This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: The Risks of Financial Institutions

Volume Author/Editor: Mark Carey and René M. Stulz, editors

Volume Publisher: University of Chicago Press


Volume URL: http://www.nber.org/books/care06-1

Conference Date: October 22-23, 2004

Publication Date: January 2007

Title: Introduction to "The Risks of Financial Institutions"

Author: Mark Carey, Rene M. Stulz

URL: http://www.nber.org/chapters/c9603
About twenty years ago, the intellectual and practical dynamics of understanding and managing the risks of financial system distress began to change. The consensus view, which was that runs on solvent banks were at the heart of banking panics—and that panics were the main problem—ironically began to unravel around the time Diamond and Dybvig (1983) published their theory of runs. The consensus was challenged by a series of events, including the emerging-market debt crisis of the early 1980s, the 1987 and 1989 stock market crashes, waves of failures of U.S. savings and loan associations (S&Ls) and banks in the late 1980s and early 1990s, the junk bond and U.S. municipal bond meltdowns of the early 1990s, the Long-term Capital Management (LTCM) crisis, and a new wave of emerging-market crises. Bank runs played a negligible role in most of these events. While new financial instruments (such as derivatives), new participants (e.g., hedge funds), and new technologies (like electronic trading), typically have improved the informational efficiency of markets and have facilitated the matching of savings with investment opportunities, they have also changed the speed with which new information is incorporated into prices, often giving little time for institutions to adjust to new information before they see their financial soundness imperiled by new balance sheet weaknesses or by liquidity problems.

The traditional public policy prescription also became less satisfactory. The prescription was that financial system distress can be prevented or managed by a combination of banking supervision and regulation (to pre-
serve bank solvency and to permit central banks to identify solvent banks in a panic), lender-of-last-resort advances (to solvent banks experiencing liquidity problems during a panic), and deposit insurance. But none of the new crises fit the old mold. Some of the new events featured sharp movements in asset prices and sharp contractions in market liquidity. Others featured massive credit losses due to concentrations of poorly underwritten loans or failure to appreciate credit risk concentrations. Moreover, numerous emerging-market countries experienced banking crises, but deposit insurance does not seem to have reduced the probability of banking crises and perhaps even contributed to them (see Demirgüç-Kunt and Detragiache [2001]).

Facing events at variance with the prevailing intellectual framework, policymakers were forced to feel their way toward crisis solutions and toward new preventive measures. Developments in capital markets, especially the growth in derivatives markets, increased the tools available to firms to take on and manage risks. These developments also made traditional accounting numbers, which regulators used to assess financial institutions and executives used to manage such institutions, much less relevant to measurement of exposures to various risks. Through the trading of derivatives, for instance, a bank can take large risks that are nearly invisible when investors look at its balance sheet. For instance, banks would traditionally take interest rate exposures by taking deposits or making loans and buying bonds. However, with derivatives, a bank can use an interest-rate swap to take the same interest rate risk as if it bought a bond, but the acquisition of the swap, in contrast to the acquisition of the bond, is not recorded on the balance sheet at inception because the value of a swap at inception is zero. After inception of the swap, mark-to-market accounting requires the bank to record the market value of the swap, but that market value provides little information about the bank’s interest rate exposure. Moreover, bank managers discovered that they could boost traditional accounting performance measures through trading, which requires little funding capital. A traditional measure of performance such as return on equity would improve through trading revenue or revenue from fees because such activities typically required little incremental equity. However, such activities can sharply increase the risks

1. Of course, runs on solvent banks might have occurred had authorities been less vigilant or credible, and insolvency rates might have been worse without bank supervision and regulation. Our argument is that the traditional intellectual foundations do not seem to predict many of the problems that have occurred, and that some of the policies these foundations imply may no longer be appropriate.

2. Banks experiencing runs in Demirgüç-Kunt and Detragiache’s (2001) sample of crises were often insolvent at the time of the runs and thus such crises, while very important, did not fit the Diamond and Dybvig paradigm of runs on solvent banks.


4. A swap is an exchange of cash flows. With an interest rate swap, one party pays a fixed rate on a notional amount and receives a floating rate on the same amount.

5. See Merton and Perold (1993) for an early discussion of this issue and an analysis of the role of risk capital in financial firms.
taken by the institution, and broker-dealers and investment banks traditionally backed such activities with substantial capital. These developments forced both bank regulators and market participants to focus on approaches that would capture the risks borne by institutions in a way that accounting numbers could not.

Market participants chose to address these changes in markets, and the increased frequency and variety of financial crises that threatened their investments and earnings, by developing formalized, quantitative risk measurement and management technologies. It was becoming increasingly clear that prevailing, mostly informal, seat-of-the-pants ways of managing risk were inadequate. The goal of the new measurement technologies is to produce realistic conditional forecasts of the distribution of returns to a financial institution, especially of the tail of the distribution corresponding to adverse outcomes. Given such forecasts, the institution can make informed decisions about its portfolio and capital structure and can also design internal incentive and control systems to help ensure that decisions are implemented properly. It has become typical for up-to-date, large financial institutions to take into account the impact of each activity on their overall risk when they evaluate the profitability of activities. Typically, a firm identifies a charge for an activity proportional to some measure of the impact of that activity on the firm’s risk. In principle, risks associated with financial crises can be incorporated in the modeling. Such new technologies are having a profound impact on financial institution risk and financial system risk and have already made it necessary to develop new ways of thinking about such risk and new public policy regimes. Pressure for such developments will increase in the future.

An example may help illustrate how the new techniques are being used. Suppose a bank is considering an expansion of lending to investment-grade, large corporate borrowers. Such loans pay relatively low interest rate spreads, but loss rates are very low in a typical year, so profit margins may appear positive and overall accounting profits may seem boosted by large volumes of such lending. Traditionally, senior bank managers might make a strategic decision to expand such lending, and implement the decision by rewarding loan officers based on the volume of loans made. Many new loans would be individually large.

More recently, the bank would make decisions based on how the new loans contribute to the risk of its portfolio of credits in relation to their contribution to the bank’s expected profits. It would measure the risk of a portfolio of credits by estimating the distribution of the portfolio’s aggregate loss, focusing in particular on the loss that might be expected to be exceeded rarely—say, once in 200 bank years (the 99.5th percentile). To produce such estimates, the bank would use a portfolio credit risk model. The reason for the focus on such loss rates is that their distribution is crucially important for maximization of franchise value, since the distribution of tail losses directly impacts a financial institution’s probability of financial
distress. Such tail-loss forecasts are often generically referred to as “value-at-risk” (VaR) measures. Value-at-risk measures for credit portfolios are generally referred to as credit VaRs. In our example, the new loans, particularly if they are large, may have a material impact on the firm’s credit VaR. If they do, the risk of bank insolvency increases. To keep the probability of insolvency unchanged, the bank would have to allocate extra equity capital, which it would typically call risk capital, as protection. Though finance theories that assume markets to be frictionless find that there are no deadweight costs to equity finance, finance theories that take into account information asymmetries and agency problems find equity to be an expensive source of finance (see, for instance, Myers 1977, Myers and Majluf 1984). Consequently, even if mean loss rates on the new loans are low, the loans might still be unprofitable, because their spreads might be too small to cover both expected losses and the required return to the extra equity that is needed.

Portfolio models can also be useful in implementing decisions. Lending officers can be provided incentives based on the marginal profit flowing from a new loan rather than on volume. The models can be used to include in measures of marginal profit the costs of allocated risk capital as well as expected credit losses and other costs. Particularly where portfolio models include fine-grained diversification effects (where the model correlates the risks posed by individual new loans with the risks of individual loans already in the portfolio), such risk-adjusted profitability measures can (in principle) be embedded in internal control and incentive systems in such a way that the bank’s target risk posture is almost automatically maintained. Such systems are especially important to the operations of very large financial institutions, where many operational decisions must be decentralized.

Although the example focuses on credit risk, the approach is used by financial institutions for other risks as well. For instance, the risks assumed by a trading desk can be evaluated by estimating the VaR of the trading desk as well as the contribution of these risks to the market risk of the financial institution or to its enterprise-wide risk.

The new risk measurement and management techniques are associated with, and in some cases are driving, a number of important changes in financial systems, including:

- A better appreciation of the types of risk to be considered and of the relationships among them.
- A better understanding of the drivers and dynamics of each type of risk and of how to model and manage risk.
- New instruments and markets that support risk transformation and risk shifting, such as securitization and derivative products.
- Changes in the industrial organization of financial systems:
  — Larger financial institutions can be more efficiently managed, adding impetus to trends toward greater concentration.
New kinds of institutions, such as hedge funds and boutique securitization sponsors.

A blurring of traditional classifications of types of institutions by the type of risk borne, aided by new instruments and by entry into each other’s markets.

- Greater attention to legal, accounting, regulatory, and other financial infrastructure. The new techniques flourish in environments that support good data and enforceable contracts.

- Changes in the nature and incidence of risks that affect the stability and soundness of the financial system—so-called systemic risks.

- Changes in the appropriate structure of regulatory and central bank policy.

Taken together, such developments are likely to change risks of distress and crisis for individual financial institutions and for national and international financial systems.

The papers and their discussants’ remarks in this volume make new contributions to the understanding, measurement, and management of financial institution risk. While some papers focus on the determinants and measurement of risks at the level of individual institutions, others focus on the determinants of systemic risk in a world where individual financial institutions measure and man age risk using approaches developed over the last twenty years. Perhaps more importantly, taken together, the papers and remarks demonstrate how interrelated the changes that are in progress are, and support the importance of continuing efforts to understand them. Another contribution, felt most forcefully by the conference participants, is the utility of bringing together academic researchers, market participants, regulators, and central bank people. All have much to contribute, and progress is particularly tangible when they are brought together.

The order in which the papers appear in the volume is somewhat arbitrary. Each paper makes contributions to an understanding of more than one of the issues in the previous list, so many different orderings can be imagined. In the remainder of this introduction, we discuss in a bit more detail each of the issues and show how the papers in this volume contribute. We hope this will help readers to better understand the overall contribution of this volume and to place these papers in a more general context. We also hope readers focusing on one or a few issues will be able to more easily find contributions of particular interest to them.

**Risk Management and Firm Value Maximization**

Financial institutions choose the level of risk that maximizes the objectives of those who run them, subject to constraints and penalties imposed by those who regulate them and by capital markets. If the incentives of managers are well aligned with the interests of shareholders, managers
maximize shareholder wealth. Observers who emphasize the moral hazard created by deposit insurance sometimes conclude that deposit insurance leads banks to take as much risk as regulators will let them take. It is now clear that such a view is much too simple.

Many financial institutions have substantial franchise value that could be lost if they are viewed as too risky. As has been emphasized by Merton (1993) and others, risk management is uniquely important for financial institutions because, in contrast to firms in other industries, their liabilities are a source of wealth creation for their shareholders. For instance, a financial institution that writes long-dated derivatives would usually be shut out of the market if the credit rating of the vehicle it uses to write such derivatives fell below an A rating. Another example is a life insurance company writing policies on its general account. Its customers would disappear if its rating fell to a junk rating, and most likely before that (few life insurance companies have ratings below A–). Because its franchise value depends on the risk of its insolvency, a financial institution has an optimal level of risk that maximizes its value for its shareholders. Risk minimization is never optimal, because there cannot be a franchise value without taking risks, so that the firm always faces costs and benefits when its risk level increases.6

To maximize shareholder wealth, managers of financial institutions therefore have to be able to measure and manage the risk of their institution. In principle, they would want to take into account the whole distribution of firm value. In practice, they focus on measures of downward risk, because adverse outcomes are those that endanger franchise value. Value-at-risk is a measure of downward risk: it measures the maximum value loss at some confidence level. For instance, a firmwide daily VaR of $100 million at the 95 percent confidence level means that in five days out of a hundred, the bank expects to have a loss that exceeds $100 million. Cash flow at risk or earnings at risk are similar measures of downward risk for cash flows and earnings. For instance, cash flow at risk is the shortfall in cash flow at a given percentile of the cash flow distribution, such as the 95th percentile.

The level of risk that has to be measured and managed is the level of risk for the whole institution. In practice, this has proved difficult. Initially, firms focused mostly on the risk of specific activities and on specific types of risks. However, lately, firms are increasingly focusing on aggregating risks firmwide.

Once a firm has measured its level of risk, it has to decide whether it is optimal for that level to be maintained, increased, or decreased. Taking on risk enables a firm to make profits, but it also endangers franchise value. To take more risks, a firm therefore has to protect franchise value by holding

6. For more details on this tradeoff see Stulz (2002).
on to more capital or by hedging. Both are costly, so firms that can manage risks better are more profitable.

With this logic, risk management may lead a financial institution to hold more capital than required by its regulators because it maximizes the wealth of its shareholders by doing so. However, the ability to manage risks also enables financial institutions to take complex risks that will be hard to detect by regulators. If the downside of such risks is likely to materialize in states of the world where governments will be tempted to bail out the financial institution, such risks may be taken even where they nominally endanger franchise value. Safety nets can therefore lead to inefficient risk taking.

Understanding the Range and Types of Risks

Managing expected firmwide risk, though necessary, is hard to do in practice. Measuring risk at the firm level would be drastically simplified if risk managers could simply model firmwide cash flow or firmwide value using time-series or cross-sectional data for these variables, or cross-sectional information. There is some evidence that time-series models of the lower tail of aggregate profit-and-loss (P&L; used as a benchmark in Berkovitz and O'Brien, 2002), and measures of cash flow risk based on comparables (Stein et al. 2001) can be reasonably successful. However, such approaches are difficult to implement in a way that appropriately reflects the risks of the financial institution at the time the measure is computed. They can be misleading if risks have changed significantly in the recent past. More importantly, such measures are not useful for the purpose of actually managing a firm’s risk because they cannot be used to evaluate how various actions by the firm change its risk. Nor are they useful for monitoring risk taking, because they do not reveal which risks are large and which are small.

Instead, firms have focused on measuring risk from the bottom up, starting at the level of individual positions, business units, and individual trading desks. As a result, risk measurement is organized according to a taxonomy of risk types that has become richer as risk management has matured, but that remains incomplete.

Established Risks: Market, Credit, and Operational

Before the late 1980s, only interest rate risk was modeled quantitatively at the portfolio level. The modeling was usually crude, often consisting of simple interest rate sensitivity measures such as a one-year duration gap, but it was sufficient to keep most institutions out of trouble in an environment when most assets and liabilities were straight debt. As interest rate derivatives became more important, simulation of changes in portfolio value in response to different interest rate scenarios became more widespread.

Market risk modeling grew up in response to the stock market crashes of
the late 1980s, to high-profile losses suffered by institutions victimized by “rogue traders,” to the expanding scope of trading and market-making activities, and to the growing importance of derivatives positions. Market risks are generally defined to be risks associated with fluctuations in prices of traded financial instruments. Increasingly, interest rate risk is often thought of as just one form of market risk that has an impact on the balance sheet that goes beyond its impact on the trading book.

As banks acquired more exposures to currencies, equities, and commodities through market-making in the spot markets and through derivatives trading, focusing most of their risk measurement efforts on their exposure to interest rates was no longer appropriate. They had to find ways to measure exposure to other factors and to aggregate their market-risk exposures to different factors. To do so, following the lead of Bankers Trust and JP Morgan, firms started using portfolio risk measures for their trading books. However, the standard portfolio risk measure—volatility—was not adequate, because the distributions of returns for portfolios, including derivatives, are generally not symmetric, so that volatility might hide substantial downside risk. To assess downside risk directly, banks focused on forecasting the VaR of a portfolio. With that approach, the VaR at the 5 percent probability level is the loss that will be exceeded with probability 0.05. Because trading books can change so quickly in liquid markets, most banks measured VaR over a one-day horizon. Though in principle all trading-book positions (and perhaps some less-liquid positions) could be included in VaR measures, achieving this goal was often difficult because different traders and trading groups had different computer systems and data architectures.

Today, market VaR models are ubiquitous at all kinds of financial institutions, especially those that actively trade. They are used to assess portfolio risk, allocate capital internally, and evaluate alternative investment strategies. They are also part of internal control systems designed to detect excessive risk taking by individual units or traders, and often are part of incentive systems designed to optimize the level of risk taken by individual units or traders.

Portfolio credit risk modeling was only five years or so behind market risk modeling in the timing of the explosive phase of its adoption, but it represented a much larger cultural innovation in the financial community. Quantitative analysis of investment portfolios, based on financial theories such as the capital asset pricing model, became common decades before VaR models, and thus the growth of VaR models represented an expansion

7. Although a few VaR systems were implemented in the 1980s, the main watershed events were in the 1990s, including a Group of Thirty (1993) report and JP Morgan’s 1994 launch of its RiskMetrics model.

8. Harmonizing data management across operations within large financial institutions, even if limited to trading activities, often involves extremely large IT expenditures.
of the toolkit rather than a wholesale change. In contrast, even through the early 1990s, credit risk was generally managed using intuition and rough approximations. Most commercial bank managers were aware that credit risk is the big gorilla for commercial banks, completely dominating other risk types as a source of bank insolvencies. But most of their efforts were focused on traditional analysis of financial statements to support appraisals of the default risk of individual borrowers. Perhaps this was because portfolio credit risk is far more difficult to model than market risk. Much of the important variation is at relatively low business-cycle frequencies, and data are sparse and harder to obtain than in the case of market risk. Moreover, distributions of returns on credit portfolios are highly skewed.

Early adopters of portfolio credit risk modeling in the United States were motivated by their near-death experiences during the 1990–91 recession. Others began to seriously incorporate credit risk modeling into their operations at least partly in response to the Basel Committee’s (1999, 2001, 2004) proposals to embed credit VaR techniques in bank capital regulations. Some firms focused on measuring losses associated with default events, thus focusing on default rates and loss-given-default (so-called default-mode modeling). Others focused on measuring changes in the mark-to-market value of credit portfolios caused by any event.9 Both approaches remain in widespread use.

Operational risk is a relative newcomer to the taxonomy. At this point, there is not even a generally accepted definition of operational risk. Some practitioners call operational risk all the risks that are not market and credit risks. Others follow the Basel II definition of operational risk: “the risk of direct or indirect loss resulting from inadequate or failed processes, people and systems or from external events.”10 Operational risk has become an important part of financial institution risk management efforts, partly because it was highlighted by the Basel Committee (2001), because of Section 404 of Sarbanes-Oxley and related regulations about internal controls, and because of the disruptions associated with the September 11, 2001, attacks. Though some still doubt whether it is material or even can be measured, financial institutions increasingly allocate capital to operational risk. For instance, a survey by Oliver Wyman and Company of ten large international banks found that they allocate 53 percent of their economic capital to credit risk, 21 percent to market risk and asset-liability rate risks, and 26 percent to operational and other risks.11 One contribu-

---

9. The earliest models were CreditMetrics from JP Morgan, which focused on the mark-to-market value of credits, and CreditRisk + from CSFB, which focused only on losses associated with default. See Stulz (2002), chapter 18, for a description of these models. Basel II is a prominent example of the default-mode approach.
10. See Basel Committee (2001), page 2.
tion of this volume is de Fontnouvelle, Jordan, and Rosengren’s evidence that operational risk is material. In addition to methodological contributions described subsequently, their evidence implies that operational-risk VaR is on the order of market-risk VaR for typical commercial banks, consistent with the survey results just cited.

Still-Early Days: Liquidity, Strategic, and Business Risk

Taken together, do market, credit, and operational risks represent the entirety of risk in financial institutions? If operational risk is defined so that it includes everything that is not market and credit risk, the three types of risk would represent the entirety of risk in financial institutions. But as a practical matter, operational risk modeling has come to focus on a subset of event types that are susceptible to internal measurement by individual financial institutions. Regulators have also chosen a narrow definition of operational risk. The definition of operational risk in the Basel II accord excludes risks such as strategic risks, reputational risks, and liquidity risks. Though operational risk includes many facets of what people would call business risk, many definitions of operational risk do not include the business cycle and fluctuations of the fee income of banks.

If one presumes that anything can be bought and sold for a price, an implication follows that financial institutions can raise liabilities or sell assets as needed, so liquidity risk would be subsumed by market risk. Periods characterized by liquidity problems would simply be periods when prices move a lot, and a good market risk model would capture the risk of such price movements. Such a view would be correct if the only dimension of liquidity risk is changing bid-ask spreads.\(^\text{12}\) In this case, risk management modeling of changes in bid prices for long positions and ask prices for short positions would properly take into account liquidity. However, in general, this view is flawed, because when liquidity is imperfect the price at which an asset or a liability can be quickly sold depends on the quantity sold.\(^\text{13}\) In practice, sometimes assets cannot be sold, and liabilities cannot be raised, at any price close to fundamental value in a timely fashion. Perhaps more unnerving, worries about future liquidity can lead to crashes as investors rush for the exits.\(^\text{14}\) Commercial and central banks have worried about liquidity risk for centuries, and have evolved various mechanisms to deal with it. Indeed, Gatev, Schuermann, and Strahan’s article in this volume offers evidence that the core business lines of banks (deposits and lines of credit) act as a kind of automatic stabilizer for the whole financial system during periods of stress, with liquid deposits flowing in from some clients just at the time when other clients need to make drawdowns on their lines.

However, at the level of individual financial institutions, to our knowledge, liquidity risks have not yet been quantitatively analyzed in the same manner as market, credit, or operational risk. Perhaps because liquidity shortages are relatively rare and often are associated with other events, data are difficult to obtain and conceptual models are lacking. Thus, progress toward VaR-like models of liquidity risk or toward a careful incorporation of liquidity risk in market risk models may be slow.¹⁵

Business risk and strategic risk modeling are a little bit further along. Measures that focus on cash flow at risk (CaR) or earnings at risk (EaR) capture business risk.¹⁶ Though similar to VaR measures in that the loss rate at a percentile of a loss distribution is measured, CaR or EaR measures assume that a firm’s cash flows or earnings provide the correct measure of its capacity to finance investments and repay debt, whereas VaR measures implicitly assume that all the assets and liabilities included in the measure are liquid.¹⁷ The modeling horizon of these measures is different for business and for strategic risk. For business risk, the horizon is usually a single accounting period; for example, a quarter. But strategic decisions cannot be evaluated in the context of one accounting period. Instead, one has to look over time to see how decisions will contribute to the value of the firm and how they will affect the risk of the firm. More generally, quantification of the risk of strategic decisions forces firms to make their assumptions precise and to more directly understand the risks involved in making such decisions.

Model Risk and Systemic Risk

A final part of the risk taxonomy—model risk—is a consequence of the growth of the new risk technologies. Model risk denotes the risk institutions face because of model errors. These errors can have a wide variety of causes. For instance, a pricing model could have a coding error, could have an assumption that leads to substantial biases in some states of the world, or wrong data could have been used as input. Concerns about model risk have been raised at both individual-institution and systemic levels. In the former case, the concern is that by building models into its management and control systems a financial institution may be led by a bad model to take large risks that it would never have taken in the absence of the model. This is a legitimate concern, but the practical solutions are obvious: human review of strategies and positions, use of multiple models, and simulation of the impact of hypothetical model errors.

¹⁶. Unfortunately, the alphabet soup used in risk management is not standardized. As a result, at some banks CaR means “capital-at-risk,” essentially VaR, and not cash-flow-at-risk. At some other banks, EaR is the abbreviation used for P&L at risk, which again is VaR, not earnings-at-risk in the sense used here.
Systemic model risks have recently received more attention. The most common concern is that if financial institutions adopt a common risk-modeling framework, their tendency to herd will be amplified and markets may be destabilized (see Basak and Shapiro 2001, Danielsson, Shin, and Zigrand 2002, Persaud 2000, and Scholes 2000 for this and related ideas). Existing risk management models treat the risks of positions as exogenous, and are therefore of little use to financial institutions in evaluating the risks created by model-driven behavior, either their own behavior or that of other institutions. The current volume makes contributions to this debate on both sides. Adding to the concerns in the literature are Allen and Gale’s model, which might be interpreted as raising concerns about inefficient regulatory use of risk management models. The papers of Jorion and Berkowitz and O’Brien assuage such concerns. Their papers show that there is little evidence that commercial bank market VaR forecasts are highly correlated, that banks take large exposures to market risks, or that P&L exposures to risk factors are highly correlated across banks—except, perhaps, for interest rate risks. The findings are strong enough and robust enough to support rejection of hypotheses that the use of VaR measures, either internally or for regulatory purposes, will be automatically destabilizing. This is an extremely important finding because it strengthens the case for moving forward with use and improvement of risk management techniques.

Similarly, Chan, Getmansky, Haas, and Lo do not find strong evidence of commonalities in the sensitivities of hedge fund indexes to risk factors, even though risk measures like VaR are used widely among hedge funds. However, they do find evidence that bank stock returns are correlated with hedge fund returns, suggesting that further investigation into channels of contagion is needed.

Measuring Firmwide Risk

Although analyzing each type of risk in isolation allows measures to be customized to suit the properties of the risk, and thus improves their quality as stand-alone measures, at some point the different risk measures must be combined to give a view of risk for a whole financial institution—a firm-wide risk measure. As noted previously, this has proved challenging. Often, financial institutions attempting to measure firmwide risk found that they had information systems that could not talk to one another, that they had little computer-readable historical data (except in their trading activities), and that they had no records at all of information important to the assessment of risks. Even partial solutions to these problems can require huge investments in information technology.

More fundamentally, however, financial institutions find it difficult to aggregate firmwide risks for three important reasons:
1. The shapes of distributions differ for different types of risk, so that the analysis of the aggregated risks is not straightforward. Whereas distributions for market risks are typically close to symmetric but with fat tails, distributions for credit risks and for operational risks are extremely skewed. With debt, the most the financial institution can receive is the promised payments—but it can lose the whole position. With operational risks, the high-frequency losses are typically small, but there is also the potential for extremely large losses, which have a low probability of occurring. Such differences in risk distribution typically make it inappropriate to use simple portfolio risk formulas to aggregate market, credit, and operational risks because means, variances, and covariances are not sufficient statistics for these risk distributions.

2. Conditional correlations of different types of risk are hard to measure with confidence. For instance, the historical record suggests that bad-tail market and credit risk outcomes are correlated—but not perfectly—and historical data do not cover enough potential states of the world. Further, correlation may not be the appropriate measure of dependence between these various types of risks because of their fat tails. In particular, it is possible that tail outcomes of different types of risks are more highly correlated than other outcomes. Chan, Getmansky, Haas, and Lo in this volume discuss the phenomenon of “phase-locking,” meaning states of the world where many variables become very highly correlated that otherwise tend not to be.

3. As discussed previously, risk tends to be measured over different horizons for different types of risks, but to aggregate risks at the firm level they need to be forecasted over comparable periods. For market risk, the focus is generally on days; for operational risk and credit risk, it is often on one budget year. We lack clear foundations for existing choices of horizon. They appear to be empirical compromises, driven by the nature of the positions being modeled, the needs of internal control systems, and the nature of available data. But we have little idea of how to do things differently. One approach to the problem is a framework advocated by a consulting firm, Algorithmics, named Mark-to-Future. This framework differs from traditional VaR calculations in that the simulations are computed over multiple periods and allow for actions by firms to be path dependent. However, in practice, implementation of such a framework faces a multitude of obstacles.

Firms and regulators have often approached the firmwide risk aggregation problem by using ad hoc assumptions about correlations. An example is the National Association of Insurance Commissioners’ (NAIC) risk-based capital regime for insurance companies, in which risks are aggregated by a formula based on relatively simple but rather arbitrary cor-

---

18. The existence of fat tails explains the growing importance of the copula measure of dependence used in the extreme-value literature. See Rosenberg and Schuermann (2006).
relation assumptions. Another possibility is to stop short of aggregating bottom-up risk measures, turning instead to alternative measures of risk at the whole-institution level.

**How to Model Risks, Including Systemic Risks**

Especially if model-induced herding is less of a concern, it seems obvious that better measurement is good. Both practitioners and researchers seem to have agreed with this view over the past couple of decades. Much attention has been given to details of measurement, but much remains to be done. The papers in this volume make a number of new contributions.

**Market Risk**

The RiskMetrics approach proposed by JP Morgan became especially popular as JP Morgan made the methodology and the daily data freely available. This approach forecasts volatilities and correlations for a number of risk factors, assuming returns to be conditionally normal, and uses exponential weighting for the forecasts. The risks of positions are then represented in terms of exposures to the risk factors, so that the return of the portfolio becomes a weighted average of the returns of the risk factors. The volatility of a portfolio can then be computed using the formula for the variance of a portfolio. The approach is mostly focused on forecasting the risk of the portfolio over the next day, making the assumption that expected returns equal zero reasonable. Under such assumptions, the VaR at the 95 percent confidence interval is simply 1.65 times the volatility of the portfolio. Andersen, Bollerslev, Christoffersen, and Diebold in this volume discuss some of the weaknesses of this approach (and other approaches that ignore serial correlation in volatilities) and show ways to overcome them.

Early users of RiskMetrics soon began to focus more on simulations of portfolio values, because portfolio risk formulas could not handle well the risks of derivatives. Moreover, the normal distribution that was used almost exclusively in early implementations of VaR proved flawed for market risk because relevant empirical distributions have fat tails. However, parametric distributions that could be used to replace the normal distribution were generally viewed as impractical for large portfolios. The basic idea of simulation methods is to estimate portfolio value in a realistic array of circumstances, conditional on the details of portfolio positions. This led firms to either simulate risk factors using estimated distributions for each factor, an approach involving Monte Carlo simulation, or to use so-called historical simulation, wherein portfolio returns are simulated from historical realizations of risk factors. For a detailed description of these approaches, see Jorion (2000).
cially practical way to address the problem of the inadequacies of the normal distribution.

In practical applications, the historical simulation approach is often insufficiently conditional—that is, it does not take sufficiently into account the recent past, so that sharp increases in volatility that will persist in the near future are not given sufficient weight.\textsuperscript{20} Existing evidence on the performance of VaR models at large banks shows that they had an unusual number of days where the VaR was exceeded (“exceedances”) in August and September 1998 (see Berkovitz and O’Brien 2002), demonstrating that the models fail to adequately capture the changes in the joint distribution of returns that took place during that period. Chan, Getmansky, Haas, and Lo point out that inferences about risk can be acutely sensitive to the sample period used to generate risk measures. As a result, quiet periods will lead to low VaRs. Andersen, Bollerslev, Christoffersen, and Diebold show how important it is to recognize time-varying volatility and correlation in VaR estimation. They demonstrate how this can be done using parametric generalized autoregressive conditional heteroskedasticity (GARCH) modeling, filtered historical simulation, and high-frequency data.

The model risks of risk measurement make it essential for institutions to use additional risk measures and not focus on VaR only. One contribution of Chan, Getmansky, Haas, and Lo is to provide an array of alternative risk measures in the context of their analysis of hedge fund risk.

In recent years, practitioners and regulators have put much emphasis on the use of stress tests as an alternative to VaR. Stress tests measure the impact on portfolio value of shocks to key risk factors. For example, a stress test might investigate how a bank would perform if an earlier market disruption, such as the events of August and September 1998, were repeated; any scenario, however, including one outside the boundaries of historical experience, can be used.\textsuperscript{21} Stress test methods essentially make no use of statistical and econometric theory. They became popular because of generic concerns about model inadequacies and especially because it is difficult to model volatility and correlation behavior in times of market stress. Much has been made in the literature of correlation breakdowns in such times; the principals at LTCM are on record in stating that their correlation assumptions fell apart in August and September 1998.\textsuperscript{22}

After describing the problems that arise in capturing time-varying volatility when there are a large number of factors, Andersen, Bollerslev, Christoffersen, and Diebold show how new techniques in multivariate time-series estimation could be usefully brought to bear to address some of the problems that pushed banks toward historical simulation and stress test-

\textsuperscript{20} Pritsker (2001) analyzes this issue in detail.
\textsuperscript{21} The Committee on the Global Financial System (2001) provides a detailed survey of stress tests across financial institutions.
\textsuperscript{22} See McKenzie (2003).
ing. They argue for an evolution of market-risk modeling of asset-return volatility and correlations away from both parametric (RiskMetrics-like) and historical-simulation methods. Instead, where feasible, they suggest the use of nonparametric volatility measurement, using high-frequency data, paired with parametric volatility models designed to support computationally efficient solutions to high-dimensional problems. Strikingly, they propose the development of risk management systems with a limited number of risk factors (less than thirty), but for which intraday data would be available and hence volatility and correlation forecasts more reliable. It remains to be seen whether such an approach could capture the risks that financial institutions now model using a much larger number of factors.

Andersen, Bollerslev, Christoffersen, and Diebold also point out the practical problems of dimensionality that arise when the number of positions is large and large numbers of factors must be used, as is common at the largest banks today. A large bank active in trading may use more than 1,000 risk factors and have more than 100,000 positions, each of which must be repriced for each draw of the underlying factors. Computational burdens of repricing are high because of nonlinear sensitivities of prices to factors. Even with RiskMetrics-style parametric modeling using a normal distribution, 1,000 factors requires modeling over 500,000 variances and covariances. Dimension reduction methods help, but they also introduce estimation errors and still require a large number of parameters to be estimated.

Credit Risk

Using a multifactor portfolio credit risk model that includes explicit dynamic modeling of macroeconomic dynamics, Pesaran, Schuermann, and Treutler offer evidence of considerably larger benefits of credit diversification than are implied by current workhorse models. The simplest portfolio credit risk model, which is widely used, has only a single systematic factor and is a model in which all borrowers have the same exposure to the factor (Gordy 2003). Commonly used model implementations, such as CreditMetrics, are multifactor to some extent (multiple equity indexes may be included as factors and each firm may have different factor loadings), but for large portfolios an overall average equity factor often drives model results. Moreover, examination of implications of intuitively generated scenarios can be difficult. In contrast, Pesaran, Schuermann, and Treutler’s setup features explicit and observable macroeconomic and industry factors and has a built-in small macroeconomic model. It can be used to study the implications of a variety of shock types. The model implies that credit VaR for a globally and industrially diversified portfolio is quite a bit smaller than credit VaR from one of the standard models for the same portfolio.
Operational Risk

As noted previously, de Fontnouvelle, Jordan, and Rosengren offer evidence that operational risk is a quantitatively important element of the risk management taxonomy. They also examine the properties of different estimators of operational VaR. They consider parametric approaches to estimating quantiles of the operational loss distribution and find that fat-tailed distribution functions perform well in some respects but not in others (thin-tailed functions perform poorly in almost every respect). In contrast, a technique from the extreme value theory (EVT) literature performs well in the most important respects.

Systemic Risk

Practitioners rarely model crisis events or systemic risk, preferring to turn to scenario analysis when they consider such events at all. In contrast, public policymakers are most concerned with such events. Hartmann, Straetmans, and de Vries offer methods and evidence that should be useful to both audiences. Using techniques from EVT, they measure bad-tail co-movements of equity returns of major banks in the United States and in euro-area countries. Where the amplitude of such comovements is large but not associated with catastrophic deteriorations in bank condition, the comovements can be thought of as a form of systematic risk that is of particular interest to credit-risk modelers at bank counterparties and also to market-risk modelers with significant major bank exposures in their portfolios. Where the comovements are very large or indicative of bank distress, the authors’ measures can be thought of as indicators of systemic risk. Though surely not capturing all aspects of systemic risk, any such measures remain useful to students of a subject that has proved resistant to empirical analysis.

Risk Shifting, Risk Transformation, and the Industrial Organization of Finance

It is axiomatic that diversification in portfolios is good, and thus that new opportunities to cost-effectively diversify portfolios are desirable. The creation of new instruments, and entry of formerly specialized financial institutions into each other’s markets (insurance companies into syndicated loans, banks into investment banking, etc.) are to some extent a result of better measurement, which has revealed previously underappreciated opportunities for diversification. Such developments also are a result of the greater transparency and greater feasibility of new instruments that better risk measurement confers.

However, axioms that diversification and innovation are good are sub-
ject to qualification. Diversification of activities as opposed to diversification of portfolios has costs. There is now a large literature that shows that diversified firms are valued less than specialized firms. Recent evidence on diversification within the financial industry shows that it is not clear that shareholders benefit from diversification. At the same time, however, the new ways of managing risk have an impact on the optimal size of institutions. First, modern risk management involves large fixed costs. For example, once a risk measurement and monitoring system is in place to measure the risk of a trading desk, the cost of the system is mostly unaffected by the scale of the positions of the desk. Second, to the extent that a cost of conglomereration is that it is harder to manage a multidivision firm than a single-division firm, the new practices in risk management make it easier to measure and manage risks in conglomerates.

Rather than taking on diversifying activities, firms can shed risks and take on risks within existing activities to increase their level of diversification within these activities. However, managing risks through risk transfer has beneficial systemic effects only to the extent that those who take on the risks are in a better position to bear them than those who shed them. It is not always clear that this is the case when risk transfer is motivated mainly by regulation. Further, the amount of risk transferred may be less than meets the eye because of implicit commitments and because of structures that lack transparency.

In this volume, Allen and Gale’s paper shows that inefficient regulation can lead to risk-transfer activity that is focused on evasion of regulation, and that such activity can increase systemic risk. Gorton and Souleles offer evidence that credit card securitizations do not transfer as much risk as a literal interpretation of such structures might imply, because sponsors enter into an implicit contract to make up the losses suffered by external investors in many states of the world. Considerable tail risk is still transferred, because sponsors will default on the implicit contract when they are near insolvency themselves. However, one can imagine scenarios involving serially correlated shocks to the sponsor’s solvency in which support of securitizations early in the game weakens the sponsor enough that later shocks push it into insolvency.

Franke and Krahnen offer evidence that European securitizations increase the systematic risk exposure of sponsoring banks. Sponsors retain the equity tranche, which absorbs the first losses on the securitized pool of assets. A large fraction of the default risk is retained by the sponsor. The net effect of a securitization on a sponsor’s risk posture depends on the associated investment behavior. If, in a true sale, the sponsor reinvests the proceeds in risk-free assets or to pay down debt, then systematic risk will fall, because the bank has less asset risk or less leverage. If, however, the

sponsor reinvests the proceeds in risky loans of comparable quality, then
the sponsor’s systematic risks increases, because it has a similar portfolio
to the one it had before plus exposure to first losses, which is a high beta
asset. By examining changes in bank betas, Franke and Krahnen offer evi-
dence that systematic risk rises. Although systematic risk is not the same
as systemic risk, so it is not clear that there is a public policy concern, their
finding implies that common assumptions that securitization is risk reduc-
ing for the sponsor may need to be qualified.

As noted previously, modern risk management is providing some of the
impetus for changes in the industrial organization of finance. This volume’s
only study that touches upon consolidation is Beck, Demirgüç-Kunt, and
Levine’s examination of the relationship between systemic stability and
concentration. We discuss their work later rather than here because it does
not examine the effect of risk management on concentration. But Chan,
Getmansky, Haas, and Lo’s paper illuminates the increasing role of hedge
funds, a type of institution that has grown dramatically in recent years. The
evolution in risk measurement no doubt had a role in the growth of hedge
funds. A better understanding of the implications of proprietary trading
risks within diversified financial institutions probably made it less attrac-
tive for such institutions to bear some of these risks. Improvements in risk
measurement also made it easier for stand-alone hedge funds to borrow,
because their lenders could better monitor the risk in hedge fund positions.

Legal Regime, Regulation, Disclosure, and Systemic Stability

The changes in capital markets and the crisis events that led to the new
risk management techniques, as well as the techniques themselves, both
depend upon and influence the legal and regulatory environment in which
they are used. They depend on the environment, because good data and
enforceable contracts are essential to risk measurement and to the engi-
neering of new financial products. They are influencing the environment by
changing how regulators and central bankers think about systemic risk, and
by supporting the development of more risk-sensitive regulatory regimes,
such as Basel II.

Possible Unintended Effects of Regulation and Disclosure

The dramatic progress in financial engineering technology has made reg-
ulation that simplistically specifies required capital for specific positions
increasingly ineffective. As discussed in the context of the Allen and Gale
paper, it may well be that such regulation leads to more rather than less
systemic risk. It also forces the regulators to constantly play catch-up. As
a result, regulation has evolved so that capital requirements depend on
measures of the overall risk taken by an institution rather than on positions
taken by that institution. The obvious difficulty then becomes how risk can
be measured for the purpose of setting capital requirements. Since financial institutions measure risk, it made sense for regulators to try to use their risk measures to set capital. Regulators did so first for market risk with the market risk amendment to the Basel Accord, which came into effect in 1998. Now, with Basel II, they will make some use of internal measures for credit and operational risks.

If banks can use their own risk models, there is a risk that they will manipulate them to lower their capital requirements. Similarly, if risk measures become a part of an institution’s public disclosures, incentives arise to choose measures that window-dress the institution’s risk posture. There is danger that the dialectic between external users and internal incentives can cause risk measures to be less effective tools for management of the institution.

For example, to prevent manipulation of measures that drive capital requirements for market risk, bank regulators have introduced many safeguards, including mandatory backtesting of the VaR model used for regulatory capital. Banks that are too optimistic in their VaR forecasts are penalized, giving banks an incentive to be pessimistic, rather than rewarding precision in VaR estimates by penalizing banks for being too pessimistic as well. Moreover, as discussed in Jorion’s paper, for the measurement of market risk, regulators specify the dataset banks can use as well as how observations can be weighted. For credit risk, under Basel II, regulators will not permit banks to use their internal credit VaR models to set capital, but instead will let them use some inputs to their models as inputs to a simplified regulatory credit VaR model. In none of these cases are banks required to use the regulatory measures for internal management purposes, so perhaps undesirable side effects of regulatory use of internal measures in these cases are modest.

But supporting the general concern, Berkovitz and O’Brien (2002) show that the market risk VaR measures of very large banks in the United States seem to be systematically too conservative. While conservative risk measures might please regulators, since they mean that banks face higher capital requirements, such measures are less useful for managing institutions, since they do not provide an unbiased estimate of risk.

Use of internal measures in regulation also creates concern that the use of similar risk models to satisfy regulators as well as for the management of firms, perhaps as a means of limiting risk-management costs, will stifle innovation in risk management and make risk models less useful. For instance, some of the techniques advocated by Andersen, Bollerslev, Christoffersen, and Diebold would not meet current regulatory requirements, because they put too much weight on recent observations, even though the evidence marshaled by the authors shows that such techniques produce superior measures of risk.

Cost pressures to adopt regulatory measures for internal use could be
especially material if regulators specify types of measures not used internally. Although VaR measures are useful for measuring and managing the distress probability of firms, they may not be the best measures for use in regulation. To see this, consider a firm that faces financial distress if its equity capital falls below a threshold, say $10 billion. If that firm wants its probability of financial distress to be 0.05 percent at the end of the fiscal year, then it should have equity capital above its threshold at least equal to its 0.05 percent VaR at the beginning of its fiscal year. The firm then has a probability of losing its buffer of 0.05 percent and hence a probability of financial distress of 0.05 percent. However, VaR measures are questionable as instruments to set capital requirements. After all, two banks with the same VaR could have vastly different expected losses if the VaR is exceeded. These expected losses have led to a new risk measure, the expected tail loss if VaR is exceeded, called conditional VaR, or CVaR. If two banks have a VaR of $100 million but one bank has a CVaR of $1 billion while the other has a CVaR of $1 million, these two banks would not pose equal threats to the financial system. In many ways, CVaR would be a better risk measure from the perspective of measuring potential systemic threats. However, to date regulators have stuck with measures similar to those in common use at financial institutions, perhaps because CVaR is harder to estimate than VaR, since it requires an estimate of the whole tail of the firmwide loss distribution.

Financial institutions have been prodded toward greater risk transparency by bank regulators, but nonbank institutions have also chosen to be more transparent. Risk transparency has considerable benefits, since it makes it easier for outsiders to monitor the safety of financial institutions and create incentives for those who run them to manage risk well. Unfortunately, transparency has costs also. Rather than focusing institutions on producing unbiased and precise risk measures, it may give them incentives to produce conservative but less useful risk measures. Such an attitude leads to the odd development, in the context of scientific risk measurement, of having institutions declare victory when the number of VaR breaches is too low compared to the expected number.

Stepping back to examine welfare, Pelizzon and Schaefer’s paper reveals that optimal safety-and-soundness regulatory design is sensitive to the sophistication of available risk management and regulatory monitoring and intervention technologies, and to the ability of banks to quickly shift their portfolio risk posture. The relationship between the economic environment and the nature of optimal regulation is not simple. In what might be called the prerisk-management environment, say thirty or more years ago in the United States, bank risk postures were relatively transparent to regulators but also were hard to change, so that a bank in trouble could not quickly shed risk in order to increase the chance of staying solvent. Capital regulations, though crude by modern standards, were relatively hard to
evade. Pelizzon and Schaefer’s results imply that in the old environment, the existence of risk-based capital requirements was more important to welfare than precise calibration of the requirements, and the welfare benefit of early bank-closure rules was not obvious, so that the regulatory environment of the time was perhaps appropriate. But then the new risk measurement and management techniques changed the world drastically. In what one might call a 1990s environment, in which banks could use new instruments and risk-management technologies to easily evade archaic capital regulations, capital regulation itself arguably was welfare-reducing, but early-closure rules were importantly welfare-enhancing (and were implemented in the United States). It is difficult to know the practical implications of Pelizzon and Schaefer’s results for the coming Basel II environment. Much depends on whether the new capital regulations are sufficiently responsive to risk so that they once again become difficult to evade, and also on whether banks will be able to quickly shed risk at low cost in response to an early intervention by regulators.

Systemic Stability

Beck, Demirgüç-Kunt, and Levine’s paper examines the relationship between bank concentration at the national level and systemic stability, wherein instability is measured by the incidence of banking crises. They find that concentration is associated with more stability, not less, as is often claimed, but that the relationship is not a result of the competitive environment. Neither do features of the bank regulatory regime influence stability. They speculate that larger banks are more diversified and thus systems composed of large banks are more stable. We believe their results are also consistent with a view that if risk measurement and management techniques make management of very large banks more feasible, and more such banks appear, systemic stability will be enhanced. But our hypothesis cannot be tested with their data because few banks in the emerging-market nations that dominate their sample employed modern risk management techniques during the sample period.

As noted previously, other papers in this volume contribute to an understanding of the relationship between risk management, regulation, and systemic stability. Jorion offers evidence that practitioner and regulatory use of market VaR measures is not likely to be destabilizing. Berkovitz and O’Brien show that exposures to market risk are typically limited and not highly correlated across firms. Gorton and Souleles, Franke and Krahnen, and Allen and Gale offer evidence and models showing that the details of how securitizations are structured and used are important to their net effect on bank insolvency risk and systemic risk. Hartmann, Straetmans, and de Vries offer measures of the size of systematic relationships among United States and euro-area banks.
Some Speculations about Ways Forward

We believe the new risk measurement and management technology is best viewed as a kind of machinery that, overall, improves welfare by improving the efficiency of financial institutions and by reducing systemic risks in the financial sector. It should be neither feared nor deified. Like any machinery, methods and models that work poorly in the sense of being unrealistic might be harmful in that they might lead to decisions inferior to those associated with better methods and models. It is likely that much more time and experience will be needed before many kinds of crisis events can be adequately captured in risk measures, and it is even possible that risk models will, in effect, malfunction during some crises. These will be growing pains. The way forward is to diagnose weaknesses in measures, models, and management methods, fix them as necessary, and improve them when possible.

The machinery we have discussed creates risks also, however. An institution that is well equipped to measure and manage risks can increase risk, as well as decrease it, more efficiently than an institution that is not well equipped. While risk transparency would seem to make it harder for institutions to take on too much risk, risk measures can be manipulated and transparency has costs. There is always a danger that measuring risk carefully, with well-defined risk measures, just pushes risk where it is not measured. More-detailed regulations of risk measurement are unlikely to prevent these problems, as the resourcefulness of financial engineers knows few bounds. Ultimately, a financial institution’s governance plays a key role in ensuring that its risk position is optimal from the perspective of its owners. As long as regulations do not make excessive risk taking optimal and as long as financial institutions are well governed, we would expect improved risk measurement and management to enhance welfare.

Better risk measurement should be a continuing part of the research agenda, as well as better understanding of how to optimize the legal environment and regulatory policies and practices. Though VaR and stress tests dominate risk management now, the motivation for the use of these tools is mostly practical. In principle, risk management should help firms take risks that make money for them and shed those that do not. It is not clear that VaR and stress tests are the best solution for profit maximization.

Though regulators and central banks have been ready to deal with the classic systemic crisis involving bank runs, they are acquiring new roles as they are called upon to ensure systemic liquidity in all kinds of crisis situations and to act as coordinating agents in the diagnosis and repair of systemic problems. Often such coordination does not involve regulation but rather a fostering of technical and institutional advances. In the past decade or so, both regulators and market participants seem to have be-
come increasingly comfortable with such a role for the official sector, and the official sector has played a significant role in the developments we have discussed. Yet this role of the official sector does raise a troubling issue. If the official sector is an instrument of progress in risk management, why is it that private firms could not make such progress on their own? Is it because the official sector values risk management more than private firms because of externalities, so that there is an implicit subsidy from the public sector to the development of risk management? Or is it that many private firms value risk management too little because of governance failures? Or that free-rider problems interfere with the uncoordinated development of certain kinds of risk management innovations?

References


