
COMMITTEE ON THE GLOBAL FINANCIAL SYSTEM

April 2000

**STRESS TESTING BY LARGE FINANCIAL INSTITUTIONS:
CURRENT PRACTICE AND AGGREGATION ISSUES**

BANK FOR INTERNATIONAL SETTLEMENTS
Basel, Switzerland

Copies of publications are available from:

Bank for International Settlements
Information, Press & Library Services
CH-4002 Basel, Switzerland

Fax: +41 61 / 280 91 00 and +41 61 / 280 81 00

This publication is available on the BIS website (www.bis.org).

© *Bank for International Settlements 2000.*

All rights reserved. Brief excerpts may be reproduced or translated provided the source is stated.

ISBN 92-9131-601-6

COMMITTEE ON THE GLOBAL FINANCIAL SYSTEM

April 2000

**STRESS TESTING BY LARGE FINANCIAL INSTITUTIONS:
CURRENT PRACTICE AND AGGREGATION ISSUES**

BANK FOR INTERNATIONAL SETTLEMENTS
Basel, Switzerland

Table of contents

Executive summary	1
Chapter 1 The current use of stress tests by financial firms	6
What is a stress test?	6
Stress testing and value-at-risk	7
How do risk managers use stress tests?	8
Scenario analysis methodologies	12
Recent trends in the use of stress tests	13
Limitations of stress tests	14
Chapter 2 Aggregate stress tests	16
What is an aggregate stress test?	16
How might risk managers use aggregate stress tests?	16
How might central banks and financial regulators use aggregate stress tests?	18
Methodology of aggregate stress tests	18
Limitations to the potential usefulness of aggregate stress tests	23
Chapter 3 Discussion and recommendations	25
Annex 1 Bibliography	27
Annex 2 Market liquidity, one-way markets and stressful shocks	29
Annex 3 Market dynamics in times of stress	33
Members of the Working Group on Macro Stress Testing	34

Executive summary

The activities of large, internationally active financial institutions have grown increasingly complex and diverse in recent years. This increasing complexity has necessarily been accompanied by a process of innovation in how these institutions measure and monitor their exposure to different kinds of risk. One set of risk management techniques that has attracted a great deal of attention over the past several years, both among practitioners and regulators, is “stress testing”, which can be loosely defined as the examination of the potential effects on a firm’s financial condition of a set of specified changes in risk factors, corresponding to exceptional but plausible events.

This report represents the findings of a Working Group on Macro Stress Testing established by the Committee on the Global Financial System. The group was asked to investigate the current use of stress testing at large financial institutions, in line with the Committee’s overall mandate to improve central banks’ understanding of institutional developments relevant to global financial stability. The term “macro” in the group’s name indicates another element of the group’s mandate, namely to explore the possibility that aggregating financial firms’ stress test results might produce information that is of use to central banks, other financial regulators, and private-sector practitioners.

Members of the group interviewed risk managers at more than twenty large, internationally active financial institutions, both in their home countries and as a group at a meeting hosted by the Banque de France. From these interviews, the group gained a substantial base of knowledge on the current “state of the art” in the design and implementation of stress tests and on the role of stress testing in risk management decisions at the corporate level.

Drawing on this knowledge, the group then considered some of the issues relating to the aggregation of the results of stress tests conducted at different financial firms. The group concluded that, under ideal circumstances, aggregate stress tests could potentially provide useful information in a number of areas. Aggregate stress tests might be used by financial firms to help make *ex ante* assessments of market liquidity risk under stress when evaluating the riskiness of a trading strategy. Central banks and financial regulators might use them to more effectively monitor broad patterns of risk-taking and risk-intermediation in financial markets. However, the group also noted that it is as yet unclear whether such ideal circumstances prevail. In particular, it is unclear whether an appropriate reporting population can be assembled, whether the stress tests currently conducted by financial firms are compatible with one another, and whether the information obtained would justify the reporting burden.

The report concludes that stress testing is likely to remain an important element of the risk management strategies of large financial firms, and that further information about stress testing practices could prove informative regarding the vulnerabilities faced by the financial system. Accordingly, the report recommends that a one-off survey of the scenarios currently used by risk managers be conducted.

The report is organised as follows. The first chapter summarises current practices in stress testing and discusses some of its limitations. Chapter 2 of the report examines the potential usefulness of aggregate stress tests, discusses the methodologies that could be used to construct aggregate stress tests and gives a preliminary discussion of the trade-offs that would be involved in an aggregate stress testing program. Chapter 3 presents and discusses the proposed census of scenarios. There are three annexes to the report: a bibliography, a conceptual discussion of what aggregate stress tests might tell us about market liquidity risk, and a discussion of the issues raised by the dynamic aspects of market behaviour under stress.

1. The current use of stress tests by financial firms

What is a stress test?

“Stress testing” has been adopted as a generic term describing various techniques used by financial firms to gauge their potential vulnerability to exceptional but plausible events. The most common of these techniques involve the determination of the impact on the portfolio of a firm or business unit of a move in a particular market risk factor (a *simple sensitivity test*) or of a simultaneous move in a number of risk factors, reflecting an event which the firm’s risk managers believe may occur in the foreseeable future (*scenario analysis*). The scenarios are developed either by drawing on a significant market event experienced in the past (*historical scenarios*) or by thinking through the consequences of a plausible market event which has not yet happened (*hypothetical scenarios*). Other techniques used by some firms to capture their exposure to extreme market events include a *maximum loss approach*, in which risk managers estimate the combination of market moves that would be most damaging to a portfolio, and *extreme value theory*, which is the statistical theory concerned with the behaviour of the “tails” of a distribution of market returns.

Stress testing and value-at-risk

In most of the interviewed firms, stress tests supplement value-at-risk (VaR). VaR is thought to be a critical tool for tracking the riskiness of a firm’s portfolio on a day-to-day level, and for assessing the risk-adjusted performance of individual business units. However, VaR has been found to be of limited use in measuring firms’ exposures to extreme market events. This is because, by definition, such events occur too rarely to be captured by empirically driven statistical models. Furthermore, observed correlation patterns between various financial prices (and thus the correlations that would be estimated using data from ordinary times) tend to change when the price movements themselves are large. Stress tests offer a way of measuring and monitoring the portfolio consequences of extreme price movements of this type.

How do firms use stress tests?

Stress tests enable managers to track a firm’s exposure to price changes during events that are considered plausible, without obliging them to develop a statistical model for such events. This, in turn, allows senior management and business-unit heads to determine whether the firm’s exposures correspond to its risk appetite. Because of their intuitive appeal, stress tests are thought to facilitate the dialogue between risk managers, senior managers and business-unit heads about the risks taken by the firms and methods for monitoring and managing those risks. From this process decisions emerge regarding such matters as the limits set on proprietary position-taking, capital charges on traders and trading units, and the appropriateness of the risk-managers’ modelling assumptions. It should be emphasised that stress tests are typically only one element of the process through which a financial firm develops its quantitative and qualitative risk-management policies.

Consolidated stress testing at the firm level was introduced in response to the amendment to the Basel Capital Accord in 1996, which made approval of the “models approach” to a firm’s market risk capital requirement conditional on the presence of a firm-wide stress test program. However, individual trading desk managers at many firms had, at their own initiative, put in place stress test programs well before the creation of corporate-level programs. Particular interest has been devoted to stress testing since the financial crises in Asia in 1997 and the turbulent events of the autumn of 1998. The group found that most of the interviewed firms are increasing the resources devoted to developing stress tests.

Scenario analysis methodologies

The interviewed firms tended to use both historical and hypothetical scenarios. Historical scenarios are easier to formulate and to understand intuitively. Hypothetical scenarios allow risk managers to challenge the common tendency to pay more attention to past events than future dangers. In both

cases, scenarios are chosen with an eye to markets and business segments in which the firm is highly involved. The usefulness of the scenarios is enhanced when they are run at periodic intervals, allowing the firm's exposure to be tracked over time. Corporate-level stress tests are facilitated by the presence of firm-wide information-technology systems which offer risk managers an up-to-date, usable, accurate source of information on exposures.

Limitations of stress testing

A stress test estimates the exposure to a specified event, but not the probability of such an event occurring. In addition, numerous decisions in the specification of a stress test must be made that rely on the judgement and experience of the risk manager. There is thus no guarantee that the risk manager will choose the "right" scenarios or interpret the results effectively. Stress tests also impose a high computational cost, particularly in collecting the data from diverse business units and from the need to revalue complex options-based positions. A further limitation is that, at present, firms cannot integrate market and credit risks in a systematic way in their stress tests, although some interviewed firms are engaged in efforts in this direction.

2. Aggregate stress tests

What is an aggregate stress test?

An aggregate stress test is a measure of the risk exposure of a group of reporting firms to a specified stress scenario. Each reporting firm would provide information on its own exposure under the stress scenario. The responses would be aggregated by a central co-ordinator. The aggregation could produce a single number capturing the combined exposure of all reporting firms. In addition, information on the distribution of exposures among firms as well as across markets and risk factors could also be captured.

The potential uses of aggregate stress tests by risk managers

The primary area of risk management in which financial firms might use aggregate stress test information would be to assess market liquidity risk. When a firm enters into a trading strategy, it makes an explicit or implicit judgement about the market liquidity risk involved, drawing on its knowledge of the strategy's liquidity needs and its estimation of possible market liquidity conditions under stress. The firm incorporates the resulting judgement in formulating its corresponding risk management strategy, for example by its selection of the presumed holding periods for the positions involved. One of the ways in which firms might use aggregate stress tests might thus be to set these assumed holding periods more accurately.

There are many channels through which a market shock could lead to a rapid evaporation of liquidity. Aggregate stress tests could potentially be informative about the presence of one of these channels, namely the emergence of a "one-way market" in which price adjustment is hindered because of an imbalance between sellers and buyers. The report explores factors that might influence the informativeness of aggregate stress tests regarding the potential emergence of such a market. Aggregate stress tests might be useful to market participants as a complement to existing sources of information, such as order flows through their trading desks, public and proprietary news sources, and informal contacts with other participants.

The sum of reported stress test exposures would indicate the intensity of trading that might be triggered by a market shock. An aggregate stress test would be more informative than a compilation of static aggregate position figures to the extent that non-linear (e.g. options-based) positions are present. The usefulness of the results would then depend on the composition of the reporting population relative to the market as a whole. If the reporting firms are considered more likely to adjust their positions following a market shock than most non-reporters a larger reported stress test exposure could indicate a greater risk of a one-way market. This might be the case, for example, if non-reporters tend

to be households and pension funds who adjust their positions slowly, while reporters include dealing firms who need to rebalance their positions rapidly in volatile markets.

Information on the distribution of stress exposures across participants would provide further insights into the extent to which different kinds of trading behaviour may be present in a given market. Distributional information could also be useful if it provides insights into the concentration of exposures in a market. If a few participants have very large exposures relative to the size of the market, then a shock, if unanticipated, would be more likely to trigger a drying up of liquidity.

The potential uses of aggregate stress tests by central banks and financial regulators

Central banks and other public authorities already collect a wide range of data, such as the BIS consolidated banking and derivatives market statistics, that are relevant for monitoring risk-taking and risk-intermediation in financial markets. They also monitor markets through discussions with participants and through their market operations. Aggregate stress tests could potentially complement these information sources by offering “forward-looking”, rather than historical, information on aggregate risk taking behaviour. A high level of aggregate exposure to a stress scenario, relative to exposures that had been observed in the recent past, could, in combination with information from other sources, indicate a potential systemic vulnerability.

Methodology of aggregate stress tests

Aggregate stress tests, like those conducted at the firm level, can be based on historical or on hypothetical scenarios. One way to identify likely candidate scenarios would be to survey firms on the scenarios they use. Even if this is done, however, there still exists the danger that market participants could misinterpret the central banks’ interest in a particular scenario as containing information on central banks’ views of what shocks are likely to occur. This concern might be alleviated if a process were adopted that assigned private sector market participants responsibility for specifying scenarios. Another issue that would have to be addressed would be the degree and nature of public disclosure of aggregate stress test results.

There is a danger that, as with firm-level stress tests, the chosen aggregate scenarios might become less relevant to the exposures that are of concern to the intermediaries as time passes. A second concern is that not enough aggregate scenarios would be chosen to adequately capture the set of risks faced by the financial system. A third concern might be that, even under the high-burden alternative of asking all firms to use identical scenarios, their valuation methods might differ so much as to make the results non-comparable.

As a framework for better understanding the trade-off between the information gained from an aggregate stress test and the reporting burden, the report identifies four hypothetical reporting populations that might participate in an aggregate stress exercise, ranging from very broad (all actual and potential frequent participants in financial markets) to very narrow (only dealers), and discusses the kinds of data which might be collected from each population. For a narrower reporting population, an aggregate stress test would be less difficult to organise but also less useful.

Limitations to the potential usefulness of aggregate stress tests

The limitations of aggregate stress tests for firms, central banks and financial regulators should also be recognised. They would not be informative regarding the risks facing individual banks, or, given the limits of current risk measurement technology, about the interaction of market and credit risk. The presence of aggregate scenarios may make banks less inclined to develop scenarios specific to their own circumstances. They would be unlikely to give an adequately timely picture of current aggregate risk exposures, given the time needed for compilation and the speed with which exposures change. Aggregate stress tests would also not necessarily be informative about the dynamic aspects of the responses of market participants to extreme events, a topic considered further in an annex to the report.

3. Conclusion and recommendations

At this time, the Working Group does not believe it has enough information to judge whether aggregate stress tests would be useful to firms, central banks, or financial regulators. Given the specific shortcomings that are discussed in the report, as well as the broader uncertainties about the usefulness of aggregate stress tests, it is the view of the Working Group that it does not currently have enough information to make an informed judgement as to whether aggregate stress tests would provide value added relative to current methods of monitoring financial markets.

However, the Working Group believes that now is a fruitful time for central banks and other financial regulators to rethink what data are most useful for monitoring the functioning of financial markets. In response to innovations in the science of risk management and in information technology, firms have developed information systems for collecting firm-wide risk data. Perhaps at some future date, as financial institutions and risk management systems evolve, aggregate stress tests will be found to be a way to tap into this new data source that would provide forward-looking information about aggregate risk exposures that would be of use to financial firms, central banks, and other financial regulators.

As a way to add to the overall transparency of the risk-management process and allow firms to improve information sharing, at a relatively low cost in terms of reporting burden, the report recommends that a one-off census of scenarios in use at dealer firms be conducted. A census of the stress tests used for internal risk management by large dealers would indicate the risks dealers are concerned about. Such a census would also indicate the markets and risk factors in which firms' risk-taking activities are concentrated.

Chapter 1

The current use of stress tests by financial firms

The activities of large, internationally active financial institutions have grown increasingly complex and diverse in recent years. This increasing complexity has necessarily been accompanied by a process of innovation in how these institutions measure and monitor their exposure to different kinds of risk. Stress testing refers to a set of risk management techniques that has attracted a great deal of attention over the past several years, both among practitioners and regulators.

This report is organised as follows. The remainder of this chapter summarises current practices in stress testing. It is based on interviews conducted by the Working Group with banks and investment banks.¹ The chapter first defines different types of stress tests. The rationales for employing stress tests are also discussed. The chapter concludes with a discussion of some limitations of stress testing.

Chapter 2 of the report explores the potential usefulness of aggregate stress tests. Aggregate stress tests could potentially be used by financial firms to help make ex ante assessments of market liquidity risk under stress when evaluating the riskiness of a trading strategy. The methodologies that could be used to construct aggregate stress tests are also discussed. Chapter 2 also provides a preliminary discussion of the trade-offs that would be involved in an aggregate stress testing program. Chapter 3 recommends a census of scenarios in use at large financial institutions as a way to improve understanding about the uses of stress tests and about perceived systemic vulnerabilities. There are three annexes to the report: a bibliography, a conceptual discussion of what aggregate stress tests might tell us about market liquidity risk, and a discussion of the issues raised by the dynamic aspects of market behaviour under stress.

What is a stress test?

“Stress testing” has been adopted as a generic term describing various techniques used by financial firms to gauge their potential vulnerability to exceptional, but plausible, events. The term covers many different techniques. The four discussed here are listed in Table 1 along with the information typically referred to as the “result” of that type of a stress test.

A *simple sensitivity test* isolates the short-term impact on a portfolio’s value of a series of predefined moves in a particular market risk factor. For example, if the risk factor were an

Table 1
Stress testing techniques

Technique	What is the “stress test result”?
Simple sensitivity test	Change in portfolio value for one or more shocks to a single risk factor
Scenario analysis (hypothetical or historical)	Change in portfolio value if the scenario were to occur
Maximum loss	Sum of individual trading units’ worst-case scenarios
Extreme value theory	Probability distribution of extreme losses

¹ Members of the group met with selected private-sector risk managers in their respective home countries in the first half of 1999. At its February 1999 meeting in Paris, the Working Group met with a group of risk managers assembled with the help of the Institute of International Finance. No attempt was made to develop a representative or comprehensive sample of banks, or even of large banks. Instead, interviewees were contacted on an informal basis if their range of activities seemed to call for sophisticated risk management methods or if they were reputed to be able to offer a useful perspective on stress-test methodology.

exchange rate, the shocks might be exchange rate changes of +/- 2 percent, 4 percent, 6 percent and 10 percent.²

A *scenario analysis* specifies the shocks that might plausibly affect a number of market risk factors simultaneously if an extreme, but possible, event occurs. It seeks to assess the potential consequences for a firm of an extreme, but possible, state of the world. A scenario analysis can be based on an historical event or a hypothetical event. Scenario analysis is currently the leading stress testing technique. In the Working Group's interviews, we discovered that almost all of the interviewed firms were increasing resources devoted to scenario analysis.

A *maximum loss* approach is in use at a few firms interviewed by the Working Group. This technique assesses the riskiness of a business unit's portfolio by identifying the most potentially damaging combination of moves of market risk factors. Interviewed risk managers who use such "maximum loss" approaches find the output of such exercises to be instructive but they tend not to rely on the results of such exercises in the setting of exposure limits in any systematic manner, an implicit recognition of the arbitrary character of the combination of shocks captured by such a measure.

A small number of firms are exploring the use of *extreme value theory* (EVT) as a means to better capture the risk of loss in extreme, but possible, circumstances.³ EVT is the statistical theory on the behaviour of the "tails" (i.e., the very high and low potential values) of probability distributions. Because it focuses only on the tail of a probability distribution, the method can be more flexible. For example, it can accommodate skewed and fat-tailed distributions. A problem with the extreme value approach is adapting it to a situation where many risk factors drive the underlying return distribution.⁴ Moreover, the usually unstated assumption that extreme events are not correlated through time is questionable.⁵ Despite these drawbacks, EVT is notable for being the only stress test technique that attempts to attach a probability to stress test results.

Stress testing and value-at-risk

In most of the interviewed firms, stress tests supplement value-at-risk (VaR). VaR is used to provide a probability-based boundary on likely losses for a specified holding period and confidence level (for example, the maximum loss that is likely to be experienced over one day with a 99 percent level of confidence). Firms employ VaR prospectively, to assess the risk of potential portfolio allocations, and retrospectively, to assess the risk-adjusted performance of individual business units.

Firms recognise the limited ability of statistical models such as VaR to accurately capture what happens in exceptional circumstances. In part, this is due to modelling assumptions that make it easier to compute VaR. However, there is a more fundamental problem with using statistical models like VaR for assessing risks in exceptional circumstances. By definition, exceptional circumstances occur rarely, and statistical inference is imprecise without a sufficient number of observations. Stress tests partially fill this gap, and thus complement VaR, by offering a quantitative picture of the exposure associated with a possible extreme event. In the absence of a reliable statistical measure of the probability of such an event, stress testing calls on the informed judgement of risk managers and senior executives to assess whether, and to what degree, the firm should move to limit or modify such an exposure.

² A slight variant is to simultaneously shock blocks of linked market risk factors, such as the level and volatility of an equity index.

³ See McNeil (1999) for an accessible overview of the potential use of EVT by risk managers. See also Danielsson and de Vries (1997) and Longin (1999).

⁴ McNeil (1999) and Longin (1999) examine models with multiple risk factors.

⁵ See Van den Gorbergh and Vlaar (1999) for a critical evaluation of the application of the extreme value approach to distributions of stock returns.

Even if a statistical model could be built that accurately captured risk in extreme circumstances, risk managers and senior management appear likely to prefer to continue using stress tests, because the assumptions underlying such a statistical model would not be transparent. This observation is supported by the Working Group's finding that only a few of the interviewed firms were exploring new statistical models of extreme circumstances (such as extreme value theory), while nearly all interviewed firms plan to devote more resources to methods of stress testing not based on statistical models.

How do risk managers use stress tests?

Stress tests *produce information summarising the firm's exposure to extreme, but possible, circumstances*. Risk managers at interviewed firms frequently described their roles within firms as assembling and summarising information to enable senior management to understand the strategic relationship between the firm's risk-taking (such as the extent and character of financial leverage employed) and risk appetite.⁶ These risk managers were aware that senior managers want their firms' decisions on risk taking to be influenced by the size of losses that could occur in extreme, but possible, circumstances. This applies to decisions taken at a senior level on the firm's overall risk exposure and to decisions taken at the individual business unit level.⁷

Typically, the results of a small number of stress scenarios are computed on a regular basis and monitored over time. Time series of stress test results are one of the key components of the risk information provided to senior management. In firms that restrict proprietary trading, the time series of stress test results tends to confirm the unchanging character of the firm's major exposures. On the other hand, for firms with sizeable proprietary trading books, regular stress testing reveals the variability of the firm's large exposures. One implication of this difference is that members of the latter group find stress tests to be especially valuable since the time series results can be readily interpreted as indicators of changes in risk appetite.

In some interviewed firms, the corporate-level risk management function has little influence on risk-taking behaviours within the firm. In some of these, the results of stress tests are dispersed within the firm to inform, but not to control, risk-taking, and the results are not shared with senior management. In others, stress test results are shared with senior management but are not viewed as a component of firm-wide risk control. In one of these firms, the decision not to develop stress test capacities has been linked to senior management's policy of controlling risk-taking through management and staff compensation systems that impose ex post penalties on excessive risk-taking.

Some of the specific ways stress tests are used to influence decision-making that were mentioned by the risk managers interviewed by the Working Group are to:

- manage funding risk
- quantify tail risk
- provide a check on modelling assumptions
- set limits for traders
- determine capital charges on trading desks' positions

In most of the interviewed firms, senior managers use stress tests to help them *make decisions regarding funding risk*. Managers have come to accept the need to manage risk exposures in anticipation of unfavourable headlines. This is done to minimise the impact of such news on the firm's

⁶ A thorough discussion of various concepts of financial leverage can be found in the report of the Counterparty Risk Management Policy Group (1999).

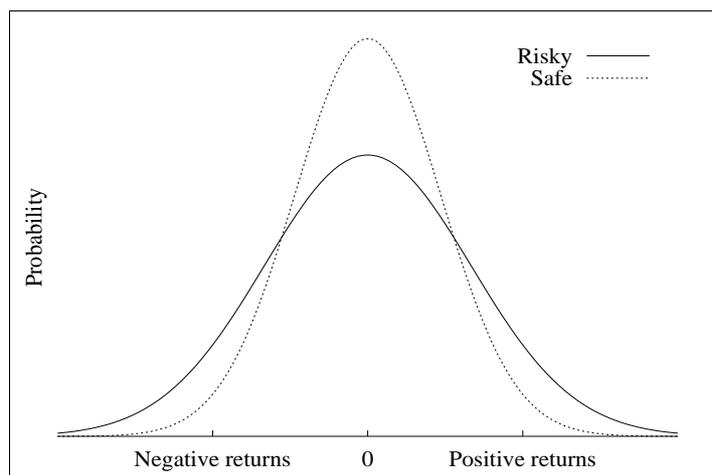
⁷ See Schachter (1998), an overview discussion of the role of the stress test program in a US money center bank's market risk management efforts.

access to funding and other markets.⁸ Several influential events in recent years have illustrated how news of large losses can impact on a bank's cost of funding or its access to particular sources of funds.

The significance of such information will vary according to a firm's exposure to funding or liquidity risk. Portfolio managers who mark-to-market frequently will seek to avoid exposures whose value is sensitive to stress losses. Others will be less inclined to avoid such exposures.

A second use of the information derived from stress tests is to *quantify tail risk*.⁹ "Tail risk" is defined and illustrated in the following example.

Figure 1
Two bell-curve-shaped distributions



A firm's risk positions give rise to a probability distribution of future returns. It is this distribution that a risk manager must summarise. Figure 1 shows the probability distribution for two portfolios, labelled "Risky" and "Safe." For these simple portfolios, a risk manager could summarise the choice between the two portfolios by presenting just one set of numbers: the standard deviation of returns or the VaR.¹⁰ Stress tests would be superfluous.

In reality, a risk manager's job is complicated by the unlimited variety of financial instruments available for risk-taking. The distribution of future returns need not follow a smooth bell curve as in Figure 1 – it can take on any conceivable shape.

Figure 2 shows two future return distributions with equal mean and standard deviation.¹¹ Clearly, Distribution 1 has a greater risk of large losses than Distribution 2. Distribution 1 can be said to have more "tail risk."

If large losses carry an especially heavy cost to the firm, senior management may use stress tests to guide the firm away from risk profiles with excessive tail risk. Alternatively, the firm could charge its customers higher premia in exchange for bearing this risk.

Both simple sensitivity tests and scenario analyses are thought to be particularly well-suited for revealing the vulnerability of portfolios to tail risk that might go undetected were the firm to rely exclusively on other risk measures such as VaR. Simple sensitivity tests could reveal if a firm was exposed to a large move in a market risk factor (perhaps because it had sold "out of the money" option positions to customers). Scenario analysis on the other hand is used to examine the potential knock-on effects of a large market move. For example, in response to a surge in the sovereign credit risk of a major Asian country, what reactions should be anticipated for other risk factors such as exchange rates or the credit risk spreads of the traded sovereign debt of neighbouring countries? Such a shock to credit risk may have happened often enough in the past to provide an indication of the impact on

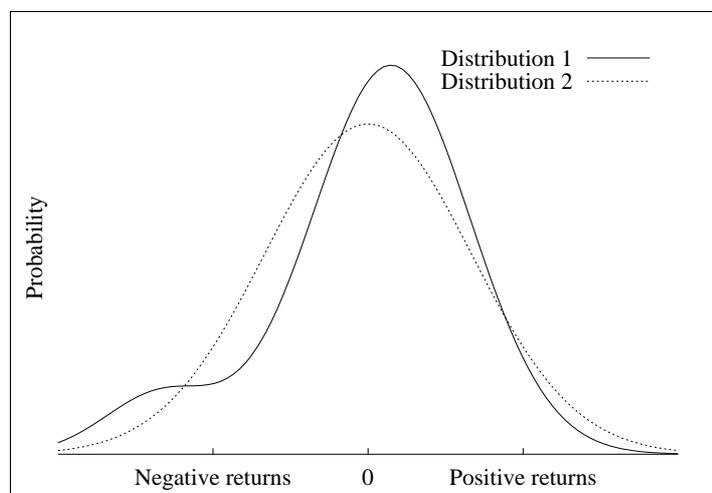
⁸ In a similar vein, senior management also is supplied with stress test results concerning the impact of potential market moves on the firms' assets that can be posted as collateral.

⁹ One risk manager interviewed described this as "identifying the black holes" in the portfolio.

¹⁰ For normally distributed portfolios like those in the figure, the VaR will simply be a multiple of the standard deviation.

¹¹ It would be easy to construct an example where the two portfolios had equal value-at-risk instead of equal standard deviation, and the same argument would apply.

Figure 2
Two future-return distributions with equal means, equal standard deviations, but different skewness



spreads for neighbouring countries, but not often enough to provide forecasts that are reliable in a statistical sense. Instead, a scenario is explicitly developed drawing on the judgement of economists, lending officers and financial analysts.

In a discussion of one firm's use of stress tests, the risk manager emphasised the detection of "gamma holes," a reflection of the firm's business strategy of being a leading dealer in option products.¹² The risk manager noted that risk measures extrapolated from small changes in the underlying risk factors would not detect gamma holes because dynamic hedging (continuously adjusting a hedge position in order to manage the risk associated with options) can

effectively insulate the value of the consolidated position against small shocks. Accordingly, to detect gamma holes, risk measures should be based on large moves in market risk factors, for example through the conduct of stress tests. This illustrates the general point that knowing what business lines and trading strategies the firm has pursued informs the risk manager as to which stress test results would be most appreciated by senior management.

In the same manner, knowing the assumptions employed by the firm in its VaR modelling is relevant information to focus stress testing on the likely sources of "model risk." For example, a firm whose VaR model does not fully capture the "fat tails" found in empirical distributions of many market risk factors may rely heavily on stress testing to control tail risk.

Scenario analysis is also used to *highlight the role of particular correlation and volatility assumptions* in the construction of firms' portfolios of market risk exposures. In this case, scenario analysis can be thought of as a means through which firms check on the portfolio's sensitivity to assumptions about the extent of effective portfolio diversification. For example, it has been argued that correlations between international equity returns are higher during bear markets than during bull markets. This would suggest that the benefits of international diversification are less impressive than conventionally assumed.¹³

In some firms, stress tests are used to *set limits*. Simple sensitivity tests may be used to put hard limits on business units' market risk exposures.¹⁴ Such hard limits effectively set the boundaries for firms'

¹² A gamma hole can be present in a firm's portfolio when the firm has sold options and reduces the risk of the short position by dynamic hedging. The firm dynamically hedges by holding a position in the underlying asset that is continuously adjusted so that the value of the consolidated portfolio (option plus hedge) is insensitive to a small move in the underlying risk factor. A large negative value of gamma (which measures the rate at which the hedge must be adjusted), which typically occurs when at-the-money options are near expiration, signals a situation where the hedge portfolio must be immediately adjusted and access to market liquidity is especially important to the firm. This situation is called a gamma hole.

¹³ Statistical models have been developed which explicitly incorporate assumptions about how volatility and correlation vary over time, but the usefulness of such models in a risk management context has yet to be proven. This is particularly so given the difficulty of modelling very rare events, as noted earlier.

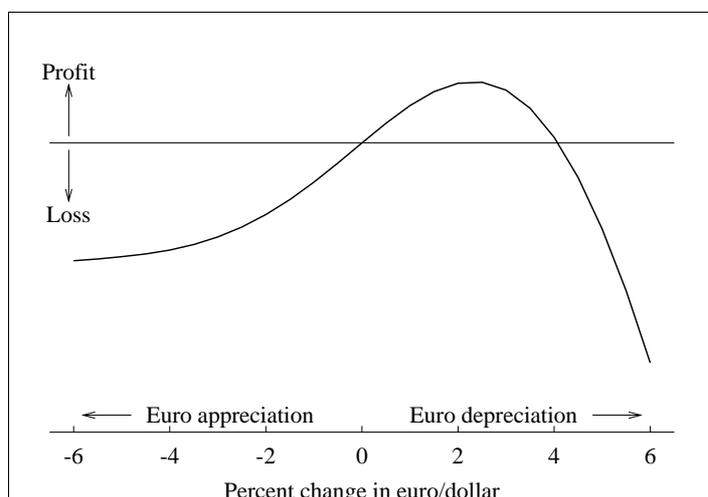
¹⁴ In such firms, VaR is typically used to set *soft* limits, that is, trigger points for discussions between risk managers and business unit heads.

risk exposures.¹⁵ A number of interviewed firms argued that the setting of hard limits through a stress test exercise was regarded as prudent in light of the typical staff compensation package of a financial trading firm which is designed to generously reward successful risk-taking through the payment of unit bonuses. Large losses, on the other hand, are absorbed by the firm's capital and typically not factored into future bonuses.

Stress testing plays a key role at firms that routinely enter into over-the-counter (OTC) options and other financial contracts with non-linear payoffs, according to the interviews. Such firms place a heavy emphasis on stress-test-based limits. The following example illustrates how stress tests can generate information that might be used to set limits.

Consider a hypothetical firm holding a portfolio of OTC options.¹⁶ For simplicity, the portfolio is assumed to consist of long and short positions in euro/dollar call options. Figure 3 shows how the value of the firm's portfolio depends on movements in the exchange rate. Appreciation of the euro would generate losses at a decreasing rate. Depreciation of the euro would produce a profit, but only up to a point. For larger depreciations of the euro the portfolio loses money at an increasing rate.¹⁷

Figure 3
Change in the value of a portfolio of a €/\$ option portfolio as a function of the change in the exchange rate



A portfolio like that shown in Figure 3 could lose an unlimited amount of money if the euro depreciates far enough. Conducting stress tests would reveal the potential for large losses. Stress-test-based limits would prevent a firm from holding such a portfolio above the size threshold imposed by the stress test limit.

In one firm, *capital charges* of individual business units are based on hypothetical losses under certain stress scenarios. The losses are measured at the level of the individual transaction. The capital charges are deducted from each business unit's bonus pool. This procedure was designed to provide each business unit with an economic incentive to reduce the risk of extreme losses.

The capacity to identify individual transactions at this bank was attributed to its corporate data management strategy. To implement such a strategy, a firm develops a central database (or "data warehouse") to capture and store details of individual transactions. For the example portfolio from Figure 3 above, the database would store details of the option contracts that would account for the bulk of the loss in the event of a depreciation of the Euro. The centralised storage of such detailed data lets the firm's risk manager "drill down" from firm-wide scenario results to the underlying exposures (contracts) that would account for

¹⁵ In its 1998 annual report, UBS, a major international bank, commented that it employs stress-test-based limits to "limit the scale of losses which would occur on the days when the VaR limits are exceeded." The Counterparty Risk Management Policy Group (1999) report suggests that in the future more firms will supplement existing risk limits with stress test limits.

¹⁶ Some of a firm's option positions could be embedded in other financial contracts, such as structured notes purchased and structured notes issued.

¹⁷ Practitioners refer to this situation as "negative gamma." Gamma is the practitioners' term for the second derivative of the curve shown in Figure 3.

outsized losses. The banks that have made these investments in centralised databases have also tended to invest heavily in the detailed construction of scenarios.

Scenario analysis methodologies

In the interviews, two standard ways of developing scenarios were identified: historical and hypothetical scenarios. The interviewed firms use a mixture of the two.

Historical scenarios employ shocks that occurred in specific historical episodes. A simple way to do this is to identify days in the past that were “stressful” and use the observed changes in market risk factors on those days. For example, a portfolio of market risk exposures could be stress tested by seeing how its value would change given the changes recorded for market risk factors for a day, or over a longer period. The selection of the day, or period, is typically based on “headline” disturbances, such as the widening of credit risk spreads in financial markets in the autumn of 1998.

One advantage of this technique is that the structure of market factor changes is historical rather than arbitrary. The fact that the market moves used are historical enhances the credibility of the exercise from the point of view of senior management. By contrast, “worst-case” exercises which rely on aggregations of worst-case market factor draws tend to be disregarded because they are viewed by senior managers and other monitors of the firm’s risk-taking as not being plausible when taken together.

A second advantage of historical scenarios is their transparency. A statement like “if the October 1987 stock market crash happened tomorrow, the firm would lose X million dollars” is easy to understand. Since one of the roles of stress tests is to facilitate a firm’s internal “conversation” about the relationship between its risk-taking and its risk appetite, stress scenarios need to be understood by all participants in this conversation.

One disadvantage of historical scenarios is that firms may (consciously or unconsciously) structure their risk-taking to avoid losing money on shocks that have occurred in the past, rather than anticipating future risks that do not have a precise historical parallel. This could represent a conscious choice on the part of traders, if firms give traders an incentive to minimise exposure to stress tests through limits or capital charges. It could also represent an unconscious choice, if traders overestimate in their own minds the likelihood of shocks that they have first-hand experience of, relative to shocks that have never happened before. For example, bond traders who experienced the 1994 bond market crash may make sure they are hedged against a simultaneous increase in both risk-free and risky interest rates (no change in credit spreads), putting less importance on being hedged against a drop in the risk-free rate along with an increase in the risky interest rate. One reason why firms choose to do hypothetical as well as historical stress tests is that hypothetical stress tests allow them to assess the extent to which the conventional wisdom (based on the history of recent market moves) may be driving position-taking.

A second disadvantage of historical scenarios is that they may be difficult to apply to products which did not exist at the time of the historical event in question, or to risk factors whose behaviour has changed in a significant way since that event.

Hypothetical scenarios use a structure of shocks thought to be plausible in some foreseeable, but unlikely circumstances for which there is no exact parallel in recent history (though the assumptions that go into constructing these scenarios often draw on historical experience). One example of a scenario that has been routinely run is a so-called “flight to quality.” Typically, flight-to-quality scenarios involve shocks to credit spread relationships, such as that between high-yield US corporate debt and US Treasury bonds. The specification of such scenarios has subsequently been tweaked, based on what happened to liquidity premia in the fall of 1998. Emphasis has shifted to so-called deleveraging scenarios rather than simple flight-to-quality scenarios. The new scenarios place more emphasis on liquidity-related movements in spreads, such as the spread between on- and off-the-run US Treasury securities.

When specifying scenarios, risk managers rely on firm-specific information. Knowing what risks are in the firm's portfolio is relevant information for identifying likely sources of extreme risks. For example, one interviewed firm active in OTC commodity derivatives uses a "Gulf War" scenario to stress test its commodity-related exposures, while firms that are not active in commodity-related businesses would be unlikely to employ such a scenario.

In their construction of scenarios, firms attempt to incorporate "contagion" effects that are reflected in the size of assumed shocks to related markets and in the assumed correlation structure among these shocks. The values used by firms for contagion effects are generally based on judgement and historical experience, rather than on formal models of market behaviour.

Risk managers also draw on past experience in accounting for the risk of a sudden drop-off in market liquidity.¹⁸ In their adjustments for liquidity effects, firms tend to incorporate the judgement that in the aftermath of the shock they will be unable to access markets for some period, so that over this period, they will face cumulative changes in market risk factors rather than the usual daily changes. An historical scenario might incorporate the observed cumulative change in a given market risk factor over a number of trading days, rather than for one trading day. For example, if in a previous episode a price index fell 10 percent in one day and 30 percent for a week, the stress test would be based on the 30 percent value to take into account the assumed inability to use the market for the full week.

Firms produce firm-wide scenario analyses by aggregating information from individual business units. One approach is to revalue each business unit's risk positions under the exhaustive set of market risk factor shocks defined under the scenario, and then simply add up the business units' results to get the firm-wide stress exposure. Another practice encountered begins by asking each business unit to run a series of simple sensitivity tests. For example, swap curves might be shocked at 5 basis point intervals.¹⁹ Individual business units' outputs are then aggregated by a firm's risk manager to produce the firm-wide stress exposure. The choice of this procedure reflects the high costs to the firm of conducting and assembling firm-wide stress tests.

The relative merits of the various methodologies for defining stress tests is open to debate. The interviews conducted by the Working Group revealed little standardisation in the market. On the other hand, the Working Group did discover considerable agreement on the character of scenarios to be run. For example, many firms reported running scenarios with common titles such as "1987 stock market crash," "1994 bond market crash," and a "hypothetical flight-to-quality." However, the interviews conducted by the Working Group did not produce the detailed information needed to assess the comparability of scenarios run under common titles. The information required could be gathered through a census of the scenarios that are currently run by firms, as recommended in Chapter 3 of this report.

Recent trends in the use of stress tests

Through its interviews with major banks and securities firms, the Working Group confirmed that many firms have upgraded their capacity to conduct firm-wide stress tests of market risk exposures in the wake of the Asian financial crisis and financial market events in the autumn of 1998.²⁰ Because of the extreme moves in market risk factors seen in the last two years, senior management is now more inclined to incorporate such information in the setting of trading limits. These events, rather than the

¹⁸ Liquid markets are those in which trading can occur with the least effect on price.

¹⁹ Business units may run such simple sensitivity tests for their own risk management purposes.

²⁰ Discussions of the potential significance of the Asian financial crisis for the risk management activities of international banks can be found in Institute of International Finance (IIF) (1999), Committee on the Global Financial System (CGFS) (1998), and BIS (1999b). The IIF's report was of particular interest to this Working Group since it incorporated the view of risk management practitioners that the Asian financial crisis demonstrated the need for bank's senior management to know how extreme market situations might affect portfolio values.

complexity of positions *per se*, have intensified senior management's desire to have consolidated information at the firm level to help them make decisions concerning their firm's strategic risk profile. For all but a few of the interviewed firms, corporate-level stress test programs were first established in response to the 1996 market risk amendment to the Basel Capital Accord, that is, in response to a regulatory mandate.²¹ However, nearly all interviewed firms reported substantial increases in budgeted resources devoted to these programs since their establishment, indicating that the programs' outputs have become valued by firms' senior managers.

At most of the interviewed firms, individual trading desk managers had, on their own initiative, put in place stress test programs well before the creation of corporate-level programs. One exceptional firm had set up a consolidated firm-wide stress testing program in 1992 for a then-separate subsidiary which conducted a specialised business involving the sale of complex, long-dated option products. In choosing to set up the stress testing program, senior management of the parent firm reportedly focused on a need for information that could be used to ensure that the subsidiary's activities conformed to the approved limits on risk-taking and did not expose the consolidated firm to losses that could threaten its viability. A significant investment was made in state-of-the-art computer systems that permitted firm-wide risk exposures to be disaggregated (for example, by business line, legal entity, risk type and customer) in order to identify sources and concentrations of risk.²²

Limitations of stress tests

Stress testing can appear to be a straightforward technique. In practice, however, stress tests are often neither transparent nor straightforward. They are based on a large number of practitioner choices as to what risk factors to stress, how to combine factors stressed, what range of values to consider, and what time frame to analyse. Even after such choices are made, a risk manager is faced with the considerable tasks of sifting through results and identifying what implications, if any, the stress test results might have for how the firm should manage its risk-taking activities.

A well-understood limitation of stress testing is that there are no probabilities attached to the outcomes. Stress tests help answer the question "How much could be lost?" The interviewed practitioners fully appreciated that the answer to this question is not as informative as would be the answer to the question "How much is likely to be lost?" The lack of probability measures exacerbates the issue of transparency and the seeming arbitrariness of stress test design, a point readily acknowledged by the risk managers interviewed by the Working Group. On the other hand, risk managers found it easier to base an internal dialogue on risk management issues on stress tests rather than VaR, which was felt to be even less transparent because of its reliance on hidden assumptions about probabilities.

²¹ The 1996 market risk amendment to the Basel Capital Accord conditions supervisory approval of a bank's use of its internal model on the bank's having put in place a rigorous and comprehensive stress testing program. The program is to cover all three major types of bank risk exposures—credit, market and operational. Banks are directed to engage in stress testing to identify "events or influences" that need to be taken into account in their assessments of their capital positions. In this regard, the Accord advises that banks need to cover in their tests those factors that can have a major impact on the value of their trading portfolios or can make the control of trading portfolio risks very difficult.

Bank supervisors have monitored compliance with the stress-test requirement of the market-risk amendment since its adoption. In doing so, bank supervisors have not required the adoption of a uniform approach. This absence of supervisory guidelines perhaps accounts for the findings of two informal surveys in 1997 and 1998, respectively, by the Bank of England's Traded Market Team. The surveys were conducted to assess current market practice in stress testing and found widely different approaches to stress testing across banks. Finally, in its June 1999 report on the performance of the market risk amendment, the Models Task Force of the Basel Committee on Banking Supervision (1999b) found that, for a significant number of surveyed banks, stress tests had not been good predictors of the results of the extreme market moves experienced in the autumn of 1998. In the interviews conducted by the Working Group, risk managers did not mention quantitative predictive accuracy as an important criterion for the usefulness of stress tests.

²² See Gibson (1997) for a discussion of the hurdles faced by financial trading firms to overcome organisational, as well as technological barriers, in the organisation of firm-wide risk information systems.

Stress tests may be carried out infrequently because of their high computational cost. Systems incompatibilities across business units make frequent stress testing costly for some firms, reflecting the limited role that stress testing had played in influencing the firm's prior investments in information technology. Many interviewed firms noted that their stress testing relies on full revaluation of option positions, while VaR uses approximations.²³ Risk managers noted how the heavy computational burden imposed by full revaluation of portfolios of complex options limits the frequency of their stress tests.

None of the interviewed firms systematically integrate market and credit risk in stress testing.²⁴ The integration is limited to considering credit-related changes in the pricing of traded instruments. Knock-on effects from market risk factors to loan default and recovery rates are not considered. Although most firms interviewed commented positively on the desirability of integrating credit and market risk assessments, there was substantial variation across firms as to whether they assigned a high priority to developing techniques in this area.

²³ Because there is no simple formula for the value of a complex option, such an option has to be revalued with computationally intensive techniques, such as Monte Carlo techniques.

²⁴ In a recent report, the Basel Committee on Banking Supervision (1999a) concluded that little progress has been made in developing techniques to implement credit risk stress tests. By contrast, in the Counterparty Risk Management Policy Group (1999) report, a group of 12 internationally active commercial and investment banks commented that some firms are conducting targeted stress tests of their counterparty credit exposures. These efforts are based on conjectures concerning the statistical relationship between market risk factors and the quality of a firm's counterparty credit risk exposure.

Chapter 2

Aggregate stress tests

This chapter discusses the potential uses of aggregate stress tests and the methodologies that could be used to construct them. Compared with Chapter 1, this chapter's conclusions should be viewed as less definitive, because they are not directly based on discussions with risk managers. At the time of those discussions, the Working Group felt it would have been premature to ask firms how they could integrate aggregate stress test information with other information (public and firm-specific) now used to support their risk management efforts.

The Working Group feels it has made some progress on coming to a common understanding as to how market liquidity in times of stress is determined and how risk managers could potentially use aggregate stress test information to manage some aspects of that risk better. The first part of this chapter discusses the potential usefulness of aggregate stress tests for risk managers, central banks and financial regulators, drawing on a conceptual framework that is discussed further in Annex 2.²⁵ The second part gives a sense of the technical issues involved and the choices to be made in actually aggregating firms' stress tests. The phrase "potential usefulness" reflects the Working Group's view that the usefulness of aggregate stress tests, as well as the burdens they entail, can only be discussed meaningfully when one specifies the details of what an aggregate stress test would consist of (who would report, what information would they report, etc.).

What is an aggregate stress test?

An aggregate stress test is a measure of the risk exposure of a group of reporting firms to a specified stress scenario. As part of an aggregate stress test exercise, each reporting firm would provide information on its own exposure under the scenario. The responses would then be aggregated by a central co-ordinator. The aggregation would produce a single number capturing the combined exposure of all reporting firms. This single number could be an index (analogous to an index of macroeconomic indicators), a total net gain or loss figure, or the net gain or loss as a fraction of reporting firms' equity capital. In addition to the combined exposure, information on the distribution of exposures across firms could also be captured. Determining whether a given aggregate stress test exposure is "high" or "low" would be easier once some history of the stress exposure had been accumulated. For this reason, an aggregate stress test would ideally be run repeatedly over a period of time. Through such repetition, an historical record would be created that linked actual portfolio decisions with their implications in terms of outcomes if a defined scenario had occurred.

How might risk managers use aggregate stress tests?

An important area of risk management in which firms could use aggregate stress test information might be to assess market liquidity risk under stress. This risk is treated here as distinct from market liquidity risk in normal times.²⁶

²⁵ Annex 2 is based on a note contributed by Michael Gibson of the staff of the Board of Governors of the Federal Reserve System.

²⁶ Market liquidity risk in normal times can be thought of as the risk that a given transaction will incur an additional cost because of day-to-day fluctuations in liquidity in the relevant market. Market liquidity risk under stress can be thought of as the risk that liquidity will decline sharply and unexpectedly, and for a significant period of time, because of turbulent market conditions. Measuring the former typically involves studying fluctuations in bid-ask spreads and the market impact of large trades using day-to-day market data. The report of the CGFS study group on market liquidity (BIS

An assessment of market liquidity risk under stress enters into a financial firm's ex ante judgement of the riskiness of a trading strategy. The firm makes such an assessment by combining its knowledge of the strategy's liquidity needs with an assumption about possible market liquidity conditions under stress. The latter will be based on whatever information the firm has at hand. Holding other factors – such as expected returns, market risk and credit risk – constant, a firm's willingness to engage in a particular trading strategy is likely to be lower when market liquidity risk is higher. If a firm becomes aware that it has been underestimating the market liquidity risk of a particular strategy, it could scale back its use of that strategy, for example through liquidations of positions in related securities.

There are many reasons why market liquidity can dry up suddenly and unexpectedly. The possible mechanisms include a re-evaluation of the credit risk of an important class of counterparties, the mutually reinforcing effects of broad shocks to credit quality and market prices, co-ordination problems related to the need for participants to have consistent behavioural expectations, and doubts about the integrity of settlement systems. Another way that liquidity might dry up would be if new information or a shock to prices leads to a severe imbalance between buyers and sellers, i.e. a "one-way" market. This is thought to have been a key factor in certain liquidity-shortage episodes in the past, such as the October 1987 equity market reversal. In what follows, it will be argued that aggregate stress tests would be useful to the extent that they help us anticipate the emergence of a one-way market of this kind. Even if they can help to anticipate such situations, however, it must be emphasised that they might be of little use in anticipating other kinds of liquidity shocks. For instance, the turmoil that affected many markets in the autumn of 1998 appears to have resulted primarily from counterparty credit concerns and the need to reduce excessively leveraged positions.²⁷

Risk managers use a variety of sources of information to anticipate the possible emergence of a one-way market. These include public sources, proprietary information sold by outside data providers, and private information gathered by observing customers' order flow and by discussions with the firm's market contacts, customers and counterparties.²⁸ At present, market participants do not appear to be satisfied that they have enough information from these various sources to do a good job assessing market liquidity risk, both in normal times and under stress.²⁹ To be worth the effort involved, information from aggregate stress tests would need to help them to do this job better.

To understand how aggregate stress test information might be useful in assessing the likelihood of the emergence of a one-way market, and what limitations it might have in this regard, we need to know what factors could cause the kinds of market behaviour that would produce a one-way market. These factors are notoriously difficult to model in a satisfactory way and to measure empirically. Annex 2 offers a set of considerations that the working group found useful in its efforts to understand these phenomena.

The analysis in Annex 2 suggests how aggregate stress test information could be used to help assess the risk of a one-way market. The sum of reported stress test exposures would indicate the intensity of trading that might be triggered by a market shock. The usefulness of the results would then depend on the composition of the reporting population relative to the market as a whole. If the reporting firms are considered more likely to adjust their positions following a market shock than most non-reporters, a larger reported stress test exposure could indicate a greater risk of a one-way market. This might be

(1999a)) is primarily concerned with the measurement and dynamics of market liquidity in normal times, though two papers prepared for the study group – Muranaga and Shimizu (1999) and Miyanoya (1999) – analyse market liquidity under stress.

²⁷ See BIS (1999b) for an account of this episode.

²⁸ CGFS (1998) cited an example: a risk manager at an international bank learned from her bank's traders that a certain group of market participants had accumulated a large position in Brazilian debt, raising the likelihood that the market for Brazilian debt could become illiquid in the wake of a stressful shock.

²⁹ For example, the report of the Counterparty Risk Management Policy Group (1999) urges that firms pay more attention to market liquidity risk. The report does not discuss what information firms could use to do that.

the case, for example, if non-reporters tend to be households and pension funds who adjust their positions slowly, while reporters include dealing firms who need to rebalance their positions rapidly in volatile markets in a way that makes them “positive feedback traders”. If the reporters are pure intermediaries who attempt to maintain a balanced position at all times, then their rebalancing activities are likely to be less. Yet even if reporters maintain neutral positions with respect to the directions of market prices, they may be exposed to changes in volatility through their options-writing activities.

Another situation in which a one-way market may be especially likely is when a small number of market participants are responsible for a large fraction of the outstanding exposures in the market. Knowing the distribution of exposures across firms or groups of firms as well as the aggregate exposure could therefore also be a useful outcome of an aggregate stress test, and would help firms manage their market liquidity risk. If exposures are concentrated among a few agents, a large shock to demand might be more likely to trigger a drying up of market liquidity. For example, if the holder of a large long position in a market chooses to sell, it might be easier for the market to absorb this smoothly if there are many potential buyers than if there are only a few. It might also be important to know whether large aggregate losses are likely to be borne disproportionately by “strong” or “weak” firms (measured, for example, by their Tier I capital).

How might central banks and financial regulators use aggregate stress tests?

In addition to their use by risk managers, aggregate stress tests could potentially help central banks monitor risk-taking and risk-intermediation in financial markets. Central banks already collect some information that is relevant for market monitoring: for example, the BIS consolidated international banking statistics, the BIS semi-annual OTC derivatives statistics, and the triennial central bank survey of foreign exchange and derivatives market activity. However, these are historical data; they do not give “forward-looking” information on risk-taking, nor do they reveal anything about the risks associated with rare, extreme events.³⁰ Aggregate stress tests could potentially help fill this gap.

A high level of aggregate exposure to a stress scenario could indicate a potential systemic vulnerability.³¹ Currently, central banks assess systemic vulnerabilities by compiling and assessing statistical reports as well as by reviewing anecdotal evidence obtained from market contacts, including information generated from their own market operations. Aggregate stress tests could complement these information sources. It is natural to look to stress tests for information on risks of extreme outcomes, since (as discussed in Chapter 1) the current state-of-the-art among financial firms is to use stress tests as a way to understand their exposure to such risks.

Methodology of aggregate stress tests

This section describes possible ways that stress tests could be aggregated across firms. They reflect the diverse techniques in use at private sector firms, as discussed in Chapter 1. In all cases, there would be a link between the number of aggregate stress tests conducted and the range of risks that could be covered. The various approaches differ in terms of how easy they are to interpret, how much information could be produced for a given amount of reporting burden, and what share of the burden would fall on reporting firms relative to the central co-ordinator.

³⁰ These limitations have been pointed out in two previous documents addressing the information needs that are necessary for central banks to perform their macroprudential role, BIS (1995) and BIS (1996).

³¹ As discussed above, judging whether the stress test exposures are “high” or “low” would be easier when a historical time series of exposures is available for comparison.

Historical and hypothetical scenarios

Aggregate stress tests could be based on historical or hypothetical scenarios. Chapter 1 discussed the important issues that firms face when setting stress scenarios. The same issues would apply to setting aggregate stress tests.

One way to identify useful scenarios would be to survey firms on the historical and hypothetical stress scenarios they use and to focus on those in use at a large number of firms. Since firms orient their stress testing towards the risks that are material in their own portfolio, this approach would identify scenarios that are relevant to a significant number of reporting firms' portfolios. If firms and central banks participate together in the scenario-setting process, the resulting scenarios may be acknowledged by all parties to be useful. If aggregate stress tests are conducted periodically, the scenarios would need to be updated to ensure that they are still relevant to the risks being taken in the market.

Any participation of a central bank in the scenario-setting process, however, poses the risk of sending the wrong signals to the private sector. Market participants could misinterpret the central banks' interest in a particular historical or hypothetical scenario as containing information on central banks' views of what shocks are likely to occur or how central banks might respond to such shocks. A process that assigned private sector market participants responsibility for setting scenarios, hypothetical scenarios in particular, could be a way of avoiding this pitfall.

How stress tests could be aggregated

The Working Group discussed two ways an aggregate stress test could be computed. The two methods differ in what information would be collected from reporting firms, how costly it would be for reporting firms to assemble the information, and how the information would be used to compute an aggregate stress test result.

Revalue the portfolio under an exhaustive set of shocks

A conceptually straightforward approach used by the majority of firms is to define a stress scenario as an exhaustive set of shocks to all the market risk factors (prices, rates, volatilities) that affect the value of the firm's portfolio and to revalue the portfolio under that scenario. To apply this approach to an aggregate stress test, a central co-ordinator would determine the exhaustive set of shocks that applies to each stress scenario and ask each reporting firm to revalue its portfolio assuming that those shocks occur. The reporting firms' responses would be directly comparable across firms, since each firm's scenario would use exactly the same set of shocks. The responses could be summed or their distribution could be reported directly.

This approach can be summarised as "high burden, high accuracy." The accuracy is high because results are directly comparable across firms. The burden falls on reporting firms to reprogram their risk management systems to run the aggregate stress tests in addition to the stress tests they run for their own purposes. Two factors make this burden potentially large. First, an exhaustive set of shocks might require setting shocks for as many as 30,000 risk factors (to use the figure cited in one interview), and each firm will have to map the scenario definition into a form that its internal risk systems can understand. Second, as discussed in Chapter 1, revaluing the entire portfolio for a stress test is a burdensome exercise for firms, because of systems incompatibilities across business units within a firm and because complex positions, particularly those involving options, require a lot of computational power to fully reprice. Simply collecting and aggregating data on the size of reporting firms' positions might be a less burdensome alternative to full revaluation, but would come at a cost in terms of accuracy to the extent that non-linear (e.g. options-based) positions are present.

Collect and aggregate stress test information that firms already produce

Rather than force reporting firms to conform to a centrally defined exhaustive set of shocks, an alternative would be to collect results of stress tests that firms already run.³² This approach relies on the fact, documented in Chapter 1, that many firms reported running stress scenarios based on the same set of broad market moves (e.g. the 1987 equity market crash, the 1992 ERM crisis). Although the stress tests of different firms will not be exactly comparable, because the underlying set of many thousands of market risk factor shocks will surely be different at every firm, they may be similar enough that the information can be aggregated in some way. A central co-ordinator would be responsible for coming up with a meaningful method of aggregating stress tests that measure similar risk factor exposures but are not identical. (This problem is similar to the problem of constructing an index of leading economic indicators from disparate variables like stock prices, money supply, business confidence, etc.)

This approach can be summarised as “low burden, low accuracy.” The reporting burden is low because firms would report the stress tests they are already running for their own purposes. The accuracy is low because different firms’ stress tests will not be directly comparable. In this approach, the central co-ordinator has the burden of coming up with a way to aggregate results from the participants’ disparate stress tests.

Disclosure of aggregate stress test results

Another issue that would have to be addressed would be the degree and nature of public disclosure of aggregate stress test results. One danger is misinterpretation: it is an open question whether the point could be clearly made to the market that scenario results are descriptions of risk exposures at a point in time, not predictions of likely outcomes. Another concern is the possible revelation of proprietary firm-level information, which is more likely if the exercise is limited to a small number of institutions or if information on the distribution of exposures is disclosed. If results are only revealed to participants, this could give them an unfair informational advantage vis-à-vis non-participants. Such issues would need to be carefully considered and resolved in the design of any aggregate stress exercise.

The reporting population

The information that could be gained from an aggregate stress test depends crucially on the number and characteristics of the firms that would participate in the exercise. Four hypothetical reporting populations are identified in Table 2, along with options for the information that might be requested in each case. For the first three populations, the risk information to be reported is simply the firm’s exposure to a stress scenario. The fourth specification considers asking for more information than simple exposure. The specifications range from relatively high usefulness/high reporting burden (option 1) to relatively low usefulness/low burden (option 4). As noted above, these assessments of the likely usefulness and burden are based on the Working Group’s own discussions, not on discussions with risk managers. This list is idealised and is not meant to be exhaustive.

³² Firms could report all their internally generated stress tests, or just a subset (such as the “ten worst case scenarios”).

Table 2
Potential specifications for aggregate stress tests

	Who would report	What information	Usefulness/burden
1	Actual and potential frequent participants in financial markets	Stress test exposure and type of trading strategy followed	High
2	Frequent participants in financial markets	Stress test exposure and type of trading strategy followed	↓
2a		Stress test exposure	
3	Regulated frequent participants in financial markets	Stress test exposure and type of trading strategy followed	
3a		Stress test exposure	
4	Dealers	Stress test exposure	Low
4a		Their own stress test exposure, and that of their counterparties (covering cash and derivatives markets)	Unknown, because not feasible with current risk systems

1. **Full universe.** This represents the ideal set of reporters, comprising any current or potential future frequent participant in financial markets. They would each report their exposure to the stress scenario, along with the type of trading strategy they follow. (In terms of the conceptual framework developed in Annex 2, this would allow one to completely fill in the rightmost column of Table A-1.) Market participants could use this information to judge the risk that market liquidity will collapse in that type of a stress event. To the extent that behavioural patterns of different classes of market participants are stable and well understood, risk managers would find the information produced by this type of an aggregate stress test to be useful, because the information would cover the entire universe of participants in financial markets. Of course, this specification of an aggregate stress test is a totally implausible extreme, but it is presented here to help in elucidating the issues involved.

2. **Actual frequent participants.** We think of actual frequent participants in financial markets as comprising dealer banks, securities firms, insurance companies, mutual funds, hedge funds, asset managers, and pension funds. Compared with option 1, this group excludes the “potential” traders who currently are not frequent participants in financial markets and who could never, in practice, be identified. Actual frequent participants also have the greatest need to manage market liquidity risk, since they rely on market liquidity as part of their business strategies. This may make them more willing participants in an aggregate stress test exercise.

Under this option, these firms would report their exposure to the stress scenario along with a characterisation of their trading strategy. That would allow one to construct an analogue to Table A-1 covering only actual frequent participants. Some fraction of the total exposure would be captured.

Under a variant of this option (option 2a), these firms would report their stress exposure but not their trading strategy. It may be difficult to categorise agents by trading strategy, and disclosing the breakdown of stress exposures by trading strategy may be problematic if it requires the disclosure of information that firms consider to be highly proprietary because of its commercial value. In this option, the total stress exposure of all reporting firms, along with the distribution of exposures across firms, could be aggregated and disclosed. This information would be less useful than option 2, because the identification of stress exposures with a trading strategy is omitted.

3. **Regulated, frequent participants.** Compared with option 2, this group excludes non-regulated financial firms, such as hedge funds and some asset managers. Regulated firms may be more likely to participate in an aggregate stress test exercise.

Again, one can think of two possible variations (options 3 and 3a), which reflect the same choice of whether or not to report the trading strategy followed. One can conclude that the information produced would be less useful than under option 2, because the coverage of likely rebalancing and rehedging in stress is smaller.

4. **Dealer firms.** Compared with option 3, this group includes only the very large banks and securities firms, totalling perhaps a few dozen worldwide, that are the “core” intermediaries in global financial markets. Having information on only dealers’ stress exposures would be less useful than option 3, for two reasons. First, other regulated firms, who would certainly do some rebalancing and rehedging in the wake of a stressful shock, are not included. Second, dealer firms, if they truly just intermediate market risk for their customers, may not be very exposed to market risk stress events (except as it affects their credit exposures with their customers). An exception would be the dynamic hedging of written options, which is only one of the five factors discussed in Annex 2.

Another option (option 4a) would be to take advantage of dealers’ position as intermediaries and ask them for information on their customers’ aggregate market risk exposures to a stress scenario (both cash and derivatives market exposures). This would be an indirect way of obtaining the stress exposures of non-reporting firms, which would be obtained directly in options 1–3 above.

Consider Table 3, which is a modified version of Table A-1. The table breaks down the rightmost column of Table A-1, the stress test exposure of each type of trader, into those arising from derivatives positions with dealers, from cash positions financed with dealers, and from cash positions funded internally. Dealers have information on the boxes marked with an “X”, comprising their own stress exposures and those of their customers.

Option 4a would produce more complete information on stress exposures than option 4, which only captures dealers’ own exposures. The burden on dealers to produce this kind of information would be prohibitively high using currently available risk management practices and systems, because it would require combined stress testing of market and credit risk exposures. While dealer firms have indicated they want to move in the direction of combining market and credit exposure measurement, firms we interviewed gave the impression that achieving this would be a medium-term goal, not a short-term priority.³³ Both the methodology and the information systems infrastructure for jointly evaluating market and credit risk exposures need a great deal of development before joint stress testing can become standard market practice. However, firms are moving in this direction, so this option will become more feasible over time. Furthermore, even if dealer firms produced combined market and credit risk stress tests that captured their counterparties’ stress exposures, there are additional questions about how those stress tests could be meaningfully aggregated. The Working Group did not try to answer these questions; we flag them as potentially interesting for further exploration.

³³ The Counterparty Risk Management Policy Group (1999) urges that “when measuring exposure to stress events, [financial intermediaries] should estimate both market and credit risks.”

Table 3

**Financial institutions grouped according to trading strategy,
with each firm's exposure to the stress event according to the type of position**

(An "X" indicates information that could potentially be gathered from dealers alone.)

Strategy	Rebalancing/ rehedging demand in stress	Exposure to the market stress event arising from:		
		derivatives contracts with dealers	cash positions financed with dealers	cash positions funded internally
Asset manager strategy #1	Cannot be observed directly	X	X	
Asset manager strategy #2		X	X	
–		X	X	
Relative value trader strategy #1		X	X	
Relative value trader strategy #2		X	X	
–		X	X	
Dealers		X	X	X
Market-wide total		X	X	

Limitations to the potential usefulness of aggregate stress tests

Besides the question of the reporting population discussed above, a key issue for aggregate stress tests is the relevance of the chosen historical or hypothetical scenarios for capturing the large risks taken in the market at any point in time. An individual firm chooses its stress scenarios to cover the important risks taken at that firm, using knowledge of the firm's portfolio. The choice of scenarios is reviewed periodically and, if the portfolio has changed significantly, the set of scenarios may be changed as well. For aggregate stress tests, it would be harder to gauge when the set of scenarios has become so "stale" that changes are needed.

Aggregate stress test results would not be meaningful if firms' valuation models did not produce comparable results. If the results were not comparable, aggregating stress test results across firms would incorporate valuation errors of an uncertain magnitude, rendering the aggregated results uninterpretable. The comparability of firms' valuation models will depend, of course, on the composition of the portfolio being valued. Complex derivatives and new products are more likely to be areas where firms differ in their valuations. The relevance of this argument is an empirical question, that could be answered by an examination of firms' valuation models.

Another argument against pursuing aggregate stress tests is a concern that a limited number of stress tests would not adequately capture the range of risks present in the markets and could give a misleading picture of the true nature of risk-taking by participating firms. Because stress tests are computationally burdensome, involving time-consuming calculations to value non-linear positions such as options, the number of aggregate stress tests that could be done is small. As a result, it would not be possible to stress every market risk factor that has the potential to affect dealer firms' portfolios in a significant way.

Conducting aggregate stress tests might also in and of itself lead to an emphasis on managing certain kinds of risk to the exclusion of others. For example, while it is argued above that aggregate stress tests might allow firms to anticipate market liquidity risk arising from the patterns of trading behaviour present in a market, it is also noted that another reason why market liquidity could vanish relates to counterparty credit risk. Aggregate stress tests might distract firms from devoting resources

to gathering the information that would enable them to manage this risk better, such as information on counterparties' market risk exposures.

The limits to the usefulness of aggregate stress tests for central banks and financial regulators should also be recognised. While aggregate stress tests may provide useful information for market monitoring, they will not be useful for such tasks as setting capital charges against market risk. Those tasks require risk information that is tailored to an individual firm's portfolio. Some members of the Working Group voiced a concern that specifying scenarios for aggregate stress tests could make risk managers and supervisors less inclined to develop firm-specific scenarios that would be more relevant for an individual bank. At present, aggregate stress tests would not capture the interaction of market and credit risk exposures. This interaction would undoubtedly be important for assessing important potential systemic vulnerabilities such as funding risk. Finally, although aggregate stress tests could provide information on the level of risk exposures taken on by core financial intermediaries, they can give only an imperfect indication about the likely response of those intermediaries to extreme events.³⁴

Finally, there is the issue of the time horizon at which the results of an aggregate stress test would be useful to the private and public sectors. In a time of crisis, frequent and timely – perhaps even daily – information on exposures might be called for, but daily aggregate stress tests would be infeasible even with the most advanced information systems. For this reason, an aggregate stress test program would have to be evaluated on its ability to anticipate potential systemic pressures at a monthly or quarterly frequency, rather than on its usefulness to those charged with managing a crisis that is already underway.

³⁴ This topic is discussed further in Annex 3 to this report, which is based on a note contributed by Tokiko Shimizu of the Bank of Japan.

Chapter 3

Discussion and recommendations

At this time, the Working Group does not believe it has enough information to judge whether aggregate stress tests would be useful to firms in their management of market liquidity risk under stress. Given the specific shortcomings outlined at the end of the previous chapter, as well as the broader uncertainties about the usefulness of aggregate stress tests, it is the view of the Working Group that it does not currently have enough information to make an informed judgement as to whether aggregate stress tests would provide value added relative to current methods of monitoring financial markets.

However, the Working Group believes that now is a fruitful time for central banks and other financial regulators to think more broadly about what data are most useful for monitoring the functioning of financial markets. In response to innovations in the science of risk management and in information technology, firms have developed information systems for collecting firmwide risk data. Perhaps at some future date, aggregate stress tests will be found to be a way to tap into this new data source that would provide forward-looking information about risk exposures in aggregate that would be of use to financial firms, central banks, and other financial regulators.

This process would correspond to a longstanding pattern, in which innovations in management science lead firms to assemble new data for their own business needs and, subsequently, public policy makers exploit the new data for their own purposes. An example of such an innovation is capital budgeting, the process of systematically managing a firm's investment spending on a multi-year basis. Capital budgeting was developed beginning around 1910 as a way to more effectively manage the newly emerging large-scale manufacturing corporations. By the 1940s, capital budgeting was "best practice" management in all large corporations in the United States. Public policy makers concerned with business cycle fluctuations, including the Federal Reserve System, recognised that firms' capital budgets, if aggregated, would be an accurate and timely predictor of future investment spending.

The interviews conducted by the working group established that financial firms face a trade-off between the computational burden of a stress test and its accuracy, and are still choosing to increase the resources they devote to stress testing. A *census of scenarios in use at firms* would be a relatively inexpensive way of obtaining information regarding which risk factors are the focus of firms' risk management scrutiny, given the fact that they can only afford to focus on a limited number of scenarios at a given time.

Such a one-off survey would be worthwhile for at least four reasons. First, by helping central banks to better understand the role that stress testing plays in risk management, the census would better equip central banks to understand risk-management procedures and thereby deal with future bouts of financial instability. Second, since individual firms tailor their stress tests to the risks that are important for their portfolio, knowing the risks that firms are focusing on could help central banks in their efforts to identify and monitor potential systemic vulnerabilities. Third, the degree of similarity of stress tests among dealer firms will convey information on the heterogeneity of risk-taking at a point in time. Fourth, given the heightened importance placed on stress testing by risk managers, central banks, and financial regulators in the wake of recent episodes of turmoil in financial markets, a public discussion of the census of stress scenarios would increase the transparency of the risk-management process. As noted by the Counterparty Risk Management Policy Group report, at

present, “firms are only beginning to share information about how they set stress tests, and no consensus has yet emerged.”³⁵

The burden on a firm of participating in a census of stress scenarios would be low. A firm would need to make a one-time effort to come to an understanding with the census-takers on how to map the stress tests it uses internally into a common reporting framework. Efforts should be made to preserve this low reporting burden and to maximise the benefits of the census to participating firms. Comments and suggestions should be solicited from participating risk managers at an early stage. Efforts should also be made to coordinate with other regulatory and central bank projects requesting information from the same group of risk managers. The results of the census should be disseminated in a timely manner. Concerns about confidentiality of proprietary information should be addressed.

³⁵ Counterparty Risk Management Policy Group (1999), p. A-4.

Annex 1

Bibliography

- Bank for International Settlements (1995): *Issues of Measurement Related to Market Size and Macprudential Risks in Derivatives Markets*, Basel (February).
- Bank for International Settlements (1996): *Proposals for Improving Global Derivatives Market Statistics*, Basel (July) (can be downloaded from <http://www.bis.org>).
- Bank for International Settlements (1999a): *Market Liquidity: Research Findings and Selected Policy Implications*, Basel (May) (can be downloaded from <http://www.bis.org>).
- Bank for International Settlements (1999b): *A Review of Financial Market Events in Autumn 1998*, Basel (October) (can be downloaded from <http://www.bis.org>).
- Basel Committee on Banking Supervision (1999a): *Credit Risk Modelling: Current Practices and Applications*, Basel (April) (can be downloaded from <http://www.bis.org>).
- Basel Committee on Banking Supervision (1999b): *Performance of Models-Based Capital Charges for Market Risk: 1 July-31 December 1998*, Basel (September) (can be downloaded from <http://www.bis.org>).
- Committee on the Global Financial System (1998): *On the Use of Information and Risk Management by International Banks*, Basel (October).
- Counterparty Risk Management Policy Group (1999): *Improving Counterparty Risk Management Practices* (June) (can be downloaded from <http://www.crmgroup.org>).
- Danielsson, Jon and Casper G de Vries (1997): "Value-at-Risk and extreme returns", working paper, London School of Economics and University of Iceland (September) (can be downloaded from <http://www.hag.hi.is/~jond/research/>).
- Gibson, Michael S (1997): "Information systems for risk management", International Finance Discussion Paper No. 585, Federal Reserve Board (July) (can be downloaded from <http://www.federalreserve.gov/pubs/ifdp/>).
- Institute of International Finance (1999): *Report of the Task Force on Risk Assessment*, Washington, DC.
- Longin, François M (1999): "From Value at Risk to stress testing: the extreme value approach", Discussion Paper No. 2161, Centre for Economic Policy Research (May).
- Miyanoya, A (1999): "Price Discovery Functions in Japan's corporate bond market: An Event Study of the Recent Fall 1997 Financial Crisis", in BIS (1999a).
- McNeil, Alexander J (1999): "Extreme value theory for risk managers", working paper, Department of Mathematics, ETH Zentrum (May) (can be downloaded from http://www.math.ethz.ch/~mcneil/pub_list.html).
- Muranaga, Jun and Tokiko Shimizu (1999): "Expectations and market microstructure when liquidity is lost", Discussion Paper 99-E-15, Institute for Monetary and Economic Studies, Bank of Japan (May) (can be downloaded from <http://www.imes.boj.or.jp>).
- Schachter, Barry (1998): "The value of stress testing in market risk management", *Derivatives Risk Management Service*, Ed T Haight (March) (can be downloaded from <http://pw2.netcom.com/~bschacht/var/drms.html>).
- Shleifer, Andrei and Robert W Vishny (1997): "The limits of arbitrage", *Journal of Finance* 52:1, pp 35–55 (March).

Tuckman, Bruce and Jean-Luc Vila (1992): “Arbitrage with holding costs: A utility –based approach”, *Journal of Finance* 47, pp 1283-1302.

Van den Gorbergh, Rob W J and Peter J G Vlaar (1999): “Value-at-Risk analysis of stock returns: historical simulation, variance techniques or tail index estimation?” Research Memorandum WO&E 579, De Nederlandsche Bank (can be downloaded from <http://www.dnb.nl>).

Annex 2

Market liquidity, one-way markets and stressful shocks

As noted in Chapter 2, aggregate stress tests could be useful to the extent that they help market participants anticipate the evaporation of liquidity because of the emergence of a “one-way market”. This annex offers a conceptual framework, drawing on the theoretical literature on market microstructure, that may be of use in understanding why one-way markets occur, the impact they may have on market liquidity, and why an aggregate stress test might be of use in anticipating them.

Trading behaviour and one-way markets

To help us understand how a one-way market might come about, we can separate market participants into two groups according to how they respond to a stressful shock. We label the two groups “negative feedback traders” and “positive feedback traders.” Negative feedback traders buy when the price falls, while positive feedback traders sell when the price falls. These labels are purely for convenience; a given individual or entity could be a negative feedback trader at one time and a positive feedback at another, or simultaneously a negative feedback trader in one market and a positive feedback trader in a different market. Anyone who sells after a price decline, for whatever reason, is thus a positive feedback trader according to this categorisation.³⁶ Market liquidity is then a function of the aggregation of negative and positive feedback trading.

In a liquid market, negative feedback traders outweigh positive feedback traders, so that market price fluctuations are damped and liquidity does not dry up in the wake of a stressful shock.³⁷ Liquidity is at risk of drying up when positive feedback traders become more important, relative to negative feedback traders. A collapse of market liquidity occurs when the “market share” of positive feedback traders increases enough to outweigh that of negative feedback traders. The market becomes one-sided and market liquidity vanishes.

A risk manager’s assessment of market liquidity risk under stress is undertaken before the hypothetical “stressful shock” has occurred. Whether a market is considered to be “vulnerable” to the kind of phenomenon being discussed depends on whether the risk manager forecasts that positive feedback trading would be important following a hypothetical large decline in the asset price. To measure this kind of market liquidity risk under a hypothetical stress event, a risk manager would thus need information on the relative share of positive feedback traders that would exist after a market stress event, and the amount of trading they are likely to do.

What types of traders are in the market?

For such a mechanical model of market liquidity to be useful in practice, it must be combined with a classification of the various types of agents in the market. We focus on agents’ attributes that matter for market liquidity, ignoring other attributes (such as whether they trade on their own or someone else’s behalf). The emphasis will be on entities that trade frequently, rather than relatively infrequent market participants such as households and commercial firms. Frequent traders are most likely to

³⁶ The use of the term “positive feedback trader” here covers any agent who sells following a decline in the price, for any reason. It is not restricted to those traders who follow mechanical trading strategies, such as dynamic hedging of written options.

³⁷ If the initial shock reflects a decline in the fundamental value of the asset, agents with full information on the fundamental value may engage in negative feedback trading if the asset price overshoots its fundamental value. These fully informed agents can play a stabilising role, as described in standard models of risk arbitrage such as Tuckman and Vila (1992).

stressful shock is more likely, the greater is leverage of those invested in the market.⁴⁰

Limited arbitrage

Some asset managers and relative value traders normally engage in negative feedback trading, i.e. investing in assets whose price has fallen below the trader's perception of its future value. According to a model of "performance-based arbitrage," they may be less able to do so in the wake of a large shock.⁴¹ In the model, negative feedback trading is conducted by a small number of agents using funds provided by investors. Because investors cannot distinguish low returns due to bad luck from low returns due to bad investing strategies, agents with poor performance will face withdrawals by investors. As a result, following a negative shock to asset returns, the funds available to the negative feedback traders will decline, so negative feedback trading falls.

Sharpe-ratio-based trading

Some asset managers and relative value traders choose among alternative investment strategies by comparing Sharpe ratios.⁴² Market participants understand that, empirically, large negative moves in some markets tend to be followed by high volatility in those markets. All other things equal, this empirical relationship tends to produce a mechanical (but real) decline in an asset's Sharpe ratio following a decline in the price of the asset. Traders who base their asset allocation on Sharpe ratios will tend to rebalance away from stressed markets, leading them to be positive feedback traders.

Dynamic hedging

The exact trading strategy followed by a dynamic hedger depends on the portfolio of options being hedged. A typical dealer is a net seller of options and will be a positive feedback trader. The amount of underlying assets sold as part of a dynamic hedging strategy following a negative shock grows with the size of the shock, so positive feedback trading of such dealers rises in stress.

When a risk manager makes an *ex ante* assessment of market liquidity risk under stress, she must attempt to gauge the amount of rebalancing and rehedging that each type of trading strategy listed in Table A-1 would call for. Ideally, she would use information such as that contained in the third column of Table A-1, that is, the trading demand resulting from each type of strategy. The market-wide total would allow the risk manager to see whether market liquidity risk under this stress scenario is high or low. If we adopt the convention that a positive n_i indicates an increase in exposure and a negative n_i indicates a reduction in exposure, market liquidity risk under stress would be inversely related to the market-wide total change in exposure, other factors held constant.

An accurate assessment of each strategy's rebalancing and rehedging demand would call for different information for each of the five factors discussed above.⁴³

⁴⁰ Of course, the existence of leverage does not always have a negative effect on financial market functioning. For example, financial market stability in stress would be enhanced if negative feedback traders were able to use increased leverage to increase their trading activity.

⁴¹ The model is due to Shleifer and Vishny (1997).

⁴² The Sharpe ratio measures a strategy's reward relative to its risk. It is the ratio of expected return on the strategy in excess of the risk-free interest rate to the standard deviation of the strategy's returns. A Sharpe ratio thus reflects a trader's estimate of a strategy's risk, which incorporates an estimate of the volatility of relevant asset prices.

⁴³ Some of the same information could be used to assess contagion risk. For example, stop-loss rules triggered in one market could lead traders to sell off their risky positions in other markets. Note that, in addition to the information used to assess market liquidity risk, information on common positions across markets is needed to assess contagion risk.

Stop-loss rules	Use information on how far away from his or her stop-loss limit each trader in the market is to begin with, along with the size of each stop-loss trader's exposure in the market. The latter also affects the amount of selling that would occur after the stop-loss limit is hit.
Leverage and funding risk	To know how likely financial-distress-related liquidation is, use information on the leverage (or more generally the risk profile, including funding risk) of asset managers and relative value traders with exposures in this market, along with information on their exposures to the market.
Limited arbitrage	This factor will be more significant, the larger share of the market is held by arbitrageurs before the shock. Use information on the size of arbitrageur's exposures.
Sharpe-ratio-based trading	Use information on the size of such traders' exposures to know the potential significance of their rebalancing away from the market.
Dynamic hedging	Use information on the size of dealers' short option positions in the market to know how large their dynamic hedging needs would be.

Of course, for all five factors such precise information on all of the relevant players in a market is not available to a risk manager. However, the relevant information will be related to the size of agents' exposures under stress. In other words, the quantity of agents' rebalancing and rehedging in stress is likely to be positively related to the size of their exposures. This suggests that, while it would not be the ideal set of information presented in the third column of Table A-1, information on agents' exposures to market stress as shown in the fourth column of Table A-1 could be useful.⁴⁴ If tracked over time, such information could help a risk manager better anticipate the emergence of patterns of trading behaviour that would cause a one-way market to develop.

⁴⁴ Using math terminology, we could say that $n_i = f_i(e_i)$.

Annex 3

Market dynamics in times of stress

The importance of considering market participants' behaviour

What one would eventually want to produce through macro stress testing is information that helps predict market behaviour in a stress situation. By adding dynamic information, such as market participants' behaviour in response to a market shock, to static information such as the expected loss amount under a stress scenario, it may be possible to obtain a better prediction of overall market behaviour. Such testing which takes into account the reaction of market participants will be referred to as "dynamic macro stress testing."

What would be the advantage to market participants of participating in dynamic macro stress testing? When market participants are confronted with a stress situation they respond in three steps: First, they recognise the magnitude of the loss which they passively incur. Second, they come up with a plan to handle the loss. And, third, they consider whether they can in fact smoothly re-balance or re-hedge their portfolio based on this plan. While aggregate stress testing as described in Chapter 2 of this report aims at collecting information concerning the first step, dynamic macro stress testing collects information for the second step so as to provide market participants with information necessary for the third step.

Market participants use their best guess of all three steps when they make an ex ante evaluation of the riskiness of a given strategy (as discussed in Chapter 2). In principle, such information cannot be collected unless the behaviour of each market participant is aggregated. However, through dynamic macro stress testing we may be able to obtain some information on market behaviour under stress, specifically information implying rapid changes in liquidity in the market in response to which market participants can perhaps conduct hedging or funding operations in advance of a stressful shock.⁴⁵

The framework of dynamic macro stress testing

How should dynamic macro stress testing which incorporates the behaviour of market participants be formulated? If we consider dynamic stress testing as a simultaneous equation macro model to describe market behaviour, the necessary inputs would include a common stress scenario about how risk factors change and the expected loss and reaction of individual participants under the scenario. While the expected loss and reaction of individual participants under a stress scenario can also be expressed in the form of behavioural functions, estimating these functions is quite difficult. We could obtain such information through interviews. The output obtained from the model is information about expected market behaviour.

The behaviour of market participants and measures to forecast its impact on the market

While it is important to account for the re-balancing and re-hedging operations of market participants, it is quite difficult to forecast them (for both market participants themselves and others, including central banks). In this regard, in order to forecast market behaviour, i.e. the aggregate of market participants' behaviour, information collected through macro stress testing would play an important role.

⁴⁵ To accurately forecast the macro impact, we need to take into account not only the behaviour of market participants but also contagion via expectations. Such contagion refers to the process by which a market participant determines his/her behaviour based on expectations regarding the behaviour of others. However, it should be noted that dynamic macro stress testing does not provide sufficient information to forecast this kind of contagion. To come up with a dynamic macro stress test methodology which can even forecast such contagion remains a future task.

Members of the Working Group on Macro Stress Testing

Chairperson Board of Governors of the Federal Reserve System	Mr Allen Frankel
Bank of Canada	Mr Richard Black
Banque de France	Mr Pierre Cailleteau Mrs Isabelle Strauss-Kahn
Deutsche Bundesbank	Mr Stefan Hohl Mr Stefan Rehsman
Banca d'Italia	Ms Antonella Foglia
Bank of Japan	Ms Tokiko Shimizu Mr Satoshi Kawazoe
De Nederlandsche Bank	Mr Peter J G Vlaar
Sveriges Riksbank	Mr Tor Jacobson
Swiss National Bank	Mr Christof Stahel
Bank of England	Ms Anne Vila
Federal Reserve Bank of New York	Mr John Kambhu Mr Jim Mahoney
Board of Governors of the Federal Reserve System	Mr Michael Gibson
European Central Bank	Mr Jukka Vesala
Bank for International Settlements	Mr Benjamin Cohen

ISBN 92-9131-601-6