day t. The daily volatility of returns over a period of N days, with all observations weighted equally, can be calculated as follows:

$$\sigma = \sqrt{\sum_{i=1}^{N} (R_i - \mu)^2},$$

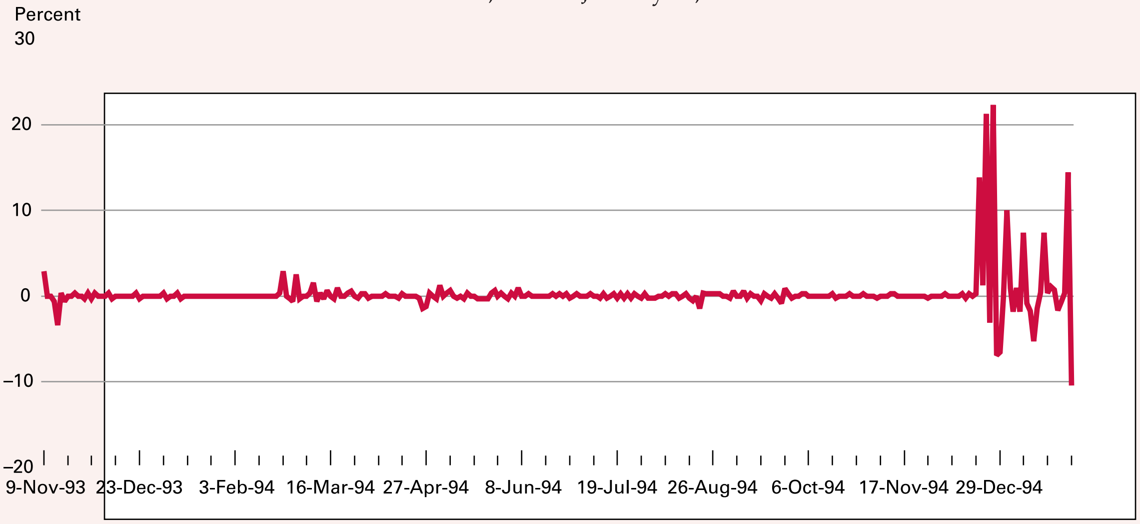
where σ is the volatility of daily returns and μ is the mean return.

[2] Taleb (1997) eloquently summarizes the case against VAR.

Figure 3

Daily Percent Change in the Peso/Dollar Exchange Rate

November 9, 1993 to January 31, 1995



The volatility during the sample period was found to be 0.4704 percent. The VAR calculation is shown in Table 1. To calculate the VAR of this position for two weeks (or 10 trading days), we multiply the daily volatility by $\sqrt{10}$. Then we multiply the resulting two-week volatility by 2.33 for the 99th percent confidence level and by the dollar value of the portfolio ($10 million). VAR = $.4704\% * \sqrt{10} * 2.33 *$ $10 million = $347,000. This means that the bank would expect to lose no more than $347,000 within the next two weeks, 99 times out of a 100.

The subsequent events are shown in Figure 3. Within the next eight days, the peso depreciated 65.7 percent, when the exchange rate went from 3.47 pesos to the dollar on December 19 to 5 pesos to the dollar on December 27. The bank would have lost $6.5 million on its $10 million position, compared to its $347,000 VAR estimate.

Table 1

Calculation of VAR

Position	Daily Volatility	10-day Volatility	99% Confidence	2-week VAR
$10 million	.4704%	1.5%	3.47%	$347,000

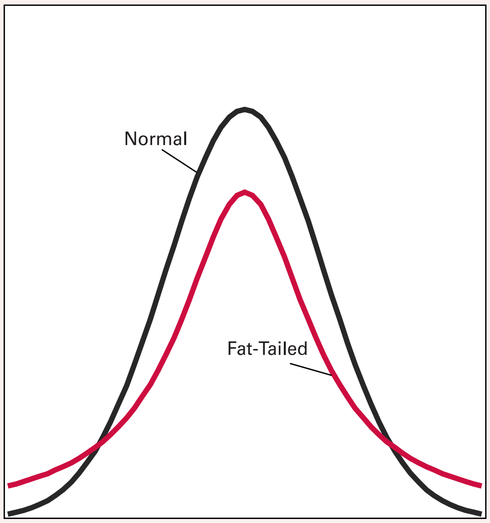
IV. Fat Tails

Why did the risk estimate prove so inadequate in the case of the Peso Crisis? Having first verified that the exchange rate data were correct and that the arithmetic calculations were also done correctly, we are left to examine the assumptions underlying the model. To arrive at this VAR estimate, we adopted the most commonly used model of asset returns. We assumed that returns are independently distributed over time and identically distributed over time (IID), and that they are normally distributed. These assumptions make the model analytically tractable but they are not necessarily close to reality. The assumption of temporal independence allowed us to scale the daily volatility by the square root of time. Furthermore, the assumption of the normal distribution allowed us to infer that a loss 2.33 standard deviations away from the mean would happen no more than one time out of a hundred. The Peso Crisis was a 44-standard-deviation event, something that might happen to a normally distributed random variable once in many billions of years.[3] Nevertheless, it did happen. Nor

[3] The exact probability of such an event is not reported here because none of the statistical programs tried by the author could distinguish such a small number from zero.

Figure 4

Normal and Fat-Tailed Distributions



was the Peso Crisis of 1994 unique. Crashes happen regularly in financial markets, and they do so much more frequently than would be predicted by the normal model.

While the normal distribution is used routinely in modeling asset returns, it has been widely recognized for many years that financial markets exhibit significant non-normalities. In particular, asset returns exhibit "fat tails," meaning that more of their probability is to be found at the tail ends of the distribution and less at the center. A fat-tailed distribution and a normal distribution are illustrated in Figure 4. A fat-tailed distribution makes extreme outcomes such as crashes relatively more likely than does the normal distribution.

How to Measure Fat Tails

The degree of tail-fatness of a distribution can be measured in two ways. The first is kurtosis, or the fourth moment of a distribution. Kurtosis for a sample of daily returns can be estimated as follows:

$$K = \frac{1}{T\sigma^4} \sum_{t=1}^{T} (R_t - \mu)^4,$$

where K is kurtosis, T is the number of days in the sample, μ is the mean return, σ is the standard deviation of returns, and R_t is the return on day t. The kurtosis of a normally distributed variable is 3. By contrast, the sample kurtosis of the daily returns on the S&P 500 index between January 1, 1980 and March 31, 1995 was 128.38. For the most part this high kurtosis can be explained by the stock market crash of 1987. The kurtosis was 2.62 after the crash and 0.98 before it. The high kurtosis was entirely due to the period between October 1 and December 31 of 1987 (Fortune 1996, p. 33).

Another measure of tail-fatness is the frequency of outcomes that are a given number of standard deviations away from the mean. For example, for a normal distribution, outcomes that are farther than 2.33 standard deviations away from the mean in either direction have a 1 percent probability of occurring and events 2.58 standard deviations away from the mean have only a 0.5 percent probability. Thus, one can measure tail-fatness of returns of a given asset by counting how often such extreme outcomes occurred in the sample.

The empirical evidence of fat tails in asset returns is extensive. It was first demonstrated for daily stock-market returns by Fama (1965) and subsequently confirmed in the academic literature for many asset returns. Recently, Duffie and Pan (1997) measured the kurtosis and tail probabilities of daily returns between 1986 and 1996 for 33 return series, including 17 equity indexes representing 13 countries, exchange rates between the dollar and 12 foreign currencies, and 3 commodity series. They found evidence of fat tails across all 33 series they examined. Interestingly, they found the returns on the U.S. dollar/Mexican peso exchange rate to have the fattest tails of all, with kurtosis of 217.5, or 20 moves beyond 5 standard deviations and 5 moves beyond 10 standard deviations.

Fat Tails and Options Pricing

Fat-tailed distributions not only play havoc with Value at Risk, but can also lead to errors in pricing options if the standard Black-Scholes model is used. A critical determinant of the stock option price is the distribution of the terminal price of the underlying stock, that is, its price at the time the option expires. If the returns on the stock are normal, then the terminal stock price has a lognormal distribution. If, on the other hand, the returns distribution has fatter tails than the normal, then the terminal price distribution also has fatter tails than the lognormal. This introduces a bias into the Black-Scholes price of deeply in-the-money and out-of-the-money options,[4] in a way that causes the model to underprice both.

Consider, for example, a deeply out-of-the-money call option (such as an option that gives the holder the right to buy a stock for $50 at the end of the month, while the stock is currently worth only $40). This option will expire with a positive value only if the stock price increases by a large amount. The probability of such an increase depends on the right tail of the price distribution at the end of the month. If the tail is fatter than assumed under the Black-Scholes model, then this probability is higher and the option is more valuable than the Black-Scholes model would lead one to believe.

Fat-tailed distributions not only play havoc with Value at Risk, but can also lead to errors in pricing options if the standard Black-Scholes model is used.

An analogous argument applies to an out-of-the-money put, which will have a positive value only if the stock price drops significantly. The fatter the left tail of the price distribution, the more likely the price drop. Thus, the Black-Scholes model underprices the out-of-the money put. The pricing bias also holds for deep-in-the-money options because of the put–call parity, which is independent of the price distribution. If the call is out of the money, then the corresponding put is in the money, and vice versa. Thus, an in-the-money put will have the same pricing bias as an out-of-the-money call, and an out-of-the-money put the same bias as an in-the-money call.

These pricing biases in option pricing exist if the distribution of returns is fat-tailed but symmetrical, that is, the left and right tails are of equal thickness. However, this need not be the case. The returns distribution can be skewed, with one tail thicker than the other. In particular, it has often been observed that in equity markets, crashes occur more frequently than sudden sharp increases in stock prices. This would imply that the stock returns distribution has a fat left tail but a normal right tail. In this case, only out-of-the-money puts and in-the-money calls will be underpriced by the Black-Scholes model (Hull 1997, p. 494).

V. Three Recipes for Fat Tails

We can try to improve on the normal model by explicitly modeling fat-tailed distributions and using these distributions for pricing options and for calculating the probabilities of losses of certain magnitudes. The challenge is in choosing a fat-tailed distribution that is both mathematically tractable and empirically relevant. Among the most widely used approaches to modeling fat-tailed distribution are stable distributions, jump diffusion, and stochastic volatility.

Stable Distributions

Stable distributions (also known as Pareto-Levy or stable-Paretian), first described by Levy (1924), are a class of bell-shaped symmetrical distributions of which the normal is a special case. Generally, non-normal stable distributions have narrower peaks at the mean and fatter tails than the normal distribution. The fat-tailed distribution illustrated in Figure 4 alongside the normal is Cauchy distribution, which can be described as follows:

$$f(x) = \left(\frac{1}{\pi}\right) \frac{\gamma}{\gamma^2 + (x - \delta)^2},$$

where $\gamma = 1$ and $\delta = 1$.

Stable distributions imply long-range correlations in asset returns and are characterized by trends, cycles, and discontinuous changes. The application of stable distributions to financial asset returns was first proposed by Mandelbrot (1963) and gave rise to a large literature on the subject.[5]

The most striking feature of stable distributions is that the variance and all higher moments do not exist. This feature also makes their use in modeling very difficult, because of the crucial role the concept of variance plays in finance theory. Moreover, the historical record of asset returns does not fit some empirical implications of stable distributions. For example, sam-

[4] An option is in the money if it would lead to a positive cash flow if it were exercised immediately. Similarly, an out-of-the-money option would lead to a negative cash flow if exercised immediately.

[5] See Campbell, Lo, and MacKinlay (CLM) 1997, p. 17 for more detail and a list of references.

ple estimates for variance drawn from random variables that follow stable distributions would not converge to a single number as the sample gets larger, but would keep increasing without limit. In practice, estimates of variance of sample returns usually converge as the sample gets larger. In addition, the use of a stable distribution to model short-term returns implies that long-term returns would also follow the same distribution. In practice, unlike short-term returns, long-term returns are often well-approximated by the normal distribution (CLM 1997, p. 19). As a result, the use of stable distributions in finance has always been controversial and recently has been supplanted by other models such as jump diffusion and stochastic volatility, which produce fat tails while preserving the finite variance.

Jump Diffusion

The jump diffusion model was introduced by Merton (1976). This model assumes only that returns are IID. In particular, the returns usually behave as if drawn from a normal distribution but are periodically "jumped" up or down by adding an independent normally distributed shock. The arrival of these jumps is random and their frequency is governed by the Poisson distribution with a given expected frequency.

Fortune (forthcoming) estimates the jump diffusion process for the daily returns on the S&P 500. He finds that the volatility of returns arises both from the normal process and from the jumps, with the jumps accounting for 80 percent of overall return volatility. He also compares intraweek daily returns with returns over weekends and finds that jumps occur more frequently on weekdays and are, on average, positive. In contrast, weekend jumps are, on average, less frequent, of larger size, and negative. This is consistent with companies' announcing bad news over weekends, hoping to minimize its impact on their stock prices when the markets next open.

The advantage of the jump diffusion model is that it can make extreme events appear more frequently. Duffie and Pan (1997) illustrate the impact of jumps on risk measurement by comparing a normal distribution of returns with a constant standard deviation of 15 percent, to an otherwise identical process that has jumps that occur with an expected frequency of once a year and come from a normal distribution with a zero mean and a standard deviation of 10 percent. Despite the low frequency of jumps, this is equivalent in risk to a standard normal distribution of returns with a standard deviation of 185 percent. The jump process makes extreme events far out in the tail more likely. For example, in the above model, one can expect to lose overnight at least a quarter of the value of one's position. In contrast, as Duffie and Pan point out, in the normal model "one would expect to wait far longer than the age of the universe to lose as much as one quarter of the value of one's position overnight."[6]

Stochastic Volatility

Stochastic volatility means that the volatility is not constant but undergoes random changes over time. This results in a fat-tailed unconditional distribution of returns, while the conditional distribution (that is, the distribution of returns at time t conditional on all previous returns) remains normal. The most popular class of stochastic volatility models is GARCH, and the most popular type of GARCH is GARCH (1,1). GARCH stands for Generalized Autoregressive Conditional Heteroskedasticity. The model is autoregressive because it involves regressing the variance on its own past values; heteroskedasticity simply means changing variance. The first version of this type of model, known as ARCH, was introduced by Engle (1982) and the generalized, or GARCH, version, was developed by Bollerslev (1986).

GARCH (1,1) relates the variance of asset returns at time t to the variance of asset returns at time $t-1$ and the excess return at time t. The equation is:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2,$$

where σ_t is the volatility of asset price at time t, ϵ_t is the excess return, that is, the difference between the actual return on the asset at time t and the average return, and ω, α, and β are constants. This model is referred to as GARCH (1,1) because volatility at time t depends only on the return and the volatility at time $t-1$ and not on the path they have taken in the past. In this model, β is the persistence parameter, which determines how much carryover effect the previous period's volatility has on today's volatility. The three parameters can then be estimated from the historical returns data.

Implied Volatility—Smiles and Scowls

A different approach to volatility estimates calculates them from option prices. This can be done

[6] Duffie and Pan credit Mark Rubinstein with the use of this language to describe the expected frequency of an overnight loss of this magnitude in the normal model.

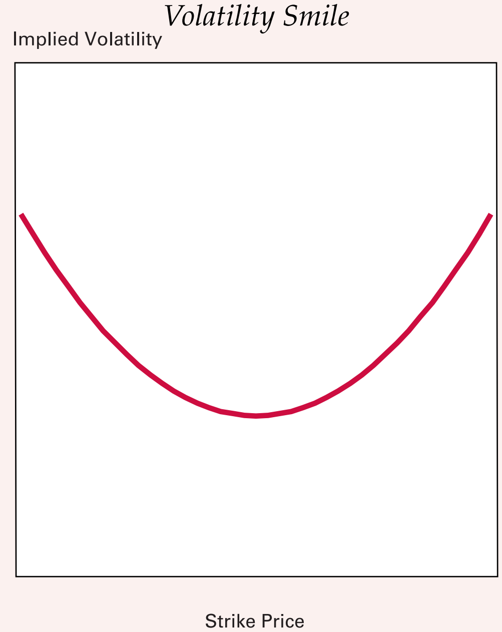
because in the Black-Scholes model a one-to-one correspondence exists between the price of any option traded in the market and the volatility of the underlying security that went into the calculation of that option price. Thus, if the option price is known, the volatility can be "backed out" from the Black-Scholes formula. (The formula cannot be solved analytically, but the volatility can be found by plugging in different values for it until one gets arbitrarily close to the option price.)

The volatility calculated from the option price is known as "implied" volatility. If one knows the price of an option on a certain underlying security and its implied volatility, one could, in principle, calculate the price of any other option on the same underlying security. Moreover, one could also use the implied volatility in value-at-risk calculations. The advantage claimed for using implied volatility rather than historical volatility in pricing options and measuring value at risk is that it is forward-looking, reflecting actual expectations of market participants rather than the past. The advantages of implied volatility, however, are more apparent than real. The main reason for this are the simplifying assumptions in the Black-Scholes model. Recall that the model assumes that the volatility of the underlying security is constant. Thus, if traders really were using the Black-Scholes to price options, the implied volatility for any underlying security would be identical for all options on it. In reality, nothing could be further from the truth. Option prices on the same underlying security imply different volatilities, depending on the option's expiration date and strike price. Implied volatility first tends to decrease with the strike price and then to increase, resulting in the so-called volatility "smile." Figure 5 shows a hypothetical example of a volatility smile. Its symmetry is usually characteristic of currency options (Hull 1997, p. 503).

Implied volatility plots for stock index options resemble crooked "scowls" rather than smiles. Stock index option volatilities are highest for deeply out-of-the-money options and they keep decreasing, the more options get into the money. Figure 6 shows several implied volatility plots for call options on the S&P 500 index traded on June 10, 1997. The S&P 500 index traded at 862.88 when these volatilities were calculated. Each plot shows the variation in implied volatility for options with different strike prices but the same expiration dates. The different curves represent different months of option expiration. The steepest slope is for the nearest expiration month (in this case, June) with the more distant expiration months

Figure 5

Volatility Smile

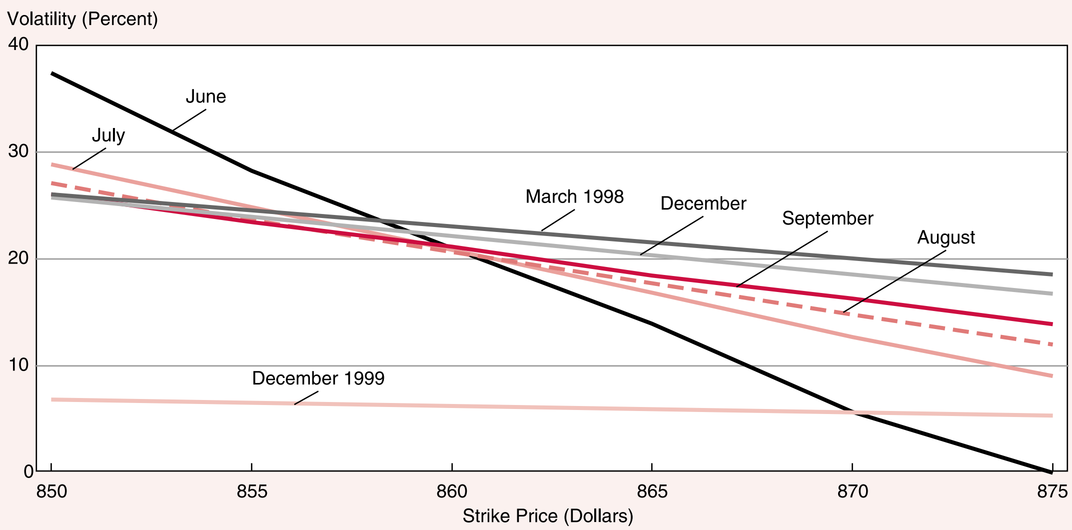


having progressively flatter slopes. The relationship between implied volatility and time to expiration is known as the term structure of volatility, by analogy with the term structure of interest rates.

The differences in implied volatilities among different strike prices and expiration dates reflect not only market views of volatility, but other considerations. They may depend, in part, on differences in option premiums caused by variations in supply and demand for different kinds of options. This is particularly true for deeply in- and out-of-the-money options, as they tend to be infrequently traded. Implied volatilities for different times to expiration tend to converge for near "at-the-money" options because of the relative depth of their markets. Ultimately, however, it must be kept in mind that despite their ubiquity, different implied volatilities for the same underlying securities calculated from the Black-Scholes formula do not have a well-defined economic meaning. In fact, such volatilities reflect a logical inconsistency, since they are calculated from a model whose fundamental assumptions are violated by their very existence. Implied volatilities can be useful, however, if they are estimated from an options-pricing model that allows for time-varying volatility, provided that we know which model was used to price

Figure 6

Volatility Smile for SPX June Calls



the option, and provided that this model was correct. These conditions make implied volatilities more appropriate for trading options than for more general value-at-risk calculations. Risk managers concerned with value-at-risk often prefer historical data because of their desire to have independent models rather than relying on pricing models used by traders.

VI. Choosing the Model

The use of historical data to model fat-tailed distributions of asset returns has gained ground in modeling value at risk. In particular, jump diffusion and stochastic volatility models are appealing because they can predict "tail events" more accurately than the normal distribution. Nevertheless, the normal distribution continues to be used despite its frequently proclaimed and well-documented flaws. The appeal of the normal model lies in its simplicity. This may not seem like a good excuse, given its less-than-reassuring track record; however, the appeal of the normal becomes more obvious once we consider the practical difficulties in implementing the alternatives. To use the normal distribution, one needs to estimate only one unobserved parameter from the historical data, namely the variance. In contrast, the two alternative models described above, jump diffusion and GARCH (1,1), each require estimation of three unobserved parameters. The GARCH model requires the estimates of ω, α, and β. The jump diffusion model requires the estimates of the variance of returns, the variance of jumps, and the expected frequency of jumps. Thus, it is not enough to choose a model structure that happens to describe the stochastic process generating the data reasonably well. It is also necessary to estimate its parameters correctly, and if one or more parameters of the model are misspecified, the resulting model is not going to be an improvement over the normal model.

VAR and the Risk of Catastrophic Events

Would a VAR measure based on a stochastic volatility model such as GARCH have done any better than the normal model in estimating the probability of a catastrophic event such as the Peso Crisis? The answer depends on the amount and appropriateness of the historical data used for estimating the model. It will be recalled that the volatility of the peso/dollar exchange rate was very low during the year leading up to the crisis. It was low no matter how it was measured and what sort of structure was imposed on the data. The peso/dollar exchange rate simply did not change very much in the year leading to the dramatic devaluation, and no model can find the variability that is not there. The obvious choice is to use historical data further in the past. In particular, Mexico has had currency devaluations since the middle 1970s. Possibly, 20 years of daily exchange rate data would provide an estimate of the probability of the latest devaluation, and one model might do so better than another. However, using 20 years of daily data in every market in which a bank is trading is impractical, and in general 20-year-old data usually would not be relevant.

The preceding observations are not meant to suggest that no information was available in 1994 that suggested a possible devaluation. Such information, however, was not to be found in the volatility of the exchange rate but in the political climate and macroeconomic policies of the Mexican government. Recently there has been renewed interest in the causes of currency crises, spurred in part by the events in Mexico. The basic model as applied to currency crises, however, was developed by Krugman in 1979. It attributes speculative pressures on currencies to the inconsistency of the government's monetary and fiscal policies with its exchange rate target. Specifically, such inconsistency can arise if a country runs budget deficits while its central bank finances them by expansionary monetary policy. The expanding supply of local currency relative to the U.S. dollar tends to depress the value of the local currency. If the central bank wants to defend the value of its currency and keep it pegged to the U.S. dollar, it must buy it with its foreign currency reserves. But its foreign currency reserves are not infinite, and once they are exhausted, the currency can no longer be defended. Rational investors will foresee that the reserves will eventually be exhausted and seek to profit immediately by selling the currency before the reserves are depleted. Krugman's theory predicts that the collapse of the exchange rate peg will occur as soon as investors become convinced that the government's policies are unsustainable and mount a speculative attack on the currency.

The basic model as applied to currency crises was developed by Krugman in 1979. It attributes speculative pressures on currencies to the inconsistency of the government's monetary and fiscal policies with its exchange rate target.

Mexican experience in 1994 is consistent with Krugman's model of a speculative attack. At the time, Mexico had a large and growing current-account deficit and relied on an overexpanding money supply and short-term borrowing. These weaknesses were masked for a time by large capital inflows from foreign investors which allowed Mexico to maintain the value of the peso and maintain its foreign exchange reserves. In 1994, however, Mexico experienced a great deal of political instability, notably the peasant uprising in Chiapas and the assassination of presidential candidate Luis Donaldo Colosio. Political turmoil led to a loss of confidence by foreign investors, who withdrew their money. As the Mexican government sought to defend the value of the peso, its foreign exchange reserves were rapidly depleted, forcing the devaluation.

Catastrophic Events and Stress Testing

While no method of modeling catastrophic events like the Peso Crisis is universally accepted, it is clear that VAR is not sufficient and must be supplemented by other methods. These usually involve stress testing using scenario analysis. Stress testing involves creating a set of realistic scenarios and evaluating their impact on all current trading positions. Besides the Peso Crisis, other examples of catastrophic market events in the 1990s include the European currency crisis precipitated by the decision of Great Britain and Italy to leave the Exchange Rate Mechanism (ERM) in September of 1994, the sharp fall in the U.S. and European bond markets in April of 1994, and the Kobe earthquake in February of 1995. In addition to using scenarios based on actual past catastrophic events, stress testing can be supplemented by views of management about possible future events or based on econometric models of market trends. Stress testing can then be used to define the limits on market positions the traders are permitted to take, and to determine the level of diversification necessary to bring the impact of catastrophic events to acceptable levels.

Thus, it is useful to make a distinction between the risk of events that occur somewhat rarely (like 1 out of 100 times) but are still a part of business as usual, and catastrophic events that happen *very* rarely and whose probability cannot be estimated with accuracy. The former risk is measured by VAR and the latter by stress testing. These approaches are complementary, and both are necessary for modeling risk. The main difference between VAR and stress testing is that, unlike VAR, stress testing does not assign probabilities to the scenarios being tested, which makes it difficult to compare results to VAR or to each other.

VII. Conclusion

At the beginning of this article, we distinguished between four types of models—macro models, micro models, valuation models, and risk management models such as Value at Risk. The bulk of the article has focused on valuation and risk management models and their crucial first step—modeling the statistical processes of asset returns. We observed that modeling asset returns is complicated by changes in the parameters of the distribution, such as the volatility. These parameter shifts can be due to changing economic conditions, technical innovations, and fiscal and monetary policies. These structural shifts can inform our stress tests and scenario analyses, but the real challenge is formally incorporating them into the models of asset returns. Up to now, such attempts have not resulted in a significant improvement in the predictive powers of the models of asset returns. The most fruitful approaches so far have been the modeling of asset returns as fat-tailed distributions, including stable distributions, jump diffusion, and stochastic volatility models. However, a trade-off almost always exists between the realism and the analytical tractability of the model. Striking the right balance in the face of this trade-off, and maintaining it through changing market conditions for different financial instruments, is more art than science and requires considerable experience and judgment. Bank regulators recognize this and as a result place more emphasis on qualitative aspects of risk management than on the content of models themselves.

References

Black, Fischer and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81 (May–June), pp. 637–59.

Bollerslev, Tim Peter. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, vol. 31, pp. 307–27.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay. 1997. *The Econometrics of Financial Markets.* Princeton, NJ: Princeton University Press.

Duffie, Darrel and Jun Pan. 1997. "An Overview of Value at Risk." *The Journal of Derivatives*, vol. 4, no. 3 (Spring), pp. 7–49.

Engle, Robert F. 1982. "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of UK Inflation." *Econometrica*, vol. 50, pp. 987–1008.

Fama, Eugene F. 1965. "The Behavior of Stock Market Prices." *Journal of Business*, vol. 38 (January), pp. 34–105.

Fortune, Peter. 1996. "Anomalies in Option Pricing." *New England Economic Review*, March/April, pp. 17–40.

———. Forthcoming. *New England Economic Review*.

Group of Thirty. 1993. *Derivatives: Practices and Principles*. Washington, DC.

Hull, John C. 1997. *Options, Futures, and Other Derivatives.* Third Edition. Upper Saddle River, NJ: Prentice Hall.

Krugman, Paul. 1979. "A Model of Balance of Payments Crises." *Journal of Money, Credit and Banking*, vol. 11, pp. 311–25.

Levy, P. 1924. "Théorie des Erreurs. La Loi de Gauss et Les Lois Exceptionelles." *Bull. Soc. Math.*, vol. 52, pp. 49–85.

Mandelbrot, Benoit. 1963. "The Variation of Certain Speculative Prices." *Journal of Business*, vol. 36, pp. 394–419.

Merton, Robert C. 1976. "Option Pricing When Underlying Stock Returns Are Discontinuous." *Journal of Financial Economics*, vol. 3 (March), pp. 125–44.

Paul-Choudhury, Sumit. 1997. "This Year's Model." *Risk*, vol. 10, no. 5 (May), p. 19.

Taleb, Nassim. 1997. Interview in *Derivative Strategies*, vol. 36, December/January, pp. 37–40.