
Credit Risk Rating at Large U.S. Banks

William F. Treacy, of the Board's Division of Banking Supervision and Regulation, and Mark S. Carey, of the Board's Division of Research and Statistics, prepared this article.

Internal credit ratings are becoming increasingly important in credit risk management at large U.S. banks. Banks' internal ratings are somewhat like ratings produced by Moody's, Standard & Poor's, and other public rating agencies in that they summarize the risk of loss due to failure by a given borrower to pay as promised.¹ However, banks' rating systems differ significantly from those of the agencies (and from each other) in architecture and operating design as well as in the uses to which ratings are put. One reason for these differences is that banks' ratings are assigned by bank personnel and are usually not revealed to outsiders.²

For large banks, whose commercial borrowers may number in the tens of thousands, internal ratings are an essential ingredient in effective credit risk management.³ Without the distillation of information that ratings represent, any comparison of the risk posed by such a large number of borrowers would be extremely difficult because of the need to simulta-

neously consider many risk factors for each of the many borrowers. Most large banks use ratings in one or more key areas of risk management that involve credit, such as guiding the loan origination process, portfolio monitoring and management reporting, analysis of the adequacy of loan loss reserves or capital, profitability and loan pricing analysis, and as inputs to formal portfolio risk management models. Banks typically produce ratings only for business and institutional loans and counterparties, not for consumer loans or other assets.

In short, risk ratings are the primary summary indicator of risk for banks' individual credit exposures. They both shape and reflect the nature of credit decisions that banks make daily. Understanding how rating systems are conceptualized, designed, operated, and used in risk management is thus essential to understanding how banks perform their business lending function and how they choose to control risk exposures.⁴

The specifics of internal rating system architecture and operation differ substantially across banks. The number of grades and the risk associated with each grade vary across institutions, as do decisions about who assigns ratings and about the manner in which rating assignments are reviewed. In general, in designing rating systems, bank management must weigh numerous considerations, including cost, efficiency of information gathering, consistency of ratings produced, staff incentives, the nature of the bank's business, and the uses to be made of internal ratings.

A central theme of this article is that, to a considerable extent, variations across banks are an example of form following function. There does not appear to be one "correct" rating system. Instead, "correctness" depends on how the system is used. For example, a bank that uses ratings mainly to identify deteriorating or problem loans to ensure proper monitoring may find that a rating scale with relatively few grades is adequate. In contrast, if ratings are used in computing

1. For example, bonds rated Aaa on Moody's scale or AAA on Standard & Poor's scale pose negligible risk of loss in the short to medium term, whereas those rated Caa or CCC are quite risky.

2. For additional information about the internal rating systems of large and smaller banks, see Thomas F. Brady, William B. English, and William R. Nelson, "Recent Changes to the Federal Reserve's Survey of Terms of Business Lending," *Federal Reserve Bulletin*, vol. 84 (August 1998), pp. 604-15; see also William B. English and William R. Nelson, "Bank Risk Rating of Business Loans" (Board of Governors of the Federal Reserve System, April 1998).

For information about the rating systems of large banks and about credit risk management practices in general, see Robert Morris Associates and First Manhattan Consulting Group, *Winning the Credit Cycle Game: A Roadmap for Adding Shareholding Value Through Credit Portfolio Management* (1997).

For a survey of the academic literature on ratings and credit risk, see Edward I. Altman and Anthony Saunders, "Credit Risk Measurement: Developments over the Last 20 Years," *Journal of Banking and Finance*, vol. 21 (December 1997), pp. 1721-42.

3. See the Federal Reserve's Supervision and Regulation Letter SR 98-25, "Sound Credit Risk Management and the Use of Internal Credit Risk Ratings at Large Banking Organizations" (September 21, 1998), which stresses the importance of risk rating systems for large banks and describes elements of such systems that are "necessary to support sophisticated credit risk management" (p. 1). SR Letters are available on the Federal Reserve Board's web site, <http://www.federalreserve.gov>.

4. Credit risk can arise from a loan already extended, loan commitments that have not yet been drawn, letters of credit, or obligations under other contracts such as financial derivatives. This article follows industry usage by referring to individual loans or commitments as "facilities" and overall credit risk arising from such transactions as "exposure."

internal profitability measures, a scale with a relatively large number of grades may be required to achieve fine distinctions of credit risk.

As with the decision to extend credit, the rating process almost always involves the exercise of human judgment because the factors considered in assigning a rating and the weight given each factor can differ significantly across borrowers. Given the substantial role of judgment, banks must pay careful attention to the internal incentives they create and to internal rating review and control systems to avoid introducing bias. The direction of such bias tends to be related to the functions that ratings are asked to perform in the bank's risk management process. For example, at banks that use ratings in computing profitability measures, establishing pricing guidelines, or setting loan size limits, the staff may be tempted to assign ratings that are more favorable than warranted.

Many banks use statistical models as an element of the rating process, but banks generally believe that the limitations of statistical models are such that properly managed judgmental rating systems deliver more accurate estimates of risk. Especially for large exposures, the benefits of such accuracy may outweigh the higher costs of judgmental systems. In contrast, statistical credit scores are often the primary basis for credit decisions for small lending exposures, such as consumer credit.

Although form generally follows function in the systems used to rate business loans, our impression is that in some cases the two are not closely aligned. For example, because of the rapid pace of change in the risk management practices of large banks, their rating systems are increasingly being used for purposes for which they were not originally designed. When a bank applies ratings in a new way, such as in risk-sensitive analysis of business line profitability, the existing ratings and rating system are often used as-is. It may become clear only over time that the new function has imposed new stresses on the rating system and that changes in the system are needed.

Several conditions appear to magnify such stresses on bank rating systems. The conceptual meaning of ratings may be somewhat unclear, rating criteria may be largely or wholly maintained as a matter of culture rather than formal written policy, and corporate databases may not support analysis of the relationship between grade assignments and historical loss experience. Such circumstances make ratings more difficult to review and audit and also require loan review units in effect to define, maintain, and fine-tune rating standards in a dynamic fashion.

This article describes internal rating systems at large U.S. banks, focusing on the relationship

between form and function, the stresses that are evident, and the current conceptual and practical barriers to achieving accurate, consistent ratings. We hope to promote understanding of this critical element of risk management—among the industry, supervisors, academics, and other interested parties—and thereby promote further enhancements to risk management.

This article is based on information from internal reports and credit policy documents for the fifty largest U.S. bank holding companies, from interviews with senior bankers and others at more than fifteen major holding companies and other relevant institutions, and from conversations with Federal Reserve bank examiners. The institutions we interviewed cover the spectrum of size and practice among the fifty largest banks, but a disproportionate share of the banks we interviewed have relatively advanced internal rating systems.⁵

THE ARCHITECTURE OF BANK INTERNAL RATING SYSTEMS

In choosing the architecture of its rating system, a bank must decide which loss concepts to employ, the number and meaning of grades on the rating scale corresponding to each loss concept, and whether to include “watch” and “regulatory” grades on such scales. The choices made and the reasons for them vary widely, but on the whole, the primary determinants of bank rating system architecture appear to be the bank's mix of large and smaller borrowers and the extent to which the bank uses quantitative systems for credit risk management and profitability analysis. In principle, banks must also decide whether to grade borrowers according to their current condition or their expected condition under stress. Although the rating agencies employ the latter, “through the cycle,” philosophy, almost all banks have chosen to grade to current condition (see the box “Point-in-Time vs. Through-the-Cycle Grading”).

Loss Concepts and Their Implementation

The credit risk of a loan or other exposure over a given period involves both the *probability of default* (PD) and the fraction of the loan's value that is likely to be *lost in the event of default* (LIED). LIED is always specific to a given facility because it depends

5. Internal rating systems are typically used throughout U.S. banking organizations. For brevity, we use the term “bank” to refer to consolidated banking organizations, not just the chartered bank.

Point-in-Time vs. Through-the-Cycle Grading

A common way of implementing a long-horizon, through-the-cycle rating philosophy involves estimating the borrower's condition at the worst point in an economic or industry cycle and grading according to the risk posed at that point. Although "downside" or "borrower stress" scenarios are an element of many banks' underwriting decisions, every bank we interviewed bases risk ratings on the borrower's current condition. Rating the current condition is consistent with the fact that rating criteria at banks do not seem to be updated to take account of the current phase of the business cycle. Banks we interviewed do vary somewhat in the time period they have in mind when producing ratings, with about 25 percent rating the borrower's risk over a one-year period, 25 percent rating over a longer period such as the life of the loan, and the remaining 50 percent having no specific period in mind. How closely raters adhere to time horizon guidelines at banks that have them is not clear.

In contrast to bank practice, both Moody's and S&P rate through the cycle. They analyze the borrower's current condition at least partly to obtain an anchor point for determining the severity of the downside scenario. The borrower's projected condition in the event the downside scenario occurs is the primary determinant of the rating. Only borrowers that are very weak at the time of the analysis are rated primarily according to current condition. Under this philosophy, the migration of borrowers' ratings up and down the scale as the overall economic cycle progresses will be muted: Ratings will change mainly for those firms that experience good or bad shocks that affect long-term condition or financial strategy and for those

whose original downside scenario was too optimistic. The agencies' through-the-cycle philosophy probably accounts for their considerable emphasis on a borrower's industry and its position within the industry. For many firms, industry supply and demand cycles are as important or more important than the overall business cycle in determining cash flow.

In interviews, we did not discuss the reasons that banks rate to current condition, but two possibilities are the greater difficulty of the agency method and differences in the investment horizon of banks relative to that of users of agency ratings. Consistency of ratings across a wide variety of credits may be easier to achieve when the basis is the relatively easy-to-observe current condition. Also, greater difficulty means through-the-cycle grading entails greater expense, and for many middle-market credits the extra expense might render such lending unprofitable for banks.

Regarding investment horizon, the rating agencies' philosophy may reflect the historical preponderance of long-term, buy-and-hold investors among users of ratings. Such users are naturally most interested in estimates of long-term credit risk. That banks should naturally have a short-term orientation is not clear, especially as the maturity of bank loan commitments has increased steadily over the past decade or two. If it were not for the considerations of feasibility and cost, as well as the fact that many banks use ratings to guide the intensity of monitoring of borrowers, the banks' choice of point-in-time grading would be more debatable.

on the structure of the facility. PD, however, is generally associated with the borrower, the presumption being that a borrower will default on all obligations if it defaults on any.⁶ The product of PD and LIED is the *expected loss* (EL) on the exposure in a statistical sense. It represents an estimate of the average percentage loss rate over time on a group of loans all having the given expected loss. A positive expected loss is *not*, however, a forecast that losses will in fact occur on any individual loan.

The banks at which we conducted interviews fall into two categories with regard to loss concept. About 60 percent have one-dimensional rating systems, in which ratings are assigned to facilities. In such systems, ratings approximate EL. The remaining

40 percent have two-dimensional systems, in which the borrower's general creditworthiness (approximately PD) is appraised on one scale while the risk posed by individual exposures (approximately EL) is appraised on another; invariably the two scales have the same number of rating categories.⁷

A number of banks would no doubt dispute our characterization of their single-scale systems as measuring EL; in interviews, several maintained that their ratings primarily reflect the borrower's PD. However, collateral and loan structure play a role in grading at such banks both in practical terms and in the definitions of grades. Moreover, certain specialty loans—such as cash-collateralized loans, those eligible for government guarantees, and asset-based loans—can receive relatively low risk grades, a distinction reflecting the fact that the EL for such loans

6. Admittedly, PD might differ across transactions with the same borrower. For example, a borrower may attempt to force a favorable restructuring of its term loan by halting payment on the loan while continuing to honor the terms of a foreign exchange swap with the same bank. However, for practical purposes, estimating a single probability of any default by a borrower is usually sufficient.

7. The policy documents of banks we did not interview indicate that they also have one- or two-dimensional rating systems, and our impression is that the discussion of loss concepts above applies equally well to these banks.

1. Example of a two-dimensional risk rating system using average LIED values

Grade	Borrower scale: borrower's probability of default (PD) (percent) (1)	Assumed average loss on loans in the event of default (LIED) (percent) (2)	Facility scale: expected loss (EL) on loans (percent) (1 × 2)
1—Virtually no risk ...	0	↑ 30 ↓	0
2—Low risk1		.03
3—Moderate risk3		.09
4—Average risk	1.0		.30
5—Acceptable risk	3.0		.90
6—Borderline risk	6.0		1.80
7—OAEM ¹	20.0		6.00
8—Substandard	60.0		18.00
9—Doubtful	100		30.00

1. Other Assets Especially Mentioned.

is far less than for an “ordinary” loan to the same borrower. Such single-grade systems might be most accurately characterized as having an ambiguous or mixed conceptual basis rather than as clearly measuring either PD or EL. Although an ambiguous basis may pose no problems when ratings are used mainly for administrative and reporting purposes and when the nature of the bank’s business is fairly stable over time, a clear conceptual foundation becomes more important as quantitative models of portfolio risk and profitability are used more heavily and during periods of rapid change.

In two-dimensional systems, one grade typically reflects PD and the other EL. Banks with such systems usually first determine the borrower’s grade (its PD) and then set the facility grade equal to the borrower grade unless the structure of the facility is such that LIED is substantially better or worse than “normal.” Implicitly, grades on the facility scale measure EL as the PD associated with the borrower

grade multiplied by a standard or average LIED (table 1). In this way, a two-dimensional system can promote precision and consistency in grading by separately recording a rater’s judgments about PD and EL rather than mixing them together.

A few banks said they had plans to shift to a system in which the borrower grade reflects PD but the facility grade explicitly measures LIED. The rater would assign a facility to one of several LIED categories on the basis of the likely recovery rates associated with various types of collateral, guarantees, or other considerations associated with the facility’s structure. EL for a facility would be calculated by multiplying the borrower’s PD by the facility’s LIED.⁸

Rating Scales at Moody’s and S&P

At the agencies, as at many banks, the loss concepts (PD, LIED, and EL) embedded in the ratings are somewhat ambiguous. Moody’s states that “ratings are intended to serve as indicators or forecasts of the potential for *credit loss* because of failure to pay, a delay in payment, or partial payment.” Standard & Poor’s states that its ratings are an “opinion of the general creditworthiness of an obligor, or . . . of an obligor with respect to a particular . . . obligation . . .

8. Systems recording LIED rather than EL as the second grade can promote precision and consistency in grading. PD-EL systems typically impose limits on the degree to which differences in loan structure permit an EL grade to be moved up or down relative to the PD grade. Such limits can be helpful in restraining raters’ optimism but, in the case of loans with a genuinely very low expected LIED, such limits can materially limit the accuracy of risk measurement. Another benefit of LIED ratings is the fact that raters’ LIED judgments can be evaluated over time by comparing them to loss experience.

2. Moody’s and Standard & Poor’s bond rating scales and average one-year default rates

Category	Moody’s		Standard & Poor’s	
	Grade	Average default rate (PD) per year, 1970–95 (percent)	Grade	Average default rate (PD) per year, 1981–94 (percent)
Investment grade	Aaa	.00	AAA	.00
	Aa, Aa1, Aa2, Aa3	.03	AA+, AA, AA–	.00
	A, A1, A2, A3	.01	A+, A, A–	.07
	Baa, Baa1, Baa2, Baa3	.13	BBB+, BBB, BBB–	.25
Below investment grade (“junk”)	Ba, Ba1, Ba2, Ba3	1.42	BB+, BB, BB–	1.17
	B, B1, B2, B3	7.62	B+, B, B–	5.39
	Caa, Ca, C	n.a.	CCC, CC, C	19.96
	Default	D

NOTE. Grades are listed from less risky to more risky, from top to bottom and from left to right.
 n.a. Not available.
 . . . Not applicable.

SOURCE. Moody’s Investors Service Special Report, *Corporate Bond Defaults and Default Rates 1938–1995* (January 1996). Standard & Poor’s Creditweek Special Report, *Corporate Defaults Level Off in 1994* (May 1, 1995).

based on relevant risk factors.” On balance, a close reading of Moody’s and Standard & Poor’s detailed descriptions of rating criteria and procedures suggests that the two agencies’ ratings incorporate elements of PD and LIED but are not precisely EL measures.⁹

Risk tends to increase nonlinearly on both bank and agency scales. For example, on the agency scales, default rates are low for the least risky grades but rise rapidly as the grade worsens (table 2).

Administrative Grades

All the banks we interviewed maintain some sort of internal “watch” list as well as a means of identifying assets that fall into the “regulatory problem asset” categories (table 3). Although watch and regulatory problem-asset designations typically identify high-risk credits, they have administrative meanings that are conceptually separate from risk per se. Special monitoring activity is usually undertaken for watch and problem assets, such as formal quarterly reviews of status and special reports that help senior bank management monitor and react to important developments in the portfolio. However, banks may wish to trigger special monitoring for credits that are not high-risk and thus may wish to separate administrative indicators from risk measures (an example would be a low-risk loan for which an event that might influence risk is expected, such as a change in ownership of the borrower).

Among the fifty largest banks, all but two have grades corresponding to the regulatory problem-asset categories Other Assets Especially Mentioned (OAEM), Substandard, Doubtful, and Loss (some omit the Loss category).¹⁰ All other assets are collectively labeled “Pass” by regulators. The bank supervisory agencies do not specifically require that banks maintain regulatory categories on an internal scale but do require that recordkeeping be sufficient to ensure that loans in the regulatory categories can be quickly and clearly identified. The two banks that use procedures not involving internal grades appear to do so because the regulatory asset categories are not consistent with the conceptual basis of their own

9. Moody’s Investors Service, *Global Credit Analysis* (IFR Publishing, 1991), p. 73 (emphasis in the original); Standard & Poor’s, *Corporate Ratings Criteria* (1998), p. 3. Other rating agencies play important roles in the marketplace. We omit details of their scales and practices only for brevity.

10. A few break Substandard into two categories, one for performing loans and the other for nonperforming loans.

3. Regulatory problem asset categories

Category	Regulatory definition	Recommended specific reserve (percent)
Special Mention (OAEM) ¹ . . .	Has potential weaknesses that deserve management’s close attention. If left uncorrected, these potential weaknesses may, <i>at some future date</i> , result in the deterioration of the repayment prospects for the credit.	No recommendation
Substandard	Inadequately protected by current worth-paying capacity of obligor or collateral. Well-defined weaknesses jeopardize liquidation of the debt. Distinct possibility that bank will sustain some loss if deficiencies are not corrected.	15
Doubtful	All weaknesses inherent in substandard, AND collection/liquidation in full, on basis of currently existing conditions, is highly questionable or improbable. Specific pending factors may strengthen credit; treatment as loss deferred until exact status can be determined.	50
Loss	Uncollectible and of such little value that continuance as bankable asset is not warranted. Credit may have recovery or salvage value, but not practical/desirable to defer writing it off even though partial recovery may be effected in future.	100

NOTE. Assets that do not fall into one of these categories are termed Pass by the federal banking regulators.

1. Other Assets Especially Mentioned.

grades.¹¹ Moreover, banks and regulators may sometimes disagree about the riskiness of individual assets that fall into the various regulatory grades.¹²

Watch credits are those that need special monitoring but do not fall in the regulatory problem-asset grades. Only about half the banks we interviewed include a watch grade on their internal rating scales. Others add a watch flag to individual grades, such as 3W versus 3, or simply maintain a watch list separately, perhaps by adding an identifying field to their computer systems.

11. Although the definitions are standardized across banks, our discussions and inspection of internal documents imply that banks vary in their internal definition and use of OAEM. Among the regulatory categories, OAEM in particular can have an administrative dimension as well as a risk dimension. Most loans identified as OAEM pose a higher-than-usual degree of risk, but some loans may be placed in this category for lack of adequate documentation in the loan file, which may occur even for loans not posing higher-than-usual risk. In such cases, once the administrative problem is resolved, the loan can be upgraded.

12. Examiners review problem loans and evaluate whether they have been assigned to the proper regulatory problem-asset grades and also review a sample of Pass credits. Examiners heretofore have generally not attempted to validate or evaluate internal ratings of Pass credits.

Number of Grades on the Scale

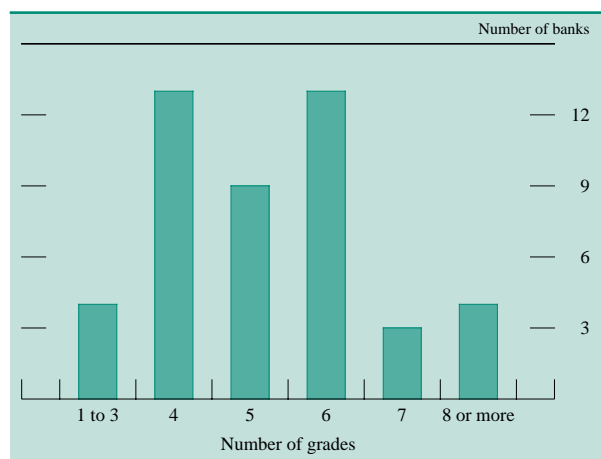
The number of grades on internal scales varies considerably across banks. In addition, even where the number of grades is identical on two different banks' scales, the risk associated with the same grades (for example, two loans graded 4) is almost always different. Among the fifty largest banks, the number of Pass grades varies from two to the low twenties. The median is five Pass grades, including a watch grade if any (chart 1). Among the ten largest banks, the median number of Pass grades is six and the minimum is four. As noted, the vast majority of large banks also include three or four regulatory problem-asset grades on their internal scales.

Internal rating systems with larger numbers of grades are more costly to operate because of the extra work required to distinguish finer degrees of risk. Banks making heavy use of ratings in analytical activities are most likely to choose to bear these costs because fine distinctions are especially valuable in such activities (however, at least a moderate number of Pass grades is useful even for internal reporting purposes). Banks that increase their analytical use of ratings may persist for a while with a relatively small number of Pass grades because the costs of changing rating systems can be large. Nonetheless, those banks that have recently redesigned their rating systems have all increased the number of grades.¹³

The proportion of grades used to distinguish among relatively low risk credits versus the proportion used

13. The average number of grades on internal scales appears to have increased somewhat during the past decade. See Gregory F. Udell, *Designing the Optimal Loan Review Policy: An Analysis of Loan Review in Midwestern Banks* (Prochnow Reports, Madison, Wis., 1987), p. 18.

1. Fifty largest U.S. banks, distributed by number of Pass grades



NOTE. Shown are the forty-six banks for which this measure was available.

to distinguish among the riskier Pass credits tends to differ with the business mix of the bank. Among banks we interviewed, those that do a significant share of their commercial business in the large corporate loan market tend to have more grades reflecting investment-grade risks. The allocation of grades between the investment-grade and below-investment-grade categories tends to be more even at banks doing mostly middle-market business.¹⁴ The differences are not large: The median middle-market bank has three internal grades corresponding to agency grades of BBB-/Baa3 or better and three riskier grades, whereas the median bank with a substantial large-corporate business has four investment grades and two junk grades. Such a difference in rating system focus is sensible in that an ability to make fine distinctions among low-risk borrowers is quite important in the highly competitive large-corporate lending market. In the middle market, fewer borrowers are perceived as posing AAA, AA, or even A levels of risk, so such distinctions are less crucial.

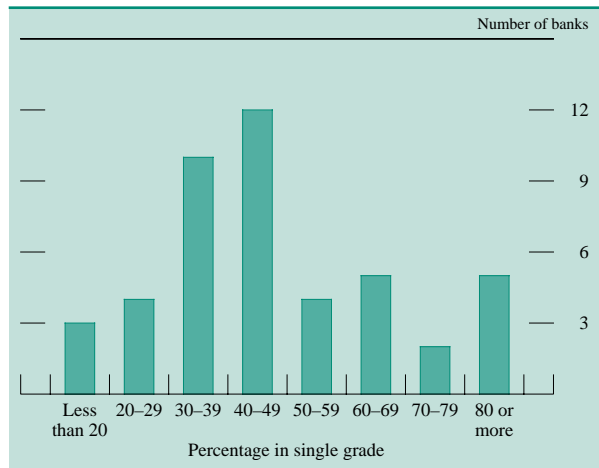
However, a glance at table 2 reveals that a good distinction among risk levels in the below-investment-grade range is important for all banks. For example, the range of default rates spanned by the agency grades BB+/Ba1 through B-/B3 is orders of magnitude larger than the risk range for, say, A+/A1 through BBB-/Baa3, and yet the median large bank we interviewed uses only two or three grades to span the below-investment-grade range, one of them perhaps being a watch grade. More granularity—finer distinctions of risk, especially among riskier assets—can enhance a bank's ability to analyze its portfolio risk posture and to construct accurate models of the profitability of its broader business relationships with borrowers.

Systems with many Pass categories are less useful when loans or other exposures tend to be concentrated in one or two grades. Among large banks, sixteen institutions, or 36 percent, assign half or more of their rated loans to a single risk grade (chart 2). Such systems appear to contribute little to the understanding and monitoring of risk posture.¹⁵

14. The term "large corporate" includes nonfinancial firms with large annual sales volumes as well as large financial institutions, national governments, and large nonprofit institutions. Certainly the Fortune 500 firms fall into this category. Middle-market borrowers are smaller, but the precise boundary between large and middle-market and between middle-market and small business borrowers varies by bank.

15. Such failure to distinguish degrees of risk was recently cited in Federal Reserve examination guidance as a potentially significant shortcoming in a large institution's credit risk management process. See Supervision and Regulation Letter SR 98-18, "Lending Standards for Commercial Loans" (June 23, 1998). For additional information

2. Fifty largest U.S. banks, distributed by percentage of outstandings placed in the grade with the most outstandings



NOTE. Shown are the forty-five banks for which this measure was relevant.

The majority of the banks that we interviewed (and, based on discussions with supervisory staff, other banks as well) expressed at least some desire to increase the number of grades on their scales and to reduce the extent to which credits are concentrated in one or two grades. Two kinds of plans were voiced: Addition of a +/- modifier to all existing grades, and a split of existing riskier grades into a larger number of newly defined grades, leaving the low-risk grades unchanged.¹⁶ The +/- modifier approach is favored by many because grade definitions are modified rather than completely reorganized. For example, the basic meaning of a 5 stays the same, but it becomes possible to distinguish between a strong and a weak 5 with grades of 5+ and 5-. This approach limits the disruption of staff understanding of each grade's meaning (as noted below, such understanding is largely cultural rather than being formally written).

THE OPERATING DESIGN OF RATING SYSTEMS

In essentially all cases, the human judgment exercised by experienced bank staff is central to the

about current bank lending practices, see William F. Treacy, "The Significance of Recent Changes In Underwriting Standards: Evidence from the Loan Quality Assessment Project," *Federal Reserve System Supervisory Staff Report* (June 1998); and U.S. Comptroller of the Currency, *1998 Survey of Credit Underwriting Practices* (National Credit Committee, 1998).

16. At the time of the interviews, however, the majority of the banks voicing plans to increase the number of their grades had no active effort in progress. Many of those institutions actively moving to increase the number of their Pass grades do not now have concentrations in a single category.

assignment of a rating. Banks thus design the operational flow of the rating process in ways that are aimed at promoting the accuracy and consistency of ratings while not unduly restricting the exercise of judgment. Balance between these opposing imperatives appears to be struck at each institution on the basis of cost considerations, the nature of the bank's commercial business lines, the bank's uses of ratings, and the role of the rating system in maintaining the bank's credit culture.

Key operating design issues in striking the balance include the organizational division of responsibility for grading (line staff or credit staff), the nature of reviews of ratings to detect errors, the organizational location of ultimate authority over grade assignments, the role of external ratings and statistical models in the rating process, and the formality of the process and specificity of formal rating definitions.

What Exposures Are Rated?

At most banks, ratings are produced for all commercial or institutional loans (that is, not consumer loans), and in some cases for large loans to households or individuals for which underwriting procedures are similar to those for commercial loans. Rated assets thus include commercial and industrial loans and other facilities, commercial lease financings, commercial real estate loans, loans to foreign commercial and sovereign entities, loans and other facilities to financial institutions, and sometimes loans made by "private banking" units. In general, ratings are applied to those types of loans for which underwriting requires large elements of subjective analysis.

Overview of the Rating Process in Relation to Credit Approval and Review

Ratings are typically assigned (or reaffirmed) at the time of each underwriting or credit approval action. The analysis supporting the ratings is inseparable from the analysis supporting the underwriting or credit approval decision. In addition, the rating and underwriting processes, while logically separate, are intertwined. The rating assignment influences the approval process in that underwriting limits and approval requirements depend on the grade, while approvers of a credit are expected to review and confirm the grade. For example, an individual staff member typically proposes a risk grade as part of the pre-approval process for a new credit. The proposed grade is then approved or modified at the same time

that the transaction itself receives approval and must meet the requirements embedded in the bank's credit policies. In nearly all cases, approval requires assent by individuals with requisite "signature authority" rather than by a committee. The number and level of signatures needed for approval typically depend on the size and (proposed) risk rating of the transaction: In general, less risky loans require fewer and perhaps lower-level signatures. In addition, signature requirements may vary according to the line of business involved and the type of credit being approved.¹⁷

After approval, the individual that assigned the initial grade is generally responsible for monitoring the loan and for changing the grade promptly as the condition of the borrower changes. Exposures falling into the regulatory grades are an exception at some institutions, where monitoring and grading of such loans becomes the responsibility of a separate unit, such as a workout or loan review unit.

Who Assigns and Monitors Ratings, and Why?

Ratings are initially assigned either by relationship managers or the credit staff. Relationship managers (RMs) are lending officers (line staff) responsible for the marketing of banking services. They report to lines of business that reflect the strategic orientation of the bank.¹⁸ All institutions evaluate the performance of RMs—and thus set their compensation—on the basis of the profitability of the relationships in question, although the methods of assessing profitability and determining compensation vary. Even when profitability measures are not risk-sensitive, ratings assigned by an RM can affect his or her compensation.¹⁹ Thus, in the absence of sufficient controls, RMs may have incentives to assign ratings in a manner inconsistent with the bank's interests.

The credit staff is responsible for approving loans and the ratings assigned, especially in the case of larger loans; for monitoring portfolio credit quality and sometimes for regular review of individual exposures; and sometimes for directly assigning the ratings of individual exposures. The credit staff is

genuinely independent of sales and marketing functions when the two have separate reporting structures (that is, "chains of command") and when the performance assessment of the credit staff is linked to the quality of the bank's credit exposure rather than to loan volume or business line or customer profitability. Some banks apportion the credit staff across specific line-of-business groups. Such arrangements allow for closer working relationships but in some cases lead to linkage of the credit staff's compensation or performance assessment with profitability of the business line; in such cases, incentive conflicts like those experienced by RMs can arise. At other banks, RMs and independent credit staff produce ratings as partners and are held jointly accountable. Whether such partnerships are effective in restraining incentive conflicts is not clear.

The primary responsibility for rating assignments varies widely among the banks we interviewed. RMs have the primary responsibility at about 40 percent of the banks, although in such cases the credit staff may review proposed ratings as part of the loan approval process, especially for larger exposures.²⁰ At 15 percent of interviewed banks the credit staff assigns all initial ratings, whereas the credit staff and RMs rate in partnership at another 20 percent or so. About 30 percent of interviewed banks divide the responsibility: The credit staff has sole responsibility for rating large exposures, and RMs alone or in partnership with the credit staff rate middle-market loans. In principle, both the credit staff and RMs use the same rating definitions and basic criteria, but the different natures of the two types of credit may lead to some divergence of practice.

A bank's business mix appears to be a primary determinant of whether RMs or the credit staff are primarily responsible for ratings. Those banks we interviewed that lend mainly in the middle market usually give RMs primary responsibility for ratings. Such banks emphasized informational efficiency, cost, and accountability as key reasons for their choice of organizational structure. Especially in the case of loans to medium-size and smaller firms, the RM was said to be in the best position to appraise the condition of the borrower on an ongoing basis and thus to ensure that ratings are updated in a timely manner. Requiring that the credit staff be equally well informed adds costs and may introduce lags into the process by which ratings of such smaller credits are updated.

17. If those asked to provide signatures believe that a loan should be assigned a riskier internal rating than initially, additional signatures may be required in accordance with policy requirements. Thus, disagreement over the rating can alter the approval requirements for the loan in question.

18. Lines of business may be defined by the size of the business customer (such as large corporate), by the customer's primary industry (such as health care), or by the type of product being provided (such as commercial real estate loans).

19. For example, because loan policies often include size limits that depend on ratings, approval of a large loan proposed by an RM may be much more likely if it is assigned a relatively low risk rating.

20. At most banks, RMs have signature authority for relatively small loans, and the credit staff might review the ratings of only a fraction of small loans at origination.

The institutions at which an independent credit staff assigns ratings tend to have a substantial presence in the large corporate market. Placing the rating process primarily in the hands of the credit staff offers greater assurance that grading will be purely on the basis of risk, without coloration by possible ramifications for customer or business line profitability. In addition, because the credit staff is small relative to the number of RMs and is focused entirely on risk assessment, it is in a better position to achieve consistency in its ratings (that is, to assign similar grades to similarly risky loans, regardless of their other characteristics). Moreover, the costs of having the credit staff perform all analysis are small relative to the revenues generated by large corporate loan transactions. In contrast, such costs can be large relative to the transaction revenues for middle-market loans.

Our impression is that middle-market lending represents a much larger share of the business of banks we did not interview. If the pattern described above holds, the proportion of all large banks using RM-centered rating processes is probably higher than among our interviewees. Unfortunately, policy documents for those we did not interview generally do not reveal details of this aspect of the process.

Almost all the banks we interviewed are at least experimenting with consumer-loan-style credit scoring models for small commercial loans. For exposures smaller than some cutoff value, such models are either a tool in the rating process or are the sole basis for the rating. If, however, models are the sole basis, performing loans are usually assigned to a single grade on the internal rating scale rather than making grade assignments sensitive to the score value.

How Do They Arrive at Ratings?

Both assigners and reviewers of ratings follow the same basic thought process in arriving at a rating for a given exposure. The rater considers both the risk posed by the borrower and aspects of the facility's structure. In appraising the borrower, the rater gathers information about its quantitative and qualitative characteristics, compares them with the standards for each grade, and then weights them in choosing a borrower grade. The comparative process often is as much one of looking across borrowers as one of looking across characteristics of different grades: That is, the rater may look for already-rated loans with characteristics close to those of the loan being rated and then set the rating to the grade already assigned to such borrowers.

Models and Judgment

Although in principle the analysis of risk factors may be done by a mechanical model, in practice the rating process at almost all banks relies heavily on judgment. We suspect most banks are hesitant to make models the centerpiece of their rating systems for three reasons: (1) Different models would be required for each asset class and perhaps for different geographic regions; (2) data to support estimation of such models is currently rarely available; and (3) the reliability of such models would become apparent only over time, exposing the bank to possibly substantial risks in the interim. Those few banks moving toward heavy reliance on models appear to feel that models produce more consistent ratings and that, in the long run, operating costs will be reduced in that less labor will be required to produce ratings.

As part of their judgmental evaluation, most of the banks we interviewed either use statistical models of borrower default probability as an input (about three-fourths do so) or take into consideration any available agency rating of the borrower (at least half, and probably more, do so). Such use of external points of comparison is common for large corporate borrowers because they are most likely to be externally rated and because statistical default probability models are more readily available for such borrowers. In addition, as described further below, many banks use external ratings or models in calibrating their rating systems and in identifying likely mistakes in grade assignments.

Factors Considered

Bank personnel base their decisions to assign a particular rating on the criteria that define each grade, which are articulated as standards for a number of specific risk factors. For example, a criterion for assignment of a grade "3" might be that the borrower's leverage ratio must be smaller than some value. Risk factors include the borrower's financial condition, size, industry, and position within the industry; the reliability of the borrower's financial statements and the quality of its management; elements of transaction structure (for example, collateral); and miscellaneous other factors. The risk factors are generally the same as those considered in deciding whether to extend a loan and are similar to the factors considered by rating agencies. Banks vary somewhat in the particular factors they consider and in the weight they give each factor. What follows is a

description of the factors considered by a typical bank among those we interviewed.²¹

Financial statement analysis is central to appraising the likely adequacy of future cash flow and thus the ability of the borrower to service its debt. The focus of analysis is on the borrower's debt service capacity, taking account of its free cash flow, the liquidity of its balance sheet, and the firm's access to sources of finance other than the bank. Historical (and to a lesser extent, projected) earnings, operating cash flow, interest coverage, and leverage are typically analyzed, with exact definitions of financial ratios used in the analysis varying across banks and, in some cases, across borrowers or loan types. The analysis yields an assessment of the difference between current or projected performance and liquidity on the one hand and projected debt service obligations on the other. The larger the cushion, in general, the more favorable the rating.

As a context for financial statement analysis, the characteristics of the borrower's industry are often considered (such as cyclical, general volatility, and trends in cash flow and profitability). Indeed, the financial analysis often includes a formal comparison of the borrower's financial ratios to prevailing industry norms.²² Firms in declining industries are considered more risky, as are those in highly competitive industries, whereas firms with diversified lines of business are viewed as less risky. A related factor, the borrower's position in its industry, is also an important factor in determining ratings. Those borrowers with substantial market power or that are perceived to be "market leaders" in other respects are considered less risky because they are thought to be less vulnerable to competitive pressure.

One of the most important reasons that rating is usually a judgmental process is that the details of financial statement analysis vary with the borrower's other characteristics. In contrast, statistical models of default probability tend to analyze fixed sets of financial ratios and to apply fixed weights to each ratio in arriving at a default probability, perhaps with some variation in weights by industry. Subjective factors play at most a minimal role. This relative inflexibility

of models leads most banks to regard their results only as generally suggestive of an appropriate rating. When internal ratings are produced primarily by models, several models may be needed for different borrowers or loan types and continual tuning of the models is likely to be required.

Raters also appraise the quality of financial information provided by the borrower. For example, raters have much more confidence in financial statements that are audited by a major accounting firm than in those that are compiled or unconsolidated or that are audited but accompanied by important qualifications. When statement quality is poor or uncertain, financial analysis may produce a distorted view of the borrower's condition, adding substantially to risk.

A primary difference between banks and public rating agencies is whether the financial analysis is keyed to a downside (or "stress") scenario or to a "base" (or "most likely") case. As noted previously, banks assign ratings on the basis of the borrower's current condition and most likely outlook, whereas the rating agencies assign grades on the basis of a downside scenario.

In another departure from practice at the rating agencies, most banks formally consider both firm size (sales revenue or total assets) and the book or market dollar value of a firm's equity in assigning ratings. Interviewees noted that small firms—including many that would be considered middle market—usually have limited access to external finance and often have few or no assets that can be sold in an emergency without disrupting operations. In contrast, larger firms were characterized as having more ready access to alternative financing, more saleable assets, and a more firmly established market presence. For these reasons, many banks require that small borrowers be assigned relatively risky grades even if their financial characteristics might suggest a more favorable rating.

Almost all internal rating systems cite the borrower's management as an important consideration in assigning the risk grade. Such assessments are necessarily subjective and may reveal weaknesses in a number of areas related to competence, experience, integrity, or succession plans. Vulnerability of management to the retirement or departure of key individuals is usually considered. Some institutions (similar to the rating agencies) appear to give considerable weight to the rater's appraisal of management's ability and willingness to manage the firm to achieve a high level of financial performance throughout the business cycle and to its attitude toward protecting the interests of lenders.

21. We reviewed the *written* criteria for those banks among the fifty largest that we did not interview. Our experience with interviewed banks indicates that conclusions should be drawn with care from written documents alone. However, the description of risk factors herein is probably representative of the factors used by almost all large banks.

22. Staff at the banks interviewed appeared to be well aware of the potential pitfalls of such comparisons. For example, a borrower with a five-year history of stable cash flow might still be considered rather risky if the particular five-year period contained no recession and the borrower's industry is highly cyclical.

The borrower's country of domicile or operations is an important determinant of the rating in some cases. Especially when transfer risk or political risk is substantial, general practice seems to be that a borrower's grade may be no less risky than the grade assigned to the borrower's country by a special unit in the bank. Such country grades can be significantly affected by the country risk grade assigned by regulators as part of an annual cycle.

Ratings may also be influenced by exposure to event risks, such as litigation, environmental liability, or changes in law or national policy.

A handful of considerations reflecting the structure of the transaction being rated also enter into consideration because they can affect LIED. Adequate collateral can in many cases improve the rating, particularly if that collateral is in the form of cash or easily marketed assets such as U.S. Treasury securities.²³ Guarantees can generally enhance the rating as well, but not beyond the rating that would be assigned to the guarantor if it were the borrower. The term to maturity of the loan is a factor in grade assignments at only a few large banks. Similarly, few banks adjust the risk grade on the basis of other elements of the loan structure, such as financial covenants.

Written and Cultural Definitions

Large banks' written definitions of ratings specify risk factors to be used in assigning ratings, but usually the discussion is brief and broadly worded, and gives virtually no guidance regarding the weight to place on each factor.²⁴ According to interviewees, such brevity arises partly because some factors are qualitative but also because the specifics of quantitative factors and the weights on factors can differ a great deal across assets. Some noted that the number

23. Different rules are often used in grading certain classes of transactions, especially asset-based lending. At best, asset-based borrowers would be only marginally acceptable risks for banks in the absence of the detailed field audits of collateral that asset-based lenders demand. With such close monitoring, which typically includes some degree of bank dominion over accounts receivable and inventory, the expected loss associated with a default is dramatically reduced, and a more favorable rating can be assigned.

24. Written definitions are intended to address a broad range of credit classes and borrower types. At a few banks, a supplementary grid of nonbinding quantitative standards or financial ratios is provided (for example, for leverage or debt service coverage), but guidance is generally sketchy as to how such ratios should be weighted against each other or against more qualitative considerations. Interviewees indicated that even when reference grids are provided, the ratios and standards are generally not binding. Similarly, some banks provide supplemental descriptions of risk factors to be considered for particular business lines or loan types, but such supplements often closely resemble the core risk rating definitions.

of permutations is so great that attempting to write them down would be counterproductive. Instead, raters learn to exercise judgment in selecting and weighting factors through training, mentoring, and especially by experience. The actual meanings of *written* rating definitions and the specifics of assigning ratings take the form of common, *unwritten* knowledge embedded in the bank's credit culture.

Formality of Procedure

Most banks require some sort of written justification of the grade as part of the loan approval package, but a few employ forms or grids on which the rater identifies the relevant factors. Such forms or grids may also suggest a structure for the rating analysis and serve to remind the rater to consider a broad set of risk factors and to weight them appropriately. The stated motivation for such formalism is better consistency across asset types and geographic regions.

Reviews, Reviewers, and the "Keepers of the Flame"

Reviews of ratings are threefold: Monitoring by those who assign the initial rating of a transaction, regularly scheduled reviews of ratings for groups of exposures, and occasional reviews of a business unit's rating assignments by a loan review unit. Monitoring may not be continuous, but it is intended to keep the rater well enough informed to recommend changes to the internal risk grade in a timely fashion as needed. All institutions interviewed emphasized that failure to recommend changes to risk grades in a timely fashion when warranted is viewed as a significant performance failure for the relationship manager, the credit staff, or both, and can be grounds for internally imposed penalties.²⁵

Most institutions also conduct annual or quarterly reviews of each exposure, which may be in addition to those that are part of the credit approval process at the time facilities are renewed. The form of regular reviews ranges from a periodic signoff by the relationship manager working alone to a committee review involving both line and credit staff. Banks with substantial large-corporate portfolios tend to review all exposures in a given industry at the same time, with reviews either by the credit specialist for that industry or by a committee. Such industry

25. Updates to the risk grade usually require approvals similar to those required to initiate or renew a transaction.

reviews were said to be especially helpful in revealing inconsistently rated credits.

Ratings are also checked by banks' independent loan review units, which usually have the final authority to set grades. Such departments examine each business unit's underwriting practices, and its adherence to administrative and credit policies, on a one- to three-year cycle. Not unlike bank examiners, the loan review staff typically inspects only a sample of loans in each line of business. Although the sampling procedures used by different institutions vary somewhat, most institutions weight samples toward loans perceived to be riskier (such as those in high-risk loan grades), with the primary focus on regulatory problem-asset categories. In general, however, an attempt is made to review some loans made by each lender in the unit being inspected.²⁶

At a few banks, the loan review unit inspects internal ratings assigned to Pass loans only to confirm that such loans need not be placed in the watch or regulatory grades. Thus, as a practical matter, the loan review unit at these banks has little role in maintaining the accuracy of assignments within the Pass grades. In this regard, the loan review staff at these banks follows the same pattern as bank examiners. These banks tend to make relatively little use of Pass grade information in managing the bank.

Because operational rating definitions and procedures are embedded in bank culture rather than written down in detail, the loan review function at most institutions is critical to maintaining the discipline and consistency of the overall rating process. The loan review unit, as the principal entity looking at ratings across business lines and asset types, often bears much of the burden of detecting discrepancies in the operational meaning of ratings across lines.

Because the loan review unit at most institutions has the final say about ratings, it can exert a major influence on the culturally understood definition of grades.²⁷ Typically, when the loan review staff finds grading errors, it not only makes corrections but works with the relevant staff to find the reasons for

the errors. Misunderstandings are thus corrected as they become evident.²⁸

Loan review units generally do not require that all ratings produced by the line or credit staff be identical to the ratings that loan review judges to be correct. At almost all banks we interviewed, loan review units treat only two-grade discrepancies for individual loans as warranting discussion. With a typical large bank having four to six Pass categories, such a policy permits large discrepancies for individual exposures, potentially spanning two or more whole letter grades on the Standard & Poor's scale. However, most institutions interviewed indicated that a pattern of one-grade disagreements within a given business unit—for example, a regional office of a given line of business—can lead to a quick and decisive response.

All interviewees emphasized that the number of cases in which the loan review staff changes ratings is usually relatively small, ranging from essentially none to roughly 10 percent of the loans reviewed, except in the wake of large cultural disruptions such as mergers or major changes in the rating system. A low percentage of discrepancies does not imply that the loan review function is unimportant but rather that, in well-functioning systems, the cultural meaning of ratings tends to remain stable and widely understood. One element of a well-functioning system is the rater's expectation that the loan review staff will be conducting inspections.

The interviews also indicated that differences of opinion tend to become more common when the number of ratings on the scale is greater, creating more situations in which "reasonable people can disagree." More direct linkage between the risk grade assigned and the incentive compensation of relationship managers also tends to produce more disagreements. In both cases, resolution of disagreements may consume more resources.

Loan review units usually have a role apart from inspections in maintaining rating system integrity. For example, when a relationship manager and the credit staff are unable to agree on a rating for a new loan, they will consult with the loan review unit on how to resolve the dispute. In its consultative role, the loan review staff guides the interpretations of rating definitions and standards and, in novel situations, establishes and refines the definitions.

26. For an analysis of the broader role of loan review units, see Udell, *Designing the Optimal Loan Review Policy*; and Gregory F. Udell, "Loan Quality, Commercial Loan Review, and Loan Officer Contracting," *Journal of Banking and Finance*, vol. 3 (July 1989), pp. 367–82.

27. Interviews and discussions with supervisory staff suggest, however, that the notion of "final say" is murkier than suggested by written policy and stated practice. Important informal elements of rating processes, such as negotiation among various organizational units, may lead to a consensus rating or understanding. Such negotiation would not compromise the integrity of the rating system so long as loan review retains its independence and objectivity. Such informal understandings might make it more difficult, however, for an outsider to understand (much less validate) the ratings being assigned.

28. The loan review staff generally uses the same definitions of risk grades, at the same level of detail, as relationship managers and the independent credit staff. At a few banks, however, loan review also relies on older policy documents that are far more detailed than current policies. Thus, the older, more specific policies remain essentially in effect.

Because of its central role in maintaining the integrity of the rating system, the loan review unit must have both substantial independence and staff members who are well versed in the bank's credit culture and the meaning of ratings. All loan review units at banks we interviewed report to the chief auditor or chief credit officer of the bank, and many periodically brief the board (or a committee thereof) on the results of their reviews.

Loan review units may be less critical to the integrity of rating systems at banks that are primarily in the business of making large corporate loans and at which all exposures are rated by a relatively small, highly independent credit staff. Although few banks currently fit this description, they provide an interesting contrast. Such banks' credit units tend to conduct the annual industry-focused reviews mentioned previously and thus are likely to detect rating discrepancies. Having such reviews conducted by broadly based committees rather than only by industry specialists tends to restrain any drift in the meaning of ratings as applied to different industries. In such circumstances, the small credit staff is in a good position to function as the "keeper of the flame" with regard to the credit culture because it essentially carries out the key rating oversight functions of traditional loan review units.

Rating Systems and Credit Culture

"Credit culture" refers to an implicit understanding among bank personnel that certain standards of underwriting and loan management must be maintained, even in the face of constant pressures to increase revenues and bring in new business. Maintenance of a credit culture can be difficult, especially at very large banks serving many customers over a wide area. Of necessity, substantial authority must be delegated to mid-level and junior personnel, and undue relaxation of standards may not appear in the form of loan losses for some time.

At some of the banks we interviewed, senior managers indicated that the internal rating system is at least partly designed to promote and maintain the overall credit culture. At such banks, relationship managers are held accountable for credit quality partly by having them rate all credits, including large exposures that might be more efficiently rated by the credit staff. Strong review processes aim to identify and discipline relationship managers who produce inaccurate ratings. Such a setup provides strong incentives for the individual most responsible for negotiating with the borrower to assess risk properly

and to think hard about credit issues at each stage of a credit relationship rather than relying entirely on the credit staff. An emphasis on culture as a critical consideration in designing the rating system was most common among institutions that had suffered serious problems with asset quality in the past ten or fifteen years.

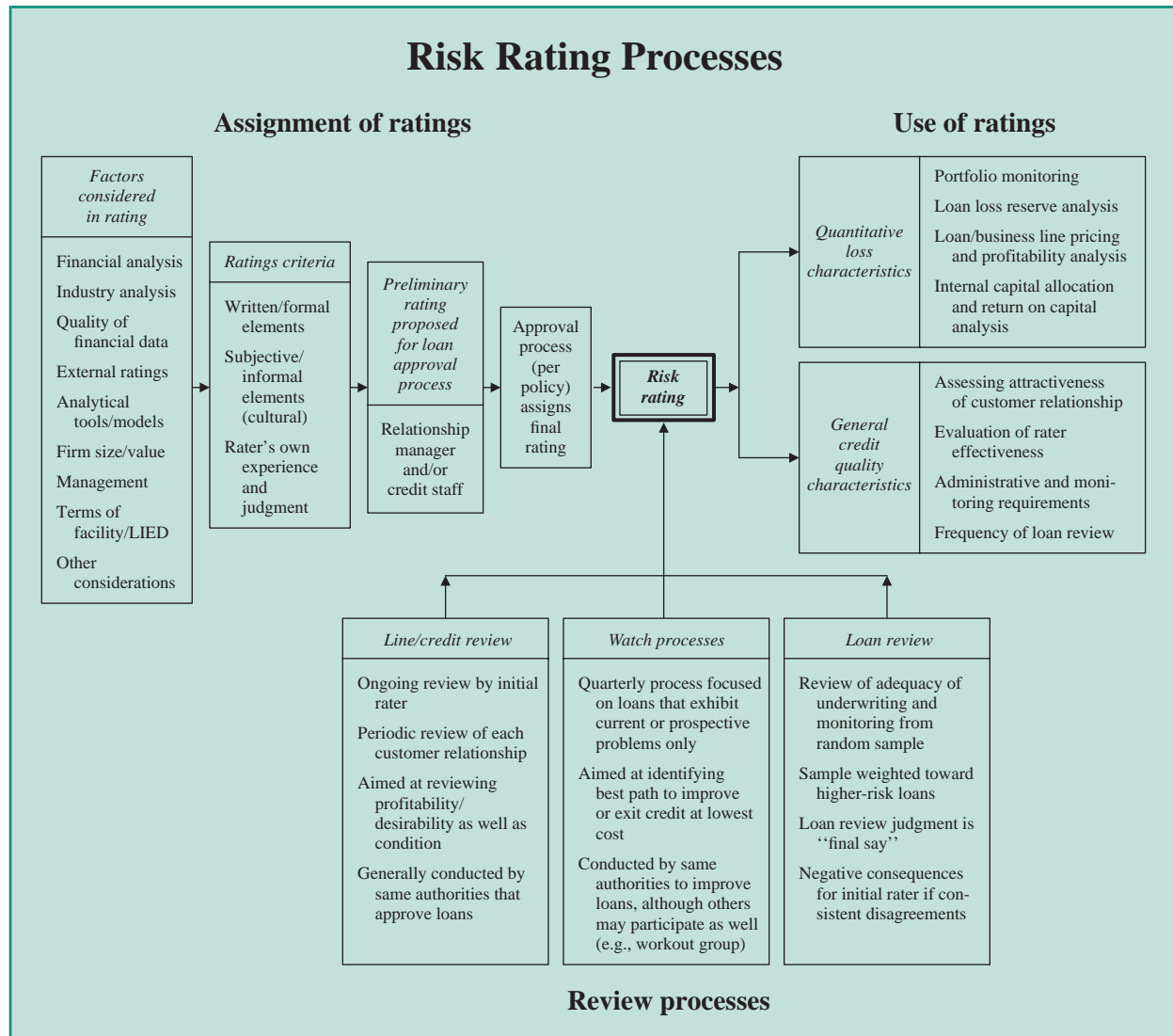
Tensions can arise when rating systems both maintain culture and support sophisticated modeling and analysis. As noted, the latter applications introduce pressures for architectures involving fine distinctions of risk, and the frequency of legitimate disagreements about ratings is likely to be higher when systems have a large number of Pass grades. If not properly handled by senior management and the loan review unit, a rating system redesign that increases the number of grades may make cultural norms fuzzier and the rating system less useful in maintaining the credit culture.

Mergers and Expense Pressure

Some of our interviews involved banks that had recently been involved in mergers, and the discussions clearly indicated that mergers can cause upheaval in credit processes and systems, credit culture, and traditional sources of rating discipline. After a rating system architecture is chosen for the combined institution, mechanical issues of converting the predecessor banks' ratings to the new scale can be challenging, especially when the predecessors' ratings of the same borrower suggest differing assessments of that borrower's risk. Cultural disruptions arising from the merger are usually even more problematic than the mechanical issues because, as noted, the operational definitions of ratings are a matter of culture. Even if the architecture of one of the predecessors is used as-is, the staff of the other bank must absorb and adjust to the new culture.

Merging institutions face a difficult choice between moving very quickly to convert the ratings of all assets to the new system, in which case stresses are high, and converting the ratings over time, which reduces the intensity of stress but also can reduce the reliability of internal rating information during the longer transition. In one version of the slower transition, which is especially common when a large bank acquires a much smaller bank, all of the acquired bank's performing loans are assigned to the riskiest nonwatch Pass grade. Each loan is then reassigned as appropriate at the time of its next review. Although such a practice may be viewed as conservative, it masks the true risk posture of the bank during the

Diagram 1



transition period. Regardless of the speed of transition, loan review units are under substantial pressure during and immediately after the transition.

Expense control has also been a focus of the banking industry in recent years. The emphasis on economy naturally puts pressure on the resources devoted to operating and maintaining the rating system, and especially to reviews. Although reviews can be curtailed or eliminated in the short run without apparent damage to rating system integrity, inadequate review activity may lead to biased and inconsistent ratings over the longer term. Another possible expense-reduction strategy is to rely more heavily on statistical models in assigning ratings, reducing the degree of judgment and, thus, the amount of labor required to produce each rating. The long-run success of such

a strategy depends on the adequacy of the models, including their ability to incorporate subjective factors and their robustness over the business cycle. Our impression is that, at present, such adequacy is uncertain.

Summary Observations on Operating Design

The rating process has many interlinked elements, as illustrated in diagram 1. At almost all large banks, internal rating systems rely importantly on the judgment of staff operating with relatively little written guidance. The operational definition of each grade is largely an element of credit culture that is determined and communicated by informal means.

Review activities, especially those conducted by loan review units, are crucial for maintaining the culture in that the feedback they give is critical to common understanding and discipline. The credit culture can be disturbed or unbalanced by changes in the incentives faced by the staff; such changes typically arise whenever the rating system is required to support additional functions or uses. The systems of banks at which all ratings are assigned by credit staff are relatively immune to such shocks, but the important role of middle-market loans in most banks' portfolios often makes rating assignment by relationship managers cost-effective. In the latter case, the rating system's resilience to shocks depends to a considerable extent on the loan review unit's ability to detect and correct problems in a timely manner. Strong support of loan review by senior management and boards of directors appears to be quite important.

Points of external comparison, such as agency ratings or results of statistical models of borrower default probability, can be helpful in maintaining the integrity of internal ratings. A few banks are moving toward models as the primary basis for internal ratings. Such an operating design largely removes the problems of culture maintenance and conflicting incentives that make management of judgmental rating systems challenging. However, the ability of models to produce sufficiently accurate ratings for the broad range of assets on the typical large bank's balance sheet remains in question.

BANK SYSTEMS RELATIVE TO RATING AGENCY SYSTEMS

Credit risk ratings have played an important role in capital markets for most of the twentieth century. Ratings of publicly issued bonds were first produced during the early 1900s by predecessors of the current rating agencies Moody's and Standard & Poor's. In the decades after 1920, other agencies, both domestic and foreign, were formed and commenced publication of ratings. Today a variety of instruments are rated, such as commercial paper, bank certificates of deposit, commercial loans, and hybrid instruments.

Agency and bank rating systems differ substantially, mainly because rating agencies themselves make no investments and thus are not a party to transactions between borrowers and lenders. Their revenue comes from the sale of publications and from fees paid by issuers of debt. Such fees can be substantial: S&P's fee for rating a public corporate debt issue ranges from \$25,000 to more than \$125,000, with the usual fee being 0.0325 percent of the face amount of the issue. Fees are a reflection of the substantial

resources the agencies typically devote to producing each rating, especially the initial rating.

At banks, the costs of producing ratings must be covered by revenues on credit products. Thus, although a bank might expend resources at a rate similar to that of the rating agencies when underwriting and rating very large loans, the expenditure of so much labor for middle-market loans would make the business unprofitable.

Agency ratings are used by a large number and variety of parties for many different purposes. To ensure wide usage (and thus their ability to collect fees), the agencies strive to be deliberate, accurate, and evenhanded. They also produce relatively fine distinctions of risk on rating scales having architectures and meanings that are stable over time. Accuracy and evenhandedness are crucial to the rating agency business—for example, an agency suspected of producing the most favorable ratings for those that pay the highest fees would soon be out of business: Investors would cease paying attention to its ratings, and issuers would thus have no incentive to pay.

Similarly, changing the rating scale can confuse the public and at least temporarily degrade the value of an agency's product. The agencies also have incentives to be relatively open about their process and to produce written explanations of each rating assignment or change. Clarity helps investors use the ratings and helps assure issuers that the process is as objective as possible.

At banks, ratings are kept private, and the costs and benefits of rating systems are internal; hence, pressures for accuracy, consistency, and fine distinctions of risk are mainly a function of the ways in which ratings are used in managing the portfolio. Moreover, the rating system can be tailored to fit the requirements of the bank's primary lines of business and can be restructured whenever the internal benefits of doing so exceed the costs.

Agencies and banks both consider similar risk factors, and both rely heavily on judgment and cultural elements rather than on detailed and mechanical guidance and procedures. However, the agencies publish supplementary descriptions of rating criteria that are much more detailed than banks' internal guidance, partly because agency ratings must be understood by outsiders. In addition, the agencies track the financial characteristics of borrowers receiving their ratings and publish both default histories for each grade and financial profiles of the "typical" borrower in each grade, thus providing additional referents to outsiders seeking to understand the meaning of their ratings.

Agencies have nothing comparable to a bank's loan review unit. The rating culture at agencies is

maintained instead by a combination of market discipline and a committee system. Market discipline arises because the agencies stand between investors and issuers, with the former typically preferring conservative ratings and the latter preferring optimism. Thus, the agencies quickly hear from investors or issuers about any perceived tendency toward excessive optimism or pessimism. Although a single agency analyst is primarily responsible for proposing a rating, committees make the final determinations. The membership of a committee changes from one rating action to the next so that agency staff members participate in many rating decisions and a cultural understanding of the meaning of each grade is maintained.

BANKS' ATTEMPTS TO MEASURE LOSS CHARACTERISTICS BY GRADE

Consistent and accurate rating assignments and reliable quantitative estimates of the risk associated with each internal grade are useful in a bank's efforts to analyze risk posture, establish its appetite for risk, and evaluate the effectiveness of its risk rating criteria. At most banks, however, the primary demands for quantitative information about PD, LIED, and EL have come from those involved in the loan loss reserve process and from credit modeling groups (those building and implementing quantitative models of portfolio risk, capital allocation, profitability, and pricing). Internal ratings are key inputs into such processes. Empirical analysis of loss characteristics by grade appears to be an area where industry practice is developing rapidly.

Problems in Evaluating the Accuracy and Consistency of Ratings

If internal ratings are to be accurate and consistent in terms of the system's loss concepts (that is, PD, LIED, or EL), different assets posing a similar level of risk should receive the same grade. Such quantities are not observable *ex ante*, however, and thus rating systems rely on criteria that are thought to predict loss. Accuracy and consistency require that rating criteria be adjusted as necessary to ensure that exposures posing similar risk are grouped together (diagram 2 illustrates what is involved in the adjustment process).

As a practical matter, alignment of the *ex ante* rating criteria to achieve accuracy and consistency in the economic meaning of each rating—that is, quan-

titative loss characteristics—is a difficult task. Two problems arise: How to ensure that criteria are calibrated so that different assets of the same general type in the same grade have the same loss characteristics, and how to address diversity among asset types. Within a narrowly defined asset class, such as loans to large commercial firms in the same industry, comparisons across firms are relatively manageable, so the main problem is defining the boundaries of rating categories and inferring the default or loss rates for each category. That by itself is not easy, but the problem becomes much more difficult when very different types of assets must be compared. For example, how would a loan to a well-established commercial real estate developer, featuring a 70 percent loan-to-value ratio, compare with a term loan to a firm in a relatively stable manufacturing industry with a current debt to equity ratio of 1:1 and an interest coverage ratio of 3?

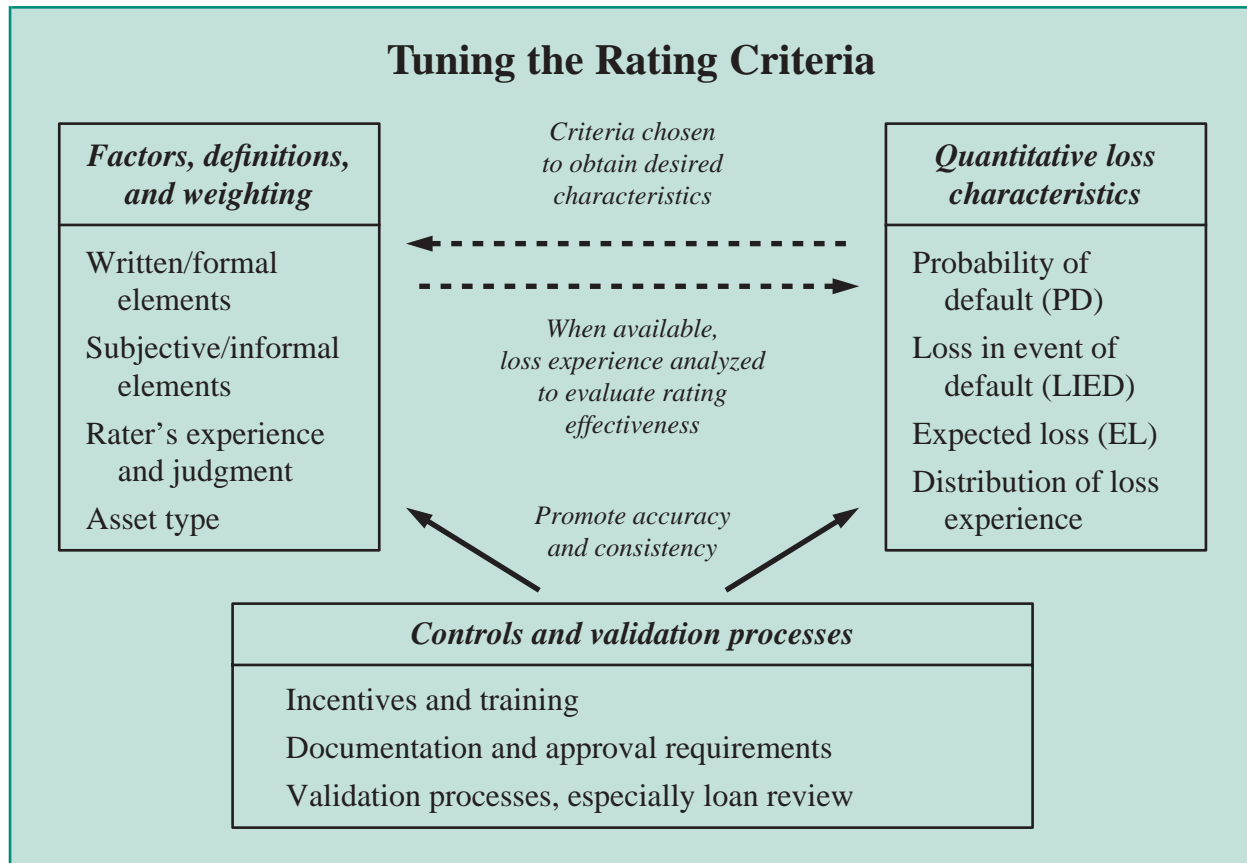
Because the rating criteria differ so greatly for different asset classes, some information about the relationship of borrower and asset characteristics to historical loss experience would appear to be necessary. Especially with loss experience data covering a fairly long period of time, say a couple of credit cycles, it would be possible to make at least rough inferences about relative risks across asset classes.

Unfortunately, to the best of our knowledge, few if any banks have available the necessary data, especially for a variety of asset classes. At a minimum, information on the performance of individual loans and their rating histories is required. Because rating criteria have changed over time at most large institutions, information about borrower and loan characteristics is also required, so that the risk implications of different rating criteria can be assessed.

Historically, banks have retained performance data by loan type (for example, data provided on Call Reports) or by line of business in the aggregate, but not by risk grade. Because of mergers, even at banks that have tracked performance by grade, data may not cover the whole of the current institution but rather only one predecessor institution. Mergers often cause upheaval not only in rating processes but also in data systems and, in particular, contribute to the loss or obsolescence of historical data.

Although data collection is costly, many large banks have recognized its importance and have begun projects to build databases of loan characteristics and loss experience. However, the costs of extracting from archival files historical data on the performance of individual loans appear to be prohibitively high. Thus, those banks that are collecting data indicated that they are several years away from having data

Diagram 2



sufficient to support empirical analyses of their own portfolios that are comparable to the studies being done for publicly issued bonds.²⁹

In the absence of data, our impression is that the traditional means of tuning both rating criteria and underwriting standards relies heavily on the judgment and experience of the senior credit staff with long tenure at their institution. Over a period encompassing multiple credit cycles, these staff members accumulate an individual and collective memory of

the credit problems experienced by the institution and of the implications for risk of various borrower and loan characteristics. Such experience is likely sufficient to support meaningful tuning of rating systems that have small numbers of Pass grades (each covering a broad band of risk) and that are used to rate traditional banking assets. The precision with which systems involving a large number of Pass grades can be tuned by experience alone is not clear.

Mapping to Agency Grades as a Partial Solution

Because little information is available internally, many banks have estimated the quantitative loss characteristics of their ratings by using the extensive data available on the loss performance of publicly issued bonds. As noted, rating agencies and others frequently publish studies covering many years of bond default and loss experience by grade, and publicly available databases of bond issuer characteristics make it possible to relate loss experience to potential

29. The situation is somewhat better with respect to loss in the event of default (LIED) in that historical studies require information only on the bad assets. Often their number is small enough that gathering data from paper files is feasible, and thus many banks are beginning to accumulate LIED information from their own portfolio experience. A few publicly available studies have also appeared. Estimating PD and EL requires much more data in that information on both performing and nonperforming assets are required. Studies with LIED statistics include Lea V. Carty and Dana Lieberman, *Special Report: Defaulted Bank Loan Recoveries* (Moody's Investors Service, 1996); Elliot Asarnow and David Edwards, "Measuring Loss on Defaulted Bank Loans: A 24-Year Study," *Journal of Commercial Lending*, vol. 77 (March 1995), pp. 11–23; and Society of Actuaries, *1986–92 Credit Risk Loss Experience Study: Private Placement Bonds* (Society of Actuaries, Schaumburg, Ill., 1996).

Mappings and the Problem of Different Architectures

Both banks and rating agencies assign ratings based on criteria that are predictive of a borrower's probability of default (PD) or a loan's expected loss (EL). However, because no mechanical formula exists that converts criteria into values of PD or EL for each grade, such values must be obtained from historical loss experience. As noted, banks rarely have databases of such experience, but the major rating agencies do. A mapping of internal grades to agency grades permits a bank to use statistics from the agencies' bond default studies to assign values of PD to each of its internal grades.

For simplicity, we focus here only on PD. Four problems can cause a mapping to lead to a materially inaccurate estimate of PD for internal grades:

- (1) A bank's rating system may place loans with widely varying levels of PD into the same grade and similar levels of PD into different grades. In this case, grades bear little relation to PD values and thus mapping will not provide good estimates of PD.
- (2) Default rates on publicly issued bonds may differ systematically from loan default rates.
- (3) The mapping exercise may simply associate the wrong agency grades with internal grades.
- (4) The implications of differences between banks' point-in-time and agencies' through-the-cycle rating philosophies may not be taken into account.

Even when the first three problems do not apply, the fourth, which is a characteristic of the most common mapping approach, can produce materially biased estimates of PD for internal grades. Such bias can confuse attempts to tune rating criteria and can seriously distort internal analysis of business line profitability, loan loss reserves, and capital allocation.

Bias arises in the most common approach to mapping because bank internal ratings change as the borrower's condition changes, whereas the PD associated with each internal grade is stable. In contrast, agency ratings tend to stay the same, while default probabilities for each rating vary with the economic cycle. Thus, mapping exercises should take into account the current point of the economic cycle and should draw default rates from the agencies' historical studies for similar points in prior cycles.

The fourth problem is illustrated here with an example: Suppose that a hypothetical large bank, BigBank, has an

I. BigBank's Pass rating scale

Grade	True PD for rating system, but precise values not known to bank (percent)
1—Virtually no risk	0
2—Low risk10
3—Moderate risk25
4—Average risk	1.00
5—Acceptable risk	2.00
6—Borderline risk	5.00

internal rating system with six Pass grades, and suppose it has two hypothetical borrowers, OK Corp. and Less-OK Corp. To focus on the point-in-time vs. through-the-cycle issue, suppose we know that BigBank's rating criteria and rating system will always group borrowers with similar values of PD into the same grade, that the "true" PD for each grade is as shown in table I, and that BigBank does not know the values of PD associated with its grades. Similarly, as shown in the top section of table II, the true PD for OK Corp. is 1 percent in upturns and 2 percent in downturns, whereas Less-OK Corp.'s true PD is 3 percent during upturns and 6 percent in downturns. However, because neither BigBank nor the rating agencies know these true PD values, they rate on the basis of observable borrower characteristics.

Having no data on its historical loss experience, BigBank maps its internal grades to agency grades simply by identifying the agency ratings assigned to those borrowers with such ratings in each internal grade. BigBank then uses the corresponding *long-term historical average one-year default rate* identified in agency default studies as an estimate of the *expected one-year default rate* for all loans in each internal grade.

II. Borrowers used for mapping, and their characteristics

Characteristic	Borrower	
	OK Corp.	Less-OK Corp.
PD in upturns	1 percent	3 percent
PD in downturns	2 percent	6 percent
BigBank rating in upturns	4—Average risk	5—Acceptable risk
BigBank rating in downturns	5—Acceptable risk	6—Borderline risk
Agency ratings (stable through cycle)	BB or Ba	B+ or B1

rating criteria. Indeed, S&P occasionally publishes tables of indicative or average financial ratio values by grade (while noting that many other factors enter into its rating decisions).

To use data on bond loss experience, a bank must develop or assume some correspondence between agency ratings and its own internal grades. Inter-

views suggest that the basis of such mappings is threefold: (1) The internal grades assigned to borrowers who have also issued publicly rated bonds; (2) analysis of the "typical" financial characteristics of bank borrowers in each internal grade vis-à-vis the characteristics of the firms with bonds in each agency grade; and (3) subjective analysis.

Mappings and the Problem of Different Architectures—Continued

Because it rates on a point-in-time basis, BigBank does not allow the PD values for each grade to vary through the economic cycle; loans whose one-year PDs increase in cyclical downturns are downgraded to a riskier internal grade. As shown in the middle section of table II, BigBank assigns ratings that are appropriate for varying risk: It rates OK Corp. a 4 in upturns and a 5 in downturns, and it rates Less-OK Corp. one grade worse—a 5 in upturns and a 6 in downturns. The rating agencies are similarly accurate in their assessment of risk (bottom section of table II), but because they rate through the cycle (that is, according to the borrower's condition when under stress), they rate OK Corp. as BB/Ba and Less-OK Corp. as B/B in both upturns and downturns.

Suppose that BigBank conducts its mapping exercise during an upturn. As shown in the top section of table III, it will assume that its grade 5 is equivalent to the agencies' B grades because Less-OK Corp. is in relatively good shape during upturns and achieves a point-in-time internal rating of 5 even though its through-the-cycle agency grade is B. BigBank should infer the PD for grade 5 from the average default frequency of B-rated public bonds *only in upturns*, which is the good-year average of 4 percent (table III); but if it follows common practice it will use the *overall* average default frequency of B-rated bonds, which is 5.5 percent.

Next, suppose BigBank conducts its mapping exercise during a downturn. As shown in the bottom section of table III, it will assume that its grade 5 is equivalent to BB/Ba because OK Corp. will be rated 5 (Less-OK Corp. is downgraded to 6 during downturns). BigBank should infer the PD for grade 5 from the *bad-year* average PD of BB/Ba rated bonds (2 percent), but instead it uses the *overall* average of 1 percent.

III. BigBank mapping and PD estimation exercise based on borrower ratings

Period of mapping	Internal grade	Equivalent agency grade	Average one-year PD for bonds		
			Overall	Good year	Bad year
Upturn	4	BB/Ba	1.00	.75	2.00
	5	B/B	5.50	4.00	6.50
	6
Downturn . . .	4
	5	BB/Ba	1.00	.75	2.00
	6	B/B	5.50	4.00	6.50

. . . Not applicable.

In this example, BigBank's and the agencies' rating systems both do an excellent job of assigning ratings that are consistent with the borrower's true PD, but mapping without regard to the difference between point-in-time vs. through-the-cycle rating causes BigBank to badly misestimate the PD. *Using the most common mapping practices, BigBank might estimate the PD of its grade 5 at 1 percent to 5.5 percent, whereas the true PD of grade 5 is 2 percent.* If the mapping is done simplistically, as in this example, and during an upturn, BigBank likely overestimates the PD, whereas during a downturn it likely underestimates the true value. If BigBank had used average default frequencies from the agencies' studies that were appropriate to the point in the cycle at which the mapping was done, it might still have obtained inaccurate estimates, but they would have been closer to the truth. BigBank might still have been somewhat uncertain about whether to consider category 5 as equivalent to BB/Ba or B, but any such equivalence can never be exact because BigBank's scale and the agency scales have different conceptual foundations.

We consider the numbers in the example to be fairly realistic and thus the mis-mapping problem at most banks to be potentially serious. The problem of mis-estimated PDs is much more important at the higher-risk end of rating scales. Precision is especially important at that end because differences in reserve and capital allocations can be large, whereas dollar differences in allocations across different classes of low-risk assets are typically small. In addition, default studies and other analyses tend to show that variations in one-year default rates on investment-grade assets tend to be driven by idiosyncratic factors rather than the credit cycle.

Mapping processes are further complicated if, over time, a borrower's agency rating is allowed to be the dominant criterion in assigning an internal grade. In general, such a practice would tend to reduce the likelihood that a loan would be appropriately downgraded during a recession—the borrower's agency rating would not change unless its performance or prospects deteriorated more than anticipated in the agency's through-the-cycle risk analysis. This procedure could effectively turn BigBank's ratings into through-the-cycle rather than point-in-time, putting loss estimates potentially out of line with management analyses that assume point-in-time grading.

When mapping is done by comparing the internally assigned grades of publicly rated borrowers with ratings assigned by agencies, the danger of circularity arises. In most cases, agency grades are a rating criterion, and even when agency grades are not written into rating definitions, assigners of inter-

nal ratings always know the agency grade for a given borrower and have an idea of the borrower's likely position on the internal scale. Obviously, if the agency rating is the sole criterion used in assigning internal grades to agency-rated borrowers, rated and unrated borrowers within a given internal grade might

differ substantially in risk. In such circumstances the mapping is circular because borrowers are assigned to internal grades based on the agency rating, and the agency rating corresponding to each internal grade is inferred only from such rating assignments. The banks we interviewed maintain that agency ratings are used only as a starting point in their rating processes, not as the sole criterion.³⁰

Mapping and the Problems Caused by Inconsistent Architectures

Because major agencies rate borrowers with the expectation that the rating will be stable through normal economic and industry cycles, only those borrowers that perform much worse than expected during a cyclical downturn will be downgraded (will “migrate” to riskier grades). In contrast, rating systems that focus on the borrower’s current condition (virtually all bank systems) are likely to feature much more migration as cycles progress but, in principle, should exhibit somewhat less cyclical variation in default rates for each individual grade.

Though apparently subtle, this difference in architectures has important implications for mapping exercises and the inference of PD values for internal grades. Both the point in the economic cycle at which the mapping exercise is done and the exact nature of the PD statistics drawn from the agencies’ studies of long-term default history can have a dramatic effect on the mapping (see box “Mappings and the Problem of Different Architectures”). Values of PD attributed to internal grades can differ by several percentage points depending on how the mapping is done. PDs are most likely to be badly estimated for the higher-risk Pass grades, but precision is also especially important for such grades in that allocated reserves and capital are most sensitive to assumptions about riskier assets.

Obtaining reasonably accurate mappings is mainly a matter of paying attention to the stage of the cycle at which the mapping is being done and of using historical average PD values from either good-experience or bad-experience years as appropriate. However, interviews left us with the impression that few banks carefully consider cyclical issues when mapping their internal grades to agency grades.

30. Even when circularity is avoided, heavy use of bond experience data in defining criteria for each grade might lead to exclusion of criteria needed to capture the risk of unrated borrowers, such as middle-market firms.

AN AGGREGATE BANK RISK PROFILE

Mapping between internal and agency grades facilitates a bank’s quantitative loss analysis and the integration of publicly available information into rating decisions. Such mappings also make possible an estimate of the risk profile of the internally rated portion of bank loan portfolios on a standardized scale. Information about the risk profile of bank credit helps put many rating system issues in perspective.

As part of the analysis leading to this article, we reviewed internal reports showing distributions of rated assets across internal grades for the fifty largest consolidated domestic bank holding companies. In addition, we obtained mappings of internal grades to agency equivalents from twenty-six of them. The mappings allow us to allocate internally rated balances to grades on a rating agency scale. To our knowledge, this is the first time that such a characterization of the overall risk profile of a large portion of the banking industry’s commercial loan portfolio has been possible.

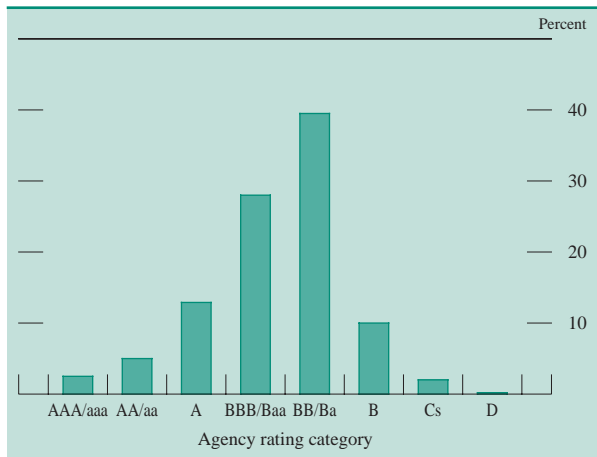
The twenty-six banks accounted for more than 75 percent of aggregate banking industry assets at year-end 1997. Rated loans outstanding at individual large banks usually represent 50 percent to 60 percent of their total loans.³¹

In general, we cannot judge whether the mappings provided by banks are correct. Inaccuracy can arise from errors or inconsistency in assigning the internal ratings themselves, problems of cyclicity or circularity in the mapping process, inconsistencies between large corporate and middle market lines of business, or other difficulties. In addition, mappings at some institutions are more precise in form than at other institutions in that they distinguish among modified agency grades, such as BB and BB+. Still, such mappings are an element of banks’ day-to-day operating procedures and analysis, which suggests that the twenty-six banks have endeavored to make them as accurate as possible within the constraints of their rating systems. It thus appears that aggregation and comparison of these mapped balances represents a reasonable—albeit crude and broad—first approximation of the actual risks in banks’ portfolios.

Chart 3 displays the aggregate weighted-average distribution of internally rated outstanding loans at year-end 1997 for the twenty-six consolidated bank holding companies. About half of aggregate rated loans pose below-investment-grade risks (were rated the equivalent of BB+/Ba1 or riskier), and about 65 percent of outstandings were concentrated around

31. Total loans includes consumer loans, which are rarely rated.

3. Percentage of aggregate internally rated outstandings placed in each agency rating category at banks mapped to agency scale, year-end 1997



NOTE. The banks are twenty-six of the fifty largest.

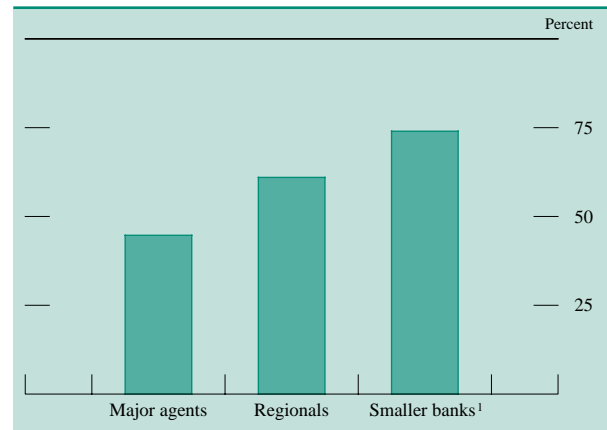
the boundary between investment and below-investment grades (rated BBB or BB).

Banks' loan loss experience during 1997 is consistent with the credit quality distribution shown in chart 3. Using the 1997 default frequencies for each grade drawn from S&P's latest annual study and an assumption that the average LIED for loans is about 30 percent, an aggregate portfolio with the quality distribution for the twenty-six banks would be expected to have an annual credit loss rate of roughly 0.20 percent. Although this rate is roughly equal to the actual loan loss experience of the banking industry's aggregate commercial loan portfolio during 1997 (0.21 percent), this simple exercise should not be taken as proof that the distribution in chart 3 is representative; nonetheless, the results are supportive.³²

Chart 4 displays the percentages of internally rated assets that are below investment grade as of year-end 1997 for twenty-six banks in three peer groupings: major loan syndication agents; smaller banks (less than \$25 billion in total assets at year-end 1997); and the remainder of the twenty-six, labeled "regionals" (many other peer groupings are possible, of course). The three peer groups display systematic differences in risk posture. On average, the major agents have 45 percent of rated assets in categories corresponding to BB and riskier, compared with

32. Actual loss experience is measured as the average annualized net charge-off rate for bank loans in the commercial and industrial, commercial mortgage, and agricultural loan categories as reported on the quarterly Report of Condition—or Call Report—filed by all banks.

4. Percentage of aggregate internally rated outstandings below investment grade at banks mapping to agency scale, by bank group, year-end 1997



NOTE. The banks are twenty-six of the fifty largest.

1. Less than \$25 billion in total assets. Regionals are those that are not major syndication agents or smaller banks.

about 60 percent for regionals and 75 percent for smaller banks.³³

USES OF INTERNAL RISK GRADES

Banks use internal ratings in two broad categories of activity: analysis and reporting, and administration. Analytic uses include reporting of risk postures to senior management and the board of directors; loan loss reserving; and economic capital allocation, profitability measurement, product pricing, and (indirectly) employee compensation. Administrative uses include loan monitoring, regulatory compliance, and credit culture maintenance. In addition, external entities such as investors or regulators may become more significant users of internal ratings information. Different uses place different stresses on the rating system and may have different implications for the internal controls needed to maintain the system's integrity (diagram 1 shows such uses).

Portfolio Reporting

Virtually all large banks report total asset balances in each of the regulatory problem-asset grades to

33. That the fraction of loans posing below-investment-grade risks is much larger at some institutions than at others does not imply a priori that such institutions are operating in an unsafe or unsound fashion. In general, provided a bank is aware of its risk posture, has adequate processes to manage risk, is pricing loans to reflect the risk, and has reserves and capital that are adequate to the risks, a portfolio with a large fraction of below-investment-grade exposures can be safe, sound, and profitable.

senior management and the board of directors. About 80 percent also internally report balances in each of their Pass grades. In the latter case, such reports appear to be used either by management or the credit staff as a means of detecting changes in portfolio mix and are only infrequently shown to boards of directors.³⁴ Balances in the regulatory grades give a sense of the share of bank assets that are troubled, whereas a profile of balances in Pass grades can provide a forward-looking sense of trends in the bank's risk posture so long as Pass grade assignments meaningfully distinguish risks; internal reports are much less informative when a large share of rated assets falls into only one or two Pass grades.

Reserving

Although many accounting and regulatory policies influence the setting of loan loss reserves and provisions, balances in the regulatory grades are integral to reserve analysis at all banks. Supervisors require a specific reserve of at least 50 percent of Doubtful loans plus 20 percent of Substandard loans; banks set the amount of additional reserves for OAEM and Pass loans according to their judgment, subject to evaluation by examiners.³⁵ Many banks develop reserve factors specific to each Pass category. According to accounting and regulatory standards, loan loss reserves are to cover losses already "embedded in the portfolio," and the generally accepted interpretation is that reserves for Pass loans should cover expected losses over a period of one year. Thus, if an institution can identify a reasonable estimate of expected loss for each Pass grade, a reserve analysis sensitive to balances in the different Pass grades provides a good estimate of embedded losses.

A significant number of the banks we interviewed do not differentiate among the Pass grades in performing reserve analysis. In such cases, a single expected-loss (EL) weight is applied to balances in all Pass grades. Such a simplification is least costly in terms of accuracy of the reserve analysis when loan

balances are concentrated in a single category or when the composition of the Pass portfolio by risk grade is very stable.

Profitability Analysis, Pricing Guidelines, and Compensation

All banks we interviewed conduct internal profitability analyses (of different business lines, for example). Some banks do not use internal ratings at all in such analyses, whereas others include a rating-sensitive expected-loss cost but no rating-sensitive capital cost. The most sophisticated analyses involve both expected-loss costs and costs of allocated capital that vary by internal rating. The higher such costs, the lower the measured profitability of a business unit or individual transaction. The use of rating-sensitive profitability analysis thus has significant implications for the design and operation of internal rating systems.

To implement rating-sensitive profitability analysis, the bank must estimate expected losses for assets in each grade as well as the amount of economic capital to allocate (if it allocates capital). Economic capital for the bank as a whole is that needed to maintain the bank's solvency in the face of unexpectedly large losses. The process of estimating the additional economic capital needed as a result of booking any given loan is complex, but as a practical matter, the loan's internal rating is a primary (if not the sole) day-to-day determinant of the capital allocations imposed by risk-sensitive profitability models.³⁶

The measured profitability of business units is an important factor in management decisions about which units should grow or shrink. When risk-sensitive profitability is appraised at the level of the individual loan or relationship, unprofitable loans are not made and unprofitable relationships are eventually dropped. At a growing number of banks, employee compensation is formally tied to profitability measured by such systems.

34. At some banks, portfolio composition is reported as a weighted-average risk grade. Such averages weight the balances by the grade's numeric designator. For example, assets in grade 4 are treated as being twice as risky as assets in grade 2. This can produce misleading averages because risk—whether PD or EL—tends to increase more than linearly with grade (table 2). At those banks we interviewed that used this measure, the staff seemed to understand that it does not reflect portfolio risk—it can indicate only whether the mix has changed.

35. Federal Financial Institutions Examination Council, *Interagency Policy Statement on the Allowance for Loan and Lease Losses* (December 1993).

36. Mechanically, one can think of economic capital for the credit risk of a whole portfolio as that amount necessary to cover (for example) 99.9 percent of the possible portfolio loss rates. Capital required to support a given loan can be viewed as that increment to total bank capital that will keep the bank insolvency probability constant if the given loan is added to the portfolio. Conceptually, total capital can be split into expected and unexpected loss portions. In an accounting sense, the loan loss reserve might be viewed as covering the expected loss and equity as covering the unexpected loss. For more details, see "Credit Risk Models and Major U.S. Banking Institutions: Current State of the Art and Implications for Assessments of Capital Adequacy," *Federal Reserve System Task Force on Internal Credit Risk Models* (Board of Governors of the Federal Reserve System, May 1998).

Interviews indicated clearly that the introduction of risk-sensitive profitability analysis puts significant new pressures on the risk grading system. Pressure to rate loans favorably arises because expected losses and capital allocations are lower for lower-risk loans. Some institutions found that many loans were upgraded shortly after the introduction of profitability analysis, although the overall degree of the shift was small. One institution specifically mentioned an upward bias of about one-half grade relative to previous rating practice. Many noted that the number of disagreements in which relationship managers pressed for more favorable ratings increased once such systems were put into place.

In addition to pressure for more favorable ratings, rating-sensitive profitability analysis also creates pressure to increase the number of rating categories. This pressure, which comes both from the business line staff and the profitability analysis unit itself, arises because some of the loans in any given grade are less risky than other loans in that grade and thus should bear smaller credit costs. Creation of more grades allows for better recognition of such risk differences. Institutions reported that the pressure to increase the number of grades has become more pronounced in recent years as competitive forces have compressed loan spreads; in this setting, reducing expected loss factors by a few basis points, or slightly reducing the amount of capital allocated to the loan, may be the difference between a transaction that meets internal profitability “hurdles” and one that does not.

These stresses place increased pressure on the loan review unit to maintain discipline and enforce consistency, stability, and accuracy. Controlling rating biases is always a challenge. As the number of grades on the scale increases and the distinctions of risk become finer, disagreements about ratings naturally arise more frequently, and the control of biases becomes even more difficult. The difficulty seems likely to be greatest just after the number of grades is increased because the loan review staff must enforce (and if necessary, develop) new cultural definitions for the grades. The latter task is somewhat easier at banks that use external referents in assigning or reviewing ratings, such as default probability models and agency ratings of borrowers; such referents give loan reviewers objective benchmarks to use in identifying problems and communicating with staff. Redesigns of the rating scale that split existing grades into smaller compartments are also easier to implement because the existing cultural definitions can be refined rather than replaced.

Risk-sensitive profitability analysis also increases the demand for internal data on loss experience and for mappings to external referents because the analysis demands relatively precise quantification of the risk characteristics of each grade. However, such analysis can also make existing data and mappings less useful, at least in the short run, because rating pressures or changes in architecture may, to some extent, change the effective meaning of grades.

Using Ratings to Trigger Administrative Actions

As noted, many banks include an internal watch grade on their scales in addition to the regulatory problem-asset grades (formally, the watch grade would be counted among the Pass grades). Reassignment of a loan to watch or regulatory grades typically triggers a process of quarterly (or even monthly) reporting and formal reviews of the loan. At institutions where the main use of ratings is for monitoring and regulatory reporting, RMs’ incentives are often the opposite of those introduced by rating-sensitive profitability analysis: Their main interest is to avoid getting caught assigning ratings that are *not risky enough*. Getting caught can have negative career implications, and thus RMs have an incentive to assign credits to the riskiest Pass grade that is not watch. For example, some banks are especially likely to penalize RMs when a loan review reassigns a Pass credit from one of the less risky grades into a regulatory grade. Penalties can be forthcoming even when a loan is reassigned from a less risky Pass grade into watch, but are likely to be less severe. Thus, in the absence of carefully designed controls, the presence of administrative grades in a rating system can reduce the accuracy of non-administrative Pass grade assignments. This sort of bias is less likely at the largest banks because the countervailing incentives of rating-sensitive profitability analysis are most likely to operate there.

However, incentives associated with rating-sensitive profitability analysis can reduce the effectiveness of administrative management of problem loans. The staff may delay assigning credits to watch or regulatory grades because of the negative implications for measured profitability. Thus, there is a certain tension in the simultaneous use of rating systems for administrative purposes and for profitability analysis. Such tension can be overcome with proper oversight, the implementation of which represents another burden on loan review functions.

Potential Uses of Internal Ratings by External Entities

Internal ratings are a potential source of information for bank investors and regulators. For example, disclosure of the profile of a bank's loans across its internal rating categories might enhance the ability of shareholders and analysts to assess bank risk.

Moreover, investors in securitizations of traditional commercial loans might benefit from information about the credit quality of the underlying assets. Some banks are reportedly considering using internal rating information in structuring such securitizations. For example, when loans in the securitized pool are paid off, the new loans replacing them may be required to be drawn from a particular internal grade. Obviously, to evaluate the attractiveness of the pool, investors (or rating agencies) must be able to understand the loss characteristics of each internal grade and must have confidence that such characteristics will remain stable over time. Thus, external validation of rating systems becomes necessary if internal ratings are to be used in securitizations. Such validation would appear to be quite difficult because each bank's rating scale is different, and the meaning of ratings is largely embedded in culture rather than in writing. Moreover, most banks do not have sufficient historical data on loss experience by internal grade to support objective measurements.

Internal ratings might also be used in bank supervision and regulation. As a banking supervisor, the Federal Reserve has long emphasized the importance of strong risk management practices at banks and has stated its desire to orient its activities more toward testing of risk management and control processes and somewhat away from testing of individual transactions. This preference allows for less intrusion into the operation of the bank and minimizes the restrictive effect of supervision on banking innovation.

Information on a bank's risk profile by internal grade and shifts in that profile over time could become a useful supervisory tool. Supervisors could use internal profile information as one consideration in evaluating the asset quality and credit risk management of large banks, probably on balance reducing the overall burden of supervision. For those institutions that map their internal ratings to external reference points, such as the S&P scale, supervisors could use the mapping to put large institutions roughly on a common scale (in a fashion similar to that shown in chart 3). While bearing in mind that this technique is very crude, analysis of risk profiles and of trends in profiles could provide valuable insights into credit

conditions and standards in the industry as well as at individual institutions. Continuing work by individual institutions to better understand the loss characteristics of loans in their own risk grades will be important to refining and interpreting such comparisons over time.

Internal risk grades could also become an explicit element in the evaluation of capital adequacy. The current risk-based capital regime (based on the 1988 Basle Accord) provides for lower capital weights on certain low-risk assets (for example, those that are government-issued or guaranteed) but applies the same capital requirement (that is, 8 percent) to essentially all loans to private borrowers regardless of the underlying risk. Internal risk grades might become one consideration in scaling capital requirements on business loans more closely to the loss characteristics of a bank's loan portfolio.

Greater supervisory reliance on internal credit risk ratings would require that supervisors be confident of the rigor and integrity of internal rating systems. Heretofore, examiners have sought to validate assignments to internal grades only as they relate to the regulatory problem-asset grades. If supervisors are to rely more heavily on Pass grade information, some degree of validation and testing would have to be extended to those grades as well.

External use of internal ratings would introduce new stresses on internal rating systems. In some respects, the stresses would parallel those associated with rating-sensitive profitability analysis. That is, incentives would arise to grade optimistically and to alter the rating system to produce more fine-grained distinctions of risk. However, new incentive conflicts would arise between outsiders on the one side and the bank as a whole on the other. Such new conflicts could overwhelm the checks and balances currently provided by internal review functions. Even in the absence of such incentive conflicts, external users might demand a greater degree of accuracy or consistency in rating assignments than that required internally. For both reasons, external reviews and validation of the rating system might be necessary. In addition, banks and external parties should both be aware that the additional stress imposed by external uses, if not properly controlled, could impair the effectiveness of internal rating systems as a tool for managing the bank's credit risk.³⁷

37. In the early 1990s, the National Association of Insurance Commissioners (NAIC) introduced a system of risk-based capital requirements for insurance companies in which requirements vary with the ratings of assets. Although such ratings are assigned by the NAIC's Securities Valuation Office (SVO), the SVO does take into account any ratings of assets published by major rating agencies. In

CONCLUDING COMMENTS

A bank's decisions about its internal rating system can have a material effect on its ability to manage credit risk. But development of internal rating system architectures and operating designs that are appropriate to the uses made of the ratings is an especially complex task. The central role of human judgment in the rating process and the variety of possible uses for ratings mean that internal incentives can influence rating decisions. Thus, careful design of controls and internal review procedures is a crucial consideration in aligning form with function.

No single internal rating system is best for all banks. Banks' systems vary widely largely because of differences in *business mix* and in the *uses* to which ratings are put. Among variations in business mix, the share of large-corporate loans in a bank's portfolio has the largest implications for its internal rating system. Banks with a substantial large corporate market presence are likely to benefit from a rating system that achieves fine distinctions among relatively low-risk credits, while other banks may find significantly less value in such distinctions. In addition, an independent credit staff is often solely responsible for rating large loans. Such an arrangement can greatly reduce potential incentive conflicts, but may involve per-loan costs that are too large to be economic for smaller loans, which are often rated by relationship managers. Smaller loans also pose less risk to bank earnings and capital, and thus grading errors and biases may be more tolerable.

Among the various uses of internal ratings, profitability analysis and product pricing models have the most significant implications for the rating system. At banks where such analysis is in place, ratings can have a material effect on the measured profitability of transactions and relationships and can directly or indirectly influence the compensation of bank staff. Thus, careful attention to review and control procedures that limit biases in ratings is important to the accuracy and consistency of internal ratings.

Profitability analysis also introduces pressures for rating systems with more risk grades. Relationship managers may press for such systems because of a desire to subdivide grades that cover broad ranges of risk, thereby allowing different expected loss and capital charges for exposures at different ends of the ranges. The groups that develop and maintain the

profitability analysis systems may also press for fine-grained distinctions in order to support better balancing of risk and return. However, internal rating systems with many grades may make review and control of grading both more difficult and more expensive because reasonable people are more likely to differ in their subjective judgments when differences between grades are small rather than large.

Our interviews indicate that certain practices can improve the quality of any internal rating system and are especially helpful to rating systems that support analytical functions such as profitability analysis and portfolio management. First, a bank with appropriate data describing its historical loss experience by internal grade and by different risk factors is better able to assess the predictive power of its ratings criteria and to estimate values of parameters needed for its analyses (such as grade-specific values of PD or EL). Second, assigning or reviewing ratings with the aid of agency ratings, statistical models of default probability, or other objective criteria helps limit the magnitude of rating biases. However, care must be used in mapping internal grades to external grades or other indicators to ensure that the desired results are achieved. Finally, internal ratings grounded in clear loss concepts are helpful in grade assignment and review because rating criteria can be clearly linked to different aspects of risk. For example, a system that has separate grades for default probability and loss in event of default can incorporate different effects for a wide variety of types of collateral. All three of these practices are likely to be helpful in refining the subjective judgments that are central to almost all rating systems.

By their nature, banks' credit cultures typically adapt slowly to changes in conditions. The rapid pace of change in risk management practice and the trend toward risk-sensitive profitability analysis has recently increased the stresses on credit cultures in general and internal rating systems in particular. Careful attention to the many considerations noted in this article can help accelerate the process of adjustment and thus the easing of stresses.

The use of internal ratings by external entities such as regulators and investors has the potential to introduce new stresses in which incentives conflicts that pit banks' interests against those of the external entities compound existing internal tensions. Use of internal ratings by entities outside the bank would probably require some external validation of the ratings and the systems that generate them. In our view, such validation is probably feasible, but careful development of a new body of practice will be required. □

the wake of this and other developments in the insurance industry, the rating agencies experienced substantial pressure from both issuers and investors (insurance companies) to assign favorable ratings to some assets, a new and difficult development for the agencies in that issuers and investors had traditionally applied opposing pressures.