Basel Committee
on Banking Supervision

Working Paper No. 19

Messages from the academic literature on risk measurement for the trading book

31 January 2011

BANK FOR INTERNATIONAL SETTLEMENTS

The Working Papers of the Basel Committee on Banking Supervision contain analysis carried out by experts of the Basel Committee or its working groups. They may also reflect work carried out by one or more member institutions or by its Secretariat. The subjects of the Working Papers are of topical interest to supervisors and are technical in character. The views expressed in the Working Papers are those of their authors and do not represent the official views of the Basel Committee, its member institutions or the BIS.

Copies of publications are available from:

Bank for International Settlements
Communications
CH-4002 Basel, Switzerland

E-mail: publications@bis.org
Fax: +41 61 280 9100 and +41 61 280 8100

ISSN: 1561-8854

Contents

Joint Working Group of the Research Task Force and the Trading Book Group of the Basel Committee on Banking Supervision

Chair: Mr Klaus Duellmann, Deutsche Bundesbank, Frankfurt

Mr Sirio Aramonte	Board of Governors of the Federal Reserve System, Washington, DC
Mr Philippe Durand	Banque de France, Paris
Mr Shun Kobayashi	Bank of Japan, Tokyo
Mr Myron Kwast	Federal Deposit Insurance Corporation, Washington, DC
Mr Jose A. Lopez	Federal Reserve Bank of San Francisco
Mr Giancarlo Mazzoni	Bank of Italy, Rome
Mr Peter Raupach	Deutsche Bundesbank, Frankfurt
Mr Martin Summer	Austrian National Bank, Vienna
Mr Jason Wu	Board of Governors of the Federal Reserve System, Washington, DC

Executive summary

This report summarises the findings of an ad hoc working group that reviewed the academic literature relevant to the regulatory framework for the trading book. This project was carried out in the first half of 2010 acting upon a request from the Trading Book Group to the Research Task Force of the Basel Committee on Banking Supervision. This report reflects the views of the individual contributing authors and should not be construed as representing specific recommendations or guidance by the Basel Committee for national supervisors or financial institutions.

The report builds on and extends previous work by the Research Task Force on the interaction of market and credit risk (see Basel Committee on Banking Supervision (2009a)). The literature review was complemented by feedback from academic experts at a workshop hosted by the Deutsche Bundesbank in April 2010, and reflects the state of the literature at this point in time. Please note that the term "value-at-risk" (VaR) should be interpreted henceforth in a broad sense as encompassing other common risk metrics, with the exception of Section 3 in which risk metrics are compared directly.

The main findings of the group are summarised below in the order of the Sections of the report.

1. Selected lessons on VaR implementation

There is no unique solution to the problem of the appropriate ***time horizon for risk measurement***. The horizon depends on characteristics of the asset portfolio (such as, market liquidity) and the economic purpose of measuring its risk; for example, setting capital or setting loss limits for individual trading desks.

Scaling of short-horizon VaR to a longer time horizon with the commonly used square-root-of-time scaling rule has been found to be an inaccurate approximation in many studies. This rule ignores future changes in portfolio composition. At present, there is no widely accepted approach for aggregating VaR measures based on different horizons.

Time-varying volatility is a feature of many financial time series and can have important ramifications for VaR measurement. Time-varying volatility can give rise to issues regarding the potential pro-cyclical effects of VaR-based capital measures. The effects of time-varying volatility on the accuracy of simple VaR measures diminish as the time horizon lengthens. In contrast, volatility generated by stochastic jumps will diminish the accuracy of long-horizon VaR measures unless the VaR measures properly account for the jump features of the data. Distinguishing between time-varying volatility and volatility changes that owe to stochastic jump process realisations can be important for VaR measurement.

Backtests that focus on the number of VaR violations have low power when the number of VaR exceptions is small. The power of backtests can be improved modestly through the use of conditional backtests or other techniques that consider multiple dimensions of the data like the timing of violations or the magnitude of the VaR exceptions. No consensus has yet emerged on the relative benefits of using actual or hypothetical results (ie P&L) to conduct backtesting exercises.

2. Incorporating liquidity

The literature distinguishes, first, between exogenous and endogenous liquidity; and, second, between normal (or average) liquidity risk and extreme (or stress) liquidity risk.

Exogenous liquidity refers to market-specific, average transaction costs and can be captured by a "liquidity-adjusted VaR" approach.

Endogenous liquidity refers to the price impact of the liquidation of specific positions. Endogenous liquidity depends on trade size and is relevant for orders that are large enough to move market prices; that is, it is the elasticity of prices to volumes. Endogenous liquidity may be easily observed in situations of ***extreme liquidity risk***, characterised by the ***collective liquidation*** of positions or, more generally, when all market participants react in the same way. Portfolios, however, may be subject to significant endogenous liquidity costs under *all* market conditions, depending on their size or on the positions of other market participants.

According to actual accounting standards, ***endogenous liquidity*** costs are not taken into account in the valuation of the trading books. A first step to incorporate this risk in a VaR measure would be to take it into account in the valuation method.

Although this last topic has attained considerable popularity in the recent literature, the practical implications for risk management have yet to be developed.

In practice, the time it takes to liquidate a position varies, depending on its transaction costs, the size of the position in the market, the trade execution strategy, and market conditions. Some studies suggest that, for some portfolios, this aspect of liquidity risk could also be addressed by extending the VaR risk measurement horizon.

3. Risk measures

VaR has become a standard risk measure in finance. Notwithstanding its widespread use, it has been criticised in the literature for lacking **subadditivity**, a property that implies that compartmentalised (say, desk-wise) risk measurement based on VaR is not necessarily conservative. The problem is relevant in practice and not only relevant for very high confidence levels of VaR.

The most prominent alternative to VaR is **expected shortfall**, which is subadditive. It is slowly gaining popularity among financial risk managers. Despite criticism focused on the complexity, computational burden, and backtesting issues associated with expected shortfall, the recent literature suggests that many issues have been resolved or have been identified as less severe than originally expected, including improvements in backtesting methodologies**.** At present**,** some financial institutions have come to more fully rely on expected shortfall metrics.

Spectral risk measures are a promising generalisation of expected shortfall. They have certain advantages over expected shortfall, including favourable smoothness properties and the possibility of adapting the risk measure directly to the risk aversion of investors. From a technical perspective, spectral risk measures require little additional effort if the underlying risk model is simulations-based.

4. Stress testing practices for market risk

Stress testing is often implemented as an *ad hoc* exercise without any estimate of the probability associated with the stress scenarios and often using modelling approaches that differ from an institution's VaR risk measurement framework. More recent research advocates the **integration of stress testing into the risk modelling framework**. This would overcome drawbacks of reconciling stand-alone stress test results with standard VaR model output.

Progress has also been achieved in theoretical research on the selection of **stress scenarios**. In one approach, for example, the "optimal" scenario is defined by the maximum loss event in a certain region of plausibility of the risk factor distribution.

The regulatory "**stressed VaR**" approach has not been analyzed in the academic literature. From a theoretical perspective it is an imperfect solution and its purpose is to reflect that current market conditions may not lead to an accurate assessment of the risk in a more stressful environment. Certain methods that could be meaningful in this context can be identified in the earlier literature on stress testing. Employing fat-tailed distributions for the risk factors and replacing the standard correlation matrix with a stressed one are two examples.

5. Unified versus compartmentalised risk measurement

Much of the risk measurement literature has focused on compartmentalised measures of risk such as interest rate, market, credit, or operational risks. In recent years, attention has shifted towards integrated or unified approaches for risk measurement that consider all risk categories jointly. From a theoretical perspective, an integrated approach is needed to capture potential compounding effects that are ignored in compartmentalised risk measurement approaches. Those approaches can underestimate risk if an asset portfolio cannot be cleanly divided into sub-portfolios along risk categories.

Empirical studies suggest that the magnitude of diversification benefits – that is, the amount by which aggregate risk is below the sum of individual risks – depends upon the level at which risks are measured. At higher levels of aggregation (eg at the holding company level), the benefits are more often detected; however, at a lower (eg the portfolio) level, risk compounding can become predominant.

The artificial distinction between a "trading book" and a "banking book" refers to positions, but these positions need not be exposed to different sets of risk. If the risks associated with these books are distinct, even if they are not independent, then adding the VaR measures of these books will be conservative. If the risks associated with the two books are not distinct, (eg if the separation is due only to accounting rules), then adding compartmentalised VaR risk measures is guaranteed to be conservative only if *all* risks relevant to *each* book are accounted for. If not, the sum of compartmentalised risk measures may understate the risk of the combined portfolio risk.

Irrespective of the separation of assets into "books", it is always questionable to calculate different risks for the *same* portfolio in a compartmentalised fashion and to hope that adding up the compartmentalised measures will be a conservative estimate of the true risk. In general, it will not be. This insight is particularly important for "backfitting packages", such as the incremental risk charge.

6. Risk management and value-at-risk in a systemic context

A number of studies are critical of VaR-based capital requirements because of their pro-cyclical nature. VaR-based capital rules require lower (higher) capital in the upswing (downturn) of the economy because volatilities of market prices of assets tend to vary over the business cycle. The procyclical nature of VaR-based capital requirements may induce cyclical lending behaviour by banks and exacerbate the business cycle. Another criticism of

VaR-based capital rules is that, under these rules, banks may face incentives to bias their models towards minimising regulatory capital charges and VaR models do not take endogeneity into account. When all banks follow a VaR-based capital rule, financial institutions may be incentivised to act uniformly in booms and busts. This tendency may create endogenous instabilities in asset markets that are typically not included when individual banks measure the risks of their trading books.

While procyclicality is often mentioned as a policy concern in the academic literature, the literature generally does not offer convincing solutions to how these concerns could be addressed in the regulatory framework, given that regulation should keep minimum capital requirements risk-sensitive in the cross-section.

Messages from the academic literature on risk measurement for the trading book

0. Introduction

This report summarises the findings of a working group (the “group”) that surveyed the academic literature that is relevant to a fundamental review of the regulatory framework of the trading book. This joint working group embraced members of the Trading Book Group and of the Research Task Force of the Basel Committee on Banking Supervision. This report summarises its main findings. It reflects the views of individual contributing authors, and should not be construed as representing specific recommendations or guidance by the Basel Committee for national supervisors or financial institutions.

The report builds on and extends previous work by the Research Task Force on the interaction of market and credit risk (see Basel Committee on Banking Supervision (2009a)). The literature review was complemented by feedback from academic experts at a workshop hosted by the Deutsche Bundesbank in April 2010 and reflects the state of the literature at this point in time.

The key findings of the group are presented in the executive summary. The structure of the remaining report is as follows:

Sections 1 to 4 address fundamental issues of a sometimes highly technical nature in current VaR-based approaches to risk measurement. More specifically, Section 1 gives an overview of implementation issues including questions on the necessity of including time-variation in volatility, the appropriate time horizon over which risk is measured and backtesting of VaR. Capturing market liquidity in a VaR framework is the key question addressed in the second section. Section 3 looks at the pros and cons of VaR as a metric for risk and considers alternative metrics put forward in the literature. Important aspects for the future evolution of stress tests are addressed in Section 4.

The last two sections 5 and 6 include management aspects, such as inter-risk aggregation and the borderline between the banking and trading books (which is discussed only briefly). They also expand the scope of this review by including macro-prudential aspects, such as systemic risk and pro-cyclicality. Section 5 is concerned with an integrated versus a compartmentalised approach to risk measurement, which has become particularly important since the recent financial crisis revealed that a focus on market risk alone may provide distorted results for a trading book. This topic draws heavily on the findings of the former working group of the Research Task Force on the interaction of market and credit risk (see Basel Committee on Banking Supervision (2009a)). Section 6 looks at the relations between and among risk measurement, systemic risk, and potential pro-cyclical effects of risk measurement.

1. Selected lessons on VaR implementation

1.1 Overview

In this section we review the academic and industry literature on VaR implementation issues, as it pertains to regulatory capital calculation. The three categories of implementation issues reviewed are: (1) time horizon over which VaR is estimated; (2) the recognition of time-varying volatility in VaR risk factors; and (3) VaR backtesting. With respect to (1), we find that the appropriate VaR horizon varies across positions and depends on the position's nature and liquidity. For regulatory capital purposes, the horizon should be long, and yet the common square-root of time scaling approach for short horizon VaR (eg, one-day VaR) may generate biased long horizon VaR (eg, ten-day VaR) estimates. Regarding (2), we find that while many trading book risk factors exhibit time-varying volatility, there are some concerns that regulatory VaR may suffer from instability and pro-cyclicality if VaR models incorporate time-varying volatility. We also sketch several approaches to incorporate time-varying volatility in VaR. As for (3), we survey the literature on VaR backtesting and discuss several regulatory issues including whether VaR should be backtested using actual or hypothetical P&L, and whether the banks' common practice of backtesting one-day VaR provides sufficient support for their ten-day, regulatory VaR.

It is worthwhile to note that some issues related to time horizons and time-varying volatility, and to a lesser extent backtesting, also pertain to risk measures other than VaR, such as Expected Shortfall (ES). A discussion of these alternative risk measures is contained in section three.

1.2 Time horizon for regulatory VaR

One of the fundamental issues in using VaR for regulatory capital is the horizon over which VaR is calculated. The 1998 Market Risk Amendment (MRA) sets this horizon to be ten days, and it allows ten-day VaR to be estimated using square-root of time scaling of one-day VaR. This approach raises three questions: (1) Is ten days an appropriate horizon? (2) Does VaR estimation based on time scaling of daily VaRs produce accurate risk measures? (3) What role do intra-horizon risks (ie P&L fluctuations within ten days) play, and should such risks be taken into account in the capital framework?

Is ten days an appropriate horizon?

There seems to be consensus among academics and the industry that the appropriate horizon for VaR should depend on the characteristics of the position. In the academic literature, Christoffersen and Diebold (2000) and Christoffersen, Diebold and Schuermann (1998) both assert that the relevant horizon will likely depend on where the portfolio lies in the firm (eg, trading desk vs CFO) and asset class (eg, equity vs fixed income), and the appropriate horizon should be assessed on an application-by-application basis. From this perspective, it appears that an across-the-board application of ten-day VaR horizon is not optimal. Indeed, one of the motivations for the Incremental Risk Charge (IRC) is to capture certain risks of credit related products at a longer horizon than ten days.

Although the literature suggests that it may be preferable to allow the risk horizon to vary across positions, Finger (2009), for instance, points out that there is no conceptual or statistical framework to justify the aggregation of a ten-day VaR and a one-year IRC. Danielsson (2002) adds that, if the purpose of VaR is to protect against losses during a liquidity crisis, the ten-day horizon at 99% refers to an event that happens roughly 25 times a decade, while a liquidity crisis is "unlikely to happen even once a decade. Hence the probability and problem are mismatched". In addition, even for the same financial product, the appropriate horizon may not be constant, because trade execution strategies depends on time-varying parameters, like transaction costs, expected price volatility, and risk aversion (Almgren and Chriss (2001), Engle and Ferstenberg (2006), Huberman and Stanzl (2005)). In addition, variation in risk aversion over the business cycle can be especially important in shortening the optimal trading horizon, potentially generating larger losses than those observable under more favourable conditions.

Danielsson (2002) questions the suitability of a ten-day horizon if VaR is to protect against a liquidity crisis, because a ten-day horizon implies a higher frequency of liquidity crisis than is observable in the data. Other authors have similarly suggested that the appropriate VaR horizon should depend on the economic purpose of VaR.[1] Smithson and Minton (1996), for instance, claim that nearly all risk managers believe a one-day horizon is valid for trading purposes but disagree on the appropriate horizon for long-term solvency or capital. Finger (2009) notes that there is "a tension between the regulatory risk horizon and the horizon at which banks manage their trading portfolios", and that the Market Risk Amendment (MRA) rules represent a compromise between regulatory and trading horizons through the introduction of the sixty-day moving average and backtesting multiplier mechanisms.

The computation of VaR over longer horizons introduces the issue of how to account for time variation in the composition of the portfolios, especially for institutions that make markets for actively traded assets like currencies (Diebold, Hickman, Inoue and Schuermann (1998)). A common solution is to sidestep the problem of changes to portfolio composition by calculating VaR at short horizons and scaling up the results to the desired time period using the square-root of time. While simple to implement, this choice may compromise the accuracy of VaR because, as discussed in the next section, tail risk is likely to be under-estimated (Bakshi and Panayotov (2010)). A second way to tackle the problem is to focus directly on calculating the portfolio VaR over the relevant horizon of interest (Hallerbach (2003)). These approaches may have limited value if the composition of the portfolio changes rapidly. Furthermore, data limitations make it challenging to study the P&L of newly traded assets. A third solution is to extend VaR models by incorporating a prediction of future trading activity, as noted by Diebold et al (1998): "To understand the risk over a longer horizon, we need not only robust statistical models for the underlying market price volatility, but also robust behavioural models for changes in trading positions."

Christoffersen and Diebold (2000) aptly characterised the issue of the optimal VaR horizon as "an obvious question with no obvious answer". Voices from the industry have suggested that a horizon longer than ten days may be necessary for regulatory capital purposes. It was also suggested that combining the liquidity horizon of individual positions with a constant level of risk may be an appropriate avenue.

Is square-root of time scaling a good idea?

Under a set of restrictive assumptions[2] on risk factors, long horizon VaR can be calculated as short horizon VaR scaled by the square root of time, if the object of interest is unconditional VaR (Kaufman (2004), McNeil, Frey and Embrechts (2005) and Danielsson and Zigrand (2006)). Unfortunately, the assumptions that justify square root of time scaling are rarely verified for financial risk factors, especially at high frequencies. Furthermore, risk management and capital computation are more often interested in assessing potential losses *conditional* on current information, and scaling today's VaR by the square root of time ignores time variation in the distribution of losses. We have not found any evidence in support of square-root of time scaling for conditional VaRs.

---

[1] For example, if VaR is expected to reduce the probability of bankruptcy, the horizon would line up with the time a bank needs to raise additional capital. If the focus is on losses while a position is being offloaded, the appropriate horizon would be more strictly related to asset characteristics.

[2] Specifically, the risk factors have to be normally distributed with zero mean, and be independently and identically distributed ("IID") across time.

The accuracy of square-root of time scaling depends on the statistical properties of the data generating process of the risk factors. Diebold et al (1998) show that, if risk factors follow a GARCH(1,1) process, scaling by the square-root of time *over-estimates* long horizon volatility and consequently VaR is over-estimated. Similar conclusions are drawn by Provizionatou, Markose and Menkens (2005). In contrast to the results that assume that risk factors exhibit time-varying volatility, Danielsson and Zigrand (2006) find that, when the underlying risk factor follows a jump diffusion process, scaling by the square root of time systematically *under-estimates* risk and the downward bias tends to increase with the time horizon. While these results argue against square-root of time scaling, it is important to acknowledge that we were not able to find immediate alternatives to square-root of time scaling in the literature. Therefore, the practical usefulness of square-root of time scaling should be recognised.[3]

Is intra-horizon risk important?

Bakshi and Panayotov (2010) discuss intra-horizon VaR (VaR-I), a risk measure that combines VaR over the regulatory horizon with P&L fluctuations over the short term, with a particular focus on models that incorporate jumps in the price process. The rationale behind intra-horizon VaR is that the maximum cumulative loss, as distinct from the end-of-period P&L, exerts a distinct effect on the capital of a financial institution. Bakshi and Panayotov (2010) suggest that VaR-I "can be important when traders operate under mark-to-market constraints and, hence, sudden losses may trigger margin calls and otherwise adversely affect the trading positions". Daily VaR does carry information on high frequency P&L but, as noted by Kritzman and Rich (2002), "Knowledge of the VAR on a daily basis does not reveal the extent to which losses may accumulate over time". Bakshi and Panayotov (2010) find that taking intra-horizon risk into account generates risk measures consistently higher than standard VaR, up to multiples of VaR, and the divergence is larger for derivative exposures.

1.3 Time-varying volatility in VaR

It is a stylised fact that certain asset classes, such as equities and interest rates, exhibit time-varying volatility. Accounting for time-varying volatility in VaR models has been one of the most actively studied VaR implementation issues. This section explores this topic, focusing on large and complex trading portfolios.

Is it necessary to incorporate time-varying volatilities and correlations?

The industry seems to think so since many firms advocate the use of fast reacting measures of risk such as exponential time-weighted measures of volatility. The reason given is that such VaR models provide early warnings of changing market conditions and may perform better in backtesting. The academic literature has also observed that time-varying volatility in financial risk factors is important to VaR, dating back to the 1996 RiskMetrics Technical document (J P Morgan (1996)). Pritsker (2006) showed theoretically that using historical simulation VaR without incorporating time-varying volatility can dangerously under-estimate risk, when the true underlying risk factors exhibit time-varying volatility.

---

[3] A concept related to square-root of time scaling is the scaling of VaR to higher confidence levels. Although we were unable to find literature on this topic, we recognize that this is an important issue particularly in situations when there are inadequate data points for one to accurately estimate risks deep into the tail. Some banks use certain reference densities (eg Student's t with six degrees of freedom) to conduct such scaling.

In contrast, some have argued that, depending on the purpose of VaR, capturing time-varying volatility in VaR may not be necessary, or may even be inappropriate. Christoffersen and Diebold (2000) observe that volatility forecastability decays quickly with time horizon for most equity, fixed income and foreign exchange assets. The implication is that capturing time-varying volatility may not be as important when the VaR horizon is long, compared to when the VaR horizon is relatively short. There are also concerns about pro-cyclicality and instability implications associated with regulatory VaRs that capture time-varying volatility. Dunn (2009), for instance, states that there is a “contradiction between the requirement for a risk sensitive metric to capture variations in volatility and correlation, and the regulatory requirement for a stable and forward looking basis for computing capital, that is not pro-cyclical”. In reference to modelling time-varying volatility in VaR, it wrote, “Some firms mentioned a trade-off in this issue, and that for some purposes such as capital allocation, risk measures with more stable properties that reflected longer historical norms were desirable.”

In summary, incorporating time-varying volatility in VaR appears to be necessary given that it is prevalent in many financial risk factors. Furthermore, many financial instruments are now priced with models with stochastic volatility features. It is logical that VaR models are constructed to account for these statistical properties. However, using VaR with time-varying volatility for regulatory capital raises the concerns of volatile and potentially pro-cyclical regulatory standards.

Methods to incorporate time-varying volatility in VaR for large, complex portfolios

Beginning with J.P. Morgan (1996), the Exponentially Weighted Moving Average (EWMA) approach has been regarded as one of the industry standards for incorporating time-varying volatility in VaR. EWMA is a constrained version of an IGARCH (1,1) model, and in the case of RiskMetrics the parameter in IGARCH was set to 0.97. An alternative and simpler approach is to weight historical data according to the weights introduced by Boudoukh, Richardson and Whitelaw (1998), where an observation from i days ago receives a weight of

$$w(i) = \frac{\theta^i (1-\theta)}{1-\theta^n}$$

Here n is the total number of days in the historical window, and θ is a number between zero and one which controls the rate of memory decay. An even simpler approach is to compute VaR with historical simulation using a short and frequently updated time series. Dunn (2009) has suggested that this method captures time-varying volatility quite well. Using simulations, Pritsker (2006) has shown that the approach of Boudoukh et al (1998) is not sensitive enough to pick up volatility changes. He advocated the use of Filtered Historical Simulation (FHS), first introduced by Barone-Adesi, Giannopoulos and Vosper (1999). Broadly speaking, FHS is based on the idea that risk factors should first be filtered through a GARCH model. The volatility is then updated using the model, and adhered to the filtered risk factors to constructed VaR.

Naturally, considerations should be given to how the above method can be applied to portfolios with large numbers of positions or risk factors. Barone-Adesi et al (1999) outlined a position-by-position FHS approach. They recommended filtering each risk factor separately, and building volatility forecasts for each factor. Analogously, EWMA and the weights introduced by Boudoukh et al (1998) can be applied the same way. However, weighting or filtering risk factors separately implicitly assumes that the correlation structure across risk factors does not change over time. Pritsker (2006) has pointed out that time-varying correlation is an important source of risk. Indeed, the recent crisis has highlighted the fact that correlations among many risk factors change significantly over time. One would need to be careful in handling time-varying volatilities as well as correlations.

Multivariate GARCH models such as the BEKK model of Engle and Kroner (1995), or the DCC model of Engle (2002) can be used to estimate time-varying volatilities as well as correlations. However, such multivariate GARCH models are difficult to estimate when there are a large number of risk factors. Some recent advances in the literature allow one to estimate a multivariate GARCH-type model when there are a large number of risk factors. For instance, Engle, Shephard and Sheppard (2007) proposed to average likelihoods before estimating the GARCH model with maximum likelihood. Engle and Kelly (2009) imposes a restriction on the correlation structure that helps facilitate estimation in large dimensions, but still allow correlations to change over time. Finally, Aramonte, Rodriguez and Wu (2010) estimates VaR for large portfolios comprising stocks and bonds by first reducing the dimension of risk factors using dynamic factor models, and then estimating a time-varying volatility model. The resulting VaR estimates are shown to out-perform historical simulation and FHS based on filtering risk factors one-by-one.

All in all, incorporating time-varying volatility in VaR measures is not straight forward when there are many risk factors. Time-varying correlations should be taken into account. Rather than using more involved methods, the industry appears to be taking less burdensome alternatives, such as using simple weighting of observations, or shortening the data window used to estimate VaR. These approaches compromise on accuracy, but are computationally attractive for large and complex portfolios. The recent academic literature offers promise that some of the sophisticated empirical methodologies may soon become practical for large complex portfolios.

1.4 Backtesting VaR models

As with any type of modelling, a VaR model must be validated. In particular, backtesting has been the industry standard for validating VaR models. This section reviews some backtesting methodologies suggested by the literature, and some issues pertaining to the application of such methodologies.

Backtesting approaches

Banks typically draw inference on the performance of VaR models using backtesting exceptions (sometimes also known as backtesting “breaches” or “violations”). For regulatory capital, the MRA imposes a multiplier on VaR depending on the number of backtesting exceptions the bank experiences.

While the MRA does not require banks to statistically test whether VaR has the correct number of exceptions, formal statistical inference is always desirable and many alternatives have been proposed in the literature. Kupiec (1995) introduced the unconditional coverage likelihood ratio tests as inference tools for whether the VaR model generated the correct number of exceptions. This methodology is simple to implement, but has two drawbacks. First, as pointed out by Kupiec (1995 and 2005), when the number of trading days used in VaR evaluation is limited (eg, one year or approximately 250 trading days), or when the confidence level is high (eg, 99% as in regulatory VaR), such tests have low power. This is not surprising, since one would expect only a small number of backtesting exceptions in most cases. Building a statistic out of a handful of exceptions, then, may induce high variance in the test statistic itself and the result may be sensitive to an incremental exception. Second, given that this test only counts exceptions, its power may be improved by considering other aspects of the data such as the grouping of exceptions in time.

Christoffersen (1998) has proposed a conditional backtesting exception test that accounts for the timing as well as the number of exceptions. The test is based on the fact that when the VaR model has conditionally the correct number of exceptions, then indicator variables representing the exceptions are IID[4] Bernoulli random variables. This test, however, may still be exposed to the low power problem. To this end, Berkowitz, Christoffersen and Pelletier (2010) provided a suite of conditional tests that have good power. These tests are based on the intuition of Christoffersen (1998) (ie correct conditional exceptions results in IID and Bernoulli exception indicators) but derive inferences from autocorrelation, spectral, and hazard rate tests.

Aside from backtesting based on the number of exceptions, a natural measure of VaR performance is the magnitude of the exceptions. Lopez (1999), for instance, formalised this idea by introducing a quadratic loss function where loss is the difference between actual P&L and VaR, when an exception occurs. Some papers, including Pritsker (2006) and Shang (2009), also consider the use of Mean-Squared-Error (MSE) as a measure of VaR performance in backtesting. Typically, one would measure the MSE between the 'true VaR' and the VaR estimate based on the model. Clearly, this method is not directly applicable to observed portfolio P&Ls, since the true VaR is never known. Nonetheless, it can be a useful validation method prior to putting a VaR model into production: one can define data generating processes mimicking those imposed by front office pricing models, simulate position P&L enough times to construct a P&L distribution, and find the 'true VaR' based on this simulated distribution. Then, the VaR model can be applied to the generated data, and the difference between 'true VaR' and estimated VaR can be analysed.

Backtesting issues

An important and yet ambiguous issue for backtesting is which P&L series to compare to VaR. Broadly speaking, the estimated VaR can be compared to either actual P&L (ie the actual portfolio P&L at the VaR horizon), or hypothetical P&L (ie P&L constructed based on the portfolio for which VaR was estimated). To complicate matters further, actual P&L may sometimes contain commissions and fees, which are not directly related to trading and trading risk. Franke, Härdle and Hafner (2008) and Berry (2009) described the relative merits of actual and hypothetical backtesting: actual backtesting has little value if the portfolio has changed drastically since VaR was estimated, but is simple to implement; hypothetical backtesting would make an 'apples-to-apples' comparison, but it comes with significant implementation burden given that hypothetical portfolios need to be constructed.

Another issue is the appropriate backtesting horizon. Banks typically backtest one-day ahead VaR and use it as a validation of the regulatory VaR, which is ten-day. The problem here is clear: a good one-day VaR (as validated by backtesting) does not necessarily imply a good ten-day VaR, and vice versa. Ten-day backtesting may not be ideal either, given the potentially large portfolio shifts that may take place within ten days. In that case, actual P&L backtesting in particular may not be very informative. While we were unable to find literature on this particular issue, it remains an important policy question.

1.5 Conclusions

We have reviewed the literature on a number of VaR implementation issues, including the appropriate time horizon, time-variation in the volatility of risk factors, and backtesting. We

---

[4] IID: independently and identically distributed.

find that the optimal way of addressing these points is idiosyncratic to the problem under consideration. For instance, when estimating long horizon VaR by scaling the short horizon counterpart by the square root of time, one may overestimate VaR if the underlying P&L process exhibits time varying-volatility, but underestimate VaR if the process has jumps.

Incorporating time-varying volatility in VaR measures appears to be important to make models more realistic although it is not straight forward when there are many risk factors. The recent academic literature offers promise in this direction. While many trading book risk factors have time-varying volatility, models that incorporate this feature may, however, generate pro-cyclical VaR and also be unstable, not least because of estimation issues.

In addition, the choice of whether to evaluate a VaR model on the basis of hypothetical or actual backtesting may be affected by the characteristics of the portfolio. Indeed, actual backtesting is less informative when the composition of the portfolio has recently changed. On the other hand, while hypothetical backtesting provides a more consistent comparison, it may impose substantial computational burdens because it requires reconstructing the history of the portfolio on the basis of its current composition.

2. Incorporating liquidity

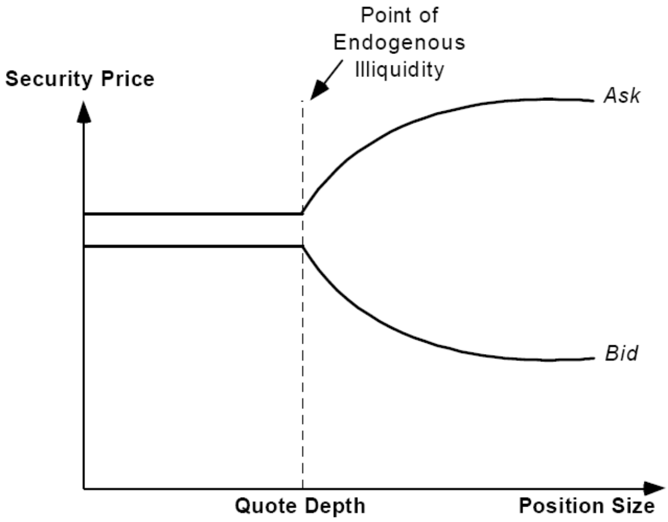
2.1 Overview

Discussing the challenging issue of how to incorporate market liquidity into a VaR model requires first of all a distinction between *exogenous* and *endogenous* liquidity. This distinction is made from the point of view of the bank, rather than in general equilibrium terms (Bangia, Diebold, Schuermann and Stroughair (1999a) and Bervas (2006)), More specifically, *exogenous* liquidity refers to the transaction cost for trades of average size, while *endogenous* liquidity is related to the cost of unwinding portfolios large enough that the bid-ask spread cannot be taken as given, but is affected by the trades themselves.

Bangia et al (1999a) give a graphical representation of *exogenous* and *endogenous* liquidity that is reproduced in Figure 1. Below a certain size, transactions may be traded at the bid/ask price quoted in the market (*exogenous* liquidity), and above this size, the transaction will be done at a price below the initial bid or above the initial ask, depending on the sign of the trade (*endogenous* liquidity).

Figure 1

Effect of position size on liquidation value



Source: Bangia et al (1999a).

The exogenous component of liquidity risk corresponds to the average transaction costs set by the market for standard transaction sizes. The endogenous component corresponds to the impact on prices of the liquidation of a position in a relatively tight market, or more generally when all market participants react in the same way, and therefore applies to orders that are large enough to move market prices (Bangia et al (1999a), Bervas (2006)). Exogenous liquidity risk, corresponding to the normal variation of bid/ask spreads across instruments can be, from a theoretical point of view, easily integrated into a VaR framework. Endogenous risk, corresponding to the impact on market prices of the liquidation of a position, or of collective portfolio adjustments, is more difficult to include in a VaR computation. Its impact, however, may be very significant, especially for many complex derivatives held in trading books of large institutions.

One way to incorporate liquidity risk into VaR measures is to include new VaR risk factors that can be used to model liquidity risks. This approach is feasible only when the parameters can be deduced from market data. Liquidity reserves taken by banks on their trading portfolio according to accounting standards correspond, more or less, to reserves for exogenous liquidity. In order to integrate this risk in the VaR computation, Bangia et al (1999a) propose to integrate the variability of the bid/offer spread for average size transactions as a risk factor.

To take into account endogenous liquidity in the value-at-risk is more difficult, as it is not even really taken into account in the valuation of trading portfolios, but its impact on both valuation and VaR should be significant. Academic literature on the subject – portfolio valuation and VaR computation – is quite rich, but very little application has been made in particular because endogenous liquidity reserves could be considered as not compliant to accounting standards.

In the following section, we first describe how, following existing literature, exogenous liquidity might be integrated into VaR measures. We then review several aspects of endogenous liquidity risk, and detail how this risk could be integrated in portfolio valuation and VaR computation. At last, we discuss on the choice of the VaR horizon when taking into account liquidity risk.

2.2 Exogenous liquidity

For the trading portfolio, following IAS rules, only exogenous liquidity risk will be taken into account in the valuation of cash assets and derivatives. Bangia et al (1999a) propose adding the bid/offer spread to characterise exogenous liquidity as a risk factor.

Their method poses that the relative spread, S = (Ask-Bid)/Mid-price, has sample mean and variance $\hat{\mu}$ and $\hat{\sigma}^2$. If the 99% quantile of the normalised distribution of S is $\hat{q}_{0.99}$, then the Cost of Liquidity is defined as

$$CoL_t = P_t \left(\frac{\hat{\mu} + \hat{q}_{0.99}\hat{\sigma}}{2} \right),$$

where P_t is today's value of the position. CoL_t is added to VaR to form a liquidity-adjusted VaR.

2.3 Endogenous liquidity: motivation

Adverse market conditions can generate a flight to liquid and high-quality assets, which reduces the ability to unwind positions in thinly-traded, low-quality assets. The effect can be compounded when the inventory of market makers becomes imbalanced, thus reducing their willingness to further accommodate sell trades, and when risk management standards for traders become tighter, reducing the probability of finding a counterparty.

Margin requirements are also a source of variation in the response of assets' prices and liquidity to fundamental shocks, because higher margins increase the probability of binding funding constraints. While the choice of margin requirements is endogenous to a security's liquidity, assets with identical payoffs can have different prices depending on margin requirements and the opportunity cost of capital.

The trading activities associated with hedging may also have an impact on the dynamics of the underlying assets. For example, delta hedging an option position entails buying the asset when its price goes up, and selling it when the price goes down: if the size of these adjustments is not negligible with respect to the volumes traded on the underlying, this strategy will increase upward and downward price movements.

Such effects will be particularly important when:

- the underlying asset is not very liquid,
- the size of the positions of the investors hedging an option is important with respect to the market,
- large numbers of small investors follow the same hedging strategy,
- the market for the underlying of the derivative is subject to asymmetric information, which magnifies the sensitivity of prices to clusters of similar trades (Gennotte and Leland (1990)).

In particular, on some specific markets driven by exotic options (eg Power Reverse Dual Callable, some CPPI[5] strategies, etc), even if a bank's trading book positions are small with

[5] CPPI: Constant Proportion Portfolio Insurance.

respect to the market, this bank may be exposed to losses due to endogenous liquidity. When many other banks have the same kind of positions and none has an opposite position[6], all these banks will have to adjust their hedging portfolio in the same way at the same time, and will then influence the market dynamics, and thus its small position may then be exposed to a significant liquidity cost.

The implications of derivative hedging have been extensively studied and derivative hedging has been identified as a potential explanation for the relation between implied volatilities and strike prices that can be observed on option markets (the so-called volatility smile). This literature includes the work of Platen and Schweizer (1998), Sircar and Papanicolaou (1998), Schönbucher and Wilmott (2000) and Subramanian (2008).

2.4 Endogenous liquidity and market risk for trading portfolios

Several authors have studied the implications of endogenous liquidity risk for portfolio valuation and on value-at-risk measures (Jarrow and Protter (2005), Rogers and Singh (2005)). In general, these authors define an optimal liquidation strategy in a finite (or infinite time horizon) model and deduce from this strategy the market value of the portfolio which is equal to the expectation of its liquidation price. The associated VaR measure, defined as a confidence interval around this expected price, implicitly incorporates market and liquidity risks.

Some studies suggest that endogenous liquidity costs should be added to position returns before carrying out VaR calculations. To that end, Bervas (2006) suggests to incorporate Kyle's Lambda or Amihud's (2002) illiquidity ratio in returns. Both measures are based on the relationship between returns and volume. Wu (2009) applies the illiquidity cost of Amihud (2002) to stock returns and calculates the sum as "liquidity-adjusted returns". VaR is then estimated by applying a GARCH type model to the adjusted returns. Francois-Heude and Van Wynendaele (2001) suggest an approach that modifies the model of Bangia et al (1999a) by using average weighted bid-ask spreads, with weights based on volume. Berkowitz (2000b) proposes to incorporate price impact of an immediate liquidation via the concept of elasticity of demand. Jarrow and Subramanian (2001) modify the mean and variance that appears in the standard parametric VaR formula to incorporate means and variances of liquidation time and liquidity discount. Botha (2008) extended Jarrow and Subramanian (2001) to the two assets portfolio level. Other notable papers include Le Saout (2002) and Hisata and Yamai (2000). Finally, Acerbi and Scandolo (2008) explore the impact of market and funding liquidity on portfolio prices and risk measure. The authors revisit the coherent measures of risk criteria introduced by Artzner et al (1999). They explain how these criteria should be interpreted; in particular they study liquidity models that lead to solutions for the valuation of portfolios constituted of analytically tractable assets.

The liquidity risk adjustments proposed in the academic literature, for the most part, have not been applied to the trading books of banks. One reason for this may be that the suggested valuation methods are not necessarily compliant with actual accounting standards.[7] Another reason academic proposals have been slow to be adopted may be the difficulty of estimating

---

[6] The clients of these banks may be investors who do not dynamically hedge their positions.

[7] For example, IAS 39 specify in AG72: "The appropriate quoted market price for an asset held or liability to be issued is usually the current bid price.… The fair value of a portfolio of financial instruments is the product of the number of units of the instrument and its quoted market price", and in AG75: "The objective of using a valuation technique is to establish what the transaction price would have been on the measurement date in an arm's length exchange motivated by normal business considerations."

model liquidity parameters, especially for OTC products. Indeed, the necessary data are not always available, and some of these parameters may be subjective. But recent discussions in academic circles regarding OTC transaction reporting could contribute to solve this problem.

A study of the impact of endogenous liquidity on the valuation of exotic derivatives, similar to the contributions of exogenous liquidity, would be especially welcome. When significant market movements materialise, traders will adjust their hedging strategies which may have an impact on the market dynamics if the volumes they have to trade are significant. Such effect has been suggested as a possible explanation for the significant trading losses that some banks have experienced during the last financial crisis.

Some authors have integrated liquidity risk with market and credit risk. For example, in order to evaluate a portfolio, Zheng (2006) studies optimal liquidation strategies, taking into account market and liquidity risk, together with the probability of default of an issuer or of a counterparty. Stange and Kaserer (2008) suggest calculating liquidity-adjusted VaR conditional on the market value of a position by incorporating bid-ask spread liquidity adjustments in returns; Qi and Ng (2009) discuss intraday liquidity risk and its impact on VaR.

2.5 Adjusting the VaR time horizon to account for liquidity risk

The recent financial crisis has provided examples where a change in market liquidity conditions alters the *liquidity horizon*, ie the time required to unwind a position without unduly affecting the underlying instrument prices (including in a stressed market). This finding was already addressed in previous work of the Research Task Force (see Basel Committee on Banking Supervision (2009a)) and it is consistent with the literature.

Lawrence and Robinson (1997), for example, suggest that the application of a unique horizon to all positions by ignoring their size and level of liquidity is undesirable. They suggest determining the temporal horizon by the size of the position and the liquidity of the market. Haberle and Persson (2000) propose a method based on the fraction of daily volume that can be liquidated without significant impact on the market price, which can be interpreted as holding the horizon fixed and determining how much can be liquidated during that horizon. The method of Jarrow and Subramanian (2001) is also relevant in this context as it requires an estimate of the average liquidation time.

Previous work of the Research Task Force suggests an *interdependence* between risk assessment and liquidity horizon: On the one hand the exposures of banks to market risk and credit risk may vary with a risk horizon that is set dependent on market liquidity. If liquidity decreases, for example, the risk horizon lengthens and the exposure to credit risk typically increases. On the other hand, liquidity conditions are also affected by perceptions of market and credit risk. A higher estimate of credit risk for example, may adversely affect the willingness to trade and thereby market liquidity (see Basel Committee on Banking Supervision (2009a)).

Liquidation horizons vary over the business cycle, increasing during times of market stress. Besides transaction costs or the size of the position relative to the market, a trade execution strategy also depends on factors like expected price volatility and risk aversion (Huberman and Stanzl (2005)). If, for instance, risk aversion increases during a crisis, an investor may choose to trade more rapidly than during normal times, thus generating higher losses than those observable under favourable economic conditions.

2.6 Conclusions

Both exogenous and endogenous liquidity risks are important; endogenous liquidity risk is particularly relevant for exotic / complex trading positions. While exogenous liquidity is partially incorporated in the valuation of trading portfolios, endogenous liquidity is typically not, even though its impact may be substantial. Although endogenous liquidity risk is especially relevant under stress conditions, portfolios may be subject to significant endogenous liquidity costs under *all* market conditions, depending on their size or on the positions of other market participants.

The academic literature suggests as a first step to adjust valuation methods in order to take endogenous liquidity risk into account. Then a VaR integrating liquidity risk could be computed. Notwithstanding academic findings on this topic, in practice, the ability to model exogenous and endogenous liquidity may be constrained by limited data availability, especially for OTC instruments.

3. Risk measures

3.1 Overview

This section compares selected risk measures that appear to be relevant for risk management purposes either today or in the future. The alternative measures considered include VaR, expected shortfall and spectral measures of risk. The key features used to decide among alternative risk measurement approaches include ease of calculation, numerical stability, the possibility to calculate risk contributions of individual assets to portfolio risk, backtesting possibilities, incentives created for risk managers, and, linked to the latter, the relation between risk measures and regulators' objectives. Although few financial institutions currently make use of VaR alternatives, those that do are often considered as technologically leading in the industry.

In the literature, risk measures are usually defined as functions of *random variables* (portfolio losses or returns in most cases). This seems to be a trivial aspect but is actually a substantial restriction because it binds the analysis to *one* point of time; while this time horizon can be varied, a *joint* analysis of a portfolio's losses at several times, which may be important for asset/liabilities management, is excluded. Risk measures being a function of random loss variables also means that these variables are not an attribute of risk measures; the probability distributions of the variables are specified in a preceding step, and the analysis of risk measures is not an analysis of whether the random variables are correctly specified.

In our discussion of alternative measures we focus on VaR because of its high relevance to the industry today and on Expected Shortfall and Spectral Measures because of their advantages and hence a potentially growing importance in the future. Other risk measures, such as variance or upper-tail moments are briefly sketched for completeness.

3.2 VaR

Concept of VaR and its problems

VaR has become a standard measure used in financial risk management due to its conceptual simplicity, computational facility, and ready applicability. Given some random loss L and a confidence level α, $VaR_{\alpha}(L)$ is defined as the quantile of L at the probability α. The quantile is not necessarily unique if there are regions where the loss distribution function

F_L does not grow. For these cases, McNeil et al (2005) define the VaR as the smallest, ie most optimistic quantile:

$$VaR_{\alpha}(L) = \inf\{l : F_L(l) \geq \alpha\}$$

Despite its prevalence in risk management and regulation, VaR has several conceptual problems. Artzner, Delbaen, Eber and Heath (1999) point out that VaR measures only quantiles of losses, and thus disregards any loss beyond the VaR level. As a consequence, a risk manager who strictly relies on VaR as the only risk measure may be tempted to avoid losses within the confidence level while increasing losses beyond the VaR level. This incentive sharply contrasts with the interests of regulators since losses beyond the VaR level are associated with cases where regulators or deposit insurers have to step in and bear some of the bank's losses. Hence, VaR provides the risk manager with incentives to neglect the severity of those losses that regulators are most interested in.

Neglecting the severity of losses in the tail of the distribution also has a positive flipside: it makes backtesting easier or possible in the first place simply because empirical quantiles are per se robust to extreme outliers, unlike typical estimators of the expected shortfall, eg (see below).

VaR is criticised for not being a *coherent* risk measure, which means that VaR lacks an axiomatic foundation as proposed by Artzner et al (1999). They set out the following four consistency rules. A risk measure R is called *coherent* if it satisfies the following axioms.

- Subadditivity (diversification) $R(L_1 + L_2) \leq R(L_1) + R(L_2)$
- Positive homogeneity (scaling) $R(\lambda L) = \lambda R(L)$, for every $\lambda > 0$
- Monotonicity $R(L_1) < R(L_2)$ if $L_1 < L_2$
- Transition property $R(L + a) < R(L) - a$

VaR is not coherent because it may violate the subadditivity criterion. For why subadditivity indeed makes sense we quote from McNeil et al (2005):

- "Subadditivity reflects the idea that risk can be reduced by diversification,… the use of non-subadditive risk measures in a Markowitz-type portfolio optimisation problem may lead to optimal portfolios that are very concentrated and that would be deemed quite risky by normal economic standards.
- If a regulator uses a non-subadditive risk measure in determining the regulatory capital for a financial institution, that institution has an incentive to legally break up into various subsidiaries in order to reduce its regulatory capital requirements.…
- Subadditivity makes decentralisation of risk-management systems possible. Consider as an example two trading desks with positions leading to losses L_1 and L_2. Imagine that a risk manager wants to ensure that $R(L)$, the risk of the overall loss $L = L_1 + L_2$, does not exceed some number M. If he uses a subadditive risk measure R, he may simply choose bounds M_1 and M_2 such that $M_1 + M_2 \leq M$ and impose on each of the desks the constraint that $R(L_i) \leq M_i$; subadditivity then ensures automatically that $R(L) \leq M_1 + M_2 \leq M$."

Remark 1: Related to non-coherency of VaR, Basak and Shapiro (2001) create an example where VaR-based risk management may possibly be problematic. They analyse optimal, dynamic portfolio and wealth/consumption policies of utility maximising investors who must also manage market-risk exposure using VaR. They find that VaR risk managers often optimally choose a larger exposure to risky assets than non-VaR risk managers and consequently incur larger losses when losses occur.

Remark 2: At first glance, subadditivity and positive homogeneity may not appear as meaningful concepts when risk measures are applied to counterparty credit risk (CCR) or other types of credit risk. For example, assume there is CCR involved with some position in the trading book. Doubling the position can *more than* double the CCR simply because not only the exposure doubles but also because the position becoming extremely profitable can *make* the counterparty go bankrupt. This appears to contradict the postulate of positive homogeneity which claims $R(2L) = 2R(L)$. However, it is not that positive homogeneity is wrong for CCR but rather that this idea reflects a misunderstanding of risk measures as functions of *positions*. Generally, the risk of the doubled position will not be $2L$ but rather a random variable with a wider probability distribution. Similar effects are possible for subadditivity; the issue is related to Section 5 on whether a compartmentalised measurement of risk is appropriate. For instance, there may exist two positions which cause individual risks L_1 and L_2, respectively, if held alone, but the risk of holding both positions may be more severe than $L_1 + L_2$ for similar reasons as in the above example. Knowing this, one might question the subadditivity property as such because it requires $R(L_1 + L_2) \le R(L_1) + R(L_2)$. However, not subadditivity is to blame but a potential misunderstanding of $L_1 + L_2$ as the risk of holding both positions together. These considerations imply two lessons:

- It may be problematic to assume that a vector of assets linearly maps into the associated vector of random losses.
- If a "risk measure" is defined as a composed mapping *from positions* (via the loss variable) to numbers, this mapping is generally not coherent. Assuming coherence can lead to an underestimation of risk.

Is VaR failing subadditivity relevant in practice?

The favourite textbook example of VaR violating subadditivity is constructed with the help of two large losses the probability of which is lower than the confidence level of the VaR. When measured separately, each loss can have zero VaR but when aggregated, the probability that either of the losses occurs may exceed the confidence level so that the VaR of the aggregated loss is positive.

As the textbook example relies on jumps in the loss distribution one might conjecture that VaR works properly if loss distributions are smooth or if discrete losses are superimposed by sufficiently large smooth ones. Whether this intuition is correct ultimately depends on the situation and particularly on the tail thickness of the loss distributions:

- McNeil et al (2005) present an example of a *continuous* two-dimensional loss distribution in which VaR violates subadditivity. While this is alarming in that it does not build on the abovementioned textbook ingredients, the example is still rather artificial. [8]

[8] Example 6.22, p 253. While the marginal distributions are even normal, the joint distribution has very unusual properties; cf Figure 6.2.

- If the joint distribution of risk factors is elliptical (multivariate normal, eg), VaR is subadditive; see McNeil et al (2005), Theorem 6.8.
- Gourier, Farkas, and Abbate (2009) give an example where the sum of some fat-tailed, continuously distributed, and independent (!) random variables has a larger VaR than the sum of individual VaRs. The example is rather exotic as one of the random variables has infinite mean. While VaR fails in that case, it must be conceded that there is no coherent *and* practicable alternative at all because any coherent risk measure must be infinite then.[9]
- Danielsson, Jorgensen, Samorodnitsky, Sarma, and de Vries (2005) prove that VaR is subadditive for a sufficiently high confidence level if the total loss has finite mean. Note, however, that this is not an "all clear" signal but an asymptotic result only. Generally it may happen that subadditivity is only achieved for impracticably high confidence levels.
- Degen, Embrechts, and Lambrigger (2007) restrict their analysis to a parametric class of distributions but gain valuable insight into the interplay of tail thickness, confidence level, and subadditivity. For example, they find the 99%-VaR to be superadditive even for very moderate tail indices above 6, which means that moments of order 6 and lower may exist.[10] These are realistic cases in market risk. The dependence structure between the individual losses generally aggravates the problem but has surprisingly low impact in the cases considered.

To sum up, while the literature provides us with conditions that assure VaR is subadditive (and thus coherent), these conditions are generally *not fulfilled* in the market risk context; for example, Balaban, Ouenniche and Politou (2005) estimate tail indices between 1 and 2 for UK stock index returns over holding periods between 1 and 10 days, meaning that these tails are substantially heavier than necessary for assuring the subadditivity of VaR in general.

3.3 Expected shortfall

Expected shortfall (ES) is the most well-known risk measure following VaR. It is conceptually intuitive and has firm theoretical backgrounds; see, eg, Dunn (2009), Artzner et al (1999), Acerbi and Tasche (2002), Sy (2006), and Yamai and Yoshiba (2005). Therefore, it is now preferred to VaR by an increasing number of risk managers in the industry.

ES corrects three shortcomings of VaR. First, ES does account for the severity of losses beyond the confidence threshold. This property is especially important for regulators, who are, as discussed above, concerned about exactly these losses. Second, it is always subadditive and coherent. Third, it mitigates the impact that the particular choice of a single confidence level may have on risk management decisions, while there is seldom an objective reason for this choice.

---

9 Gourier et al (2009) refer to Delbaen (2002) who shows in Theorem 13 that, given a continuous distribution and some continuity of the risk measure, any coherent risk measure larger or equal than the α-VaR cannot fall short of the α-expected shortfall; the latter is already infinite in Gourier's example so that no useful coherent measure of that risk can exist.

10 The higher the tail index, the thinner is the tail. Degen et al (2007) consider random variables the distribution tail of which is as thick as that of a transform $\exp\left(gZ + 0{,}5hZ^2\right)$ of a standard normal Z; eg, $g = 2.3$ and $h = 0.25$ make the VaR super-additive; the tail index is 4 in this example.

To define ES, let L be a random loss with distribution function F_L and $\alpha \in (0,1)$ a confidence level (close to 1). Recall that the α-VaR is defined as the α-quantile of F_L. The ES at level α is defined by

$$ES_\alpha \equiv \frac{1}{1-\alpha}\int_\alpha^1 VaR_u\left(L\right)du \tag{1}$$

and can thus be understood as an average of all VaRs from level α up to 1. ES is a coherent risk measure – and so subadditive. It is continuous in α and thus avoids cliff effects that may appear when the distribution has discrete components.

If the loss distribution is continuous, there is an even more intuitive representation:

$$ES_\alpha = E\left(L \middle| L \geq VaR_\alpha\right), \tag{2}$$

ie ES is then the expected loss conditional on this loss belonging to the $100(1-\alpha)$ percent worst losses. This measure has several other names like tail conditional expectation (TCE) or conditional VaR (CVaR). It is the key to simulations-based calculations of ES but care has to be taken as it does not always coincide with ES, and it is also not necessarily subadditive. The technical problem arises if the distribution function jumps from a value below the VaR confidence level to a value above it. Then, a correction term must be introduced into (2) to reconcile it with the correct ES from (1); see Acerbi and Tasche (2002).

The calculation of ES and the marginal contributions of assets to portfolio ES is more challenging than the corresponding calculations for VaR, especially for high confidence levels, because a formula for the α-quantiles of the loss distribution is often missing. Simulations need to be done in most cases. Since the introduction of expected shortfall, substantial progress has been made on computational issues, mainly through the application of importance sampling techniques (Kalkbrener, Lotter, and Overbeck (2004), Egloff, Leippold and Jöhri (2005), or Kalkbrener, Kennedy, and Popp (2007)). Research suggests that computational techniques have advanced to a point that expected shortfall is a viable risk management option for financial institutions.

Remark 3: In Remark 1, it is noted that utility optimisation using a VaR constraint can lead to perverse investment decisions. Risk measures which control the first moment of a random variable (such as ES) have been proposed to overcome this problem. However, recently Wylie, Zhang, and Siu (2010) showed that in the context of hedging both ES and VaR can give rise to discontinuous hedging behaviour that can lead investors to take extremely high-risk positions even when apparently minimising the risk measures.

Backtesting ES

Intuitively, backtesting ES is more complicated and/or less powerful than backtesting VaR because the robust statistic given by the number of VaR violations, as the most common VaR backtest statistic, must be replaced by something that accounts for the *magnitude* of VaR exceedances so that ES backtests by nature have to cope with the size of outliers.

Whether specialised ES backtests are good or not, one simple option is always available: during an ES calculation, the VaR at the same α can be generated as a by-product with low additional effort. One can backtest this VaR with traditional methods; if the VaR is rejected, the corresponding ES calculation can hardly be correct. Of course, VaR backtest acceptance does not guarantee the correctness of the ES calculation, and this would be true even if the VaR backtest were always right.

Some backtests verify if the VaR correctly adjusts for changes in risk dynamics ("conditional coverage"; see Berkowitz and O'Brien (2002)). According to Pritsker (2006), they exploit the fact that exceedances of a correctly calculated VaR "should not help forecast future exceedances. Therefore, the autocorrelation function of the VaR exceedances should be equal to 0 at all lags". It is hard to decide whether ES or VaR is verified with these backtests because VaR exceedances are the very constituents of ES.

Because backtests that are strictly focused on some historical estimator of the risk measure, like the number of VaR violations, often have low power, several authors propose to backtest the whole distribution (or at least the tail), for instance by transforming loss realisations with the forecasted loss distribution: If the latter is correct, the transformed sample must be equally distributed on $[0,1]$. This hypothesis can be tested (Berkowitz (2001)). While not all backtests of this kind could be used in regulation[11], Kerkhof and Melenberg (2004) follow this approach to develop test statistics directly applicable to VaR and ES. The test statistic for the ES involves, besides the forecasted ES and VaR, also the calculation of the ES of the *squared* loss, which would be a tolerable extra effort in practice.

Kerkhof and Melenberg (2004) show that their backtest statistics for ES perform better than those for VaR. They also derive regulatory multiplication factors for their backtests and conclude that "the resulting regulatory capital scheme using expected shortfall compares favourably to the current Basel Accord backtesting scheme". It is important to notice that, according to Kerkhof and Melenberg (2004), a comparison of an α-ES with an α-VaR is not "fair" in the context of economic or regulatory capital. Since $ES_\alpha \geq VaR_\alpha$ for the same confidence level α, they lower the confidence level α' for the ES such that $\mathrm{ES}(\alpha') \approx \mathrm{VaR}(\alpha)$. The intuition is that a regulator would require roughly the same amount of capital for a fixed portfolio, irrespective of the risk measure in use.

This aspect is important not only in the context of backtesting but also when estimation errors for ES and VaR are compared. Yamai and Yoshiba (2005) find ES estimates of fat (generalised Pareto distributed) tailed losses to be much more volatile than their VaR counterparts but they compare ES and VaR at the same confidence level. A comparison in the spirit of Kerkhof and Melenberg (2004) seems not to have been conducted so far but could easily be done.

Wong (2008) suggests another backtest statistic for ES that accounts for the small samples of VaR exceedances. The statistic is derived for normally distributed losses and turns out to perform very well under these assumptions. The test is also powerful in detecting non-normal VaR exceedances. For the case that a bank models non-normal losses when calculating the ES, Wong suggests to derive adapted saddle-point approximations for the estimator's distribution or to use the sample transform as used by Berkowitz (2001) and Kerkhof and Melenberg (2004). These results are promising, but in the context of banking regulation it must be taken into account that Wong's backtest would require that banks provide more information than they currently do for regulatory backtests. At present, past returns are compared with reported VaRs. With Wong's backtest, the bank would also have to report its

[11] Some of these tests require that the bank fully specifies the loss distribution in the tail, not just the risk measure (ES or VaR). While this should not be a problem for bank internal purposes, a fully specified tail distribution would entail a fairly complex interface between bank and supervisor.

estimates of tail thickness, which is potentially involved with weird incentives. For instance, banks might, keeping minimum capital constant, be tempted to rely on certain tail distributions under which Wong's backtest has particular low power so that it is difficult to provide firm evidence of wrong risk reporting. Whether such concerns are substantial is left to future research.

3.4 Spectral risk measures

Spectral risk measures (SRM) are a promising generalisation of ES (Acerbi (2002)). While the α-ES assigns equal weight to all β-VaRs with $\beta \geq \alpha$ but zero to all others, an SRM allows these weights to be chosen more freely. This is implemented by a weight function $w:[0,1] \rightarrow [0,\infty)$ that integrates to 1. An SRM is formally defined as

$$SRM = \int_0^1 w(u) VaR_u(L) du .$$

Expected shortfall is a special case of spectral measure, where $w(u) = (1-\alpha)^{-1} 1_{\{\alpha \leq u \leq 1\}}$. The definition of SRM is restricted to functions w that increase over $[0,1]$, which ensures that the risk measure is coherent. This restriction also implies that larger losses are taken more seriously than smaller losses and thus the function w establishes a relationship to risk aversion. The intuition is that a financial institution is not very risk averse for small losses, which can be absorbed by income, but becomes increasingly risk averse to larger losses. As there may be a level of loss where employing additional capital to absorb yet higher loss is no longer desirable, such losses should be given the highest weights from a regulator's angle because often the public would have to bear such losses. Intuitively, a weight function that increases can also be thought of as marginal costs that rise while losses become increasingly rare, ie large.

Another advantage of SRM over ES (and VaR, a fortiori) is that they are not bound to a single confidence level. Rather, one can choose w to grow continuously with losses and thereby make the risk measure react to changes in the loss distribution more smoothly than the ES, and avoid the risk that an atom in the distribution being slightly above or below the confidence level has large effects.

If the underlying risk model is simulation-based, the additional effort to calculate an SRM as opposed to the ES seems negligible; the simulated VaR realisation are just differently weighed (Acerbi (2002)).

In spite of their theoretical advantages, SRMs other than ES are still seldom used in practice.[12] However, insurers use the closely related concept of *distortion measures* (see Section 3.5). Prominent examples such as the measure based on the Wang transformation (see below) are also SRMs.

Remark 4: Leaving aside that w must be increasing to meet the definition of SRM, VaR is a limiting case of spectral risk measures: for instance, the sequence of SRMs based on the weight functions $w_n(u) \equiv 0.5n 1_{\{\alpha - n^{-1} \leq u < \alpha + n^{-1}\}}$ converges to the α-VaR.

[12] At least one reputable risk consulting company reports it is currently implementing an SRM-based risk management system for some of its clients.

3.5 Other risk measures

There also are a number of other risk measures which are briefly introduced in this subsection.

Distortion risk measures: These measures are used in actuarial risk measurement. The definition is very general; both spectral risk measures (including ES) and the VaR are nested. To define distortion risk measures, let D be any distribution function on $[0,1]$ that is right-continuous and increasing with $D(0)=0$ and $D(1)=1$. This D is called the *distortion function*. A *distortion risk measure* of loss L is defined as

$$DM(L) \equiv \int_0^1 VaR_u(L)\,dD(u).$$

Each spectral risk measure is clearly a distortion risk measure; to see this, recall that the weight function w integrates to 1 and observe that the SRM and the distortion measure defined by the antiderivative $D(u) \equiv \int_0^u w(s)\,ds$ are identical.

Distortion risk measures are not necessarily coherent; the definition allows for distortion functions with a non-monotonous derivative (this is just the weight function of the corresponding SRM), whereas Acerbi (2002) has shown that the monotonicity of w is also *necessary* for the risk measure to be coherent.[13]

The VaR has a representation as a distortion risk measure by $D_{VaR}(u) = 1_{\{u \geq \alpha\}}$.

The Wang transform (Wang (2001)) $D_\theta^{Wang}(u) = \Phi\left(\Phi^{-1}(u) + \log\theta\right)$, where Φ denotes the Gaussian distribution function and $\theta < 1$, is an interesting distortion function. The corresponding risk measure is also a spectral risk measure because the first derivative of D_θ^{Wang} is strictly increasing. Hence the Wang transform indeed implements risk aversion over the whole range of losses but particularly in the tail. It has been applied to the pricing of catastrophe insurance contracts and exotic option pricing where Black-Scholes assumptions cannot be applied.

Variance: The variance is historically the most important risk measure and widely used in practice. It has many desirable properties but at least two drawbacks from a regulatory perspective. McNeil et al (2005) state "if we want to work with variance, we have to assume that the second moment of the loss distribution exists.… [V]ariance is a good measure of risk only for distributions which are (approximately) symmetric… However, in many areas of risk management, we deal with highly skewed distributions."

The **mean deviation**, defined as $MD(L) \equiv \mathbf{E}\left|L - \mathbf{E}L\right|$, can do without second moments but suffers from the same problems with skewed distributions as the variance. It is less accessible to analytical treatment than the variance and therefore rarely used as a risk measure.

---

[13] Wang (2001) claims all smooth distortion measures are coherent. This is wrong as subadditivity is missing in general. Wang (2001) means to build on Wang, Young and Panjer (1997) which, however, state that a distortion measure is subadditive *if it is convex* (in our notation). The latter is correct and conforms to Acerbi (2002).

Upper partial moments (see McNeil et al (2005)): Given a loss distribution F_L, an exponent $k \geq 0$ and a reference point q, which could be some VaR, the upper partial moment $UPM(k,q)$ is defined as

$$UPM(k,q) = \int_q^{\infty} (l-q)^k dF_L(l).$$

Hence, for $k > 1$ an UPM measures losses beyond the threshold q with increasing weight. It is therefore related to spectral risk measures in spirit but not equivalent in analytic terms. The higher k is, the more conservative is the UPM. For $k = 1$ and continuous loss distributions, there is a close relationship with expected shortfall:

$$UPM(1, VaR_{\alpha}) = (1-\alpha)(ES_{\alpha} - VaR_{\alpha}).$$

Left-tail measure: In a similar vein of mean deviation and lower (upper) partial moment, Wu and Xiao (2002) propose a *left-tail measure*, defined as the conditional standard deviation of VaR exceedances, ie

$$LTM \equiv \sqrt{E\left\{\left[L - E\left(L \middle| L \geq VAR_{\alpha}\right)\right]^2 \middle| L \geq VAR_{\alpha}\right\}}$$

Wu and Xiao (2002) show that the left-tail measure is useful particularly for the measurement of non-normal tail risks. This risk measure has several undesirable features such as a lack of coherency and a heavy burden of calculation.

3.6 Conclusions

While VaR has been criticised for its lack of coherence, until recently it was unclear whether this flaw is relevant for real asset portfolios, particularly for risks in the trading book. Degen et al (2007) have shown that the lack of coherence can be an important problem for trading book risk measurement. A risk measurement based on VaR is thus not necessarily conservative.

The ES avoids the major flaws of VaR but its fundamental difference from VaR – that it accounts for the magnitude of losses beyond a threshold – is an equally important advantage. By this, it aligns the interests of bank managers and owners to those of the public much better than VaR.

Much of the criticism of ES that has been brought forward in defence of VaR could be refuted. Advanced simulation techniques have helped to make ES calculations stable enough, and ES and VaR backtests have similar power, if compared on the basis that both risk measures have roughly the same value.

Spectral risk measures are a promising generalisation of expected shortfall. The main advantages are improved smoothness and the intuitive link to risk aversion. If the underlying risk model is simulations-based, the additional calculation effort as opposed to ES seems negligible.

4. Stress testing practices for market risk

4.1 Overview

VaR limitations have been highlighted by the recent financial turmoil. Financial industry and regulators now regard stress tests as no less important than VaR methods for assessing a bank's risk exposure. A new emphasis on stress testing exercises derives also from the amended Basel II framework which requires banks to compute a valid stressed VaR number.

A stress test can be defined as a risk management tool used to evaluate the potential impact on portfolio values of unlikely, although plausible, events or movements in a set of financial variables (Lopez (2005)). They are designed to explore the tails of the distribution of losses beyond the threshold (typically 99%) used in value-at-risk (VaR) analysis.

However, stress testing exercises often are designed and implemented on an ad hoc compartmentalised basis, and the results of stress tests are not integrated with the results of traditional market risk (or VaR) models. The absence of an integrated framework creates problems for risk managers, who have to choose which set of risk exposures are more reliable. There is also the related problem that traditional stress testing exercises typically remain silent on the likelihood of stress-test scenarios.

A survey of stress testing practices conducted by the Basel Committee in 2005 showed that most stress tests are designed around a series of scenarios based either on historical events, hypothetical events, or some combination of the two. Such methods have been criticised by Berkowitz (2000a). Without using a risk model the probability of each scenario is unknown, making its importance difficult to evaluate. There is also the possibility that many extreme yet plausible scenarios are not even considered.

Berkowitz proposed the integration of stress testing into formal risk modelling by assigning probabilities to stress-test scenarios. The resulting risk estimates incorporate both traditional market risk estimates and the outcomes of stress tests, as well as the probabilities of each. Therefore, they provide an integrated set of risk indicators and estimates to work with.

4.2 Incorporating stress testing into market-risk modelling

Traditional stress testing exercises can be classified into three main types, which differ in how the scenarios are constructed:

1. historical scenarios;

2. predefined or set-piece scenarios where the impact on P/L of adverse changes in a series of given risk factors is simulated;

3. mechanical-search stress tests, based on automated routines to cover prospective changes in risk factors, then the P/L is evaluated under each set of risk-factor changes, and the worst-case results are reported.

All these approaches depend critically on the choice of scenarios. A related problem is that the results of stress tests are difficult to interpret because they give no idea of the probabilities of the events concerned (Berkowitz (2000a)). These criticisms can be addressed by integrating stress testing into the market risk modelling process and assigning probabilities to the scenarios used in stress testing. Once scenarios are put in probabilistic form, a unified and coherent risk measurement system is obtained rather than two incompatible ones and backtesting procedures can be applied to impose some (albeit limited) check on scenarios. Inevitably, the choice of scenarios will remain subjective, but even there, the need to assign probabilities to scenarios will impose some discipline on risk management.

Several authors have developed an integrated approach to stress testing including Kupiec (1998) who examines cross-market effects resulting from a market shock and Aragones et al (2001) who incorporated hypothetical stress events into an Extreme Value Theory (EVT) framework.

Alexander and Sheedy (2008) analysed the problem of determining the most suitable risk model in which to conduct a stress test. Obviously if the model is mis-specified, their approach is vulnerable to a considerable degree of model risk. Hence a significant part of their research is supported through backtests, which are designed to reduce the model risk in risk models that are used for stress testing. They conduct backtests for eight risk models, including both conditional and unconditional models and four possible return distributions. Their backtesting experiment suggests that unconditional historical simulation, currently the most popular VaR methodology in the industry according to Perignon and Smith (2006), is likely to be mis-specified and is therefore unsuited for stress testing purposes.

Breuer et al (2009) define an operational definition to three requirements which the Basel Committee specifies for stress tests: plausibility and severity of stress scenarios as well as suggestiveness of risk-reducing actions. The basic idea of their approach is to define a suitable region of plausibility in terms of the risk-factor distribution and search systematically for the scenario with the worst portfolio loss over this region. One key innovation of their approach compared with the existing literature is the solution of two open problems. They suggest a measure of plausibility that is not dependent to the problem of dimensional dependence of maximum loss and they derive a way to consistently deal with situations where some but not all risk factors are stressed. They show that setting the non-stressed risk factors to their conditional expected value given the value of the stressed risk factors, the procedure first suggested by Kupiec (1998), maximises plausibility among the various approaches used in the literature. Furthermore, Breuer et al (2010b) propose a new method for analyzing multi-period stress scenarios for portfolio credit risk more systematically than in the current practice of macro stress testing. This method quantifies the plausibility of scenarios by considering the distance of the stress scenario from an average scenario. For a given level of plausibility their method searches systematically for the most adverse scenario for the given portfolio.

Finally, as a general point, it must be underlined that for the purposes of calculating the P&L impact of stress shock-factors it is generally assumed that the shock occurs instantaneously, ie that traders have no opportunity to re-hedge or adjust their positions, and it is ignored the impact of declining tenors for, for example, futures and options contracts. Apart from simplifying the calculations, such an assumption could be unreasonable in some cases given the practical experience of the actions of traders during historical events, and it may generate inconsistent results by amplifying the magnitude of the losses. Such issues have not yet been addressed in the literature.

4.3 Stressed VaR

The pressing technical issue now facing financial institutions that intend to comply with the amended Basel II framework is to understand how to calculate a valid stressed VaR number. After the revisions of July 2009, banks have to calculate a VaR using the risk engine it normally uses but “with model inputs calibrated to historical data from a continuous 12-month period of significant financial stress relevant to the bank’s portfolio” (Basel Committee on Banking Supervision (2009b)).

An over-simplistic interpretation of this specification might be to increase the assumed volatilities of the securities in a portfolio. This would have the effect of lengthening the tails of the Gaussian (normal) loss distributions that underlie all standard VaR calculations.

However, in order to calculate stressed VaR accurately it is also necessary to stress the correlation matrix used in all VaR methodologies. It is a repeated observation that during times of extreme volatility, such as occurs during every market crash, correlations are dramatically perturbed relative to their ‘normal’ historical values. In general, most correlations tend to increase during market crises, asymptotically approaching 1.0 during periods of complete meltdown, such as occurred in 1987, 1998 and 2008.

One possibility is to adopt the conditional stress test approach of Kupiec (1998). In this approach, the risk factor distributions are conditional on an extreme value realisation of one or more of the risk factors. Conditional on a large move of at least one factor, the conditional factor covariance matrix exhibits much higher correlations among the remaining factors. In this approach, the apparent shift in the correlation structure is a consequence of conditioning the distribution on a large factor shock. The unconditional correlations remain unchanged. Analysing a large number of stress test results for currency portfolios over the Asian currency crisis period, Kupiec shows that the conditional stress test process performs extremely well as very few stress test violations are recorded during this crisis period.

An alternative approach to conditional correlation is to stress the unconditional correlation matrix of the risk factors. Unfortunately, this approach is not as straightforward as the conditional correlation approach or stretching the tails of the loss distributions. The VaR calculation engine requires a correlation matrix that satisfies the mathematical property of positive definiteness, which is a way of saying that all of the correlations are internally consistent with each other. Noisy or erroneous historical price data can result in matrices that are not positive definite. Perturbing the correlation matrix, which is necessary for a true stressed VaR calculation, may result in correlation matrices that also violate the internal consistency requirement. If the matrix is not positive definite the VaR calculus will fail, so methods have to be devised to modify the stressed matrix until it becomes positive definite. Kupiec (1998) discusses some practical methods that can be used to address this problem.

Besides these technical issues one may also more fundamentally consider concepts that are not covered by the current regulatory definition of stressed VaR. A more sophisticated approach might include not only linear transforms of multivariate normal risk factors but also employing ‘fat-tailed’ distributions to model the extreme loss events more accurately. Examples of those ‘extreme value theory’ distributions are the Gumbel, Generalised Pareto, Weibull, Fréchet, and the Tukey g&h distributions.

However, one should keep in mind that the stressed VaR is from a theoretical perspective an imperfect solution – its purpose is to reflect that current market conditions may not lead to an accurate assessment of the risk in a more stressful environment. Extreme value theory distributions may already incorporate extreme market conditions and could in principle make a stressed VaR redundant. In general, these distributions are flexible enough to obtain very good fits but serious robustness issues arise instead, as regulators and risk managers had to learn in the context of operational risk, for instance.

4.4 Conclusions

More recent research advocates the integration of stress testing into the risk modelling framework. This would overcome drawbacks of reconciling stand-alone stress test results with standard VaR model output.

Progress has also been achieved in theoretical research on the selection of stress scenarios. In one approach, for example, the "optimal" scenario is defined by the maximum loss event in a certain region of plausibility of the risk factor distribution.

The regulatory "stressed VaR" approach is still too recent to have been analyzed in the academic literature. Certain methods that could be meaningful in this context can be identified in the earlier literature on stress testing. Employing fat-tailed distributions for the risk factors and replacing the standard correlation matrix with a stressed one are two examples.

5. Unified versus compartmentalised risk measurement

5.1 Overview

In this section, we survey the academic literature on the implications of modelling the aggregate risks present across a bank's trading and banking books using either a compartmentalised approach – namely, the sum of risks measured separately – or a unified approach that considers the interaction between these risks explicitly. Finally, we survey the recent literature on the systemic implications of the current regulatory capital requirements that aggregate capital requirements across risk types.

In many financial institutions, aggregate economic capital needs are calculated using a two step procedure. First, capital is calculated for individual risk types, most prominently for credit, market and operational risk. In a second step, the stand-alone economic capital requirements are added up to obtain the overall capital requirement for the bank.

The Basel framework for regulatory capital uses a similar idea. As discussed by Cuenot, Masschelein, Pritsker, Schuermann and Siddique (2006), the Basel framework is based on a "building block" approach such that a bank's regulatory capital requirement is the sum of the capital requirements for each of the defined risk categories (ie market, credit and operational risk), which are calculated separately within the formulas and rules that make up Pillar 1. Capital requirements for other risk categories are determined by the supervisory process that fits within Pillar 2; see Figure 2 which is reproduced from Cuenot et al (2006). This approach is therefore often referred to as a non-integrated approach to risk measurement. An integrated approach would, by contrast, calculate capital for all the risks borne by a bank simultaneously in one single step and accounting for possible correlations and interactions, as opposed to adding up compartmentalised risk calculations.

Pressure to reconsider the regulatory compartmentalised approach came mainly from the financial industry, where it has been frequently argued that a procedure that simply adds up economic capital estimates across portfolios ignores diversification benefits. These alleged benefits have been estimated to be between 10 and 30% for banks (see Brockmann and Kalkbrener (2010)).

Capital diversification arguments and estimates of potential capital savings are partially supported in the academic literature. More recently this view and the estimates have been fundamentally challenged by the Basel Committee (Basel Committee on Banking Supervision (2009)) and by Breuer et al (2010a). These papers have pointed out that non-linear interaction between risk categories may even lead to compounding effects. This fact questions whether the compartmentalised approach will in general give a conservative and prudent upper bound for economic capital.

Is this a merely academic debate or does it have practical implications for reform considerations related to the trading book? In this section, we survey the main arguments and give a brief review of the main papers and their findings. We then discuss policy implications that might be relevant for a discussion related to potential future reform related to the trading book.

Figure 2

Overview of risk categories relevant for banking book and trading book and their consideration in Pillar 1 and Pillar 2

<table>
<tr><th></th><th>Banking book</th><th>Trading book</th></tr>
<tr><td rowspan="7">Pillar 1</td><td>Credit risk</td><td></td></tr>
<tr><td colspan="2">Counterparty credit risk</td></tr>
<tr><td></td><td>Interest rate risk
(general and specific)</td></tr>
<tr><td></td><td>Equity risk
(general and specific)</td></tr>
<tr><td colspan="2">Foreign exchange risk</td></tr>
<tr><td colspan="2">Commodity risk</td></tr>
<tr><td colspan="2">Operational risk</td></tr>
<tr><td rowspan="4">Pillar 2</td><td>Interest rate risk</td><td></td></tr>
<tr><td colspan="2">Concentration risk</td></tr>
<tr><td colspan="2">Stress tests</td></tr>
<tr><td colspan="2">Other risks (liquidity, residual, business…)</td></tr>
</table>

Source: Cuenot et al (2006).

5.2 Aggregation of risk: diversification versus compounding effects

Diversification is a term from portfolio theory referring to the mix of a variety of investments within a portfolio. Since different investments will develop differently in the future with value losses in some investment offset by value gains in another investment, the overall portfolio risk is reduced through the spreading of risk. In a similar way, the assets of a bank can be thought of as an overall portfolio that can be divided into subportfolios. If risk analysis is done by looking at risk measures at the level of the subportfolios and the risk measures are added up, the intuition of diversification suggests that we should arrive at a conservative risk measure for the bank as a whole.

So, what is wrong with this straightforward intuition about diversification between market, credit and other risk categories? The flaw in the intuition lies in the fact that it is usually not possible to divide the overall portfolio of a bank into subportfolios purely consisting of market, credit and operational risk; these risk categories are too intertwined in a modern financial institution to possibly separate in a meaningful way. In short, we cannot construct a subportfolio of risk factors. It is therefore incorrect to think of the banking book as a subportfolio of the overall bank portfolio for which only credit risk is relevant. It is also incorrect to view the trading book as another subportfolio related solely to market risk.

A simple way to summarise this argument is to consider a portfolio of loans. The interest rate risk related to such a portfolio is usually counted as a market risk, and this risk affects the bank's refinancing costs and the revaluation of these loans. If the interest rate risk is borne by the creditors in some way, this market risk suddenly may transform into a credit risk for the bank. So, do assets with a value that fluctuates with interest rates belong in a subportfolio for market risk or in a subportfolio of credit risk? They clearly belong to both, because each loan has a market risk component as well as a credit risk component simultaneously. Trading book positions with counterparty risks or positions related to carry trades fall into the same category.

Breuer et al (2010a) consider portfolios of foreign currency loans, which are loans denominated in a foreign currency extended to domestic creditors with income in domestic currency. The credit risk in these portfolios is always a function of the market risk (ie exchange rate movements), and the risk of each position in a foreign currency loan portfolio has simultaneously a credit and a market risk component. Adding up capital and hoping for an upper bound amounts to ignoring possible "malign risk interactions" as they are called in Breuer et al (2010a). This issue has been known in the market risk literature for a long time as "wrong way risk". Wrong way risk is the risk arising from the problem that the value of a trading position is inversely correlated with the default risk of some counterparty.

From these examples, we see that a formation of subportfolios along the lines of risk factors – and for that matter across banking and trading books – is usually not possible. Breuer et al (2010a) indeed show that the ability to form subportfolios along the lines of risk categories is a sufficient condition for diversification effects to occur. Since we can in general not form such subportfolios, we must anticipate the possibility that there can be risk compounding effects between the banking and the trading book. In short, while the intuition of diversification is inviting, it does not apply to the interaction of banking and trading books since there are in general no pure subportfolios of market, credit or operational risks.

This insight is important because it demonstrates that "diversification effects" that are derived from papers using a so-called "top-down" approach are often assuming what they want to derive. By construction, the assumption of splitting up the bank portfolio into subportfolios according to market, credit and operational risk assumes that this can indeed be done. If such a split were possible, it follows from the results in Breuer et al (2010a) that diversification effects must occur necessarily.

To estimate the quantitative dimension of the problem, we therefore must focus on papers working with a "bottom-up" approach. We also need to examine the results of papers based on the "top-down" approach that assumes risk separability at the beginning of the analysis. In this section, we survey several key papers that use either of these risk aggregation methods. As part of this literature survey, we provide a summary of recent papers that estimate the range and magnitude of these differences between compartmentalised and unified risk measures. Our proposed measure is a simple ratio of these two measures, as used in other papers, such as Breuer et al (2010a). In that paper, the authors adopt the term "inter-risk diversification index" for the ratio; see also the related measure in Alessandri and Drehmann (2010). Ratio values greater than one indicate risk compounding, and values less than one indicate risk diversification. In the summary tables below, we list the various papers, the portfolio analysed, the risk measures used, the horizon over which the risks are measured, and these risk ratios.

5.3 Papers using the "bottom-up" approach

As mentioned above, a common assumption of most current risk measurement models is that market and credit risks are separable and can be addressed independently. Yet, as noted as early as Jarrow and Turnbull (2000), economic theory clearly does not support this simplifying assumption.

While the reasons behind this common assumption are mostly operational in nature, some studies have used numerical simulation techniques to generate results. For example, Barnhill and Maxwell (2002) examine the economic value of a portfolio of risky fixed income securities, which they define as a function of changes in the risk-free interest rate, bond spreads, exchange rates, and the credit quality of the bond issuers. They develop a numerical simulation methodology for assessing the VaR of such a portfolio when all of these risks are correlated. Barnhill et al (2000) use this methodology to examine capital ratios for a representative South African bank. However, in these studies, the authors do not examine the differing values of their chosen risk measures using a unified risk measurement approach versus a compartmentalised approach that sums the independent risk measures.

The study by Jobst, Mitra and Zenios (2006) provides some analysis along these lines. The authors construct a simulation model, based on Jobst and Zenios (2001), in which the risk underlying the future value of a bond portfolio is decomposed into:

- the risk of a borrower's rating change (including default);
- the risk that credit spreads will change; and
- the risk that risk-free interest rates will change.

Note that the first item is more narrowly defined to represent the portfolio's credit risk, while the last item is more narrowly defined to represent the portfolio's market risk. However, the middle item is sensitive to both risks and challenges the notion that market and credit risk can be readily separated in this analysis. The authors use portfolios of US corporate bonds and one-year VaR and CVaR risk measures at the 95%, 99% and 99.9% confidence levels for their analysis.

In their analysis, the authors generate risk measures under three sets of assumptions. To concentrate on the pure credit risk contributions to portfolio losses, they simulate only rating migration and default events as well as recovery rates, while assuming that future interest rates and credit spreads are deterministic. The authors then allow future credit spreads to be stochastically determined, and finally, they allow future interest rates to be stochastically determined. Note that the latter case provides an integrated or unified risk measurement, according to our definition for this survey.[14]

The authors' results are quite strong regarding the magnitude of the risk measures across risk types and credit ratings. For Aaa-rated bonds, the authors find that the unified risk measures at all three tail percentiles are on the order of ten times the pure credit risk measures, since highly-rated bonds are unlikely to default. As the credit quality of the portfolio declines, the ratio between the unified risk measures and the risk measures for pure credit risk drops to just above one for C-rated bonds.

Table 1 presents a short summary of several papers for which we can directly examine the ratio of unified to compartmentalised risk measures for bottom-up models. As mentioned earlier, the recent work of Breuer et al (2008, 2010a) provides a leading example of how market and credit risk cannot be readily separated in a portfolio, a fact that complicates risk

[14] Note, however, that the authors do not conduct an analysis of a market risk scenario (ie deterministic ratings and stochastic credit spreads and interest rates). Thus, we cannot examine their ratio of unified to compartmentalised risk measures as discussed above.

measurement and works to undermine the simple assumptions underling additive risk measures.

In Breuer et al (2010a), the authors present analysis of hypothetical loan portfolios for which the impact of market and credit risk fluctuations are not linearly separable. They argue that changes in aggregate portfolio value caused by market and credit risk fluctuations in isolation should sum up to the integrated change incorporating all risk interactions very rarely. The magnitude and direction of the discrepancy between these two types of risk assessments can vary broadly. For example, the authors examine a portfolio of foreign currency loans for which exchange rate fluctuations (ie market risk) affect the size of loan payments and hence the ability of the borrowers to repay the loan (ie credit risk). For their empirically calibrated example, they use expected shortfall at various tail percentiles as their risk measure and examine portfolios of BBB+ and B+ rated loans. Their analysis shows that changes in market and credit risks can cause compounding losses such that the sum of value changes from the individual risk factors are smaller than the value change due to accounting for integrated risk factors.

In particular, their reported inter-risk diversification index for expected shortfall increased sharply as the tail quantile decreased, which suggests that the sum of the two separate risk measures becomes much less useful as an approximation of the total integrated risk in the portfolio as we go further into the tail. These index values also increase for all but the most extreme tail percentiles as the original loan rating is lowered. The authors argue that this example presents evidence of a “malign interaction of market and credit risk which cannot be captured by providing separately for market risk and credit risk capital”. The authors show a similar qualitative outcome for domestic currency loans (ie loans for which default probability are simply a function of interest rates), although the index values are much lower.

In Breuer et al (2008), the authors use a similar analytical framework to examine variable rate loans in which the interaction between market and credit risk can be analysed. In particular, they model the dependence of credit risk factors – such as the loans’ default probabilities (PD), exposure at default (EAD), and loss-given-default (LGD) – on the interest rate environment. A key risk of variable rate loans is the danger of increased defaults triggered by adverse rate moves. For these loans, market and credit risk factors cannot be readily separated, and their individual risk measures cannot be readily aggregated back to a unified risk measure. They conduct a simulation study based on portfolios of 100 loans of equal size by borrowers rated B+ or BBB+ over a one-year horizon using the expected shortfall measure at various tail percentiles. They find that the ratio of unified expected shortfall to the sum of the separate expected shortfalls is slightly greater than one, suggesting that risk compounding effects can occur. Furthermore, these compounding effects are more pronounced for lower-rated loans and higher loan-to-value ratios.

In contrast to this work, the paper by Grundke (2005) lays out a bottom-up model that assumes the separability of interest rate risk (ie market risk) and credit spread risk (ie credit risk). The author examines a calibrated multi-factor credit risk model that accommodates various asset value correlations, correlations between credit spreads and other model factors, and distributional assumptions for innovations. The author examines hypothetical loan portfolios of varying credit quality over a three-year horizon, both with and without the joint modelling of interest rates and credit spreads. To assess the joint impact of interest rate and credit risk, the author uses forward market interest rates instead of separate interest rate and credit spread processes. Interestingly, the reported VaR measures at various tail percentiles lead to ratios of unified VaR measures to summed VaR measures that range widely from near zero to one, which seems to be due mainly to the separability of the interest rate risk (ie market risk) and credit spread risk (ie credit risk) in the model.

Kupiec (2007) proposes a single-factor, migration-style credit risk model that accounts for market risk. This modelling approach generates a portfolio loss distribution that accounts for the non-diversifiable elements of the interactions between market and credit risks. The integrated exposure distribution of the model is used to examine capital allocations at various thresholds. These integrated capital allocations are compared to the separated assessments. The results show that capital allocations derived from a unified risk measure importantly alter the estimates of the minimum capital needed to achieve a given target solvency margin. The capital amount could be larger or smaller than capital allocations estimated from compartmentalised risk measures. Regarding specifically the Basel II AIRB approach, the author argues that the results show that no further diversification benefit is needed for banking book positions since no market risk capital is required. Thus, Basel II AIRB capital requirements fall significantly short of the capital required by a unified risk measure.

Numerically speaking, the risk measure used in this study is the amount of capital that the unified and the compartmentalised capital approaches generate as the appropriate value to assure funding costs of a certain magnitude calibrated to historical funding rates for specific credit ratings. The hypothetical portfolios of interest are corporate loans with various rating categories represented in proportion to historical data. The author examines a wide variety of alternative separated approaches with which to calculate economic capital measures, ranging from three different alternative credit risk models to several methods for measuring market risk. Correspondingly, the range of inter-risk diversification index values is quite wide for the AAA- and BBB-rated portfolios, ranging from about 0.60 to almost 4.00. In summary, the author's capital calculations show that capital allocations derived from a unified market and credit risk measure can be larger or smaller than capital allocations that are estimated from aggregated compartmentalised risk measures.

The studies discussed above examine the different risk implications of a unified risk measurement approach relative to a compartmentalised approach for specific portfolios. In contrast, Drehmann et al (2010) examine a hypothetical bank calibrated to be representative of the UK banking system as a whole. Within their analytical framework, they do not explicitly assume that market and credit risk are separable. The authors decompose the total risk in their bank scenario analysis into:

- the impact of credit risk from non-interest rate factors,
- the impact of interest rate risk (excluding the effect of changes in interest rates on credit risk), and
- the impact of the interaction of credit risk and interest rate risk.

The latter is calculated as the difference between the total impact of the scenario shock and the sum of the first two components.

Their simulations confirm that interest rate risk and credit risk must be assessed jointly for the whole portfolio to gauge overall risk correctly. In particular, the authors find in their simulations that if banks gauged credit risk by solely monitoring their write-offs, aggregate risk would be underestimated in the short term since a rate increase would also lower its net interest income and profits. Correspondingly, the bank's aggregate risk would be overestimated in the long run as net interest income and profits recover while write-offs continue to rise.

Their main variable of interest is net profits over twelve quarters after their macroeconomic stress scenario hits their representative bank, although they also report separate measures of write-offs and net interest income. They report that the interaction between interest rate and credit risk accounts for about 60% of the decline in capital adequacy for their calibrated bank. While the decline in capital adequacy does not perfectly match our other risk measures, we can still think of the diversification index here as the ratio of the capital decline for the unified risk framework relative to the capital decline that would come from separate identification of market and credit risks. Given their reported numbers, that ratio here is 100% / (100% – 60%) = 2.5, which suggests a very clear contribution of this interaction to risk management concerns.

Following up on the work of Drehmann et al (2010), Alessandri and Drehmann (2010) develop an integrated economic capital model that jointly accounts for credit and interest rate risk in the banking book; ie where all exposures are held to maturity. Note that they explicitly examine repricing mismatches (and thus market and credit risks) that typically arise between a bank's assets and liabilities.

For a hypothetical, average UK bank with exposures to only the UK and US, they find that the difference between aggregated and unified economic capital levels is often significant but depends on various bank features, such as the granularity of assets, the funding structure or bank pricing behaviour. They derive capital for the banking book over a one year horizon. For credit and interest rate risk, they define unexpected losses and thus economic capital as the difference between VaR at the specified 99% confidence level and expected losses. Note that their measures of economic capital for just credit risk and just interest rate risk do not fully disentangle these risks as the credit risk measure incorporates the effects of higher interest rates on default probabilities and the latter the effect of higher credit risk on income. The key point is that the framework represents a plausible description of how current capital models for the banking book capture these risks.

The authors examine the ratio of unified economic capital to the sum of the component measures at three VaR quantiles. For the 95% percentile of portfolio losses, unified capital measure is near zero, and thus the ratio is nearly zero as well. For the 99% percentile, the ratio is quire small at 0.03, but the ratio rises quickly to just over 50% for the 99.9% percentile. Note, however, that this result still suggests that the compartmentalised approach is more conservative than the unified approach. The authors examine certain modifications of their assumptions – such as infinitely fine-grained portfolios to increase the correlation of portfolio credit risk with the macroeconomic factors, banking funding scenarios from all short-term debt that is frequently repriced to all long-term debt that is repriced only on a yearly basis – and find some interesting difference with the base case scenario. However, the lower integrated capital charge holds.

On balance, these authors conclude that the bank's capital is mis-measured if risk interdependencies are ignored. In particular, the addition of economic capital for interest rate and credit risk derived separately provides an upper bound relative to the integrated capital level. Two key factors determine this outcome. First, the credit risk in this bank is largely idiosyncratic and thus less dependent on the macroeconomic environment; and second, bank assets that are frequently repriced lead to a reduction in bank risk. Given that these conditions may be viewed as special cases, the authors recommend that "As a consequence, risk managers and regulators should work on the presumption that interactions between risk types may be such that the overall level of capital is higher than the sum of capital derived from risks independently."

5.4 Papers using the "top-down" approach

An alternative method for determining total firm risk, primarily for enterprise-wide risk management, is to aggregate risks calculated for different business lines or different risk types using so-called "top-down" approaches. An important difference is that top-down approaches always reference an institution as a whole, whereas bottom-up approaches can range from the portfolio level up to an institutional level. With respect to market and credit risk, the top-down approach explicitly assumes that the risks are separable and can be aggregated in some way. As outlined by Cuenot et al (2006), firms may compute their market and credit risk capital separately and aggregate the two risk types by imposing some form of correlation between them. The top-down approach thus does not require a common scenario across risk types, but because the correct form of aggregation is not known, the approach "loses the advantages of logical coherence". In addition, as suggested by Breuer et al (2008, 2010a), the assumption of separable risk will generally prevent the ability to gauge the degree of risk compounding that might be present and instead typically provide support for risk diversification.

The literature is unclear on whether the combination of financial business lines within one organisation leads to an increase or decrease in risk. The literature as surveyed by Saunders and Walters (1994) and Stiroh (2004) suggests mixed results. However, as surveyed by Kuritzkes, Schuermann and Weiner (2003), several studies, including their own, suggest that reductions in economic capital arise from the combination of banking and insurance firms. The papers surveyed in Table 2 and below find this result as well for various risk combinations at the firm level.

For example, Dimakos and Aas (2004) decompose the joint risk distribution for a Norwegian bank with an insurance subsidiary into a set of conditional probabilities and impose sufficient conditional independence that only pair-wise dependence remains; the total risk is then just the sum of the conditional marginals (plus the unconditional credit risk, which serves as their anchor). Their simulations indicate that total risk measured using near tails (95%–99%) is about 10%–12% less than the sum of the individual risks. In terms of our proposed ratio, the value ranges from 0.88 to 0.90. Using the far tail (99.97%), they find that total risk is often overestimated by more than 20% using the additive method. In terms of our proposed ratio of unified risk measure to the sum of the compartmentalised risk measures, its value would be 0.80.

Similarly, Kuritzkes et al (2003) examine the unified risk profile of a "typical banking-insurance conglomerate" using the simplifying assumption of joint normality across the risk types, which allows for a closed-form solution. They use a broad set of parameters to arrive at a range of risk aggregation and diversification results for a financial conglomerate. Based on survey data for Dutch banks on the correlations between losses within specific risk categories, their calculations of economic capital at the 99.9% level is lower for the unified, firm-level calculation than for the sum of the risk-specific, compartmentalised calculations. The ratio of these two quantities ranges from 0.72 through 0.85 based on correlation assumptions across market, credit and operational risk.

Rosenberg and Schuermann (2006) conduct a more detailed, top-down analysis of a representative large, internationally active bank that uses copulas to construct the joint distribution of losses. The copula technique combines the marginal loss distributions for different business lines or risk types into a joint distribution for all risk types and takes account of the interactions across risk types based on assumptions. Using a copula, parametric or nonparametric marginals with different tail shapes can be combined into a joint risk distribution that can span a range of dependence types beyond correlation, such as tail dependence. The aggregation of market, credit and operational risk requires knowledge of the marginal distributions of the risk components as well as their relative weights. Rosenberg and Schuermann assign inter-risk correlations and specify a copula, such as the Student-t copula, which captures tail dependence as a function of the degrees of freedom. They impose correlations of 50% for market and credit risk, and 20% for the other two correlations with operational risk; all based on triangulation with existing studies and surveys.[15]

Rosenberg and Schuermann find several interesting results, such as that changing the inter-risk correlation between market and credit risk has a relatively small impact on total risk compared to changes in the correlation of operational risk with the other risk types. The authors examine the sensitivity of their risk estimates to business mix, dependence structure, risk measure, and estimation method. Overall, they find that "assumptions about operational exposures and correlations are much more important for accurate risk estimates than assumptions about relative market and credit exposures or correlations". Comparing their VaR measures for the 0.1% tail to the sum of the three different VaR measures for the three risk types, they find diversification benefits in all cases. For our benchmark measure of the ratio between the unified risk measure and the compartmentalised risk measure, their results suggest values ranging from 0.42 to 0.89. They found similar results when the expected shortfall (ES) measure was used.

Note that the authors state that the sum of the separate risk measures is always the most conservative and overestimates risk, "since it fixes the correlation matrix at unity, when in fact the empirical correlations are much lower". While the statement of imposing unit correlation is mathematically correct, it is based on the assumption that the risk categories can be linearly separated. If that assumption were not correct, as suggested by papers cited above, the linear correlations could actually be greater than one and lead to risk compounding.

Finally, Kuritzkes and Schuermann (2007) examine the distribution of earnings volatility for US bank holding companies with at least $1 billion in assets over the period from 1986.Q2 to 2005.Q1; specially, they examine the 99.9% tail of this distribution. Using a decomposition methodology based on the definition of net income, the authors find that market risk accounts for just 5% of total risk at the 99.9% level, while operational risk accounts for 12% of total risk. Using their risk measure of the lower tail of the earnings distribution, as measured by the return on risk-weighted assets, their calculations suggest that the ratio of the integrated risk measure to the sum of the disaggregated risk measures ranges from 0.53 through 0.63.

5.5 Conclusions

Academic studies have generally found that at a high-level of aggregation, such as at the holding company level, the ratio of the risk measures for the unified approach to that of the separated approach is often less than one; ie risk diversification is prevalent and ignored by the separated approach. However, this approach often assumes that diversification is present. At a lower level of aggregation, such as at the portfolio level, this ratio is also often found to be less than one, but important examples arise in which risk compounding (ie a ratio greater than one) is found. These results suggest, at a minimum, that the assumption of risk diversification cannot be applied without questioning, especially for portfolios subject to both market and credit risk, regardless of where they reside on the balance sheet.

Recent literature on the systemic implications of the current regulatory capital requirements that aggregate capital requirements across risk types suggests that this compartmentalised approach can – at least in general – be argued to contribute to the amplification of systemic risk, which is counter to its intentions.

---

[15] Note that different correlation values could lead to risk compounding, but it is not clear what those values might be and what values would be implied by the bottom-up exercises discussed here.

In terms of policy implications, the academic literature suggests that if we are able to divide risk types easily across the trading book and the banking book (as is assumed in the top-down studies), diversification benefits appear to be certain, and aggregation of capital requirements across the books is conservative. However, recent studies have shown that if this risk separation cannot be done completely, simple aggregation of compartmentalised measures may not be conservative and, in fact, may understate the total risk. Such an outcome would clearly be undesirable as the necessary amount of capital could be underestimated by a significant margin.

These conclusions seem to directly question whether separate capital requirements for the trading and banking books provide a reasonable path to setting the appropriate level of capital for the entire firm. If we retained the different capital treatments, attempts could be made to fully detail each type of risk within each book, and the subsequent aggregation might then be considered conservative. However, performing such an analysis within the current and traditional separation between a trading and a banking book would require important changes in operational procedures. An alternative approach might be to develop a system of book keeping and risk allocation that does not artificially assign positions into different books when its risk characteristics are interrelated.

6. Risk management and value-at-risk in a systemic context

6.1 Overview

In this section, we survey the research literature on the systemic consequences of individual risk management systems and regulatory capital charges that rely on them. At the time when the Basel Committee implemented the MRA in 1996, risk management and banking regulation still was a subject that had received relatively little attention in the academic literature. Perhaps the most important change brought to the Basel framework by the MRA was the ability for banks to use their own quantitative risk models for determining the capital requirements for market risk.

Both conceptually and procedurally, this amendment was a significant departure from the previous regulatory approaches to determine bank capital. The conceptual innovation was that the notion of risk on which the new regulation relied was much closer to the notions of risk that were in use in the financial, economic and statistical research literature. Procedurally the amendment amounted to an official recognition that financial institutions themselves are in the best positions to assess their risk exposures. The new regulatory approach seemed to suggest that using and relying on this knowledge might be the best way to cope with methodological problems of risk assessment in a rapidly changing economic environment.

At the time of the amendment and in the years after, the academic literature on risk management and regulation largely accepted the conceptual reasoning behind the amendment and confined itself mostly to developing the technology of quantitative risk management itself. The discussion in the economics community remained sparse and largely sceptical.

Hellwig (1995, 1996) raised several important issues related to this new regulatory approach that did not take hold very much in the regulatory community but sound very modern in the current debate about the recent financial crises: Hellwig discussed incentive problems. Banks may find it desirable to bias their model development towards the goal of minimising capital. With hindsight, we know that the practice of determining capital based on VaR models helped large and international active banks to reduce greatly the amount of capital to be held against any given asset during the pre-crisis boom years. He also pointed out the difficulties related to using statistical techniques which work under the assumption of a stationary world in a non-stationary environment like financial markets. He also criticised the separation between market and credit risk while he acknowledged that quantitative models of integrated risk measurement are subject to the general problems outlined above.

During the discussion of the new Basel II framework, in May 2001, a group of academics at the Financial Markets Group (FMG) of the London School of Economics wrote a paper that raised a concern with respect to the use of value-at-risk that is more fundamental.[16] In the report's executive summary, there is a conclusion that calls into question the conceptual construction of the 1996 amendment: "The proposed regulations fail to consider the fact that risk is endogenous. Value-at-risk can destabilise and induce crashes when they would not otherwise occur."

In the current practice of risk management and regulation, these conclusions so far have only partly lead to a serious reconsideration of the framework initiated and extended more than a decade ago. In the current regulatory discussion, the general view seems to be that the conclusions from the financial crisis call for suitable expansions and amendments to the prevailing framework. In the meantime, the conclusions derived in the FMG paper have received more substantive underpinnings from academic research, both empirically and theoretically. The papers of Adrian and Shin (2008), the book of Shin (2008a) and joint work by Danielsson, Shin and Zigrand (2009) suggest that the use of value-at-risk models in regulation intended to function as a "fire extinguisher", function in practice rather like a "fire accelerant".[17] Rather than suggesting improving the VaR-based capital regulations by various refinements and amendments to the concepts in place, this literature suggests to abandon this approach and remove a VaR-based capital requirement from the regulatory framework. It should not be ignored, however, that some of the new regulatory initiatives will likely dampen procyclical effects in the future. The *stressed VaR* introduced by the July 2009 revisions of the Market Risk Framework is a case in point: its calculation is based on estimates from bad historical periods of the economy and so acts rather "through the cycle". Admittedly, the stressed VaR is only one addend of total trading book capital.

Although the literature for this section generally refers to VaR as the risk measure at issue, it is important to bear in mind that the term VaR should be interpreted here in a wide sense since the results generally do not depend on this specific risk measure.

In the following we give a brief outline of the main arguments and explain the boom and bust amplification mechanism identified in this literature. We then go through some of the policy conclusions suggested by this analysis.

6.2 Intermediation, leverage and value-at-risk: empirical evidence

Adrian and Shin (2010) empirically investigated the relationship between leverage and balance sheet size of the five major US investment banks shortly before the financial crises. All these institutions meanwhile left the broker-dealer sector, either because they were taken over or went bankrupt or were converted to bank holding companies. A major reason why these institutions are particularly interesting is because they all show a very clear picture of how financial intermediation works in a capital markets-based financial system with active balance sheet management through risk management systems.

[16] See Danielsson et al (2001).

[17] See Hellwig (2009).

When an intermediary actively manages its balance sheet, leverage becomes procyclical because risk models and economic capital require balance sheet adjustments as a response to changes in financial market prices and measured risks. This relationship follows from simple balance sheet mechanics. The following example is taken from Shin (2008a, pp.24 ff.) Assume a balance sheet is given with 100 in assets and a liability side which consists of 90 in debt claims and 10 in equity shares. Leverage is defined as the ratio of total assets to equity, 10 in our example. If we assume more generally that the market value of assets is A and make the simplifying assumption that the value of debt stays roughly constant at 90 for small changes in A we see that total leverage is given by

$$L = \frac{A}{A-90}$$

Leverage is thus related inversely to the market value of total assets. When net worth increases, because A is rising, leverage goes down, when net worth decreases, because A is falling, leverage increases.

Consider now what happens if an intermediary actively manages its balance sheet to maintain a constant leverage of 10. If asset prices rise by 1 %, the bank can take on an additional amount of 9 in debt, its assets have grown to 110, its equity is 11 and the debt is 99. If asset values shrink by 1%, leverage rises. The bank can adjust its leverage by selling securities worth 9 and pay down a value 9 of debt to bring the balance sheet back to the targeted leverage ratio.

This kind of behaviour leads to a destabilising feedback loop, because it induces an *increase* in asset purchases as asset prices are rising and a *sale of assets* when prices are falling. Whereas the textbook market mechanism is self stabilising because the reaction to a price increase is a reduction in quantity demanded and an expansion in quantity supplied, and to a price decreases an expansion in quantity demanded and a contraction in quantity supplied, active balance sheet management reverses this self stabilising mechanism into a destabilising positive feedback loop.

Adrian and Shin (2010) document this positive relationship between total assets and leverage for all of the (former) big Wall Street investment banks. Furthermore they produce econometric evidence that the balance sheet adjustments brought about by active risk management of financial institutions indeed has an impact on risk premiums and aggregate volatility in financial markets.

6.3 What has all this to do with VaR-based regulation?

Why would a bank target a constant leverage and what is the role of value-at-risk in all of this? The book of Shin (2008a) and the papers by Shin (2008b) and Adrian and Shin (2008) as well as by Danielsson, Shin and Zigrand (2009) explore this role in more detail.

If we consider the future value of bank assets A as a random variable, the value-at-risk (VaR) at a confidence level c is defined by

$$\Pr\left(A < A_0 - VaR\right) \leq 1 - c$$

The VaR is equal to the equity capital the firm must hold to be solvent with probability c. The economic capital is tied to the overall value-at-risk.

If a bank adjusts its balance sheet to target a ratio of value-at-risk to economic capital then bank capital to meet VaR is

$$K = \lambda \times VaR,$$

where λ is the proportion of capital to be held per total value-at-risk. This proportion may vary with time. Leverage is thus

$$L = \frac{A}{K} = \frac{1}{\lambda} \times \frac{A}{VaR}$$

Since VaR per value of assets is countercyclical, it directly follows that leverage is procyclical as the data in Adrian and Shin (2008) indeed show. [18]

The systemic consequences of this built-in risk limiting technology at the level of individual institutions works in the aggregate as an amplifier of financial boom and bust cycles. The mechanism by which the systemic amplification works is risk perception and the pricing of risk, even if all network effects and complex interconnectedness patterns in the financial system are absent.[19]

Consider intermediaries who run a VaR-based risk management system and start with a balance sheet consisting of risk free debt and equity. Now an asset boom takes place, leading to an expansion in the values of securities. Since debt was risk free to begin with, without any balance sheet adjustment, this leads to a pure expansion in equity. The VaR constraint is relaxed through the asset boom and creates new balance sheet capacity to take on more risky securities or increase its debt. The boom gets amplified by the portfolio decisions of the leveraged banking system.

Put differently, in a system of investors driven by a VaR constraint investors' demand follows and amplifies the most recent price changes in the financial market. Price increases and balance sheet effects become intertwined through the active VaR-driven risk management of financial institutions.

Of course the described mechanism also works on the way down. A negative shock drives down market values tightening the VaR constraints of leveraged investors. These investors have to sell assets to reduce leverage to the new VaR constraint. By hardwiring VaR-driven capital management in banking regulation a positive feedback loop with potent destabilising force both in booms and busts has been built into the financial system.

The mechanisms described in this section have been theoretically analysed in Shin (2008a), Danielsson, Shin, Zigrand (2009) theoretically and with explicit reference to value-at-risk. They are also central in the work of Geanakoplos (2009) although there the connection with VaR is not made explicit.

6.4 Conclusions

A literature stream on the systemic consequences of individual risk management systems as the basis of regulatory capital charges has found that the mechanical link between measured risks derived from risk models and historical data and regulatory capital charges can work as a systemic amplifier of boom and bust cycles.

---

[18] This formal derivation of the procyclicality of VaR is taken directly from Shin (2008a).

[19] For this point see also Geanakoplos (2009) who has shown in a theoretical model how risk-free debt may nevertheless give rise to fluctuations in leverage and risk pricing and thus create systemic spillover effects.

The central mechanism that leads to this feedback loop works through the pricing of risk. In good times, when measured risks look benign, a financial institution that targets a regulatory capital requirement as a function of a model based risk measure has slack capacity in its balance sheet that it can either use to buy additional risky assets or to increase its debt. This means that we have a mechanism where institutions are buying more risky assets when the price of these assets is rising and where they are buying less of these assets when prices are falling. The stabilising properties of the market mechanism are turned on their head. By this mechanic link of measured risk to regulatory capital a powerful amplifier of booms and busts is created at the system level counteracting the intention of the regulation to make the system as a whole safer.

It is important to recognise that while the current system may implement a set of rules that limit the risk taken at the level of individual institutions, the system may also enable institutions to take on more risk when times are good and thereby lay the foundations for a subsequent crisis. The very actions that are intended to make the system safer may have the potential to generate systemic risk in the system.

These results question a regulatory approach that accepts industry risk models as an input to determine regulatory capital charges. This critique applies in particular to the use of VaR to determine regulatory capital for the trading book but it questions also an overall trend in recent regulation.

The amplifying mechanism identified in this section will be at work no matter how sophisticated VaR becomes, whether it is replaced by more sophisticated risk measures, like expected shortfall, or whether it goes beyond the naïve categorisation of risk classes (market, credit and operational) towards a more integrated risk measurement. These changes generally do not address the problems raised by the papers reviewed in this section. One exception is the *stressed VaR* introduced in July 2009. This new component of trading book capital acts more “through the cycle” than the “normal” VaR. Still some argue that what may be needed is a less mechanical approach to capital adequacy that takes into account a system-wide perspective on endogenous risk. The academic literature has identified many potential shortcomings in the currently regulatory approach for bank capital but it has yet to develop an alternative approach that simultaneously satisfies all the (sometimes conflicting) regulatory policy objectives.

References

Acerbi, C (2002): "Spectral measures of risk: a coherent representation of subjective risk aversion", *Journal of Banking and Finance*, vol 26, no 7, pp 1505–1518.

Acerbi, C and G Scandolo (2008): "Liquidity risk theory and coherent measures of risk", *Quantitative Finance*, vol 8, no 7, pp 681-692.

Acerbi, C and D Tasche (2002): "On the coherence of expected shortfall", *Journal of Banking and Finance*, vol 26, pp 1487–1503.

Adrian, T and H S Shin (2010): "Liquidity and leverage", *Journal of Financial Intermediation*, vol 19, no 3, pp 418–437.

——— (2008), "Financial intermediary leverage and value at risk", Federal Reserve Bank of New York, staff report no 338.

Alessandri, P and M Drehmann (2010): "An economic capital model integrating credit and interest rate risk in the banking book", *Journal of Banking and Finance*, vol 34, pp 730–742.

Alexander, C and E Sheedy (2008): "Developing a stress testing framework based on market risk models", *Journal of Banking and Finance*, vol 32, no 10, pp 2220–2236.

Almgren, R and N Chriss (2001): "Optimal execution of portfolio transactions", *Journal of Risk*, vol 3, pp 5–39.

Amihud, Y (2002): "Illiquidity and stock returns: cross-section and time-series effects", *Journal of Financial Markets*, pp 31–56.

Aragones, J, C Blanco and K Dowd (2001): "Incorporating stress tests into market risk modeling", *Derivatives Quarterly*, pp 44–49.

Aramonte, S, M Rodriguez and J Wu (2010): "Portfolio value-at-risk: a dynamic factor approach", Federal Reserve Board.

Artzner, P F, J Delbaen, J Eber and D Heath (1999): "Coherent measures of risk", *Mathematical Finance*, 203–228.

Bakshi, G and G Panayotov (2010): "First-passage probability, jump models, and intra-horizon risk", *Journal of Financial Economics*, vol 95, pp 20–40.

Balaban, E, J Ouenniche and D Politou (2005), "A note on return distribution of UK stock indices", *Applied Economics Letters*, vol 12, pp 573–576.

Bangia, A, F X Diebold, T Schuermann and J D Stroughair (1999a): "Modeling liquidity risk, with implication for traditional market risk measurement and management", Wharton working paper.

——— (1999b): "Liquidity on the outside", *Risk*, December, pp 68–73.

Barnhill, T and W Maxwell (2002): "Modeling correlated interest rate, exchange rate, and credit risk in fixed income portfolios", *Journal of Banking and Finance*, vol 26, pp 347–374.

Barnhill, T M, P Papapanagiotou and L Schumacher (2000): "Measuring integrated market and credit risks in bank portfolios: an application to a set of hypothetical banks operating in South Africa", IMF Working Paper #2000–212.

Barone-Adesi, G, F Bourgoin and K Giannopoulos (1998): “Don’t look back”, *Risk*, November, pp 100–103.

Basak, S and A Shapiro (2001): “Value-at-risk-based risk management: optimal policies and asset prices”, *The Review of Financial Studies*, pp 371–405.

Basel Committee on Banking Supervision (2009a): *Findings on the interaction of market and credit risk*, Working Paper no 16, Basel.

——— (2009b): *Revisions to the Basel II market risk framework*, http://www.bis.org/publ/bcbs158.pdf, July.

Berkowitz, J (2000a): “A coherent framework for stress-testing”, *Journal of Risk*, vol 2, pp 1–11.

——— (2000b): “Incorporating liquidity risk into value-at-risk models”, Working Paper, University of Houston, September.

——— (2001): “Testing density forecasts, with applications to risk management”, *Journal of Business and Economic Statistics*, vol 19, no 4, pp 465–474.

Berkowitz, J and J O'Brien (2002): “How accurate are value-at-risk models at commercial banks?”, *Journal of Finance*, vol 57, pp 1093–1111.

Berkowitz, J, P F Christoffersen and D Pelletier (2010): “Evaluating value-at-risk models with desk-level data”, *Management Science*.

Berry, R P (2009): “Back testing value-at-risk”, *Investment Analytics and Consulting*, September.

Bervas, A (2006): “Market liquidity and its incorporation into risk management”, Banque de France Financial Stability Review.

Botha, A (2008): “Portfolio liquidity-adjusted value at risk”, SAJEMS NS, pp 203–216.

Boudoukh, J, M Richardson and R Whitelaw (1998): “The best of both worlds”, *Risk*, May, pp 64–67.

Breuer, T, M Jandačka, K Rheinberger and M Summer (2008): “Compounding effects between market and credit risk: the case of variable rate loans”, in A Resti (ed), *The Second Pillar in Basel II and the Challenge of Economic Capital*, London: Risk Books.

——— (2009): “How to find plausible, severe, and useful stress scenarios”, *International Journal of Central Banking*, September.

——— (2010a): “Does adding up of economic capital for market and credit risk amount to a conservative risk estimate?”, *Journal of Banking and Finance*, vol 34, pp 703–712.

Breuer, T, M Jandačka, J Mencia and M Summer (2010b): “A systematic approach to multi-period stress testing of portfolio credit risk”, Bank of Spain Working Paper, June.

Brockmann, M and M Kalkbrener (2010): “On the aggregation of risk”, *Journal of Risk*, vol 12, no 3.

Campbell, S D (2005): “A review of backtesting and backtesting procedures”, FEDS Working Paper Series.

Christoffersen, P (1998): “Evaluating interval forecasts”, *International Economic Review*, pp 841–862.

Christoffersen, P and F Diebold (2000): “How relevant is volatility forecasting for financial risk management?”, *The Review of Economics and Statistics*, vol 82, no 1, pp 12–22.

Christoffersen, P, F Diebold and T Schuermann (1998): “Horizon problems and extreme events in financial risk management”, *FRBNY Economic Policy Review*, October 1998, pp 109–118.

Crouhy, M, D Galai and R Mark (2003): *Risk management*, McGraw-Hill.

Cuenot, S, N Masschelein, M Pritsker, T Schuermann and A Siddique (2006): “Interaction of market and credit risk: framework and literature review”, manuscript, Basel Committee Research Task Force Working Group.

Danielsson, J (2002): “The emperor has no clothes: limits to risk modelling”, *Journal of Banking and Finance*, vol 26, pp 1273–1296.

Danielsson, J, P Embrechts, C Goodhart, C Keating, F Muennich, O Renault and H Shin (2001): *An academic response to Basel II*.

Danielsson, J, B N Jorgensen, G Samorodnitsky, M Sarma and C G de Vries (2005): “Subadditivity re-examined: the case for value-at-risk”, FMG Discussion Papers, London School of Economics.

Danielsson, J and J Zigrand (2006): “On time-scaling of risk and the square-root-of-time rule”, *Journal of Banking and Finance*, vol 30, pp 2701–2713.

Danielsson, J, H Shin and J Zigrand (2009): “Risk appetite and endogenous risk”, mimeo, http://www.princeton.edu/~hsshin/www/riskappetite.pdf.

Degen, M, P Embrechts and D Lambrigger (2007): “The quantitative modeling of operational risk: between g-and-h and EVT”, ASTIN Bulletin, vol 37, no 2, pp 265–291.

Delbaen, F (2002), “Coherent risk measures”, lecture notes for University of Pisa lectures, draft.

Diebold, F, A Hickman, A Inoue and T Schuermann (1998), “Scale models”, *Risk*, no 11, pp 104–107.

Dimakos, X and K Aas (2004): “Integrated risk modeling”, *Statistical Modelling*, vol 4, no 4, pp 265–277.

Drehmann, M, S Sorensen and M Stringa (2010): “The integrated impact of credit and interest rate risk on banks: a dynamic framework and stress testing application”, *Journal of Banking and Finance*, vol 34, pp 713–742.

Dunn, G (2009), “Modelling market risk”, UK FSA.

Egloff, D, M Leippold and S Jöhri (2005): “Optimal importance sampling for credit portfolios with stochastic approximation”, working paper, http://ssrn.com/abstract=1002631.

Engle, R F (2002): “Dynamic conditional correlation – a simple class of multivariate GARCH models”, *Journal of Business and Economic Statistics*, pp 339–350.

Engle, R F and R Ferstenberg (2006): "Execution risk", NBER working paper 12165.

Engle, R F and B Kelly (2009): "Dynamic equicorrelation", Stern School of Business.

Engle, R F and K F Kroner (1995): "Multivariate simultaneous generalized ARCH", *Econometric Theory*, pp 122–150.

Engle, R F, N Shephard and K Sheppard (2007): "Fitting and testing vast dimensional time-varying covariance models", New York University.

Finger, C (2009), "IRC comments", *RiskMetrics Group Research Monthly*, February.

Francois-Heude, A and P Van Wynendaele (2001): "Integrating liquidity risk in a parametric intraday VaR framework", Université de Perpignan, France, Facultés Universitaires Catholiques de Mons, Belgium.

Franke, J, W K Härdle and C M Hafner (2008): "Value at risk and backtesting", in *Statistics of Financial Markets*, pp 321–332, Berlin, Heidelberg: Springer.

Geanakoplos, J (2009): "The leverage cycle", Cowles Foundation Discussion Paper no 1715R.

Gennotte, G and H Leland (1990): "Market liqudity, hedging, and crashes", *American Economic Review*, vol 80, no5, pp 999–1021.

Gourier, E, W Farkas and D Abbate (2009): "Operational risk quantification using extreme value theory and copulas: from theory to practice", *Journal of Operational Risk*, vol 4, no 3, pp 3–26.

Grundke, P (2005): "Risk measurement with integrated market and credit portfolio models", *Journal of Risk*, vol 7, pp 63–94.

Häberle, R and P Persson (2000): "Incorporating market liqudity constraints in VaR", *Bankers Markets & Investors,* vol 44, 01/01/2000.

Hallerbach, W G (2003): "Decomposing portfolio value-at-risk: a general analysis", *Journal of Risk*, vol 5, no 2, pp 1–18.

Hellwig, M (1995), "Systemic aspects of risk management in banking and finance", *Swiss Journal of Economics and Statistics*, vol 131, no 4/2, pp 723–737.

——— (1996), "Capital adequacy rules as instruments for the regulation of banks", *Swiss Journal of Economics and Statistics*, vol 132, no 4/2, pp 609–612.

——— (2009): "Brandbeschleuniger im Finanzsystem", *Max Planck Research*, vol 2, pp 10–15.

Hisata, Y and Y Yamai (2000): "Research toward the practical application of liquidity risk evaluation methods", *Monetary and Economic Studies*, pp 83–128.

Huberman, G and W Stanzl (2005), "Optimal liquidity trading", *Review of Finance*, vol 9, no2, pp 165–200.

International Accounting Standard 39, *Financial instruments: recognition and measurement*, last version 31 December 2008.

Jarrow, R and S Turnbull (2000): "The intersection of market and credit risk", *Journal of Banking and Finance*, vol 24, no 1–2, pp 271–299.

Jarrow, R and A Subramanian (2001): "The liquidity discount", *Mathematical Finance*, pp 447–474.

Jarrow, R and Ph Protter (2005): "Liquidity risk and risk measure computation", working paper.

Jobst, N J, G Mitra and S A Zenios (2006): "Integrating market and credit risk: a simulation and optimisation perspective", *Journal of Banking and Finance*, vol 30, pp 717–742.

Jobst, N J S A Zenios (2001): "The tail that wags the dog: integrating credit risk in asset portfolios", *Journal of Risk Finance*, vol 3, pp 31–43.

J P Morgan (1996): *RiskMetrics Technical Document*, http://www.riskmetrics.com/system/files/private/td4e.pdf.

Kalkbrener, M, H Lotter and L Overbeck (2004): "Sensible and efficient capital allocation for credit portfolios", *Risk*, January, pp 19–24.

Kalkbrener, M, A Kennedy and M Popp (2007): "Efficient calculation of expected shortfall contributions in large credit portfolios", *Journal of Computational Finance*, vol 11, no 2, pp 45–77.

Kaufman, R (2004), *Long-term risk management*, PhD thesis, ETH Zurich.

Kerkhof, J and B Melenberg (2004): "Backtesting for risk-based regulatory capital*", Journal of Banking and Finance*, vol 28, no 8, pp 1845–1865.

Kritzman, M and D Rich (2002): "The mismeasurement of risk", *Financial Analysts Journal*, vol 58, no 3, pp 91–99.

Kupiec, P (1998), "Stress testing in a value-at-risk framework", *Journal of Derivatives*, pp 7–24.

——— (1995), "Techniques for verifying the accuracy of risk measurement models", *Journal of Derivatives*, pp 73–84.

——— (2007), "An integrated structural model for portfolio market and credit risk", manuscript, Federal Deposit Insurance Corporation.

Kuritzkes, A and T Schuermann (2007): "What we know, don't know and can't know about bank risk: a view from the trenches", forthcoming in F X Diebold, N Doherty and R J Herring (eds), *The Known, The Unknown and The Unknowable in Financial Risk Management*, Princeton University Press.

Kuritzkes, A, T Schuermann and S M Weiner (2003): "Risk measurement, risk management and capital adequacy in financial conglomerates", Wharton Financial Institutions Center Working Paper #2003-02.

Lawrence, C and G Robinson (1997): "Liquidity, dynamic hedging and VaR", in: *Risk management for financial institutions*, Risk Publications, London, pp 63–72.

Le Saout, E (2002): "Intégration du risque de liquidité dans les modèles de valeur en risque", *Bankers Markets & Investors,* vol 61, November–December, pp 15–25.

Lopez, J (1999): "Methods for evaluating value-at-risk estimates", *Federal Reserve Bank of San Francisco Review*, 2, pp 3–17.

——— (2005), "Stress tests: useful complements to financial risk models", *FRBSF Economic Letter*, pp 119–124.

McNeil, A, R Frey and P Embrechts (2005): *Quantitative risk management*, Princeton.

Perignon, C and D Smith (2006): "The level and quality of value-at-risk disclosure by commercial banks", Simon Fraser University.

Platen, E and M Schweizer (1998): "On feedback effects from hedging derivatives", *Mathematical Finance*, vol 8, pp 67–84.

Pritsker, M (2006), "The hidden dangers of historical simulation", *Journal of Banking and Finance*, vol 30, no 2, pp 561–582.

Provizionatou, V, S Markose and O Menkens (2005): "Empirical scaling rules for value-at-risk", University of Essex.

Qi, J and W L Ng (2009): "Liquidity adjusted intraday value at risk", *Proceedings of the World Congress of Engineering*.

Rogers, L C G and S Singh (2005): "Option pricing in an illiquid market", Technical Report, University of Cambridge.

Rosenberg, J and T Schuermann (2006): "A general approach to integrated risk management with skewed, fat-tailed risks", *Journal of Financial Economics*, vol 79, pp 569–614.

Saunders, A and I Walter (1994): *Universal banking in the United States: what could we gain? What could we lose?*, Oxford University Press, New York.

Schönbucher, P J and P Wilmott (2000): "The feedback effects of hedging in illiquid markets", *SIAM Journal on Applied Mathematics*, vol 61, pp 232–272.

Shang, D (2009): "ARCH-based value-at-risk with heavy-tailed errors", London School of Economics.

Shin, H S (2008a): "Risk and liquidity", *Clarendon Lectures*, Oxford University Press, forthcoming.

——— (2008b), "Risk and liquidity in a system context", *Journal of Financial Intermediation*, vol 17, no 3, pp 315–329.

Sircar, K R and G Papanicolaou (1998): "Generalized Black-Scholes models accounting for increased market volatility from hedging strategies", *Applied Mathematical Finance*, vol 5, no 1, pp 45–82.

Smithson, C and L Minton (1996): "Value-at-risk", *Risk*, September, pp 38–39.

Stange, S and C Kaserer (2008): "Why and how to integrate liquidity risk into a VaR-framework", CEFS Working Paper.

Stiroh, K J (2004): "Diversification in banking: Is noninterest income the answer?", *Journal of Money, Credit and Banking*, vol 36, no 5, pp 853–82.

Subramanian, A (2008): “Optimal liquidation by a large investor”, *SIAM Journal of Applied Mathematics*, vol 68, no 4, pp 1168–1201.

Sy, W (2006): “On the coherence of VaR risk measures for Lévy Stable distributions”, Australian Prudential Regulation Authority.

Wang, S (2001): “A risk measure that goes beyond coherence”, University of Waterloo, Institute of Insurance and Pension.

Wang, S, V Young and H Panjer (1997): “Axiomatic characterization of insurance prices”, *Insurance: Mathematics and Economics*, vol 21, pp 173–183.

Wong, W K (2008): “Backtesting trading risk of commercial banks using expected shortfall”, *Journal of Banking and Finance*, vol 32, no 7, pp 1404–1415.

Wu, G and Z Xiao (2002): “An analysis of risk measures”, *Journal of Risk*, vol 4, no 4, pp 53–75.

Wu, L (2009): “Incorporating liquidity risk in value-at-risk based on liquidity adjusted returns”, Southwestern University of Economics and Finance.

Wylie, J, Q Zhang and T Siu (2010): “Can expected shortfall and value-at-risk be used to statistically hedge options?”, *Quantitative Finance*, vol 10, no 6, pp 575–583.

Yamai, Y and Y Yoshiba (2005): “Value-at-risk versus expected shortfall: a practical perspective”, *Journal of Banking and Finance*, vol 29, pp 997–1015.

Zheng, H (2006): “Interaction of credit and liquidity risks: modelling and valuation”, *Journal of Banking and Finance*, vol 30, pp 391–407.

Annex

Table 1

Summary of “bottom-up” risk aggregation papers in the survey

Research paper	Portfolio analysed	Horizon	Risk measure used	Ratio of unified risk measure to sum of compartmentalised risk measures
Breuer, Jandačka, Rheinberger and Summer (2010a)	Hypothetical portfolios of foreign-exchange denominated loans of rating: BBB+ B+	One year	Expected shortfall at: the 1% level the 0.1% level the 1% level the 0.1% level	 1.94 8.22 3.54 7.59
Breuer, Jandačka, Rheinberger and Summer (2008)	Hypothetical portfolios of variable rate loans of rating: BBB+ B+	One year	Expected shortfall at: the 1% level the 0.1% level the 1% level the 0.1% level	 1.11 1.16 1.06 1.10
Grundke (2005)	Hypothetical portfolios of loans with various credit ratings, asset value correlations, distributional assumptions, and correlations between the risk-free rate, credit spreads and firm asset returns	Three years	VaR at: the 1% level the 0.1% level	 0.07–0.97 0.09–1.00
Kupiec (2007)	Hypothetical portfolio of corporate loans with various rating categories calibrated to historical data	Six months	Portfolio losses at funding cost levels consistent with: AAA rating BBB rating	 0.60–3.65 0.61–3.81
Drehmann, Sorensen and Stringa (2010)	Hypothetical UK bank	Three years	Decline in capital over the horizon	2.5
Alessandri and Drehmann (2008)	Hypothetical UK bank	One year	Value-at-risk at: the 1% level the 0.1% level	 0.03 0.50

Table 2

Summary of "top-down" risk aggregation papers in the survey

Research paper	Portfolio analysed	Horizon	Risk measure used	Ratio of unified risk measure to sum of compartmentalised risk measures
Dimakos and Aas (2004)	Norwegian financial conglomerate		Total risk exposure at: the 1% level the 0.1% level	 0.90 0.80
Rosenberg and Schuermann (2008)	Hypothetical, internationally-active financial conglomerate	One year	Value-at-risk based on a normal copula at the 0.1% level. (Note: similar results using expected shortfall.)	0.42–0.89 based on different correlation assumptions between market, credit and operational risk.
Kuritzkes, Schuermann and Weiner (2003)	Representative Dutch bank	One year	Economic capital	0.72–0.85 based on different correlation assumptions between market, credit and operational risk.
Kuritzkes and Schuermann (2007)	US banking system from 1986.Q2 through 2005.Q1	—	Tail quantile of the earnings distribution at: the 1% level the 0.1% level	 0.63 0.63