OESTERREICHISCHE NATIONALBANK

Guidelines on Market Risk

Volume 5

Stress Testing

Guidelines on Market Risk

Volume 1: General Market Risk of Debt Instruments
2nd revised and extended edition

Volume 2: Standardized Approach Audits

Volume 3: Evaluation of Value-at-Risk Models

Volume 4: Provisions for Option Risks

Volume 5: Stress Testing

Volume 6: Other Risks Associated with the Trading Book

Published and produced by:

Oesterreichische Nationalbank

Editor in chief:

Wolfdietrich Grau

Author:

Financial Markets Analysis and Surveillance Division

Translated by:

Foreign Research Division

Layout, design, set, print and production:

Printing Office

Internet:

http://www.oenb.at

Paper:

Salzer Demeter, 100% woodpulp paper, bleached without chlorine, acid-free, without optical whiteners.

DVR 0031577

The second major amendment to the Austrian Banking Act, which entered into force on January 1, 1998, faced the Austrian credit institutions and banking supervisory authorities with an unparalleled challenge, as it entailed far-reaching statutory modifications and adjustments to comply with international standards.

The successful implementation of the adjustments clearly marks a quantum leap in the way banks enganged in substantial securities trading manage the associated risks. It also puts the spotlight on the importance of the competent staff's training and skills, which requires sizeable investments. All of this is certain to enhance professional practice and, feeding through to the interplay of market forces, will ultimately benefit all market participants.

The Oesterreichische Nationalbank, which serves both as a partner of the Austrian banking industry and an authority charged with banking supervisory tasks, has increasingly positioned itself as an agent that provides all market players with services of the highest standard, guaranteeing a level playing field.

Two volumes of the six-volume series of guidelines centering on the various facets of market risk provide information on how the Oesterreichische Nationalbank appraises value-at-risk models and on how it audits the standardized approach. The remaining four volumes discuss in depth stress testing for securities portfolios, the calculation of regulatory capital requirements to cover option risks, the general interest rate risk of debt instruments and other risks associated with the trading book including default and settlement risk.

These publications not only serve as a risk management tool for the financial sector, but are also designed to increase transparency and to enhance the objectivity of the audit procedures. The Oesterreichische Nationalbank selected this approach with a view to reinforcing confidence in the Austrian financial market and – against the backdrop of the global liberalization trend – to boosting the market’s competitiveness and buttressing its stability.

Gertrude Tumpel-Gugerell

Vice Governor

Oesterreichische Nationalbank

Today, the financial sector is the most dynamic business sector, save perhaps the telecommunications industry. Buoyant growth in derivative financial products, both in terms of volume and of diversity and complexity, bears ample testimony to this. Given these developments, the requirement to offer optimum security for clients' investments represents a continual challenge for the financial sector.

It is the mandate of banking supervisors to ensure compliance with the provisions set up to meet this very requirement. To this end, the competent authorities must have flexible tools at their disposal to swiftly cover new financial products and new types of risks. Novel EU Directives, their amendments and the ensuing amendments to the Austrian Banking Act bear witness to the daunting pace of derivatives developments. Just when it seems that large projects, such as the limitation of market risks via the EU's capital adequacy Directives CAD I and CAD II, are about to draw to a close, regulators find themselves facing the innovations introduced by the much-discussed New Capital Accord of the Basle Committee on Banking Supervision. The latter document will not only make it necessary to adjust the regulatory capital requirements, but also require the supervisory authorities to develop a new, more comprehensive coverage of a credit institution's risk positions.

Many of the approaches and strategies for managing market risk which were incorporated in the Oesterreichische Nationalbank’s Guidelines on Market Risk should – in line with the Basle Committee’s standpoint – not be seen as merely confined to the trading book. Interest rate, foreign exchange and options risks also play a role in conventional banking business, albeit in a less conspicuous manner.

The revolution in finance has made it imperative for credit institutions to conform to changing supervisory standards. These guidelines should be of relevance not only to banks involved in large-scale trading, but also to institutions with less voluminous trading books. Prudence dictates that risk – including the "market risks" inherent in the bank book – be thoroughly analyzed; banks should have a vested interest in effective risk management. As the guidelines issued by the Oesterreichische Nationalbank are designed to support banks in this effort, banks should turn to them for frequent reference. Last, but not least, this series of publications, a key contribution in a highly specialized area, also testifies to the cooperation between the Austrian Federal Ministry of Finance and the Oesterreichische Nationalbank in the realm of banking supervision.

Alfred Lejsek

Director General
Federal Ministry of Finance

Preface

Stress testing is gaining significance as a risk management tool. Independent of supervisory requirements, banks' top executives have been paying ever greater attention to stress testing over the past two years. The mounting importance credit institutions attach to this mode of testing has raised the quality of stress testing schemes. Interestingly enough, there is as yet no uniform, generally accepted standard in place.

This guideline sheds light on the various developments reflected in stress testing programs and presents miminum requirements applicable to Austrian credit institutions using internal models for measuring their exposure. A reference tool designed to prime institutions on how to incorporate stress tests in their risk management system, it clearly revolves around market risk, but also touches upon liquidity and credit risk.

The idea for this publication may be traced to the Oesterreichische Nationalbank's involvement in evaluating proprietary models used by banks to limit market risk. Given the OeNB's experience in this area, two aspects have come to the fore in particular, which also underscored the potential need for such a guideline. For one, given their design, stress tests may serve as a fairly simple tool for managing risk. In addition, little has thus far been written on the topic. Apart from the banks which employ internal models and are therefore required by the Austrian Banking Act to perform stress testing, such testing methods also lend themselves to any credit institution or enterprise with a treasury department. After all, stress testing may be implemented for in-house risk management purposes in a quick manner. It goes without saying that there is no limit to refining the methods used. This is where the scientific community comes in: It is desirable to investigate stress testing further, which should best be achieved via an interdisciplinary approach bringing together finance, macroeconomics, statistics and econometrics, to name just the key disciplines.

The authors would like to extend thanks to Alan Cathcart and Nick Palmer of the Financial Services Authority, London; Benjamin Cohen of the Basle Committee on the Global Financial System; Zahra El-Mekkawy of the Basle Committee on Banking Supervision as well as Stefan Walter and Kevin Clarke of the Federal Reserve Bank of New York for fruitful discussions about international stress testing practices. Credit also goes to Michael Boss and Ronald Laszlo for their comments and valuable suggestions. Special thanks are due to the head of the division, Helga Mramor, who promoted the production of this series of guidelines on market risk.

Vienna, September 1999

Thomas Breuer
Gerald Krenn

Table of Contents

1 Introduction

1.1 Legal Framework

The second major amendment to the Austrian Banking Act introduced the term stress testing into the legal risk management provisions applicable to Austrian credit institutions. This amendment, which incorporated the capital adequacy Directive (CAD) into Austrian law, entailed a change in the computation of the regulatory capital requirement credit institutions and groups of credit institutions need to hold. Credit institutions that keep a large-volume trading book are now required to calculate the regulatory capital requirement for trading book positions in line with the CAD standardized approach. Or, institutions may implement internal models for limiting the market risk, also referred to as value-at-risk (VaR) models, to determine the required capital for backing both the general and the specific position risk inherent in debt instruments and stocks contained in the trading book as well as in commodities positions and open currency positions. The use of such proprietary models for market risk management purposes was recommended by the Basle Committee on Banking Supervision in its January 1996 publication entitled "Amendment to the Capital Accord to Incorporate Market Risks." In the meantime, these recommendations have essentially been integrated into a Directive amending the CAD.

Both the Basle market risk paper of January 1996 and the EU Directive stipulate that the use of an internal model be subject to approval by the competent banking supervisory authority. What is more, both papers spell out stress testing as one of the prerequisites for model approval. In other words, bank regulators consider stress tests to be an effective and necessary tool that complements statistical models for quantifying and monitoring risk. Given their role as a control mechanism, stress tests are listed in the Austrian Banking Act among the qualitative standards. Stress testing does, however, also set high quantitative standards for risk management.

In summary, any credit institution using an internal model to calculate the regulatory capital requirement is bound by law to carry out stress tests. Likewise, all other credit institutions and financial institutions may in general benefit from integrating stress testing into their risk control. The methods underlying stress tests are easy to comprehend, and the requirements for performing stress tests are fairly low. This guideline therefore targets not just those credit institutions that use internal models, but rather all credit institutions; besides, this publication may prove useful to other institutional investors.

1.2 Why Use Stress Tests

The need for stress testing is justified by the Basle Committee on Banking Supervision (1995) as follows:

> "Banks that use the internal models approach for meeting market risk capital requirements must have in place a rigorous and comprehensive stress testing program. Stress testing to identify events or influences that could greatly impact banks are a key component of a bank's assessment of its capital position.
>
> Understanding and protecting against the vulnerabilities of a financial company's risk-taking activities is of course one of the major responsibilities of its board of directors and senior management. Banks' stress scenarios need to cover a range of factors that can create extraordinary losses or gains in trading portfolios, or make the control of risk in those portfolios very difficult. These factors include low-probability events in all major types of risks, including the various components of market, credit, and operational risks. Stress scenarios need to shed light on the impact of such events on positions that display both linear and non-linear price characteristics (i.e. options and instruments that have options-like characteristics).
>
> Banks' stress tests should be both of a quantitative and qualitative nature. Quantitative criteria should identify plausible stress scenarios to which banks could be exposed. Qualitative criteria should emphasise that two major goals of stress testing are to evaluate the capacity of the bank's capital to absorb potential large losses and to identify steps the bank can take to reduce its risk and conserve capital. This assessment is integral to setting and evaluating the bank's management strategy and the results of stress testing should be routinely communicated to senior management and, periodically, to the bank's board of directors."

As far as the consequences of stress tests go, the Committee states:

> "Stress testing alone is of limited value unless the bank is ready to respond to its results. At a minimum, the results should be reviewed periodically by senior management and should be reflected in the policies and limits set by management and the board of directors. Moreover, if the testing reveals particular vulnerability to a given set of circumstances, the national supervisors would expect the bank to take prompt steps to manage those risks appropriately (e.g. by hedging against that outcome or reducing the size of its exposures)."

Stress tests should, thus, provide credit institutions with answers to these three questions:

1. What will the loss be in the event of scenario X?
2. What are our institution's worst-case scenarios?
3. What can we do to limit the losses incurred in the worst-case scenarios?

Stress tests do not, however, provide an answer in quantitative terms to the question of how probable any given scenario is. Still, the plausibility of scenarios does play a certain role in interpreting stress testing results. Sections 2.5 and 4.4.2 discuss these points at greater length.

1.3 Stress Tests and Value-at-Risk Models

The issue of stress testing often crops up in connection with VaR models. As mentioned above, the execution of stress tests is stipulated by law for credit institutions that employ VaR models to compute their regulatory capital requirements. Basically, stress testing is to complement the internal models approach. Why do VaR models call for such complementary measures, and how come stress tests fit the bill?

The VaR methodology is fairly well-known: A *holding period* of t days and a *confidence level* of p% are given. The VaR is a statistical measure of the loss of a portfolio – as measured in monetary units – which will not be exceeded with a probability of p% given the portfolio remains constant throughout the holding period. Losses in excess of the VaR only occur with a low probability [(1-p)%]. A VaR model does not shed light on the dimension of such "heavy" losses. This is the first reason why stress testing is required as a complementary measure: stress tests serve to estimate potential extreme losses.

The second important reason why VaR calculations shall be combined with stress tests lies in the somewhat skeptical attitude towards the assumptions on which most VaR calculations are based. In the same vein, the multiplication factor applied to the value at risk in computing the regulatory capital requirement helps absorb the remaining uncertainty about the accuracy of the model.

There are first and foremost two assumptions whose validity is debatable. For one, the markets are assumed to remain constant over a given time horizon. Only in the event that future market movements mirror those of the past can models produce reliable results. Yet, there have always been breaks in market movements. They may be attributable to various causes, for instance, to full-blown crises, such as wars or environmental catastrophes, changes in the interest rate or exchange rate policies pursued by central banks, speculative attacks on currencies and the like. A stress situation shall therefore mean a break in the temporal constancy of a market. The objective of stress tests is, among other things, to assess the potential loss resulting from such breaks.

Furthermore numerous VaR models assume that changes in risk factors are normally distributed. However, changes in financial time series are, as a rule, not normally distributed. Instead, such time series are marked by fat tails. It follows that extreme changes in the risk factors are considerably more likely than is accounted for under the assumption of a normal distribution. The slump in stock prices triggered by the equity crash of 1987, for example, was reflected by 10 to 20 standard deviations. The table below shows that such a fall in prices should not be possible under the assumption of a normal distribution.

Probabilities of extreme changes under the assumption of a normal distribution

k	**Probability of a price slump of *k* standard deviations or more**
5	$6 \quad 10^{-7}$
6	$2 \quad 10^{-9}$
7	$3 \quad 10^{-12}$

Table 1

Stress tests are not based on statistical assumptions on how the changes in risk factors are distributed. This is why the results of stress tests are not distorted by fat tails.

As stress tests do not quantify the probability of occurrence of the individual scenarios, they lend themselves to verifying and complementing statistical risk measures such as the value at risk. As a monitoring tool, stress tests primarily serve to verify statistical assumptions unerlying the model. The pricing model of an internal model may not or only be partially verified via stress testing, since the portfolio valuation to be carried out during stress testing itself rests on a pricing model.

While stress tests do not put exact figures on the probability of scenarios, scenarios still need to be somewhat plausible. The evaluation of scenario plausibility calls for, at least, a rough idea of the probability with which given scenarios will occur.

1.4 Weaknesses of Value at Risk and Strenghts of Stress Tests: a Case Study

A case study presented in Gay et al. (1999) illustrates the fact that stress tests should, in particular in addition to VaR calculations, be used to measure the risk of financial transactions.

In late January 1998, the Korean investment house SK Securities Co. suffered a loss of USD 189 million traceable to a total return swap transaction. The swap was entered into at the end of January 1997 with a maturity of one year. A payment was to be effected at the end of the maturity the amount of which would depend on the exchange rates of the currencies of Thailand (baht, THB) , Indonesia (rupiah, IDR) and Japan (yen, JPY) vis-à-vis the USD. Basically, it had been agreed that SK Securities would receive the following amount once the swap came due

$$N \cdot \left[5 \cdot \left(\frac{B_0}{B_2} - 1\right) + \mathrm{Max}\left(0,\ \frac{3 \cdot R_0 - R_1 - R_2}{R_2}\right) + \mathrm{Max}\left(0,\ 1 - \frac{Y_0}{Y_2}\right) - 0.97 \right] \tag{1.1}$$

or that it would pay that amount if it happened to be negative. In the formula above, N designates the principal of USD 53 million, B_0 (B_2), R_0 (R_2) and Y_0 (Y_2) denote the USD rates of the baht, rupiah and yen at the beginning (end) of the life of the swap, and R_1 gives the USD rate of the rupiah after six months following the transaction date (all rates are given per USD 1).

Had the rates remained constant during the life of the swap transaction, SK Securities would have received a payment in the order of $N \cdot 0.03$ = USD 1.59 million. Expression (1.1) shows that a depreciation of the baht relative to the USD ($B_2 > B_0$) would have had unfavorable consequences for SK Securities. A depreciation of the rupiah would likewise not have benefited SK Securities, while the investment house would have profited from an appreciation of the baht or rupiah or a depreciation of the yen.

The decision of SK Securities to enter into the swap was based on historical rate movements and volatilities of the currencies involved. The historical data implied that the risk was relatively low. When the swap was transacted and in the years prior to that transaction, Thailand's central bank kept the baht strictly pegged to a currency basket the composition of which was never made public but which allegedly consisted of the USD (80%), JPY (12%) and DEM (8%). The Indonesian central bank targeted a limitation of the rupiah's loss in value relative to the USD to a maximum of 5% per annum. By contrast, the Japanese central bank largely refrained from intervening for the yen. The differing rate targets of the central banks are reflected in the historical volatilities of the exchange rates vis-à-vis the USD: the closer the peg to the USD, the smaller the volatility. This point is also illustrated by table 2 showing annualized historical volatilities based on an observation period of 26 weeks prior to January 29, 1997.

Annualized historical volatilities relative to the USD;
Observation period: August 6, 1996 to January 28, 1997; source: Gay et al. (1999)

Currency	THB	IDR	JPY
Volatility	1.23%	2.20%	6.88%

Table 2

Following the swap transaction, the central banks concerned continued to pursue their respective monetary policies. However, once Thailand's central bank had exhausted a large portion of its official reserves to shield the baht from speculative attacks, it decided on July 2, 1999 to discontinue those interventions in favor of improving Thailand's export opportunities. The baht promptly depreciated relative to the USD by 16%. Consequently, the currencies of other countries in the region lost on the USD as well. On August 14, 1997, Indonesia's central bank dropped its rate target. Table 3 demonstrates the losses on the USD of the currencies involved in the swap transaction in the period from end-January 1997 to end-January 1998.

Depreciation relative to the USD from January 29, 1997 to January 29, 1998;
source: Datastream

Currency	THB	IDR	JPY
Depreciation relative to USD	51.8%	77.9%	2.9%

Table 3

The VaR measures computed for the baht and rupiah positions in USD at the time when the swap was transacted and using a confidence level of 99% and a holding period of one year under the assumption of a normal distribution of the relative exchange rate fluctuations would have underestimated – based on the volatilities stated in table 2 – the actual losses 18fold and 15fold, respectively (e.g. VaR for USD 100 in baht: VaR = USD 100 $\cdot$ 0.0123 $\cdot$ 2.326 = USD 2.86; actual loss: USD 51.8).

Gay et al. (1999) demonstrate that even VaR calculations covering the entire swap at the transaction time would have drastically underestimated the actual loss incurred. A Monte Carlo simulation performed by the authors produces a VaR of USD 16 million, at a confidence level of 99%. The actual loss (USD 189 million) was 12 times as large.

In the case described above, stress testing could have been used as a simple method for analyzing the risk inherent in the transaction or for getting a feel for the risk implied. The depreciations shown in table 3 represent a scenario, i.e. precisely the scenario that then actually unfolded. Stress tests essentially revolve around defining scenarios and determining the changes in the value of a given financial instrument or a portfolio of financial instruments in the event of any one scenario.[1] Heavy-loss-producing scenarios are particularly relevant. Selecting adequate scenarios is integral to stress testing programs. Chapters 3 and 4 deal exclusively with how to identify scenarios. Based on considerations about which changes in the exchange rates could have adverse effects on the cash flow (1.1) of SK Securities, for instance three scenarios corresponding to a minor, midsize and major crisis (table 4) could have been defined, and it would have been fairly easy to calculate the resulting losses. The percentages shown in table 4 give the assumed depreciations of the currencies relative to the USD during the one-year life of the swap. The percentages in parentheses indicate the assumed IDR depreciations at a six-month cutoff.

Loss on the cash flow (1.1) in three different scenarios

	THB	IDR	JPY	Loss
Scenario 1: minor crisis	-15%	-15% (-8%)	0%	USD 58.0 million
Scenario 2: midsize crisis	-30%	-30% (-15%)	0%	USD 116.3 million
Scenario 3: major crisis	-50%	-50% (-30%)	0%	USD 183.9 million

Table 4

The results provide a considerably more drastic picture of the loss potential of the given transaction than the VaR measure of USD 16 million mentioned before. What is more, compared to the VaR figure, they are much easier to compute. Of course, the problem arises whether one believes a priori in the possible occurrence of the scenarios. A posteriori even scenario 3 seems perfectly realistic, yet the question remains whether the above scenarios would have been plausible in the eyes of SK Securities decisionmakers in early 1997. In this case study, considering the macroeconomic context would, no doubt, have put the assumption of constant exchange rate fluctuations in perspective.

1.5 Scope of this Guideline

This guideline is more or less confined to explaining stress testing as related to measuring and managing market risk. In how far such tests may account for or implictly cover liquidity crises is described in section 2.3. Credit risk is touched upon in section 2.4.

Compared to the wealth of publications on *value at risk,* the literature on *stress testing* is scarce. This will, however, most likely change in the future, not least because criticism of VaR models is mounting and stress testing is called for as an alternative or complementary measure to the internal models approach.

[1] *For details, see sections 2.1 and 2.2.*

From the banking supervisors' perspective, no concrete, international standards for stress testing are as yet in place, but various national supervisory authorities have started to pay more and more attention to this topic. At the current juncture, this guideline is designed to provide a rather general overview so as not to preclude future international developments. Chapter 5 outlines concrete requirements for Austrian credit institutions employing VaR models. Material new findings about the execution of stress tests as well as more concrete supervisory standards, once they evolve, will be considered in future editions.

There should always remain sufficient flexibility in carrying out stress testing though. A creative approach towards stress testing that builds on certain minimum requirements may only be conducive to risk management. In particular, the definition of stress scenarios is an ongoing, dynamic process that should involve experts of diverse fields. The Basle Committee on Banking Supervision (1996) clearly champions the idea of giving credit institutions adequate leeway in performing stress tests. This is why chapter 5 only lists the *minimum* requirements applicable to Austrian credit institutions. These requirements are in line with international standards.

2 General Aspects of Stress Testing

2.1 What is a Stress Test

The concept of stress testing is based on the notion that the value of a portfolio depends on *market risk factors (risk factors)*. Let us call the risk factors with an impact on the portfolio $r_1, r_2, ..., r_n$ and the function determining the value of the portfolio when the values of all risk factors are given, P. The values of the risk factors $r_1, r_2, ..., r_n$ characterize the market situation as far as it is of relevance to the portfolio. The risk factors may be combined into one single vector $\boldsymbol{r} := (r_1, r_2, ..., r_n)$ describing the market situation. In a market situation $\boldsymbol{r}$, the value of the portfolio is $P(\boldsymbol{r})$. Below, $\boldsymbol{r}_{MM}$ will stand for the vector representing the current values of the risk factors, i.e. the current market situation. MM stands for the "current market situation", $P(\boldsymbol{r}_{MM})$ therefore represents the current value of the portfolio.

A bank's portfolio may be considered to consist of its entire trading book. In such a case, stress testing may be said to be bank-wide. Under the Austrian Regulation on Internal Models for the Limitation of Market Risks, model users have to conduct stress tests quarterly and whenever a need arises. In practice, additional stress tests are frequently carried out for subportfolios at division, trading unit or dealer level or in respect of specific instruments (as in the case study in section 1.4). Lower-level stress tests are usually performed in response to specific needs and requested by the management responsible for the area concerned. The scenarios employed in such tests are customized to meet specific needs.

The choice of risk factors depends on the portfolio. Not all portfolios are influenced by the same risk factors. The number of risk factors must be chosen so as to include all parameters likely to have an impact on the value of the portfolio. One may, however, decide to use an even larger number, which may be wise as it allows the user to restructure his portfolio later without having to add more risk factors. The procedure for selecting risk factors is not clearly defined. The value of the portfolio may be understood as the function of several sets of risk factors. Where interest is concerned, for example, discount factors or interest rates may be chosen as risk factors. The function P depends on the portfolio: a different portfolio has a different valuation function. P is frequently not an explicit function of the risk factors. Particularly the values of portfolios of exotic options are usually determined in a valuation *process* rather than by means of a valuation *function*. One such valuation process would be the valuation of a portfolio or of single positions by means of a Monte Carlo simulation.

Stress tests answer the question of "What would happen if a market situation $\boldsymbol{r}$ suddenly occurred?" The scenario in this case is the sudden emergence of a market situation $\boldsymbol{r}$. Scenarios may therefore be identified with market situations and represented by vectors $\boldsymbol{r}$. In general language, a "scenario" is a potential future development. In connection with stress testing, a scenario is a possible future market situation. In this context, the term scenario therefore does not stand for a process but only for its outcome. This change in meaning is derived from the simulation of disturbances in financial markets. Such disturbances are characterized by a sudden confrontation of market participants with a changed market situation. This may have been caused, for example, by a dramatic rise in volatilities: when prices move so rapidly that market participants are unable to restructure their portfolios within the reaction time available, the portfolios have to be revalued on the basis of changed market conditions. The same effect occurs in liquidity crises: to a market participant, only those prices are of relevance at which he can rebalance his positions to the extent desired. In illiquid markets, trading close to quoted market prices is impossible. Therefore, a portfolio can be restructured only at a later time and at dramatically different prices. Even if quoted market prices fluctuate continuously, the prices relevant to market participants may still change dramatically in a liquidity crisis.

For stress testing, scenarios $\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k$ are selected according to specific criteria and calculations are made to determine the value of the current portfolio under these scenarios. These portfolio values are represented by $P(\boldsymbol{r}_1),\ldots,P(\boldsymbol{r}_k)$. By comparing them with the current value of the portfolio $P(\boldsymbol{r}_{MM})$ one can assess the losses that would be incurred if the market suddenly moved from $\boldsymbol{r}_{MM}$ to $\boldsymbol{r}_1,\ldots,\boldsymbol{r}_k$ without allowing a chance for rebalancing the portfolio.

2.2 Portfolio Valuation: Linear Approximation or Complete Revaluation

Analyzing scenarios means first of all to determine the value of a given portfolio on the assumption that the risk factors, instead of their real values $\boldsymbol{r}_{MM} = (r_{MM,1}, r_{MM,2},\ldots,r_{MM,n})$, have the values $\boldsymbol{r} = (r_1, r_2,\ldots,r_n)$ reflected by the scenario. In a *complete revaluation* of the portfolio, the valuation function is applied direct to the new values $\boldsymbol{r}$ of the risk factors. The value of the portfolio in a scenario $\boldsymbol{r}$ is then $P(\boldsymbol{r})$.

Linear approximation applies the sensitivities δ_i of the portfolio value relative to the individual risk factors. Sensitivities are numbers indicating for a specific risk factor how sensitive the value of a given portfolio is to changes in that risk factor. The higher the sensitivity, the heavier the impact of this factor on the value of the portfolio. Sensitivities are determined as follows: in a first step, "typical" changes $\Delta_1,\Delta_2,\ldots,\Delta_n$ are selected for all risk factors. Then the sensitivity δ_i is calculated for each risk factor:

$$\delta_i = \frac{P(r_1,\ldots,r_i,\ldots,r_n) - P(r_1,\ldots,r_i+\Delta_i,\ldots,r_n)}{\Delta_i}.$$

The sensitivities δ_i reflect the mean slope of the valuation function P across the distance Δ_i. They depend on the Δ_i selected if the valuation function P is non-linear in the *i*-th risk factor.

Different slopes for different Δ_i

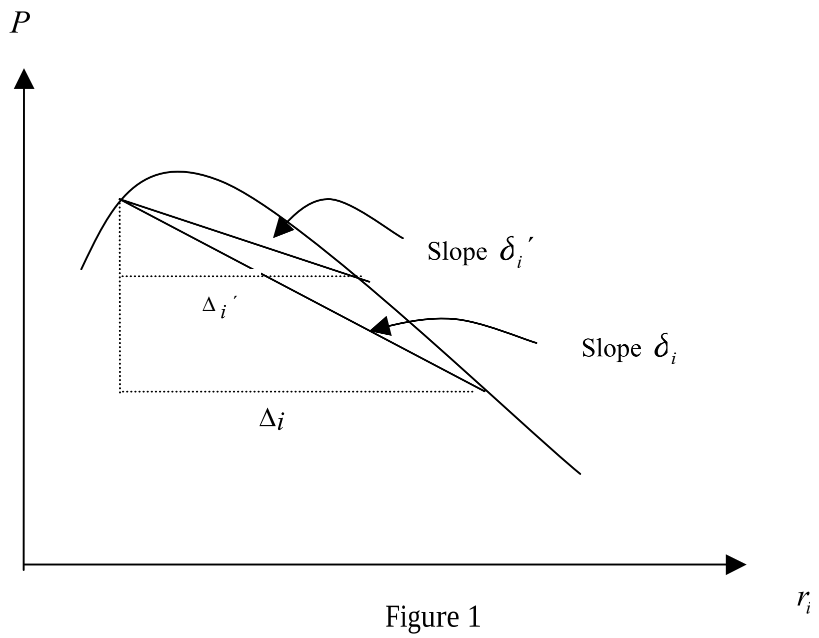

Figure 1

From the sensitivities an approximated portfolio value $\overline{P}$ is calculated by the following formula:

$$\overline{P}(r_1, r_2, ..., r_n) = P(r_{MM,1}, r_{MM,2}, ..., r_{MM,n}) + \sum_{i=1}^{n} (r_i - r_{MM,i})\delta_i .$$

$\overline{P}$ is the linear approximation of the valuation function around $\boldsymbol{r}_{MM}$. When is it permissible then to use a linear approximation of the portfolio value instead of a complete revaluation of the portfolio – and when is such an approach efficient?

Firstly, regarding the question of efficiency: calculation of the sensitivities requires n complete revaluations of the portfolio. If only a few scenarios have to be analyzed, complete revaluation is therefore more efficient and more precise than linear approximation. Approximation is more efficient only if a complete revaluation of the portfolio would require a large amount of calculations and if, beyond that, sensitivities have either been calculated before for other purposes and are therefore available without any extra effort or if the number of scenarios to be analyzed is much greater than the number n of risk factors.

Regarding the permissibility of linear approximation: the general rule is that linear approximation $\overline{P}(\boldsymbol{r})$ will not supply the correct value $P(\boldsymbol{r})$ of the portfolio in scenario $\boldsymbol{r}$ if the valuation function P is non-linear in those risk factors in which scenario $\boldsymbol{r}$ differs from the current situation $\boldsymbol{r}_{MM}$. The error in linear approximation will usually be small if for those risk factors in which the valuation function P is non-linear, the distance $r_i - r_{MM,i}$ is approximately equal to the typical distance Δ_i used in calculating sensitivity δ_i. If sensitivities are calculated specifically for linear approximation, it is therefore best to choose Δ_i for this purpose such that $r_{MM,i} + \Delta_i$ is close to r_i of the scenarios to be analyzed.

Linear approximation can be used with confidence only in scenarios in which changes from the current situation $\boldsymbol{r}_{MM}$ occur only in those risk factors on which the value of the portfolio depends linearly. Whether the value of the portfolio depends linearly on the risk factors is determined not only by the portfolio but also by the choice of risk factors. The value of a portfolio – understood as a function of specific risk factors – may indeed be linear in these factors whereas – if understood as a function of another set of risk factors – it is non-linear in that other set. There exists no portfolio that would be linear by nature.

Example:
The value of a bond depends linearly on the discount factors, but non-linearly on the underlying interest rates. If the discount factors are regarded as a risk factor, a bond portfolio is linear; if interest rates are chosen as risk factors, the bond portfolio is non-linear.

2.3 Liquidity Crises

Both the Basle Committee on Banking Supervision (1996; section B.5 no 3) and the Austrian Regulation on Internal Models for the Limitation of Market Risks (§ 7 para 2) require that liquidity crises be taken into account:

> "Stress tests should [...] incorporate both market risk and liquidity aspects of market disturbances."

Basically, one can distinguish between two types of liquidity risk: firstly, a bank may suddenly lack the financial liquidity allowing it to keep *holding* certain positions. Due to a changed market situation, it may, for example, suddenly be faced with the need to make margin payments or to provide additional security. Avoidance of this type of liquidity crisis is the responsibility of asset/liability management and will not be discussed any further in this context. Secondly, a shortfall in market liquidity may suddenly occur, preventing the bank from *closing* certain positions. When that happens, it becomes impossible to find a party willing to take up the position at the quoted market price. In such a situation the position cannot be closed at all or only with an extremely high bid-ask spread. Here we want to discuss only the second type of liquidity risk, namely the lack of market liquidity.

A lack of market liquidity may be attributable to several causes: some markets are traditionally illiquid. Other, normally liquid markets, may occasionally suffer liquidity shocks triggered, for example, by unexpected economic or political news. Finally, a market participant's exposure in a specific market may be so substantial that closing of his positions destroys the liquidity of the market.

Whatever the reason for inadequate market liquidity may be, illiquid markets do not allow any trading close to quoted market prices. Any restructuring of the portfolio – either now or later – will therefore be possible only at dramatically different prices. The only prices relevant to a portfolio manager are those at which he is able to restructure his positions to the desired extent. Even if quoted market prices are moving continuously, the prices relevant to the portfolio manager may change *dramatically* in a liquidity crisis. In a market risk crisis, the situation facing a portfolio manager is exactly the same: a dramatic rise in volatility causes prices to change so rapidly that, given his limited reaction speed, he can rebalance his positions to the desired extent only at dramatically different prices. In stress situations, both liquidity risk and market risk have the same negative consequences, namely dramatic changes in the market that make continuous restructuring of the portfolio impossible. To the portfolio manager, it does not make any difference whether the market suddenly changes overnight and he can rebalance his positions only the next day or whether, in a situation of creeping market changes, he can rebalance his positions only much later because of insufficient market liquidity.

Both situations – liquidity crisis and market risk crisis – are simulated in stress tests by revaluing a given portfolio against a background of radically changed market conditions. Liquidity stress tests therefore do not require any special methodology.

Nevertheless, the simulation of liquidity crises may call for different scenarios than the simulation of market risk crises. If, for example, in simulating a market risk crisis, historical data are used to assess the magnitude of moves in single risk factors, one will probably choose the greatest day-to-day changes or, even better, the greatest changes that occurred within the bank's response time. In simulating liquidity crises one will tend to select the scenario with the greatest change within a period of time equivalent to the duration of the liquidity crisis. The n-day drawdown defined in section 3.2 is an upper limit for moves in risk factors during a liquidity crisis of a maximum duration of n days.

2.4 Credit Risk

The Basle Committee on Banking Supervision (1996; section B.5 no 2) calls for the consideration of credit risks in stress testing:

> "Banks' stress scenarios need to cover a range of factors that can create extraordinary losses or gains in trading portfolios or make the control of risk in those portfolios very difficult. These factors include low-probability events in all major types of risk, including the various components of market, credit and operational risks."

Why should risk types, such as credit risk, that are not captured by the value-at-risk model used for market risk control be included in stress tests? Taking the combined action of market and credit risks into account is very important, as a separate consideration of market and credit risks may fail to identify some material dangers. Value-at-risk models including such a capability are still in the process of development. Hopes are therefore pinned on stress testing.

The combined action of market and credit risks can be illustrated by an example: in the first half of 1998, a number of western banks entered into ruble forward deals with Russian banks under which they agreed to buy from the Russian banks, on a specified settlement date, dollars against rubles at a specified forward exchange rate. Most of these deals were fully hedged by offsetting transactions with other western banks. *The market risk of such deals – ignoring the default risk – was therefore practically zero*. The default risk in respect of the Russian banks was limited to the difference between the agreed ruble exchange rate at which dollars were to be delivered to the western banks and the replacement cost in rubles (i.e. the spot rate on the settlement date) of dollars not delivered by the Russian banks. The agreed forward rate was usually very close to the spot rate prevailing at the time the deal was closed as exchange rates had remained unchanged for a long time and were therefore not expected to fluctuate in the future either. As long as there was no change in the ruble exchange rate, the default risk in respect of the Russian banks was close to zero as any dollars not delivered by the Russian banks could be bought in the market at very similar prices. *Therefore, the default risk of these deals – ignoring the market risk – was also practically zero*. Separate measurements of market risk and default risk show both risks to be practically zero. A look at the combined action of market risk and default risk, however, reveals the following situation: if the ruble exchange rate declines and a Russian bank defaults *at the same time*, dollars have to be bought in the market at the high ruble spot rate and delivered to the western banks at the low forward rate. A market risk was therefore created only through the default of the Russian banks. Positions that had been closed were suddenly reopened. *The combined action of market and default risks may lead to enormous losses*. This example shows the great importance of an integrated assessment of market and credit risks.

The example shows a well-known interaction between credit and market risk at work: changes in market risk factors result in changes in the values of assets and liabilities held by counterparties and thus to changes in the losses incurred in the event of default. On the other hand, the default of a large market player may also trigger strong fluctuations in market risk factors.

How can the default risk be incorporated into stress testing? For this purpose, an assessment is needed of how credit losses are influenced by market risk factors. This would in fact require an integrated credit and market risk model. A number of credit risk models, including McKinsey's CreditPortfolioView™ and KMV's PortfolioManager™, take the current state of the economy and a variety of market risk factors into account. Even these models, however, are not integrated credit and market risk models.

A relatively simple way of covering default risk in stress testing is the following: for worst-case scenarios, it is justified to use the simplifying assumption that the loss due to a counterparty's default is equal to the full market value of all assets, i.e. that nothing can be recovered from a defaulting counterparty. In selecting a credit stress scenario, two parameters have to be specified: (1) the values of the market risk factors and (2) the defaulting counterparties. The loss in such a scenario is then calculated as follows: firstly, the trading book subportfolio affected by the default is determined. For counterparties with which netting arrangements are in effect, such a subportfolio consists of all positions transacted with the respective counterparty. For counterparties with whom no netting agreements have been entered into, the subportfolio comprises all positions transacted with the counterparty concerned and having a positive market value. In a second step, the subportfolio affected by the default is valued, using the risk factor values chosen in (1).

2.5 How Tough Should Stress Scenarios Be

On the one hand, it is of course in the nature of stress tests to ask what would happen in situations that nobody expects to occur. On the other hand, test results from scenarios that are regarded as highly unlikely are not taken seriously by those to whom test reports are addressed. With this in mind, it may be helpful to run several scenarios of different degrees of severity. For the risk management of the credit institution concerned it is important to apply clear criteria in specifying scenarios and to account for these criteria in interpreting the outcome of stress testing. The recipient of a report should not be given just the mere loss figures but should also be alerted to the severity of the underlying scenarios. Where possible, senior management should participate in defining the severity of the scenarios.

2.6 Standardized Stress Tests

Many banks conduct periodic stress tests involving a revaluation of their current portfolio against certain standard scenarios. These are often standard scenarios in a dual sense: the choice of the scenarios depends neither on the bank nor on the timing of the stress test.

Thus, stress testing with standard scenarios has the advantage of guaranteeing comparability in two respects. Firstly: when several banks look at the same scenarios one can compare the outcome of stress tests of different banks. This allows the supervisor to assess the banks' exposure to those risk categories whose risk factors are changed in the standard scenarios. Secondly: when a bank always looks at the same scenarios, it can compare the results of stress tests conducted at different points in time. This enables it to monitor how its exposure to the risk categories in the standard scenarios changes over time (exposure monitoring).

Many banks use standard scenarios similar to the stress scenarios proposed by the Derivatives Policy Group (DPG). The DPG is an informal body of representatives of major American banks and investment firms. It was set up in August 1994, at the suggestion of the Securities and Exchange Commission, to formulate a code of conduct for trading in derivatives. Its rules were published in the "Framework for Voluntary Oversight."

The DPG recommends the performance of stress tests to measure the exposure of a portfolio to certain core risk factors. The DPG lists among these core risk factors

i. parallel yield curve shifts,
ii. changes in the steepness of yield curves,
iii. parallel yield curve shifts combined with changes in the steepness of yield curves,
iv. changes in yield volatilities,
v. changes in the value of equity indices,
vi. changes in equity index volatilities,
vii. changes in the value of key currencies (relative to the USD),
viii. changes in foreign exchange rate volatilities and
ix. changes in swap spreads in at least the G-7 countries plus Switzerland.

For an assessment of exposure towards the core risk factors, the DPG (1995; section 4 no 4) recommends use of the following standard scenarios in regular stress testing:

a) parallel yield curve shifts of 100 basis points up and down,
b) steepening and flattening of the yield curves (for maturities of 2 to 10 years) by 25 basis points,

c) each of the four permutations of a parallel yield curve shift of 100 basis points concurrent with a tilting of the yield curve (for maturities of 2 to 10 years) by 25 basis points,
d) increase and decrease in all 3-month yield volatilities by 20 percent of prevailing levels,
e) increase and decrease in equity index values by 10 percent,
f) increase and decrease in equity index volatilities by 20 percent of prevailing levels,
g) increase and decrease in the exchange value (relative to the USD) of foreign currencies by 6 percent, in the case of major currencies, and 20 percent, in the case of other currencies,
h) increase and decrease in foreign exchange rate volatilities by 20 percent of prevailing levels and
i) increase and decrease in swap spreads by 20 basis points.

A comparison of these DPG standard scenarios with the tables in chapter 3 listing actual maximum changes shows that some of the DPG scenarios are far removed from the maximum changes observed in the past. Therefore, they should not be regarded as reconstructions of historical crises or as worst-case scenarios.

Neither the Basle Committee on Banking Supervision nor the Austrian Regulation on Internal Models for the Limitation of Market Risks require banks to perform stress tests at regular intervals with standards scenarios like the DPG's. Nevertheless periodic stress tests with unchanged scenarios may serve as a useful instrument in monitoring exposures on an ongoing basis. The same can be said of stress test limits. Such limits specify, for a certain unchanging set of scenarios, the maximum loss acceptable with each scenario and what action to take in case the limit is exceeded.

To date, the Austrian bank supervisory authority has not specified any standard scenarios for stress testing. However, the authors would recommend credit institutions to develop their own scenarios for continuous monitoring of exposure in their respective key markets.

2.7 Interpretation of the Results of Stress Tests, Reporting and Contingency Planning

Stress tests are used primarily for the assessment of a bank's capital situation and the identification of measures designed to minimize risk. The Basle Committee on Banking Supervision (1996; section B.5 no 3) notes the following in this context:

> "Qualitative criteria should emphasise that two major goals of stress testing are to evaluate the capacity of the bank's capital to absorb potential large losses and to identify steps the bank can take to reduce its risk and conserve capital. This assessment is integral to setting and evaluating the bank's management strategy and the results of stress testing should be routinely communicated to senior management and, periodically, to the bank's board of directors."

In interpreting the results of stress tests the first question will therefore be whether the bank would be able to cope with the losses incurred in a stress scenario. A comparison of the outcome of the stress test with the bank's own capital resources may in some circumstances be misleading, however, as these funds also need to cover risks other than the market risk associated with the trading book. If at a time of market disturbance other losses were being incurred simultaneously, the bank might be in trouble even if its own capital were adequate for coping with the market crisis alone. In an alternative approach, the results of stress tests are therefore frequently compared with risk capital allocated internally for securities trading or with the regulatory capital requirements in respect of market risk associated with the trading portfolio (10-day VaR times multiplication factor).

If, in the event of a market disturbance, any loss incurred is higher than the risk capital allocated for securities trading or the regulatory capital requirements in respect of the market risk associated with the trading portfolio, the bank needs to take urgent action. In this regard, the plausibility of stress scenarios is certainly a critical factor. If a stress scenario is highly plausible, senior management will take a stress test more seriously than if it considers the stress scenario highly unlikely.

Stress tests gain practical significance only when their results are taken note of and understood by the bodies having the authority to call for a reduction of risk exposure. The Basle Committee on Banking Supervision (1996; section B.5 no 8) notes the following in this regard:

> "The results should be reviewed periodically by senior management and should be reflected in the policies and limits set by management and the board of directors. Moreover, if the testing reveals particular vulnerability to a given set of circumstances, the national authorities would expect the bank to take prompt steps to manage those risks appropriately (e.g. by hedging against that outcome or reducing the size of its exposures)."

Likewise, the Austrian Regulation on Internal Models for the Limitation of Market Risks calls for the following in § 2 para 6:

> "Where stress tests reveal vulnerability to a given set of circumstances, prompt steps shall be taken to manage those risks appropriately. The basic features of the procedure shall be outlined in the risk management handbook."

And in § 7 para 2:

> "Qualitative criteria shall be used to evaluate the extent to which the credit institution's own funds may be applied to absorb potential large losses. Furthermore, measures shall be developed whereby the credit institution can reduce its risk and avoid losses."

Stress tests conducted at regular intervals with unchanged scenarios are a suitable instrument for continuous monitoring of risk exposure. In markets or regions in which a bank's exposure is particularly large, such exposure is frequently monitored by means of periodic stress tests, usually by running worst-case scenarios for each market. Stress test limits specify the permissible magnitude of losses in each scenario and the steps to be taken when maximum allowable losses are exceeded in a stress test. In such a context, it is less important that the worst-case scenario predicts exactly how a real disturbance evolves in a given market. Rather, it is more important that the loss resulting from a supposed worst-case scenario is a good measure of the bank's exposure in the respective market.

Stress tests with scenarios in which risk factors are changed in a large number of different markets do not lead to any immediate practical consequences. With such a scenario, the mere awareness that an alarming loss may be incurred does not yet allow any conclusions to be drawn in respect of the nature of the risk factors or positions that actually cause the loss. As long as this is not understood, it remains unclear how positions might be hedged to reduce potential losses. Mere conjectures are not enough for effective risk management.

Once the risk factors contributing most heavily to losses in a worst-case scenario have been identified, it is possible to take well-targeted countermeasures. The bank can then take up positions that will make a profit when key risk factors are at their worst-case levels. Section 4.4.2 describes how key risk factors for worst-case scenarios are identified.

3 Construction of Stress Scenarios Using Historical Data

3.1 Why Use Historical Scenarios

The Basle Committee on Banking Supervision (1996; section B.5 no 6) requires the construction of stress scenarios on the basis of historical crises:

> "Banks should subject their portfolios to a series of simulated stress scenarios and provide supervisory authorities with the results. These scenarios could include testing the current portfolio against past periods of significant disturbance, for example, the 1987 equity crash, the ERM crises of 1992 and 1993 or the fall in bond markets in the first quarter of 1994, incorporating both the large price movements and the sharp reduction in liquidity associated with these events."

The Austrian Regulation on Internal Models for the Limitation of Market Risks also states in § 7 para 3 lit 2 that the Minister of Finance may obtain from the banks information on internal stress tests which test the portfolio against periods of significant market disturbances in previous years. The meaning of the provision is that the banks have to conduct stress tests based on historical scenarios, and that the Minister of Finance can request information on these tests.

One may ask now, why use reconstructions of historical crises? After all, value-at-risk models also use historical data. If stress tests use the same data as value-at-risk models, then why should the informative value of stress test results differ from that of VaR model results?

One major difference between the two methods is that value-at-risk models usually include only data from a relatively short previous period – e.g. the previous year – while stress testing can be used to reconstruct exceptional market situations which occurred at a more distant point in the past. And because VaR models use all data from the recent past, including calm market periods, peaks tend to be smoothed out. Conversely, when historical crises are modeled, only periods of dramatic market movements are taken into account, while data from uneventful periods are left out. As a result, the peaks of market movements can be modeled in full force.

One advantage of historical scenarios over worst-case scenarios is that the former describe events which have actually occurred, and the recipients of such stress test reports cannot therefore ignore the test results, arguing that such scenarios will never occur, anyway.

The construction of stress scenarios using historical data is based on the assumption that past crises are similar to future ones. The use of historical data would not make sense without this assumption, which is a version of another, more general assumption that is often applied in risk management – namely, that the future is like the past.

Generally, we cannot base plans for the future on any other evidence than that from the past. It may be dangerous, however, to take for granted the continuity of past developments. Let us consider the example of the Asian crisis, putting ourselves into the position of a risk manager in early 1997. Looking at exchange rates over the previous decade, we find the following maximum movements in relation to ATS:

Maximum absolute values of changes in exchange rates between selected Asian currencies and ATS from January 1, 1987 to December 31, 1996; source: Datastream

	Maximum one-day change	Maximum ten-day change	Maximum twenty-day change
IDR	5.3%	7.5%	11.0%
MYR	3.4%	7.6%	11.2%
PHP	7.0%	10.4%	13.6%
KRW	7.7%	8.4%	12.0%
THB	5.9%	7.3%	11.1%

Table 5

Maximum *n*-day change: See section 3.2.

This long time series would have provided no indication whatever of what was going to happen shortly thereafter. The maximum exchange rate variations over the next two years were:

Maximum absolute values of changes in exchange rates between selected Asian currencies and ATS from January 1, 1997 to December 31, 1998; source: Datastream

	Maximum one-day change	Maximum ten-day change	Maximum twenty-day change
IDR	22.6%	59.6%	70.9%
MYR	30.1%	29.5%	30.6%
PHP	10.9%	13.6%	20.4%
KRW	22.0%	34.6%	41.8%
THB	7.2%	26.8%	27.7%

Table 6

It is obvious that stress scenarios which use past maximum changes as a yardstick for prospective stress events, may seriously underestimate the potential impact of such future crises. In the above example, the reason for this is clear: For a long time, the involved currencies had been more or less closely pegged to a hard-currency basket. Consequently, exchange rates had shown very little variation in the past. But when the central banks were unable to maintain their exchange rate policies during the Asian crisis, exchange rates all of a sudden began to experience extreme movements.

Stress scenarios which use historical data model stress events based on extreme historical market movements. But this approach involves some danger: extreme scenarios are not necessarily worst-case scenarios; the maximum changes are not necessarily the worst that can happen. Certain portfolios are more dramatically affected by slight changes than by major movements. And if one pursues a straddle strategy, for example, no market movement at all is the worst that can happen.

3.2 Analysis of Time Series of One Factor

The simplest way of constructing scenarios from historical data is to determine the maximum change for each risk factor and then combine the results in a scenario.

3.2.1 Identifying Maximum Movements of Individual Factors

For the purpose of constructing scenarios from historical data, the *historical observation period* is defined as the period over which the time series is considered (e.g. 1 year, 10 years). The historical observation period is overlaid with *time windows* of equal duration (e.g. 1 day, 20 days). If for example, the duration of each time window is 20 days, the first time window comprises the period from the first to the twentieth day; the second time window the period from the second to the twenty-first day, and so on. For each time window within the historical observation period, a change parameter[2] is determined. The maximum or minimum of the change parameters for all time windows within the historical observation period is then identified as the assumed change Δr_i of the risk factor in question.

The change parameter is most commonly defined as the change from the first to the last day of the time window. This parameter is called *Start to End* (StE) in the tables below. For a time window with a duration of 20 days, for example, this is the maximum *20-day change*. Alternatively, the change parameter can also be defined as the maximum of all changes occurring between two points within the time window. Following Acar and James (1997), this parameter is called *drawdown* (DD) in the tables below.

If the minimum of the change parameters is selected, Δr_i is equivalent to the maximum reduction of the risk factor r_i. If the maximum of the change parameters is selected, Δr_i is equivalent to the maximum increase. Finally, Δr_i can also be defined as the maximum of the absolute values of the change parameters. In this case, Δr_i is equal to the value of the maximum change, regardless of whether the change was upward or downward. This appears to be useful if

[2] A parameter which measures the change of a given risk factor within the time window.

one assumes that realized changes of the risk factors might just as well have occurred in the other direction. For the tables in section 3.2.3, Δr_i was determined in this way.

Furthermore, one has to know quite clearly whether one is interested in relative or absolute changes. An extreme change in the past can be translated into different present changes, depending on whether the relative or the absolute change is considered. The general practice is to consider the relative changes, expressed as percentages, when dealing with stock prices and currency exchange rates, whereas absolute changes expressed in basis points are used in the case of interest rate movements. In this paper, we have followed this convention.

The construction of a scenario from historical data thus requires at first a selection of several parameters: the historical observation period, duration of the time window, and the change parameter. These parameters may be different for different risk factors, and there are no universally accepted standards for the selection of parameters. The same set of historical data may be used to construct quite different scenarios. No matter which parameters a bank selects, the selection will make a difference for the scenario and has to be taken into consideration in the interpretation of stress test results. We will now discuss the effect which the selection of parameters has on the resulting scenarios.

The longer the historical observation period, the more extreme are the maximum movements.
The reason for this is obvious: The biggest movement among a greater set of market movements will be bigger than the biggest movement among a *sub*set. This is illustrated by the table below, which presents the maximum absolute values of one-day changes within different periods.

Maximum absolute values of one-day changes of selected stock price indeces from January 1, 1987/January 1, 1994 to December 31, 1998; source: Datastream

	5 years back	**12 years back**
Austria	8.3%	8.9%
USA	7.2%	22.6%
Great Britain	4.4%	12.2%
Germany	8.0%	12.8%
Japan	8.0%	14.9%

Table 7

Indeces used: Austria: ATX, USA: Dow Jones Industrials, Great Britain: FTSE 100, Germany: DAX 30, Japan: Nikkei 225.

The maximum relative change is not necessarily the same as the maximum absolute change. A modest absolute change on a low level may be bigger in terms of percentage points than a bigger absolute change on a higher level.

Depending on the change parameter, one and the same time series will result in different maximum changes. The maximum one-day change will be different from the maximum ten-day or twenty-day change. The maximum ten-day drawdown will be at least as big as the maximum ten-day change. This is reflected in the table below, which presents the maximum absolute values of ATS exchange rate changes. It is noteworthy that Δr_i is biggest for the drawdown. In a small number of cases, the maximum one-day change can be bigger than the maximum ten-day change.

Maximum absolute values of changes of selected exchange rates in relation to ATS from Januar y 1, 1994 to December 31, 1998; source: Datastream

	StE 1 day	StE 10 days	DD 10 days
Hungary	8.4%	8.2%	8.9%
Czech R.	7.2%	7.6%	7.8%
Mexico	17.5%	38.7%	38.7%
Malaysia	30.1%	29.5%	30.3%
South Korea	22.0%	34.6%	37.9%
Japan	5.1%	16.8%	16.8%
USA	3.8%	6.5%	6.8%
GB	2.3%	5.9%	5.9%

Table 8

3.2.2 Integrating the Movements of Individual Factors into a Scenario

Scenarios specify values not only for one, but for all risk factors. When constructing scenarios from the maximum movements of a number of individual risk factors, one therefore has to decide what is going to happen to the other remaining risk factors.

In an attempt to create a worst-case scenario, one could subject each risk factor to the maximum change which has been found within a certain observation period through applying the above methods. So, if Δr_i is the maximum change which the *i*-th risk factor has experienced within the observation period, the resulting stress scenario is:

$$\boldsymbol{r} = (r_{MM,1} \pm \Delta r_1,\ r_{MM,2} \pm \Delta r_2, \ldots, r_{MM,n} \pm \Delta r_n).$$

A plus or minus sign must be selected for each risk factor, depending on whether its change in the stress scenario is to be upward or downward. This leads to 2^n possible results, i.e. an indeterminably large number. So which risk factors should be changed upward, and which ones downward? For example, each risk factor may be changed in the same direction in which it actually changed when it made its greatest leap. Another approach is to group together related risk factors (e.g. all stocks or stock indeces of one region with strong internal economic ties), and to change all risk factors belonging to one group in the same direction.

Another combination possibility is to subject only selected risk factors $r_{i_1}, r_{i_2}, \ldots, r_{i_w}$ to the maximum change and leave the others unchanged. This results in scenarios of the following type:

$$\boldsymbol{r} = (r_{MM,1}, \ldots, r_{MM,i_1} \pm \Delta r_{i_i}, \ldots, r_{MM,i_2} \pm \Delta r_{i_2}, \ldots, r_{MM,i_w} \pm \Delta r_{i_w}, \ldots, r_{MM,n}).$$

The process can be further streamlined if the risk factors in one group are not only assigned the same sign (all pluses or all minuses), but if they are also subjected to a change of identical size. If this is done, the group change must adequately cover the individual changes of the risk factors in the group. This also simplifies reporting.

There are two problems with the combination of extreme movements of individual factors into a scenario. Firstly, the resulting scenario is not necessarily a worst-case scenario, because the portfolio might suffer greater damage if certain risk factors were to move in the opposite direction, or if movements were less extreme. Secondly, combining all risk factors means to subject them all at the same time to the maximum change which they have undergone at some point within the observation period. As a consequence, the resulting scenario may be impossible or highly implausible.

The scenario may be impossible because, while the individual risk factors may have undergone the maximum changes at different points in time, it may be impossible for them to experience these maximum movements at the same time.

The resulting scenario will in general also be more improbable by several orders of magnitude than the movements of the risk factors taken individually. This will be the case if the resulting scenario conflicts with existing correlations or expected correlations in stress events. But this fact is not necessarily an argument against using the resulting scenario as a stress scenario. Stress tests are, after all, not supposed to provide quantitative information about the probability of the used scenarios. Rather, they should inform about the consequences of low-probability events. Scenarios which are constructed by combining the maximum changes of the individual risk factors, can be used to test for the collapse of existing correlations. However, the plausibility of the resulting scenario will be so small – especially if a large number of risk factors is involved – that the recipients of the report are unlikely to attach serious importance to the stress test outcome.

Section 3.3 presents an analysis of the simultaneous movements of several risk factors which circumvents the problem that the combination of maximum movements in a scenario may be much more improbable than the maximum movements of the factors by themselves.

Another approach is proposed by Kupiec (1998). Selected risk factors $r_{i_1}, r_{i_2}, \ldots, r_{i_w}$ can be subjected to the maximum change, while the others are selected in accordance with the prevailing correlations. The resulting scenario is more plausible than if the remaining risk factors are left unchanged or subjected to their maximum changes.

3.2.3 Tables of Maximum Changes of Individual Risk Factors

The tables below present examples of maximum absolute values of changes in individual risk factors in the risk categories stocks, foreign exchange and interest rates. The risk factors are selected stock price indeces, exchange rates vis-à-vis ATS, and interest rates. Most time series start on January 1, 1987 and end on December 31, 1998. Historical observation periods of two, five and twelve years are used, with one-day and twenty-day time windows. For the one-day time window, the maximum one-day change (StE) is stated; for the twenty-day time window, the maximum 20-day change (StE) and the maximum drawdown within 20 days (DD) are given. No difference has been made between positive and negative changes, i.e. only the absolute value of the changes has been considered. (Relative changes of more than 100%, which occur occasionally, are therefore due to growth).

This selection of parameters is not meant as a recipe to be followed by risk managers when making their own specific selection of parameters. Rather, the data is intended to provide some orientation for the selection process. Data from different data suppliers may of course yield slightly different results. What is important in any case is the reliability of the data material used, as indicators describing extreme movements are highly sensitive to outliers.

Wherever gaps appear in the tables, they were caused by missing data. Data was supplied by Datastream.

3.2.3.1 Maximum Changes of Stock Price Indeces

The table reflects two historical crisis periods: firstly, the 1987 equity crash which is contained in the 12-year observation period from January 1, 1987 to December 31, 1998. In the Western countries, maximum changes in this period are substantially larger than in the other two periods. Secondly, it is notable that the maximum changes in the two-year and five-year observation periods are the same in the Western and the Asian countries. This is due to the fact that the maximum changes in both periods can be attributed to crises that occurred in 1997 and 1998 (Asian crisis, Russian crisis).

Maximum absolute values of changes of selected stock price indeces; source: Datastream

	Jan. 1, 1997 - Dec. 31, 1998			Jan. 1, 1994 - Dec. 31, 1998			Jan. 1, 1987 - Dec. 31, 1998		
	StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D
USA (1)	7.2%	15.3%	15.3%	7.2%	15.3%	15.3%	22.6%	31.0%	34.2%
USA (2)	7.2%	17.6%	17.6%	7.2%	17.6%	17.6%	21.1%	29.5%	32.8%
Great Britain	4.4%	18.9%	19.7%	4.4%	18.9%	19.7%	12.2%	33.1%	33.4%
Germany	8.0%	23.5%	24.3%	8.0%	23.5%	24.3%	12.8%	37.0%	37.0%
Japan	8.0%	17.3%	17.3%	8.0%	17.5%	17.5%	14.9%	29.1%	32.1%
Canada	6.2%	20.2%	20.2%	6.2%	20.2%	20.2%	11.3%	27.5%	27.6%
Australia	7.2%	16.9%	17.3%	7.2%	16.9%	17.3%	25.0%	47.3%	47.3%
Austria	8.3%	21.6%	21.6%	8.3%	21.6%	21.6%	8.9%	32.0%	32.0%
Netherlands	5.9%	26.2%	28.0%	5.9%	26.2%	28.0%	12.0%	41.1%	41.1%
Italy	12.4%	25.6%	25.6%	15.0%	25.6%	25.6%	15.0%	27.9%	28.4%
Hongkong	18.8%	39.8%	40.1%	18.8%	39.8%	40.1%	33.3%	50.2%	50.2%
Indonesia	14.0%	56.1%	61.6%	14.0%	56.1%	61.6%	119.5%	161.2%	161.2%
Malaysia	23.1%	52.3%	69.4%	23.1%	52.3%	69.4%	23.1%	52.3%	69.4%
Singapore	9.2%	31.9%	31.9%	9.2%	31.9%	31.9%	9.2%	34.5%	34.5%
Switzerland	7.7%	29.6%	29.6%	7.7%	29.6%	29.6%			
France	6.3%	21.5%	24.5%	6.3%	21.5%	24.5%			
Poland	9.8%	31.6%	31.6%	15.9%	57.4%	61.7%			
Hungary	16.5%	54.4%	55.9%	16.5%	78.4%	78.4%			
Slovenia	9.3%	34.8%	40.3%	9.4%	40.2%	41.8%			
Slovakia	10.0%	16.0%	16.0%	31.7%	193.6%	193.6%			
Czech R.	6.8%	24.7%	26.4%						

Table 9

Indeces used: USA: Dow Jones Industrials (first line), S&P 100 (second line), Great Britain: FTSE 100, Germany: DAX 30, Japan: Nikkei 225, Canada: TSE 300, Australia: Australian All Ordinaries Index, Austria: ATX, Netherlands: AEX, Italy: MIB, Hongkong: Hang Seng, Indonesia: Jakarta Composite Index, Malaysia: Kuala Lumpur Composite Index, Singapore: SES All Singapore, Switzerland: Swiss Market Index, France: CAC 40, Poland: Warsaw General Index, Hungary: BUX, Slovenia: SBI, Slovakia: SAX, Czech Republic: PX 50.

Time windows: 1 day (1D) and 20 days (20D).

StE: Start to End, DD: drawdown.

3.2.3.2 Maximum Changes of Exchange Rates

As already mentioned in connection with the example in section 1.4, exchange rate movements may be heavily influenced by the exchange rate policy of the involved central banks. A currency that is pegged to a key currency may experience very limited exchange rate changes over a long time period. But if the exchange rate policy is abandoned, extreme gyrations can occur suddenly. In constructing scenarios, one should therefore consider in which way and how strongly currencies are pegged; the overall situation of the national economies in question should also be taken into account, as it may provide some indication of whether central banks are about to let hitherto pegged currencies float. The International Monetary Fund annually publishes a classification of the various exchange rate systems in its *Annual Report on Exchange Arrangements and Exchange Restrictions.* Annual Reports can be accessed at the IMF's Internet address (http://www.imf.org; see Publications).

As regards exchange rates, the choice of reference currency is important: for example, a devaluation by 50% of a foreign currency vis-à-vis ATS is equivalent to a 100% appreciation of ATS against the relevant foreign currency. The table below reflect value changes of foreign currencies vis-à-vis ATS. Because of the asymmetry of the reciprocal value, the value changes of ATS in relation to the foreign currencies would be different.

Maximum absolute values of changes in selected exchange rates in relation to ATS; source: Datastream

	Jan. 1, 1997 - Dec. 31, 1998			Jan. 1, 1994 - Dec. 31, 1998			Jan. 1, 1987 - Dec. 31, 1998		
	StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D
USD	2.9%	7.2%	7.5%	3.8%	9.3%	9.3%	3.8%	12.4%	12.4%
GBP	1.9%	7.1%	7.1%	2.3%	7.1%	7.1%	4.5%	14.7%	15.3%
JPY	5.1%	18.1%	18.1%	5.1%	18.1%	18.1%	5.1%	18.1%	18.1%
CHF	1.4%	2.7%	3.1%	1.4%	3.2%	3.2%	2.1%	4.7%	4.8%
CAD	2.8%	8.9%	9.9%	3.4%	11.0%	11.0%	3.4%	12.9%	12.9%
AUD	3.8%	8.4%	8.9%	3.8%	10.3%	10.3%	5.6%	14.7%	14.7%
HKD	2.1%	7.1%	7.3%	3.0%	9.5%	9.5%	3.8%	12.0%	12.0%
SEK	2.0%	6.8%	6.9%	2.0%	6.8%	6.9%	7.7%	13.0%	13.1%
SGD	3.6%	9.9%	11.3%	3.6%	9.9%	11.3%	3.6%	9.9%	11.3%
ZAR	6.1%	18.8%	19.1%	6.1%	18.8%	19.1%	6.1%	18.8%	19.1%
GRD	7.4%	10.8%	10.8%	7.4%	10.8%	10.8%	7.4%	10.8%	10.8%
MXP	6.6%	18.3%	18.4%	17.5%	41.3%	41.3%	22.4%	41.8%	41.8%
ARS	2.1%	7.1%	7.3%	3.0%	9.6%	9.6%			
MYR	30.1%	30.6%	30.6%	30.1%	30.6%	30.6%	30.1%	30.6%	30.6%
THB	7.2%	27.7%	28.7%	7.2%	27.7%	28.7%	7.2%	27.7%	28.7%
RUB	42.6%	72.9%	159.7%	42.6%	72.9%	159.7%			
PLZ	4.7%	9.8%	9.8%	4.7%	9.8%	9.8%			
HUF	2.1%	5.2%	5.2%	8.4%	10.4%	10.4%			
SKK	6.5%	12.0%	12.0%	6.5%	12.0%	12.0%			
SIT	2.5%	2.8%	4.0%	6.8%	8.4%	9.6%			

Table 10

Currencies are named by their ISO codes.
ARS: Argentina, AUD: Australia, CAD: Canada, CHF: Switzerland, GBP: Great Britain, GRD: Greece, HKD: Hongkong, HUF: Hungary, JPY: Japan, MXP: Mexico, MYR: Malaysia, PLZ: Poland, RUB: Russia, SEK: Sweden, SGD: Singapore, SIT: Slovenia, SKK: Slovakia, THB: Thailand, USD: USA, ZAR: South Africa.

Time windows: 1 day (1D) and 20 days (20D).

StE: Start to End, DD: drawdown.

3.2.3.3 Maximum Changes of Interest Rates

Yield curves display different dynamics, depending on debtors' credit quality. It is therefore necessary to use interest rates which fit in with debtors' credit quality when constructing stress scenarios in the area of interest rates.

For the purposes of the section below, only risk-free interest rates have been considered. As an approximation of risk-free interest rates, the table below uses interbank rates for the money market and the yield-to-maturity of benchmark bonds for the capital market. This approximation is certainly only a rough one, and it has been criticized as inadequate – particularly in extreme situations – by some authors (see Brooks and Yong Yan (1999)). It has been chosen, however, because it is simple (the relevant time series are available in Datastream) and because – as mentioned above – the tables are meant only as a general orientation for the construction of scenarios from historical data. For a more realistic picture, the actual risk-free interest rates have to be used. These rates can be calculated, for example, by applying a term structure model for interest rates on the basis of the market prices of government bonds.

One remarkable feature of the data is that much more extreme interest rate changes can be found in newly industrialized countries than in fully developed markets.

Maximum absolute values of changes of selected interest rates; source: Datastream

		Jan.1, 1997 - Dec.31, 1998			Jan.1, 1994 - Dec.31, 1998			Jan.1, 1987 - Dec.31, 1998		
	period	**StE 1D**	**StE 20D**	**DD 20D**	**StE 1D**	**StE 20D**	**DD 20D**	**StE 1D**	**StE 20D**	**DD 20D**
Austria	3 months	15	36	36	16	60	60			
	6 months	24	42	42	24	56	56			
	1 year	25	53	53	25	65	65			
	2 years	21	46	50	55	75	75	107	113	117
	5 years	38	50	51	38	81	81	67	99	99
	10 years	32	44	46	32	82	82	53	82	82
Germany	3 months	22	43	43	22	50	50	55	130	130
	6 months	23	42	42	23	50	50	60	130	130
	1 year	28	50	50	28	65	65			
	2 years	32	61	61	32	74	74	45	132	132
	5 years	29	65	65	35	86	93	35	104	104
	10 years	33	50	50	33	83	83	38	129	131
USA	3 months	20	33	33	25	88	88	69	181	213
	6 months	22	48	48	25	89	89	64	188	219
	1 year	34	59	59	38	94	94	66	194	219
	2 years	28	93	93	36	100	100	49	167	180
	5 years	24	97	97	42	97	97	62	167	179
	10 years	24	87	90	41	87	90	67	149	159

Table 11

	period	Jan.1, 1997 - Dec.31, 1998			Jan.1, 1994 - Dec.31, 1998			Jan.1, 1987 - Dec.31, 1998		
		StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D
Great Britain	3 months	19	63	63	48	63	63	150	338	338
	6 months	20	71	71	38	71	81	153	350	350
	1 year	22	81	81	31	84	84	155	353	353
	2 years	38	89	94	79	149	149	142	278	278
	5 years	31	96	101	31	99	104	62	185	222
	10 years	34	81	81	34	94	102	58	116	151
Switzerland	3 months	24	75	79	30	78	79	69	194	194
	6 months	24	66	69	31	81	81	45	156	156
	1 year	23	57	64	33	86	86	52	152	153
	2 years	43	57	59	105	105	105			
	5 years	17	51	51	42	74	74	70	90	90
	10 years	16	45	52	25	68	68	51	78	78
Japan	3 months	18	35	36	38	81	81	38	98	98
	6 months	16	30	30	25	69	69	105	98	105
	1 year	36	30	39	36	69	69	109	109	113
	2 years	39	67	68	39	83	83	43	120	120
	5 years	32	84	84	46	95	95	52	160	160
	10 years	30	84	84	30	84	84	58	150	150
France	3 months	20	32	32	150	322	322			
	6 months	41	41	47	116	247	247			
	1 year	41	49	49	75	145	150			
	2 years	20	76	76	35	99	99	111	248	248
	5 years	19	51	51	30	117	117	64	193	193
	10 years	24	48	50	24	106	106	80	181	181
Canada	3 months	170	163	175	170	242	244			
	6 months	159	149	174	159	241	244			
	1 year	121	121	140	121	227	227			
	2 years	62	104	104	70	205	205	99	205	212
	5 years	33	106	106	65	172	172	65	172	173
	10 years	21	87	87	45	127	127	67	148	169
Greece	3 months	2378	2920	2926						
	6 months	660	1005	1005						
	1 year	572	620	635						
Malaysia	3 months	575	655	712	575	655	712			
	6 months	530	555	605	530	555	605			
	1 year	225	440	440	225	440	440			
Czech Republic	3 months	841	1535	1537	841	1535	1537			
	6 months	719	932	934	719	932	934			
	1 year	482	578	578	482	578	578			

Table 11 (continued)

	period	Jan.1, 1997 - Dec.31, 1998			Jan.1, 1994 - Dec.31, 1998			Jan.1, 1987 - Dec.31, 1998		
		StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D	StE 1D	StE 20D	DD 20D
	1 year	482	578	578	482	578	578			
Slovakia	3 months	2000	2138	2289						
	6 months	3034	2300	3250						
	1 year	2767	2625	2767						
Poland	3 months	137	317	317	400	372	413			
	6 months	160	300	300						
	1 year	135	320	320						

Table 11 (continued)

All changes in basis points.
For periods of up to one year, the relevant interbank rates have been used; for periods of more than one year, the yields-to-maturity of the relevant benchmark bonds have been used.
Time windows: 1 day (1D) and 20 days (20D).
StE: Start to End, DD: drawdown.

3.3 Analysis of Time Series of Several Factors

When scenarios are constructed by subjecting the individual factors to historical maximum changes, the resulting scenarios may be extremely implausible, as we have seen in section 3.2.2. Recipients of stress test reports may therefore be unwilling to accept the resulting potential loss figures.

If we want to construct more plausible stress scenarios from historical data, the correlations between risk factors must be taken into account. That is, common movements of risk factors which actually occurred simultaneously – or at least within a very short time span – have to be combined in a scenario.

3.3.1 Simple Scenario Construction Using Time Series of Several Risk Factors

The simplest way of describing a stress event in several risk factors is probably to change each risk factor by the difference between the minimum and maximum values which the factor has reached during the stress period under consideration. If the minimum value occurred before the maximum value, a plus sign is assigned to the change; otherwise, a minus sign.

Defining the duration of the stress period is a decisive element in this approach. If the defined stress period is too short, the dimension of the crisis may be underestimated; if it is too long, the dimension of the crisis will probably be exaggerated, as long-term trends and peaks of normal market volatility are included in the scenario. Moreover, if the defined stress period is too long, the scenario will combine individual movements which actually did not at all occur simultaneously. This means ignoring the actual correlations, which will in turn reduce the plausibility of the resulting scenario.

One way of delimiting the stress period is to begin by selecting one or more risk factors which are representative for the stress event. In a second step, these risk factors can then be used to identify the stress period – for example, with the help of a graphic representation. In modeling an equity crash, for example, the representative risk factor will be a major stock price index; when a currency crisis is to be modeled, one of the involved exchange rates will be chosen as a representative risk factor. This method was used to define the stress period in section 3.3.3 below.

3.3.2 Measuring Simultaneous Changes of Several Risk Factors

To provide a sound quantitative basis for the inclusion of risk factor correlations in historical scenarios, the extent of simultaneous changes in different risk factors must be measured. This task is far from trivial, as this example shows: assuming that the risk factor r_1 drops by 15% at time t_1, and the risk factor r_2 simultaneously drops by 25%; assuming further that at time t_2, the risk factor r_1 drops by 19% and r_2 drops by 23% – when was the movement larger: at time t_1 or at t_2?

It is therefore necessary to find a measure for the simultaneous changes of risk factors. One way of doing this is to attribute equal weights to all risk factors, and to define the average of the individual risk factor changes as the measure of the common risk factor change. For the purposes of the above example, this would mean that the movement was larger at t_2, because the average change of the risk factors was 21% at t_2, but only 20% at t_1.

Once the measure of the simultaneous movements has been defined, one can proceed in the same way as in the analysis of individual time series presented in section 3.2. As a first step, a historical observation period is defined. But instead of examining the change of an individual factor between two points within the observation period, we follow the process just described to determine the measure of the common changes within the two points in time. The next step is to look for maximum changes of this measure; in doing so, we can modify the time window (e.g. 1 day, 20 days) and the change parameter (Start to End, drawdown). Once the maximum change of the measure is found, interest turns to the market states $\boldsymbol{r}_{t_1}$ and $\boldsymbol{r}_{t_2}$, between which this maximum change occurred. The absolute or relative difference between $\boldsymbol{r}_{t_1}$ and $\boldsymbol{r}_{t_2}$ is calculated, by components, for each risk factor. The vector of these changes shall be called $\Delta\boldsymbol{r}$. The resulting stress scenario $\boldsymbol{r}$ then is

$$\boldsymbol{r} = \boldsymbol{r}_{MM} + \Delta\boldsymbol{r} \ .$$

The credibility of the resulting stress scenarios is based on the fact that it includes only market movements which actually occurred in the past.

Measuring the simultaneous change of risk factors by calculating the average value of the changes of the individual factors as described above seems a plausible approach at first sight; however, it has one essential flaw: it is bound to give more weight to risk categories which are represented by many risk factors – e.g. interest rates –, than to risk categories which are represented by fewer risk factors, such as exchange rates. Equal weighting of all risk factors generally distorts the relevance of the individual factors. For this reason, two more suitable measures will be presented below. Both methods take into account historical data as well as the current portfolio.

3.3.2.1 Sensitivities

Portfolio-specific weighting of risk factors is certainly more useful than attaching equal weights to all risk factors. For example, risk factors may be weighted in proportion to the portfolio value's sensitivities δ_i to risk factor changes.

However, sensitivities depend on the scaling of risk factors. If a risk factor r_i is expressed in another unit which is, for example, one hundred times larger than the one used before, a value x in terms of the old unit will be equivalent to a value $x/100$ as expressed in the new unit; as a result, the sensitivity δ_i increases by a factor of 100. For this reason, it is generally inadmissible to say that risk factors with higher absolute values of sensitivity have a greater effect on the portfolio value than risk factors whose sensitivities are smaller in absolute terms. Sensitivities are just as arbitrary as the selection of units for the risk factors.

Example:

Consider a zero bond with a face value of CHF 100 and a residual maturity of 10 years. As risk factor r_1, we select the zero rate in CHF which fits the debtor, expressed in percentage points; we assume it is at present 2.318%. The second risk factor r_2 is the exchange rate; we assume a current exchange rate of 0.626 EUR/CHF.

The valuation function is then

$$P(r_1, r_2) = \frac{100 \cdot r_2}{(1 + r_1/100)^{10}} \ .$$

For $\Delta_1 = 1\%$ and $\Delta_2 = 0.1$ EUR/CHF, the sensitivities are $\delta_1 = -4.61$ and $\delta_2 = 79.52$. The risk factor with greater absolute value of sensitivity is r_2.

But if a risk factor $\bar{r}_1$ is selected, denoting the CHF zero rate expressed in units of 100 percentage points, a different picture results: the value of $\bar{r}_1$ is 0.02318, and the valuation function is

$$P(\bar{r}_1, r_2) = \frac{100 \cdot r_2}{(1+\bar{r}_1)^{10}}.$$

For $\overline{\Delta}_1 = 0.01$ and $\Delta_2 = 0.1$ EUR/CHF, the sensitivities are $\bar{\delta}_1 = -461$ and $\delta_2 = 79.52$. The risk factor with greater absolute value of sensitivity is now $\bar{r}_1$.

Risk factors sensitivities are meaningless if the unit of measurement in which the risk factor is measured is not stated. To determine the relative importance of risk factors for a given portfolio, the risk factors can be measured in units of the standard deviation of the time series of the risk factor in question. This means using a new risk factor

$$\bar{r}_i = \frac{r_i}{\sigma_i} \tag{3.1}$$

instead of the original risk factor r_i. The new risk factor $\bar{r}_i$ no longer depends on the scaling of the original risk factor for the following reason: if the data series $x_1, x_2, \ldots$ has a standard deviation σ_x, then the data series $10x_1, 10x_2, \ldots$ has a standard deviation $10\sigma_x$. Therefore

$$x/\sigma_x = 10x/\sigma_{10x}.$$

Irrespective of whether the original risk factor was x, $5x$ or $100x$, the new risk factor is always the same, namely $x/\sigma_x = 5x/\sigma_{5x} = 100x/\sigma_{100x}$. Please note that in this case (contrary to usual practice), σ does not denote a volatility, that is, the standard deviation of the changes in a financial time series, but the standard deviation of the financial time series itself.

Rewriting the valuation function with respect to the new risk factors would be a cumbersome process. This is not required, however, because the sensitivity for $\bar{r}_i$ in (3.1) is equal to σ_i times the sensitivity for r_i. The absolute value of $\sigma_i \delta_i$ can therefore be used as a measure for the sensitivity of the portfolio value towards changes in the risk factor r_i, because it is not influenced by the linear scaling of risk factors. This does not apply, however, to non-linear scalings of risk factors (e.g. logarithmic scaling).

Example:

We assume that in the case of the above-described CHF bond, the standard deviation σ_1 of r_1 is 0.29, and that of r_2 is $\sigma_2 = 0.0076$. For the new risk factors $\bar{r}_1 = r_1 / \sigma_1$ and $\bar{r}_2 = r_2 / \sigma_2$, the resulting sensitivities – now with $\Delta_1 = 1/\sigma_1\%$ and $\Delta_2 = 0.1/\sigma_2$ – are -1.34 and 0.60. The sensitivity for $\bar{r}_1$, which is -1.34, is equal to the product of $\sigma_1\delta_1$ and does not depend on whether $\bar{r}_1$ was found by scaling from the interest rate in percentage points, or by scaling from the interest rate in 100 percentage points. Because of $|-1.34| > |0.60|$, we can say with respect to the interest rate, regardless of the scale, that the bond value is more sensitive to the interest rate than to the exchange rate.

Consequently, risk factors should be weighted in proportion to the absolute value of $\sigma_i\delta_i$ for the purpose of measuring their common change. Given a change of n risk factors which is characterized by the fact that the i-th risk factor changes by Δr_i %, this results in

$$\sum_{i=1}^{n} |\Delta r_i \cdot \delta_i \cdot \sigma_i|$$

or

$$\sum_{i=1}^{n} |\Delta r_i| \cdot \frac{|\delta_i \cdot \sigma_i|}{\sum_{j=1}^{n} |\delta_j \cdot \sigma_j|} \tag{3.2}$$

as measures for the size of the common change of the factors. Since the weights in (3.2) sum up to 1, this measure can also be used to compare changes which concern different numbers of risk factors.

3.3.2.2 Maximum Portfolio Value Changes

Shaw (1997) proposes an alternative to the use of sensitivities in the measurement of simultaneous changes of several risk factors. This model sets out by computing the hypothetical P&Ls of the present portfolio under historical (one-day) market movements. The greatest hypothetical historical loss of the portfolio can then be identified, and one can subsequently discuss which scenarios produced these extreme losses. In this case, the P&L of the current portfolio is the measure for the size of the simultaneous factor movements. The time window can be modified again for the search for extreme losses. The maximum drawdown can also be easily taken into consideration.

This method closely resembles the historical simulation in VaR models. Both methods are based on the calculation of the hypothetical historical P&L time series for the current portfolio. But instead of looking at a relatively short past period and using "a big" (but not the biggest) loss, according to the desired confidence level, a substantially longer period of time is considered for the purposes of stress testing, and the actually biggest losses are determined. The observation period can also be deliberately defined to include a certain crisis.

The difference between this method and others discussed in this guideline is that it first calculates P&Ls and subsequently determines scenarios, whereas other methods start out by determining scenarios and then go on to calculate potential losses.

3.3.3 Table of Maximum Changes of Several Risk Factors

The table below presents the maximum changes of selected risk factors during the equity crash of 1987. The method used is the one described in the above section 3.3.1: The value of single factor changes was determined as the difference between the minimum and maximum of the respective risk factor in the stress period. The change is stated as a positive figure if the minimum value was reached earlier than the maximum, and vice versa. The observation period was defined as comprising October and November of 1987. This decision was based, inter alia, on the diagram below which illustrates the development of some stock price indeces.

1987 equity crisis: changes in stock price indeces

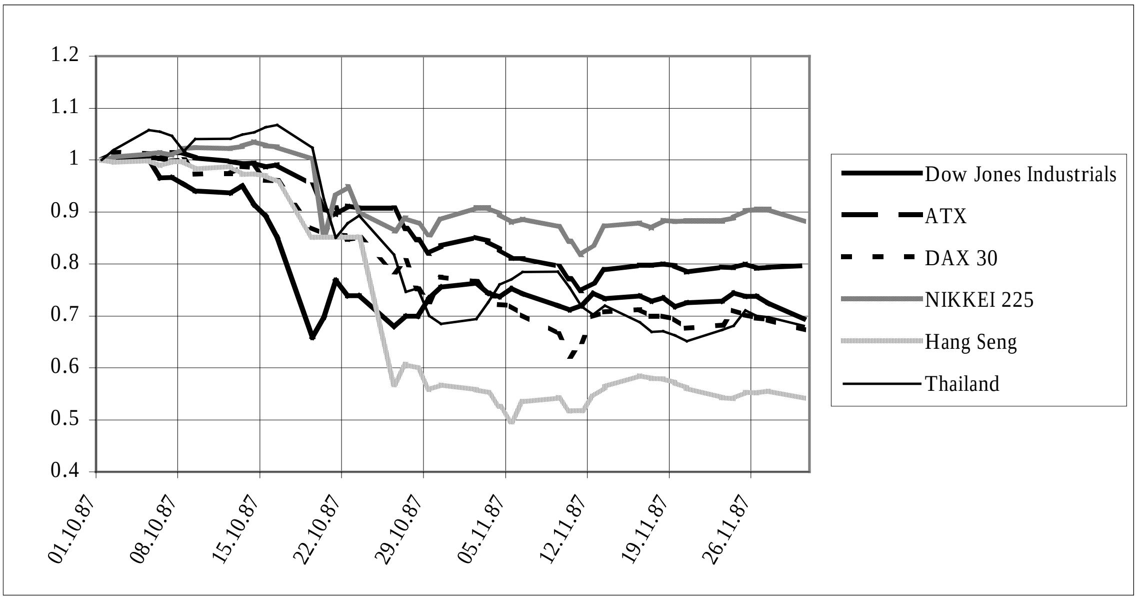

Diagram 2

In mid-October, the Dow Jones experiences a rapid downslide, followed – with varying degrees of intensity – by the other indeces. By late November, the indeces have stabilized on a lower level. Similar diagrams for the foreign exchange and interest markets show that these markets also experienced sharp swings, albeit with certain time lags, which subsided again in late November.

Changes of selected risk factors during the 1987 equity crisis;
Observation period: October 1, 1987 to November 30, 1987; source: Datastream

Relative changes of stock price indeces	Dow Jones Industrials	-34%
	ATX	-26%
	DAX 30	-39%
	Nikkei 225	-21%
	Hang Seng	-50%
	Thailand	-39%
Relative change of exchange rates	ATS/USD	-11%
	CHF/USD	-13%
	GBP/USD	-12%
	JPY/USD	-10%
	THB/USD	-2%
	HKD/USD	-1%
	GOLD/USD	8%
Yield changes of 10-year government bonds, in basis points	USA	-152
	Germany	-119
	Austria	-49
	Japan	-158
	Switzerland	-33
Increase of spreads, in basis points	USA 10 years: corporate - benchmark	171
	USA 3 months: Interbank - T-Bill	177
	Germany 10 years: Interbank - benchmark	43
	Germany Interbank: 10 years - 1 year	61
	Great Britain Interbank: 10 years - 1 year	101

Table 12

4 Identifying Portfolio-Specific Worst-Case Scenarios

4.1 Legal Basis of the Search for Worst-Case Scenarios

The Basle Committee on Banking Supervision (1996; section B.5 no 4) requires that the composition of portfolios be taken into account in selecting stress scenarios:

> "Banks should combine the use of supervisory stress scenarios with stress tests developed by banks themselves to reflect their specific risk characteristics."

and goes on to say (1996; section B.5 no 7):

> "In addition to the scenarios prescribed by supervisory authorities [...], a bank should also develop its own stress tests which it identifies as most adverse based on the characteristics of its portfolio (e.g., problems in a key region of the world combined with a sharp move in oil prices). Banks should provide supervisory authorities with a description of the methodology used to identify and carry out the scenarios as well as with a description of the results derived from these scenarios."

These provisions set an unequivocal international standard which prescribes that banks have to search for and apply worst-case scenarios. The Austrian Regulation on Internal Models for the Limitation of Market Risks also states in § 7 para 3 lit 3 that the Federal Minister of Finance can request banks to provide information on internal stress tests which measure the portfolio against potential future problem situations.

This provision means that banks have to carry out stress tests using scenarios which they regard as potential problem situations, irrespective of whether or not these situations have occurred in the past. The Minister of Finance can request information on these stress tests.

The Regulation does not specify what may be regarded as "potential future problems". But since § 7 para 3 of the Regulation is modeled on section B.5 nos 4-7 of the Basle Committee's 1996 Amendment to the Capital Accord to Incorporate Market Risks, "potential future problems" may be assumed to mean situations which the Basle paper describes as "most adverse based on the characteristics of its portfolio".

4.2 Worst-Case Scenarios versus Historical Scenarios

The search for worst-case scenarios differs from the construction of historical scenarios in two main aspects. *Firstly, past crises or scenarios constructed on the basis of historical maximum movements are not necessarily worst-case scenarios.* There may well be potential market movements which have not yet occurred, but which would result in worse consequences for the bank than the historical crises which did occur. Neither are historical maximum movements necessarily worst-case scenarios, for certain portfolios may suffer the greatest damage when risk factors move only slightly. In an attempt to identify worst-case scenarios, one does not only consider events which occurred at some point in the past, but also all potential future scenarios. For this reason, worst-case scenarios are also called "forward-looking scenarios".

Secondly, the construction of scenarios using historical data pays little attention to the characteristics of the bank's portfolio. At best, the current portfolio plays a role in the selection of the risk factors subjected to change, or when portfolio-related measures are applied to simultaneous changes of several risk factors – as, for example, the use of sensitivities discussed in section 3.3.2.1, or of P&Ls in section 3.3.2.2. Apart from these exceptions, the portfolio of the bank is of minor importance in the construction of scenarios from historical data. Conversely, the portfolio plays a central role in defining worst-case scenarios. What may be a worst-case scenario for one portfolio, may result in profits for another.

Neither the Basle Committee on Banking Supervision nor the Austrian Regulation include any provisions on how to identify worst-case scenarios. There are two fundamental options: A bank may rely on the experience and economic expertise of staff from as wide a range of fields as possible, who use their knowledge of the market, of the portfolio and of the trading and hedging strategies of the bank in an attempt to identify those market situations which could lead to particularly high losses of the bank. This approach, which is described in section 4.3 below, may be termed the subjective search for worst-case scenarios. But a bank may also use its computers to search systematically for worst-case scenarios. This may be called a systematic search for worst-case scenarios. It is described in detail in section 4.4.

4.3 Subjective Search for Worst-Case Scenarios

A subjective search for worst-case scenarios is usually based on the assumption of some surprising economic or political event which is presumed to cause particularly painful losses for the bank. Using economic and political expertise, an attempt is made to determine which further events may be triggered by the first one. Finally, the chain of events is translated in a plausible manner into changes of the risk factors. These changes of the risk factors then form the assumed worst-case scenario.

Neither of the two steps – determination of subsequent events and translation of the events into changes of the risk factors – is in any way clearly defined. The quality and plausibility of the resulting scenarios depend entirely on economic expertise and reasoning. This is why the subjective search for worst-case scenarios depends critically on the involvement of staff from as wide a range of fields as possible, with varied experience in regional economies, specific industries and banking. The search for suspected worst-case scenarios should also involve senior management. The credibility of the resulting scenario, as well as its relevance for decision-makers, depend mainly on the fact that all those involved in its construction agree that "such a thing may happen".

A subjective search for worst-case scenarios is sometimes based on the assumption of events such as an earthquake in Tokyo, the assassination of the American president, a coup in Russia, the failure of a large bank, serious budget problems in a country, or the abandonment of a fixed exchange rate under pressure. In selecting the triggering event, the specific situation of the bank must be taken into account: presumed worst-case scenarios should impact significantly on those risk factors to which the bank's portfolio is most heavily exposed. It can then be reasonably expected that such a scenario could result in particularly heavy losses for the bank.

Another important group of hypothetical scenarios is connected to the bank's model assumptions and to its investment and hedging strategies. In order to test the bank's vulnerability to a collapse of old and dear assumptions about the market, a collapse of assumptions which are essential for the bank's investment and hedging strategies is used as the triggering event in the search for hypothetical scenarios. The fact that these assumptions are often not explicitly stated should be no obstacle. Explicitly stating these assumptions about the market is an essential step towards recognizing and controlling the bank's market risk.

4.4 Systematic Search for Worst-Case Scenarios

4.4.1 Why Search Systematically for Worst-Case Scenarios

Stress tests which use historical or subjectively presumed worst-case scenarios may overlook fatal stress scenarios. They determine the potential loss only at very few points within the multidimensional space of scenarios. One difficulty with historical and suspected worst-case scenarios is that knowing which losses a portfolio can be expected to suffer under a few selected scenarios may give the bank a false sense of security, if the projected losses are manageable. The sense of security may be false because the bank does not know whether there are conceivably other scenarios which are equally plausible and result in much heavier losses. Even in a subjective search for suspected worst-case scenarios, one cannot know whether the scenarios found are actually the worst ones.

Another difficulty is that knowing about an alarming loss in a stress scenario cannot lead to practical consequences as long as it is unclear which risk factors have caused the loss. This is another question which is not adequately answered by stress tests using historical or suspected worst-case scenarios.

A systematic search for worst-case scenarios promises to remedy these problems. Its foremost objective is the reliable identification of worst-case scenarios, i.e. scenarios in which the existing portfolio will suffer particularly heavy damage. Another objective is to find out which risk factors are mainly responsible for the losses under the worst-case scenario. Once these risk factors have been found, the bank can easily identify which measures are required if it is not prepared to bear the risk of such a loss.

It will be generally impossible to describe a market state in which the portfolio has its *smallest* value, since the loss potential of a portfolio is as a rule unlimited. A simple example is that of a portfolio which consists only of a short call: its value will fall without limit as long as the value of the underlying instrument rises. For this reason, not all scenarios will be admitted; rather, the search will be for the minimum among those scenarios which meet certain plausibility conditions. The definition of such plausibility conditions is discussed in Annex A.1.

The worst-case scenario within the admissibility domain as defined by the plausibility condition can be found through using an algorithm which identifies the place of the minimum of the valuation function P within the admissibility domain. This process is discussed in Annex A.2.

4.4.2 Reporting on the Systematic Search for Portfolio-Specific Worst-Case Scenarios

The ultimate recipients of stress test reports – as, indeed, of any report on risks – are those decision-makers within a bank who are in a position to decide on a reduction of market risk exposure. Stress test reports can only serve as the basis of informed decisions if they are comprehensive and comprehensible at the same time. Some questions arise in this context.

How improbable may stress scenarios be? On the one hand, it is the nature of stress tests to ask what is going to happen in situations which nobody expects. On the other hand, test results of scenarios which are regarded as completely impossible will not be taken serious by the recipients of the test reports. The decision on how improbable stress scenarios may be, must be taken into account in the interpretation of test results.

In this situation, it appears useful to consider plausibility conditions that vary in strictness. The stricter the plausibility condition, the smaller the number of admissible scenarios, and the more harmless the worst admissible scenarios. For each plausibility condition, the stress test results show which are the most extreme scenarios which satisfy the plausibility condition, and how big losses are under these scenarios.

How can the results of a search for portfolio-specific worst-case scenarios be presented in a concise and readily understandable manner? It is certainly not enough to simply report the values of the risk factors in the worst-case scenario that has been found. For example, listing 500 risk factors of the worst-case scenario would hopelessly overtax the capacity of any recipient of the report. Consequently, reports should include only the *most important* risk factors in the worst-case scenario.

What are the "most important" risk factors of a worst-case scenario? Sensitivities are certainly not an appropriate indicator of the importance of a risk factor: sensitivities in the present market state are completely unrelated to the worst-case scenario to be characterized; and all sensitivities will be zero in the worst-case scenario if it is a local minimum.

The following approach appears more useful: The search for the key risk factors is a search for a subset of risk factors which explain the loss under the worst-case scenario up to a previously defined degree, i.e. which have a certain explanatory power. For example, an explanatory power of 80% means that we are looking for a subset of the risk factors which will be able to explain at least 80% of the loss under the worst-case scenario. This means: Let us assume that, instead of the complete worst-case scenario $\boldsymbol{r}_{WC} = (r_{WC,1}, \ldots, r_{WC,n})$, only the values of a subset of w risk factors $r_{i_1}, r_{i_2}, \ldots, r_{i_w}$ are reported. This corresponds to a simplified report scenario

$$\boldsymbol{r}_{report} := (r_{MM,1}, \ldots, r_{WC,i_1}, \ldots, r_{WC,i_2}, \ldots, r_{WC,i_w}, \ldots, r_{MM,n}),$$

where the risk factors $r_{i_1}, r_{i_2}, \ldots, r_{i_w}$ have their worst-case values $r_{WC,i_1}, \ldots, r_{WC,i_w}$, and all other risk factors have their actual values. The subset of risk factors will explain 80% of the loss suffered from the worst-case scenario if

$$P(\boldsymbol{r}_{MM}) - P(\boldsymbol{r}_{report}) \geq 0.8\,(P(\boldsymbol{r}_{MM}) - P(\boldsymbol{r}_{WC}))$$

applies. How can we find the smallest possible subset of risk factors which still has an explanatory power of, for example, 80% in relation to the total loss under the worst-case scenario? One possibility is a step-by-step approach: we first try to find a single risk factor which explains 80% of the loss. If one can be found, the objective of the exercise has been met. If not, we look for two risk factors which, taken together, explain 80% of the loss. If no two factors can do this, we search for three risk factor who can, and so on. Sooner or later, a subset of risk factors which is capable of explaining 80% of the loss can always be found.

In searching for a subset of w risk factors which can explain 80% of the loss, it is by far too cumbersome to go through all subsets with w elements. For example, if we search for a subset of 10 risk factors with an explanatory power of 80% for a worst-case scenario that is determined by 500 risk factors, the valuation function has to be evaluated $2.6\ \ 10^{35}$ times. A more efficient method is to use a minimization algorithm to find the subset $\{i_1,...,i_w\}$ for which

$$P(r_{MM,1},...,r_{WC,i_1},...,r_{WC,i_2},...,r_{WC,i_w},...,r_{MM,n})$$

takes on the lowest value. We can then ascertain whether this loss is equivalent to 80% of the loss under the worst-case scenario. This is an optimization problem in a discrete w-dimensional space. A particularly suitable approach to discrete optimization problems is the method of simulated annealing, which is discussed in Annex A.2.

The stress test report can then present the results of the search for portfolio-specific worst-case scenarios as in this model:

Model report on the systematic search for worst-case scenarios

Admissibility domain	Maximum loss within the admissibility domain	Key risk factors in the worst-case scenario	Explanatory power of the key risk factors
"cuboid with edges 3 r"	EUR 0.5bn	exchange rate EUR/USD: 0.9 6m LIBOR GBP: 5.3% 10y swap rate CHF: 3.27%	65%
"3 times enlarged ellipsoid with covariances Σ"	EUR 0.3bn	exchange rate EUR/USD: 0.95 12m LIBOR GBP: 5.42% 10y swap rate CHF: 3.27%	61%
„…"	EUR …bn	…	…%

Table 13

For a discussion of admissibility domains, see Annex A.1.

4.5 Emergency Plans for Worst-Case Scenarios

One of the main benefits which the search for worst-case scenarios provides for the risk management of a bank is that this search leads directly to an answer to the question of how the bank should respond to alarming events. Stress tests based on historical scenarios can hardly provide an answer to this question, as it is difficult to say which risk factors or which positions have caused the loss. Speculating about this question is not sufficient for effective risk management.

Conversely, stress tests based on subjectively presumed worst-case scenarios are in most cases closely focused on markets or business segments where particular vulnerability is suspected. If the stress test predicts unacceptable losses under a suspected worst-case scenario of this kind, it is clear which exposures have to be reduced.

Stress tests with historical or subjectively presumed worst-case scenarios thus have two drawbacks: a serious crisis may be overlooked, and it may remain unclear which risk factors or positions are responsible for a specific alarming result.

These two problems are solved by searching for portfolio-specific worst-case scenarios and identifying the risk factors which are most important for the worst-case scenario. If the bank knows which scenario is its worst-case scenario within a given admissibility domain, and if it knows that the loss in this scenario is manageable, then the resulting sense of security is no longer false. The bank can be sure that nothing worse can happen as long as the market state remains within the admissibility domain and the portfolio is not changed.

But if the outcome of a search for portfolio-specific worst-case scenarios is alarming, the identification of the key risk factors in this scenario also makes clear what the bank has to do: it has to open positions which will return profits when the key risk factors approach their worst-case value. The profit range of these hedging positions can be concentrated quite closely around the worst-case values. Such hedging positions with a concentrated effective range are more effective than hedging positions with a broad profit range, and will usually not cause a substantial reduction of the portfolio's profit prospects.

5 Summary of Stress Testing Requirements for Banks Using Internal Models

This section summarizes the requirements for stress testing which Oesterreichische Nationalbank takes into account in preparing evaluation reports pursuant to § 26b of the Austrian Banking Act.

5.1 Reporting and Organization

Stress tests have to be carried out regularly. The frequency of stress testing should correspond to the dynamics of the portfolio. Portfolios which are frequently rebalanced have to be subjected to frequent stress testing as well. The Austrian Regulation on Internal Models for the Limitation of Market Risks mandates quarterly stress tests for banks which use an internal model for the calculation of capital requirements to cover their market risk. In addition, interim stress testing is required for special situations, examples of which are given in the Regulation. The results of quarterly stress tests must be submitted to the Federal Ministry of Finance and to Oesterreichische Nationalbank in reports which must be at least equivalent to the reports submitted to the bank management. Furthermore, the supervisory authority can request information on special interim stress tests if this is deemed necessary.

Procedures and responsibilities relating to decisions on when interim stress tests are to be conducted, as well as regarding the selection of stress scenarios, must be laid down in the risk management handbook.

Banks must be able to carry out stress tests quickly. Like any information on risks, the results of stress tests must be quickly available to ensure that the bank can reduce its risk exposure quickly by timely responses to changing market conditions.

The risk management handbook has to define what is to be regarded as an alarming stress test result. In particular, it has to state which reference figures shall be used as a basis for comparison with potential losses found in stress tests. Emergency plans are useless if the circumstances under which they are to be applied are not clearly defined.

The risk management handbook has to define the measures which the bank will take to limit its risks adequately if stress testing reveals weaknesses. These emergency plans have to foresee measures which the bank will take in response to alarming stress test results. Such emergency plans ensure that stress testing actually serves to reduce risks and to prevent losses.

The results of stress tests have to be communicated to decision-makers which are in a position to decide on a reduction of risk exposure. The results of stress tests should be routinely submitted to the management and should by communicated periodically to the bank's supervisory board.

A feedback loop should enable managers to question stress test reports and suggest modifications. One way of doing this is to include a special column for this purpose in stress test reports. This should ensure that the management plays a significant role in planning stress tests and is able to interpret the stress test results correctly.

5.2 Scenario Selection

Stress scenarios should describe extraordinary market movements, while at the same time being plausible. Plausibility means that stress scenarios have to appear intuitively possible. If they are not, decision-makers will not attach sufficient importance to stress test results in their decision-making processes. The two requirements of extraordinary nature and plausibility are in conflict with each other. One way of solving this conflict is to consider scenarios of varying degrees of extremeness.

Banks should consider historical scenarios, and they should also search for their own worst-case scenarios. Considering only stress scenarios which are based on historical data is not enough. The selection of historical scenarios is based on the assumption that future crises will resemble past crises. The fact that they already occurred at some point in the past, lends plausibility to them and increases their acceptance. The search for worst-case scenarios, however, includes scenarios which have not yet occurred, but which are plausible. Banks have to find their specific worst-case scenarios, but they can decide for themselves whether they use a subjective or a systematic approach in doing so.

The selection of scenarios must be consistent with the risk profile of the bank. Due to their different structures, banks' portfolios have different risk profiles. The portfolio plays a central role both in a subjective and a systematic search for worst-case scenarios. Banks should also determine their vulnerability to a collapse of assumptions which are essential for their VaR models and their investment and hedging strategies. This is done through consideration of scenarios which violate such assumptions.

The identification of scenarios, especially the subjective search for worst-case scenarios, should involve the broadest possible range of departments and hierarchy levels. Staff with different macroeconomic, country-specific, industry-specific and banking expertise can contribute to the preparation of detailed scenarios. Any search for subjective worst-case scenarios should also involve senior management members. The credibility of the resulting scenarios, as well as their relevance for decisionmakers, depend mainly on the fact that all those involved in their construction agree that the resulting scenarios are plausible.

Stress tests should be conducted which consider simultaneous changes in several risk categories. Simultaneous changes in several risk categories may reveal risks which are not noticed in changes involving only individual risk categories.

Stress scenarios should also take into account aspects of liquidity crises. See section 2.3 for a more detailed discussion of this point.

It is desirable for stress testing to also take into account aspects of credit risk. In this context, it should be examined whether hedged positions exist which, due to counterparty default, could become subjected to market risk. The market risk exposure of the resulting positions should be analyzed.

In order to monitor changes of exposure in specific risk areas, certain standard scenarios should be evaluated periodically. Standard scenarios have to be defined so that they capture those risk areas where the bank's exposure is greatest. This presupposes that scenarios can be stored and reused for the valuation of the modified portfolio at a later point. If the bank changes its trading strategy, it may become necessary to introduce additional standard scenarios.

5.3 Computation

Stress testing of portfolios which contain options or other products with non-linear valuation functions should be based on a complete revaluation of the portfolio. Linear approximation using sensitivities is not sufficient. The risk of products with option characteristics is often described by the delta, gamma, vega, rho and tau factors. These factors are sensitivities of the option value to minor changes in the risk factors. For large risk factor changes, the linear approximation of the value change through the use of sensitivities increasingly loses validity. Stress tests often look at very large changes in the risk factors. For this reason, they require a full revaluation of the portfolio.

The same valuation mechanisms should be used for the purpose of stress testing as for the value-at-risk model. This ensures that stress test results can be compared to VaR results.

Computation processes should be automated as far as possible. This will keep the incidence of errors and inaccuracies as low as possible and help to shorten the response time when a stress event occurs. Position data input and valuation must be fully automated. It must be possible to enter scenarios flexibly, and to store them.

Stress tests have to take into account the impact of the scenarios on the entire trading book. Stress tests which examine subportfolios separately could overlook a loss which is still manageable within each subportfolio by itself, but not if it affects the whole portfolio. It must be noted in this context that stress tests for the banking book are not a prerequisite for the admission of a value-at-risk model. Stress tests for the banking book would presuppose a market valuation of the banking book, which is not possible in all cases. If a bank can do it, an integration in full or in part of the banking book can substantially improve the informative value of stress tests.

It should be possible to perform stress tests on any desired subportfolio level. Such levels could include divisions, trading units, traders or individual instruments. Scenarios used on lower levels should be tailored to the needs of the relevant area.

Technical Annex

A.1 Admission Criteria for Scenarios in the Systematic Search for Worst-Case Scenarios

A.1.1 Admission Criteria which Ignore Correlations

The question to be discussed is which plausibility conditions will be useful in identifying stress scenarios. A conceivable group of plausibility conditions can be defined as follows. Risk factor time series are used to determine for each risk factor the standard deviation σ_i of the relative changes. If $\boldsymbol{r}_{MM} = (r_{MM,1},...,r_{MM,n})$ denotes the present market state, i.e. the present values of all risk factors, the following condition can be defined for each positive number k:

Plausibility condition *"cuboid with edges* $2\ k\ \ r$ ":

This admits all scenarios $\boldsymbol{r} = (r_1, r_2,...,r_n)$ which satisfy $r_{MM,i}(1-k\sigma_i) \le r_i \le r_{MM,i}(1+k\sigma_i)$ for each risk factor r_i. If the risk factor r_i can have only positive values, the condition must be: $\max\left\{0, r_{MM,i}(1-k\sigma_i)\right\} \le r_i \le r_{MM,i}(1+k\sigma_i)$.

This plausibility condition admits only scenarios which are situated within an *n*-dimensional cuboid with edges $2k\sigma_i r_{MM,i}$ and center $\boldsymbol{r}_{MM}$. The larger k is, the more generous the plausibility condition "cuboid with edges $2\ k\ \ r$", the more scenarios are admitted, and the more extreme the worst of the admitted scenarios will be.

If it appears too crude to ensure the positiveness of certain risk factors in the plausibility condition "cuboid with edges $2\ k\ \ r$" simply by a cut-off, the following alternative may be used:

Plausibility condition *"cuboid in logarithmic scale"*:

This admits all scenarios $\boldsymbol{r} = (r_1, r_2,...,r_n)$ which satisfy $r_{MM,i}e^{-k\sigma_i} \le r_i \le r_{MM,i}e^{k\sigma_i}$ for each risk factor r_i.

This formula appears useful for stock prices and other risk factors which are often modeled as lognormal distributions. The advantage of this plausibility condition is that the risk factor r_i is always positive, including cases where k is big. If k is small, $e^{k\sigma_i}$ has nearly the same value as $1+k\sigma_i$; this is why the plausibility condition "cuboid with edges $2\ k\ \ r$" is nearly equal to the plausibility condition "cuboid in logarithmic scale" for small k.

Some caution is due with respect to the above described plausibility conditions for the following reason: These plausibility conditions may be fulfilled by scenarios which violate certain no-arbitrage conditions. If, for example, the three exchange rates EUR/CHF, EUR/USD, and CHF/USD are among the risk factors, the values which these three risk factors can take on in an arbitrage-free world are limited. If two exchange rates are given, the third is also fixed. Thus, the fact that a scenario fulfills the plausibility condition is not necessarily sufficient to ensure its reliability.

It may be argued, however, that no-arbitrage conditions must not necessarily be fulfilled in times of crisis, owing to the illiquidity of the markets, and that consequently, scenarios which violate no-arbitrage conditions may well be realistic stress scenarios. Ultimately, a separate decision is required for each scenario to determine whether or not it will be admitted as a stress scenario. The admission criteria for scenarios have to be taken into account when interpreting stress test results.

A.1.2 Admission Criteria which Take into Account Correlations

The plausibility condition "cuboid with edges 2 k r" admits scenarios which are as a rule much less probable than a change of an individual risk factor by $k\sigma$. For two risk factors $(k = 2)$, this effect can be illustrated as follows:

Lines of equal probability for bivariate normally distributed risk factors

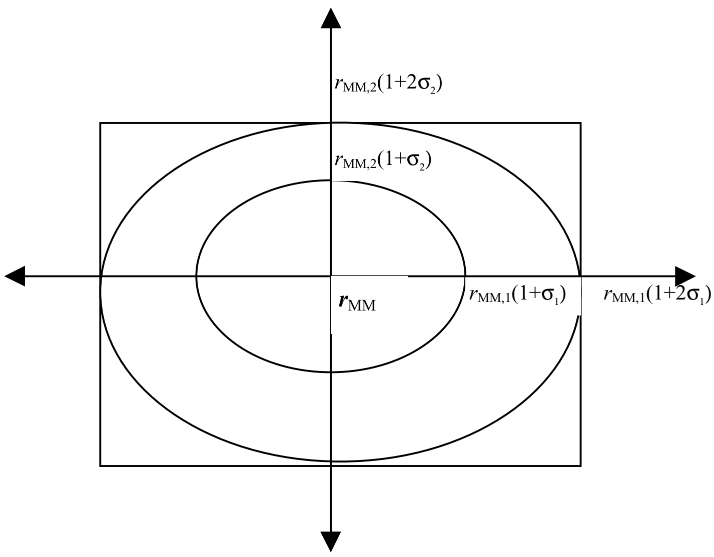

Diagram 3

The ellipses consist of scenarios which are equally probable if the correlation between the two risk factors is zero and the risk factor changes are normally distributed. The bigger rectangle is the "cuboid" with edges $4r_{MM,1}\sigma_1$ and $4r_{MM,2}\sigma_2$. Scenarios near the corners of the rectangle are less probable than a change by $4r_{MM,i}\sigma_i$ of the individual factors.

Moreover, the plausibility condition "cuboid with edges 2 k r " ignores correlations between the risk factors. If there is a strong positive correlation between the two risk factors in the above two-dimensional example, scenarios in the upper right-hand corner of the cuboid are significantly more probable than scenarios in the upper left-hand corner. A movement of the risk factors against the direction of the correlation is much more improbable than a change of the factors in the direction of the correlation.

At first sight, this effect appears not to pose any problem: Firstly, stress tests are not supposed to say anything about the probability of the scenarios used. Secondly, correlations are likely to change in stress events, anyway. It is frequently argued, for example, that during stress events, the correlations between most risk factors are close to 1 or -1.

It is still useful, however, to take into account correlations when defining plausibility conditions, given the importance of the plausibility of scenarios in the interpretation of results. Stress test results which show heavy losses for a bank will more readily lead to counter-measures if decision-makers tend to regard the scenario as plausible. Plausibility conditions should therefore be defined so as to exclude scenarios which are next to impossible and could for this reason undermine the credibility of stress test results. Neither is the change of normal correlations in stress events a valid argument against the inclusion of correlations in the definition of plausibility conditions. For if correlations are included in plausibility conditions, this can also be done for stress event correlations which differ substantially from the correlations observed in untroubled periods.

How can we include correlations in the definition of admission conditions for scenarios? Assume a variance-covariance matrix of risk factor changes,

$$\Sigma := \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix},$$

with the variances $\sigma_i^{\,2} = E\left[(\Delta r_i - \mu_i)^2\right]$ of the risk factor changes on the diagonal, and the covariances $\sigma_{ij} = \sigma_{ji} = E\left[(\Delta r_i - \mu_i)(\Delta r_j - \mu_j)\right] = \sigma_i \sigma_j \varrho_{ij}$ of the risk factor changes outside the diagonal, where μ_i denotes the mean value of the changes of the risk factor r_i.
A normal distribution of n variables with a density of

$$P(\Delta r_1,...,\Delta r_n) = const \cdot \exp(-\tfrac{1}{2}(\Delta r_1,...,\Delta r_n)^T \cdot \Sigma^{-1} \cdot (\Delta r_1,...,\Delta r_n))$$

results exactly in these correlations. If the risk factor changes were normally distributed, and given these correlations, then the scenarios $\boldsymbol{r}$, to which a leap from the present scenario $\boldsymbol{r}_{MM}$ is equally probable, would form an n-dimensional ellipsoid

$$(\boldsymbol{r}_{MM} - \boldsymbol{r})^T \cdot \Sigma^{-1} \cdot (\boldsymbol{r}_{MM} - \boldsymbol{r}) = k^2 .$$

The lengths of the major axes of the ellipsoids are k times the eigenvalues of the matrix Σ^{-1}. If the risk factor changes are normally distributed, and given covariances Σ, the probability that the market state $\boldsymbol{r}$ lies within the ellipsoid is determined by the value of the χ^2 distribution function with n degrees of freedom at k^2,

$$F_{\chi_n^2}(k^2) = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^{k^2} s^{\frac{n}{2}-1} e^{-\frac{s}{2}} ds .$$

Admission criteria for scenarios may now be specified as in the following example.

1. A confidence level p is set, for example, $p = 95\%$.
2. Tables for the χ^2 distribution function with n degrees of freedom are used to determine the k^2 for which $F_{\chi_n^2}(k^2) = p$ applies. Press et al. (1992, chapter 6) describe how k^2 may also be directly computed from the gamma function.
3. The admissibility domain is defined as the set of all scenarios $\boldsymbol{r}$ which fulfill

$$(\boldsymbol{r}_{MM} - \boldsymbol{r})^T \cdot \Sigma^{-1} \cdot (\boldsymbol{r}_{MM} - \boldsymbol{r}) \le k^2 .$$

This results in the

Plausibility condition *"k times enlarged ellipsoid with covariances* Σ ":
This admits all scenarios $\boldsymbol{r}$ which satisfy $(\boldsymbol{r}_{MM} - \boldsymbol{r})^T \cdot \Sigma^{-1} \cdot (\boldsymbol{r}_{MM} - \boldsymbol{r}) \le k^2$.

Only if the risk factor changes are normally distributed with covariance matrix Σ, is it justified to state that with a probability p one of the scenarios will be situated within the k times

enlarged ellipsoid with covariances Σ. In this case, the value at risk can also be computed from the above plausibility condition, using a minimization algorithm: the VaR is the difference between the present portfolio value and the minimum portfolio value within the ellipsoid which corresponds to a confidence level of 95% or 99%.
It must be doubted, however, that the same correlations will apply in stress periods and in untroubled periods, and that the risk factor changes are indeed normally distributed, instead of showing fat tails, for example. For this reason, one cannot as a rule say that a scenario will be situated in the ellipsoid with a probability p. Even so, the ellipsoids can serve as suitable admissibility domains for scenarios.

It should be noted again that present correlations are not the only ones which can be selected for the covariance matrix Σ. It may also be useful to use stress correlations for the correlation matrix. Stress correlations can be estimated, for example, on the basis of historical stress event data.

A.2 Methods for the Systematic Search for Worst-Case Scenarios

A.2.1 Factor Push Method

The factor push method is a relatively simple way of getting a rough idea of worst-case scenarios. The basic process is to change each individual risk factor by a given value in the direction that will most reduce the portfolio value. More specifically, one proceeds as follows:

1. For each risk factor r_i, the portfolio values which result from a positive and a negative risk factor change by a defined value, are determined. The value of the change is usually defined as a multiple of the standard deviation of the risk factor change under consideration, e.g. k times the standard deviation. So the two values $P(r_{MM,1},\dots,r_{MM,i}(1+k\sigma_i),\dots,r_{MM,n})$ and $P(r_{MM,1},\dots,r_{MM,i}(1-k\sigma_i),\dots,r_{MM,n})$ are calculated.
2. A plus or minus sign $VZ(i)$ is assigned to each risk factor, using the formula

$$VZ(i) := \operatorname{sgn}\left(P(r_{MM,1},\dots,r_{MM,i}(1+k\sigma_i),\dots,r_{MM,n}) - P(r_{MM,1},\dots,r_{MM,i}(1-k\sigma_i),\dots,r_{MM,n})\right).$$

   $VZ(i)$ is 1 if the upward change of the i-th risk factor results in a higher portfolio value than the downward change. Otherwise, $VZ(i)$ is -1.
3. The new stress scenario can be written:

$$\boldsymbol{r}_{WC} := \left(r_{MM,1}\cdot\left[1-VZ(1)\cdot k\sigma_1\right],\dots,r_{MM,n}\cdot\left[1-VZ(n)\cdot k\sigma_n\right]\right).$$

One of the main advantages of this method is that computation is easy. Only $2n$ evaluations of the valuation function are necessary to find the new scenario.

Two drawbacks of the method should be mentioned: firstly, it supplies only scenarios which are situated on the surface – or, more precisely, at a corner – of the n-dimensional cuboid

$$\left\{ \boldsymbol{r} \in R^n : r_i = r_{MM,i}(1 \pm k\sigma_i) \right\}.$$

This may be of little consequence for portfolios with linear valuation functions. But for portfolios with non-linear valuation functions, the valuation function minimum may be situated within the n-dimensional cuboid. This may easily be the case for portfolios which contain derivative instruments.

For this reason, the simple factor push method is unsuitable for portfolios containing derivatives. This can partly be remedied by applying the factor push method not only for one, but for several values of k, e.g. $k = 1/2, 1, 3/2, 2, \ldots, 10$. This produces minima on the surfaces of 20 cuboids which are stuck within each other. The shorter the distance between one cuboid surface and the next, the more precisely can we localize minima within the biggest cuboid. The amount of computation required increases in linear proportion to the number of cuboids considered.

Another drawback of the method is that it cannot even be safely assumed that it will find the minimum on the cuboid surface. One can find valuation functions whose minimum on the cuboid surface is not situated at a corner.

A.2.2 Monte Carlo and Quasi-Monte Carlo Methods

Monte Carlo and quasi-Monte Carlo methods provide an approximate value for the minimum of a valuation function within a certain range. They are relatively simple, but require substantial computation power. To find an approximate minimum within the admissibility domain defined by the plausibility condition, one could, for example, proceed as follows:

1. A transformation T from the n-dimensional unit cube

$$\left\{ (x_1, x_2, \ldots, x_n) \in R^n : x_i \in [0,1] \right\}$$

to the admissibility domain is determined. If the admissibility domain is the n-dimensional cuboid $\{\boldsymbol{r} \in R^n : r_i = r_{MM,i}(1 \pm k\sigma_i)\}$, the following function can be chosen for T:

$$T(x_1, x_2, \ldots, x_n) := \left(r_{MM,1}(1 - k\sigma_1 + 2x_1 k\sigma_1), \ldots, r_{MM,n}(1 - k\sigma_n + 2x_n k\sigma_n)\right).$$

2. A series $\{\boldsymbol{x}_j := (x_{j,1}, x_{j,2}, \ldots, x_{j,n}) \in R^n : x_{j,i} \in [0,1]\}_{j=1,\ldots,N}$ of N random vectors or quasi-random vectors which fill the space within the n-dimensional unit cube as uniformly as possible, is produced. For the purposes of this discussion, the term "random vector" will be used to denote not only genuine random vectors, but also pseudo-random vectors and quasi-random vectors. The greater the number N of the random vectors, the more uniformly will the interior of the unit cube be filled.
3. For each of the random vectors $\boldsymbol{x}_j$, the value of the portfolio $P(T(\boldsymbol{x}_j))$ at the place $T(\boldsymbol{x}_j)$ is determined. The resulting stress scenario is

$$\boldsymbol{r}_{WC} := T(\boldsymbol{x}_j)$$

where the random vector $\boldsymbol{x}_j$ produces the smallest value of $P(T(\boldsymbol{x}_j))$.

Compared to the factor push method, the most important advantage of Monte Carlo methods and quasi-Monte Carlo methods is that minima can be found not only on the surface of the n-dimensional cuboid, but also within the cuboid space. This is absolutely necessary for portfolios with heavily non-linear elements, in particular option portfolios.

The generation of the random vectors is the chief point on which an efficient implementation of these methods depends. The more uniformly and densely the n-dimensional unit cube is filled by the random vectors, the more reliable is the approximation of the minimum portfolio value. The more irregular the surface of the valuation function as a function of the risk factors, the more densely must the interior of the n-dimensional unit cube be filled by the random vectors to produce a reasonably reliable approximation of the minimum.

The efficiency of Monte Carlo methods and quasi-Monte Carlo methods is determined by the number of random vectors required. The extent of computation power required is caused by the fact that the valuation function has to be computed separately for each random vector. It is not necessary to compute the series of random vectors afresh every time; rather, it can be saved and used again for any problem with the same dimension – i.e. with the same number of risk factors.

Once the desired density of coverage is fixed, the required number of random vectors is determined by how uniformly they fill the unit cube. Discrepancy is a mathematical measure for the deviation from the greatest possible uniformity of filling. The lower the discrepancy of a series of random vectors, the smaller the series may be to reach a given coverage density and thus, a given accuracy of approximation.

What is decisive for the efficiency of Monte Carlo methods and quasi-Monte Carlo methods is not the randomness of the random vectors, but the filling density. It will therefore as a rule be more efficient to select series of random vectors with as little discrepancy as possible, even though these series may contain only quasi-random numbers, rather than genuine random numbers or pseudo-random numbers. A more detailed discussion of the efficiency of Monte Carlo methods and quasi-Monte Carlo methods can be found in Niederreiter (1992).

A.2.3 Other Loss Maximization Algorithms

Monte Carlo methods and quasi-Monte Carlo methods fill the admissibility domain defined by the plausibility condition as uniformly as possible with points (random vectors) at which the valuation function is then computed. A drawback of these methods is that many points are situated in parts of the cuboid where no valuation function minimum is expected. Other methods promise to be more efficient in minimizing the valuation function – i.e. maximizing potential loss. In practice, however, most valuation functions are so complicated that it would require too much computation power or be simply impossible to calculate valuation function derivatives with respect to risk factors. This leaves banks with minimization algorithms which require only evaluation of the function, but not of its derivatives. Descriptions of the following algorithms, including programming instructions, can be found in Press et al. (1992)

The *multidimensional simplex method* was first described by Nelder and Mead. (It should not be confused with the simplex process which is used in linear programming to find extreme values of a linear function.) A simplex in an n-dimensional space consists of a vertex and n linearly independent vectors. The simplex is the n-dimensional domain which is created if the n vectors act at the vertex. Beginning with a start simplex, the algorithm determines a series of wandering simplexes of diminishing size which approach a domain in which a local minimum of the valuation function is situated. The series can be halted if the distance between a new simplex and the preceding one gets smaller than a certain tolerance, or if the value of the valuation function diminishes by less than a given tolerance from one step to the next. The resulting scenario is the vertex of the last simplex.

The multidimensional simplex method is relatively simple, but it requires a rather high number of function evaluations. A more efficient method, but one which is more complicated to implement, is the multidimensional Powell method. This method consists of steps whereby in each step, n one-dimensional minimizations in n directions are performed. The crucial point is the determination of the n directions for the next step of n minimizations. In this, one of two strategies may be followed: One either searches for directions which correspond as closely as possible to the directions of the valleys of the valuation function, or one searches for directions with the characteristic that the minimization in one direction is not destroyed by the subsequent minimization in another direction. Implementations of both strategies can be found in Press et al. (1992; pp. 413-420).

The *simulated annealing method* has received much attention because it can be used to solve optimization problems which are notorious for their high computation requirements. ("Annealing" is a term used for the slow cooling-off process of metals which leads to a state of minimum energy.) The particular strength of this method lies in dealing with cases where the desired global minimum is hidden among many small local minima. The special feature of the method is that it proceeds from one scenario to the next not by a deterministic, but by a stochastic process. On the basis of one scenario, a candidate for a new scenario is randomly selected. Assume that the difference between the valuation function values of the would-be scenario and the old scenario is ΔP. If the valuation function has a lower value in the would-be scenario, it is realized; if the valuation function has a higher value in the would-be scenario (i.e. $\Delta P > 0$), then it is realized only with a probability of $e^{-\Delta P / T}$. The parameter T corresponds to temperature and determines the inclination of the system to go into a market state with a higher portfolio value. As the search process continues, T – and thus, the inclination of the system to realize market states with higher portfolio values – is gradually reduced. The number of searches and the extent by which the parameter T is reduced are determined in an "annealing schedule". The selection of the annealing schedule is vital for the efficiency of the algorithm.

The simulated annealing method carries a lower risk of getting stuck in a local minimum than other minimization algorithms, because the process can also move to market states with higher portfolio values. The step-by-step reduction of the parameter T corresponds to the gradual shift from rough searches to fine-tuned searches.

As the risk factors have a continuous domain, the simulated annealing method is more difficult to implement in this case than in the minimization of functions with a discrete domain. An implementation is given in Press et al. (1992, pp. 451-455). For the purposes of a search for worst-case scenarios, the risk factor domain may also be discretized, provided that a sufficiently fine-meshed grid is selected. The sharper the peaks of the valuation function, the more fine-meshed the grid to be selected.

Bibliography

Regulatory Publications

Basle Committee on Banking Supervision (1995): *An Internal Model-Based Approach to Market Risk Capital Requirements*, proposal for consultation.

Basle Committee on Banking Supervision (1996): *Amendment to the Capital Accord to Incorporate Market Risks.*

Provisions of the Austrian Banking Act as promulgated in the Federal Law Gazette No. 753/1996 and No. 757/1996.

CAD: *Council Directive 93/6/EEC of 15 March 1993 on the capital adequacy of investment firms and credit institutions*, Official Journal of the European Communities No. L 141 of 11 June 1993, pp. 1-26.

CAD II: *Directive 98/31/EC of the European Parliament and of the Council of 22 June 1998 amending Council Directive 93/6/EEC on the capital adequacy of investment firms and credit institutions*, Official Journal of the European Communities No. L 204 of 21 July 1998, pp. 13-25.

Derivatives Policy Group (1995): *Framework for Voluntary Oversight*. To be accessed at http://risk.ifci.ch.

International Monetary Fund: *Annual Report on Exchange Arrangements and Exchange Restrictions*. To be accessed at http://www.imf.org.

Federal Ministry of Finance: *Regulation on Internal Models for the Limitation of Market Risks*, Federal Law Gazette II, No. 179/1997).

Scientific Publications

Acar E. and James S. (1997): Maximum Loss and Maximum Drawdown in Financial Markets, International Conference on Forecasting Financial Markets, London, UK, May 1997.

Brooks R. and Yong Yan D. (1999): London Inter-Bank Offer Rate (LIBOR) versus Treasury Rate: Evidence from the Parsimonious Term Structure Model, *Journal of Fixed Income*, June 1999, pp. 71-83.

Gay G. D., Kim J. and Nam J. (1999): The Case of the SK Securities and J.P. Morgan Swap: Lessons in VaR Frailty, *Derivatives Quarterly*, Spring 1999, pp. 13-26.

Kupiec, P. H. (1998): Stress Testing in a Value at Risk Framework, *Journal of Derivatives*, Fall 1998, pp. 7-24.

Nelder J. A. and Mead R. (1965): *Computer Journal* **7**, pp. 308-313.

Niederreiter H. (1992): *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia PA.

Press W. H., Teukolsky S. A., Vetterling W. T. and Flannery B. P. (1992): *Numerical Recipes in C: The Art of Scientific Computing,* 2nd edition, Cambridge University Press, Cambridge, UK.

Shaw J. (1997): Beyond VAR and Stress Testing. In: *VAR – Understanding and Applying Value-at-Risk*, KPMG/Risk Publications, pp. 221-224.