



February 2001

Rating Methodology

Contact	Phone
<i>New York</i>	
Richard Cantor	1.212.553.1653
Ted Collins	
Eric Falkenstein	
David Hamilton	
Chee Mee Hu	
Chester Murray, Chair	
Sameer Nayar	
Robert Ray	
Everett Rutan	
Farisa Zarin	

**TESTING FOR RATING CONSISTENCY IN
ANNUAL DEFAULT RATES**

Product of the Moody's Standing Committee on Rating Consistency

Rating Methodology

Author

*Richard Cantor
Eric Falkenstein*

Senior Production Associate

Mark A. Lee

© Copyright 2001 by Moody's Investors Service, Inc., 99 Church Street, New York, New York 10007. All rights reserved. **ALL INFORMATION CONTAINED HEREIN IS COPYRIGHTED IN THE NAME OF MOODY'S INVESTORS SERVICE, INC. ("MOODY'S"), AND NONE OF SUCH INFORMATION MAY BE COPIED OR OTHERWISE REPRODUCED, REPACKAGED, FURTHER TRANSMITTED, TRANSFERRED, DISSEMINATED, REDISTRIBUTED OR RESOLD, OR STORED FOR SUBSEQUENT USE FOR ANY SUCH PURPOSE, IN WHOLE OR IN PART, IN ANY FORM OR MANNER OR BY ANY MEANS WHATSOEVER, BY ANY PERSON WITHOUT MOODY'S PRIOR WRITTEN CONSENT.** All information contained herein is obtained by **MOODY'S** from sources believed by it to be accurate and reliable. Because of the possibility of human or mechanical error as well as other factors, however, such information is provided "as is" without warranty of any kind and **MOODY'S**, in particular, makes no representation or warranty, express or implied, as to the accuracy, timeliness, completeness, merchantability or fitness for any particular purpose of any such information. Under no circumstances shall **MOODY'S** have any liability to any person or entity for (a) any loss or damage in whole or in part caused by, resulting from, or relating to, any error (negligent or otherwise) or other circumstance or contingency within or outside the control of **MOODY'S** or any of its directors, officers, employees or agents in connection with the procurement, collection, compilation, analysis, interpretation, communication, publication or delivery of any such information, or (b) any direct, indirect, special, consequential, compensatory or incidental damages whatsoever (including without limitation, lost profits), even if **MOODY'S** is advised in advance of the possibility of such damages, resulting from the use of or inability to use, any such information. The credit ratings, if any, constituting part of the information contained herein are, and must be construed solely as, statements of opinion and not statements of fact or recommendations to purchase, sell or hold any securities. **NO WARRANTY, EXPRESS OR IMPLIED, AS TO THE ACCURACY, TIMELINESS, COMPLETENESS, MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OF ANY SUCH RATING OR OTHER OPINION OR INFORMATION IS GIVEN OR MADE BY MOODY'S IN ANY FORM OR MANNER WHATSOEVER.** Each rating or other opinion must be weighed solely as one factor in any investment decision made by or on behalf of any user of the information contained herein, and each such user must accordingly make its own study and evaluation of each security and of each issuer and guarantor of, and each provider of credit support for, each security that it may consider purchasing, holding or selling. Pursuant to Section 17(b) of the Securities Act of 1933, **MOODY'S** hereby discloses that most issuers of debt securities (including corporate and municipal bonds, debentures, notes and commercial paper) and preferred stock rated by **MOODY'S** have, prior to assignment of any rating, agreed to pay to **MOODY'S** for appraisal and rating services rendered by it fees ranging from \$1,000 to \$1,500,000. PRINTED IN U.S.A.

Table of Contents

Executive Summary..... 5

Defining And Measuring Rating Consistency 6

 Measuring Consistency Along A Specific Dimension Of Credit Risk 6

 Measuring Consistency Based On Historical Default Rates 6

 Fluctuations In Default Probabilities Over Time..... 7

Analyzing Default Rate Consistency 9

 Using Historical Default Rates To Estimate Underlying Default Probabilities
 When Default Probabilities Are Constant Over Time 9

 Using Historical Default Rates To Estimate Underlying Default Probabilities
 When Default Probabilities *Fluctuate* Over Time 11

 Accounting For Persistence In Default Rate Shocks 13

 Significance Tests For Differences Between Historical Default Rates 15

*Three Case Studies: Bank v. Nonbanks, US v. Non-US Issuers,
And Utilities v. Nonutilities*..... 18

 Case Study I - Comparing The Default Rates Of US Speculative-Grade Banks
 And Nonbanks..... 18

 Case Study II - Comparing The Default Rates Of US And Non-US
 Speculative-Grade Issuers 22

 Case Study III - Comparing The Default Rates Of Ba-Rated US Utilities And Other
 US Ba-Rated Companies 24

Appendix: Recent Academic Research On Rating Consistency..... 27

 Inconsistencies Across Rating Agencies..... 27

 Inconsistencies Across Industries 27

 Inconsistencies Across Geographical Regions 27

 Implications For The Consistency Of Moody’s Ratings..... 28

Executive Summary

Investors, issuers, academics, and financial market regulators, alike, have been increasingly focusing their attention on rating consistency. To encourage and facilitate this scrutiny, Moody’s has for many years published historical default and debt recovery statistics. More recently, we have published commentary designed to provide guidance about the intended meanings of our bond ratings.¹ In the process, we have acknowledged historical differences in the meaning of our ratings across broad market sectors – namely, corporate finance, structured finance, and public finance. However, as discussed in prior research, due to certain steps we have taken internally, these differences are expected to diminish over time.²

This *Special Comment* addresses the issue of measuring rating consistency, and more specifically, it evaluates the reliability of historical default rates as estimates of the true underlying default probabilities associated with Moody’s ratings. The comment also provides some technical guidance to interpret the differences in historical annual default rates across bond market sectors. Finally, it concludes with three case studies that are designed to illustrate the usefulness of these methods for comparing the historical default rates of different bond market sectors.

Based upon this work, we make the following observations:

- Rating consistency is difficult to test because credit risk has multiple attributes – default probability, loss severity, and transition risk. When focusing more narrowly on just historical default rates, consistency still needs to be measured at multiple horizons; although consistency over longer investment horizons is clearly more central to the meaning of Moody’s ratings than consistency over shorter investment horizons.
- But differences in historical default rates can indeed be subjected to rigorous tests of statistical significance, as long as such tests incorporate the volatility and persistence of macroeconomic and sectoral shocks. The presence of macroeconomic and sectoral shocks means that historical default rates may vary across bond market sectors for fairly long periods of time without necessarily implying fundamental differences in underlying default probabilities. Nevertheless, the observed historical volatility of these shocks does impose limits on expected differences in observed default rates.
- For certain sectoral comparisons, such as for Banks vs. Nonbanks and for US vs. Non-US issuers, observed differences in historical speculative-grade default rates may appear significant if the presence macroeconomic and sectoral shocks is ignored. However, when the effects of these shocks are considered, the differences are no longer statistically significant. This stands in contrast to the experience of speculative-grade issuers in the Utilities sector when compared to Other Companies. In this comparison, the differences in historical default rates are statistically significant, regardless of whether or not the effects of annual default rate shocks are incorporated into the analysis.

1 For further discussion, see “The Evolving Meaning of Moody’s Bond Ratings,” *Moody’s Special Comment*, August 1999 and “Promoting Global Consistency for Moody’s Ratings,” *Moody’s Special Comment*, May 2000.

2 For example, we are placing increasingly greater emphasis on an expected-loss rating paradigm – to promote rating consistency across industries and geographical regions.

Defining And Measuring Rating Consistency

Moody's ratings provide capital market participants with a consistent framework for comparing the credit quality of debt securities. For example, securities that are currently rated single-A are generally expected to perform similarly to one another and similarly to how A-rated securities have performed in the past. Actual performance will, of course, vary significantly over time and across industries due to random sample fluctuations, irrespective of any difference in expected average loss rates.

MEASURING CONSISTENCY ALONG A SPECIFIC DIMENSION OF CREDIT RISK

Rating consistency is, however, difficult to measure with quantitative precision. A credit rating compresses three characteristics or credit quality – default probability, loss severity, and rating transition risk – into a single symbol that is meant to be relevant to a variety of potential investment horizons and user constituencies. Bonds with the same credit rating, therefore, may be comparable with respect to overall credit quality, but will generally differ with respect to specific credit quality characteristics. For example, debt issuers that are subject to greater potential ratings volatility (transition risk) are generally rated lower than issuers that share the same default probability but have lower transition risk.³ Moreover, although two issuers carry the same rating, one might have a higher short-term risk of defaulting and another might have a higher long-term risk of defaulting. For all these reasons, it may be difficult, if not impossible, to definitively measure rating consistency.

Although these analytical considerations make the task of measuring rating consistency difficult, it is possible to measure whether the ratings assigned to issuers in different industries and geographical regions have had consistent performance along specific dimensions of credit risk. For instance, one can measure whether loss severity given default has varied systematically across sectors. A recent Federal Reserve study using Moody's data concludes that the average recovery rates on defaulted bonds have historically been lower for US banks than US non-banks, whereas the average recovery rates on non-US firms have been roughly equal to those of US firms.⁴

MEASURING CONSISTENCY BASED ON HISTORICAL DEFAULT RATES

To better measure its success at achieving rating consistency, Moody's regularly disaggregates its historical data to identify time periods and sectors in which historical default rates have been higher or lower than the long-term global averages for those rating categories. For reasons outlined above, Moody's does not intend that all corporate bonds carrying the same rating will have the same historical frequency (or even the same expected probability) of default at, say, the one-year, five-year, or ten-year horizon. Nevertheless, for broadly defined regions and industries, we generally expect that future default rate experience, measured over a suitably long period of time, will be similar for bonds that carry the same rating.

³ Within the corporate sector – which includes industrials, banks, and other financial institutions of both US and non-US issuers – Moody's has applied a consistent set of rating definitions to meet the needs of potential crossover investors. However, in response to the different needs of relatively segmented investor groups, Moody's has in the past applied different meanings to ratings it assigned in other market segments, such as the US municipal and the US investor-owned utility markets. In order to meet the needs of a growing number of crossover investors, Moody's expects that the meanings of its ratings will converge between these market segments and the corporate sector, a process that has already been underway for almost a decade. However, in response to the needs of investors who are highly sensitive to default and transition risk, Moody's continues to overweight those aspects of credit risk in certain sectors. For further discussion, see "The Evolving Meaning of Moody's Bond Ratings," Moody's Special Comment, August 1999.

⁴ Ammer, John and Frank Packer, "How Consistent Are Credit Ratings? A Geographical and Sectoral Analysis of Default Risk," Board of Governors of the Federal Reserve System, International Finance Discussion Paper #668, June 2000.

Identifying differences in historical default rates is a relatively straightforward exercise. It is difficult, however, to determine which differences are significant and which are insignificant. A common approach, found in the academic literature, measures the statistical significance for the difference in two historical average default rates under the assumption that the default rates are drawn from independent binomially distributed sample populations.⁵ This test of statistical significance is valid, however, only if the sample population default rates in each sector are expected to be constant over time.

In fact, we expect and regularly see large, persistent fluctuations in annual default rates that are due to unpredictable changes in the economic environment. These annual default rate “shocks” to expected default rates come in at least two varieties – “global/macroeconomic” shocks which affect all industries and regions, and “sector” shocks which affect individual industries or regions. The presence of these shocks renders the standard significance test invalid. In particular, the standard test will tend to overstate the significance of observed differences in default rates.

FLUCTUATIONS IN DEFAULT PROBABILITIES OVER TIME

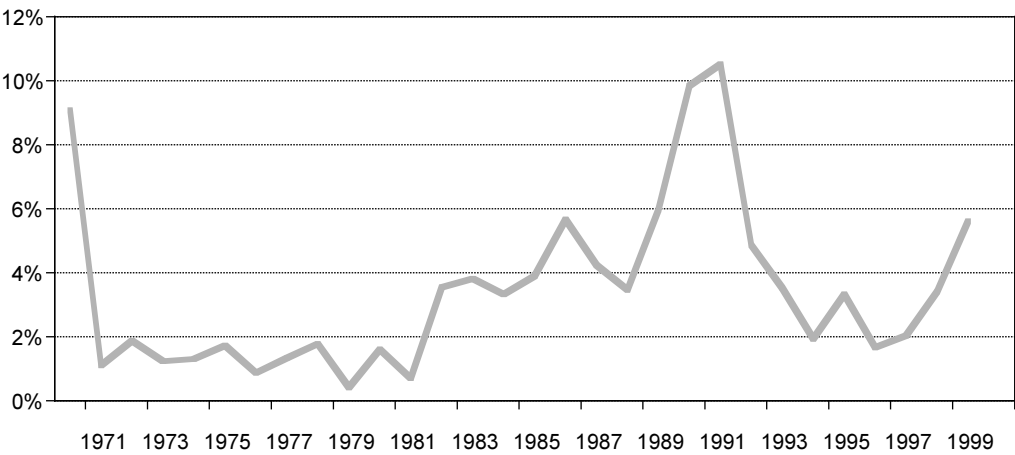
Macroeconomic Shocks Often Explain Shifts in Aggregate Default Rates

Annual default rates for specific rating categories fluctuate much more widely than would be expected if their true underlying default probabilities were constant over time. While this observation may not be surprising for particular industry sectors where sector shocks are likely to be substantial, it is also true for the broadest aggregations which, though diversified by sector, remain vulnerable to macroeconomic shocks.

The Example of Speculative-Grade Issuers in Early 1990s

Chart 1 displays the one-year default rates for Moody’s speculative-grade rating universe between 1970 and 1999. The observed fluctuations in annual default rates are largely the result of changes in macroeconomic and financial market conditions. From a statistical perspective, these fluctuations cannot be ascribed to idiosyncratic risk, i.e., unusual draws from independent, binomially distributed sample populations.

Chart 1
One-Year Default Rates For Speculative-Grade Issuers
Moody’s Global Database 1970–1999



5 See, for example, Nickell, Pamela, William Perraudin, and Simone Varotto, “Stability of Rating Transition Matrices,” Journal of Banking and Finance, 24 (1&2), 2000.

For example, the default rates in 1991 and 1996 were 9.9% and 1.6% respectively. What is the chance that these two default rates represent draws from the same underlying statistical distribution? The probability of this occurring by chance is effectively zero.⁶ A much more reasonable interpretation, however, is that junk bond financing was much tighter and the macroeconomic environment was much weaker in 1991 than in 1996. Hence, a larger fraction of speculative-grade issuers were more likely to fail in 1991 than 1996. Notice also that changes in the default rate tend to be persistent. For example, a high (9.4%) default rate was observed in 1990 and was followed by another year of high defaults (9.9%) in 1991.

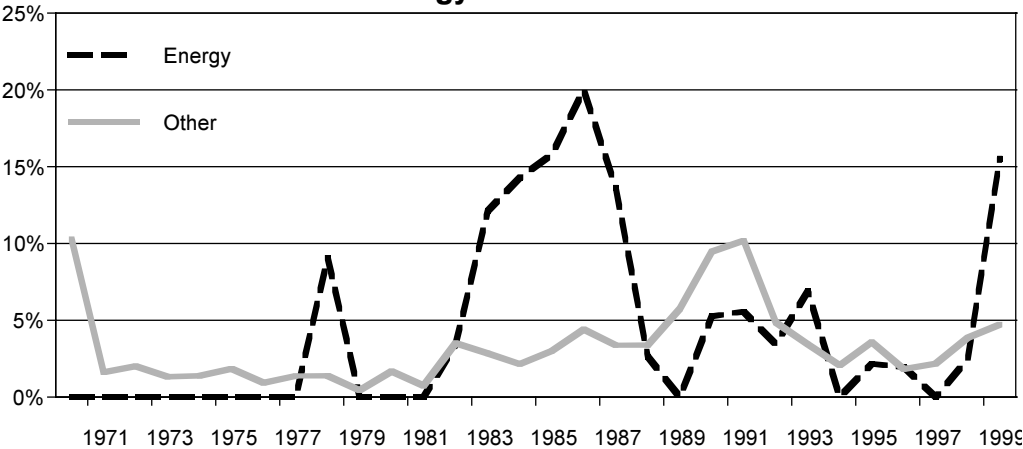
Individual Sectors Also Subject To Random Shocks

Default rates for specific bond market sectors tend to be more irregular than the aggregate default rate. One should expect greater variations in sectoral default rates than aggregate default rates, because by definition individual sectors have fewer issuers than the composite, and are thus subject to a greater degree of uncertainty. But random variation cannot explain the bulk of observed differences in realized default rates across sectors. Individual sectors are also subject to random shocks to their underlying default rates.

The Example Of US Speculative-Grade Energy Companies In Mid-1980s

For example, during the mid-1980s, the default incidence of speculative-grade companies in the US energy sector was considerably higher than that of other similarly rated US companies. One could argue that the oil and gas industry was rated too high during the early 1980s. An alternative interpretation, however, is that the oil and gas industry endured an extraordinary adverse shock. While the possibility of an oil price decline was factored into the ratings in the early 1980s, the ratings assigned to companies in the oil and gas industry also reflected the fact that the probability of the shock occurring was still fairly low.

Chart 2
One-Year Speculative-Grade Default Rates On
US Energy And Other Issuers



6 The standard test for a difference in means reveals that this 8.3% difference in default rates (=9.9%-1.6%) is eight standard deviations away from zero. There were 726 speculative-grade issuers in 1991, and 1073 in 1996, of which 72 and 17 defaulted in those respective years. Under the (null hypothesis) assumption that the underlying default probabilities were the same and equal their weighted average, which is 4.9% ($4.9\% = (72+17)/(726+1073)$), the appropriate test statistic is $8.3\% / \sqrt{4.9\% \cdot (1-4.9\%) \cdot (1/726 + 1/1073)} = 8.0$, indicating that the probability of observing this difference in default rates equals the probability of observing a normally distributed random variable 8 standard errors away from its mean.

Consider the historical data presented in Chart 2. The annual default rate was zero in the energy sector through much of the 1970s, it spiked in 1978, and went to zero again for a few years thereafter. During the mid-1980s, however, the industry experienced five years of double-digit default rates. Following the 1999 oil price shock, energy sector default rates have shot back up to 16%.

Should we expect higher default rates in this sector compared to others going forward? The answer is not obvious. Over the sample period, the energy sector’s average default rate was 6.0%, whereas the nonenergy sector’s average default rate was 3.9%. Yet, the difference is much smaller (4.5% for energy and 3.3% for nonenergy) if one calculates simple average, rather than issuer-weighted, average annual default rates. Moreover, dropping just one year, 1999, from the sample would lower the energy sector’s weighted average default rate to 4.9%. Dropping the mid-1980s from the sample (which represents only one oil price shock) would imply a lower default rate for the energy sector than that of the nonenergy sector.

Analysis Confirms Intuition – Observed Default Rates May Reflect Event Risk

This type of impressionistic data analysis provides strong evidence that macroeconomic and sectoral shocks to default rates are important and relevant to one’s interpretation of observed differences in default rates. Our thesis is simple. With at most 30 years of data used for most analyses, macroeconomic and sectoral shocks to underlying default rates materially affect the sample default rate volatility. Higher volatility, in turn, affects the statistical significance of differences in average default rates. Thus, what at first may appear to be a significant difference in default rates, often turns out to be insignificant when reasonable estimates of the variability of default rate shocks are incorporated in the analysis.

Analyzing Default Rate Consistency

USING HISTORICAL DEFAULT RATES TO ESTIMATE UNDERLYING DEFAULT PROBABILITIES WHEN DEFAULT PROBABILITIES ARE *CONSTANT* OVER TIME

Objective

This section presents the standard method for estimating the precision with which a historical annual default rate can be used to infer the underlying annual probability of default associated with a given rating level in a particular bond market sector.⁷ The method presented is, however, only valid when the underlying probability of default is constant over time.

Approach

Define the following notation:

- p = the probability of default
- n_t = the number of issuers in year t
- d_t = the number of default in year t
- dr_t = d_t/n_t = the default rate in year t
- T = the number of years in the sample
- N = $\sum_{t=1}^T n_t$ = total number of issuers over T years
- D = $\sum_{t=1}^T d_t$ = total number of defaults over T years
- DR = D/N = the weighted average default rate over T years

⁷ Extending the analysis to multi-period default rate horizons is relatively straightforward, although adjustments must be made for the use of overlapping data sets.

If the underlying default probability is constant and equal to p , then it is well known that in large samples the empirical default rates, dr_t and DR , will both be approximately normally distributed with their means equal to p and their standard deviations equal to $\sqrt{\frac{p(1-p)}{n_t}}$ and $\sqrt{\frac{p(1-p)}{N}}$, respectively.

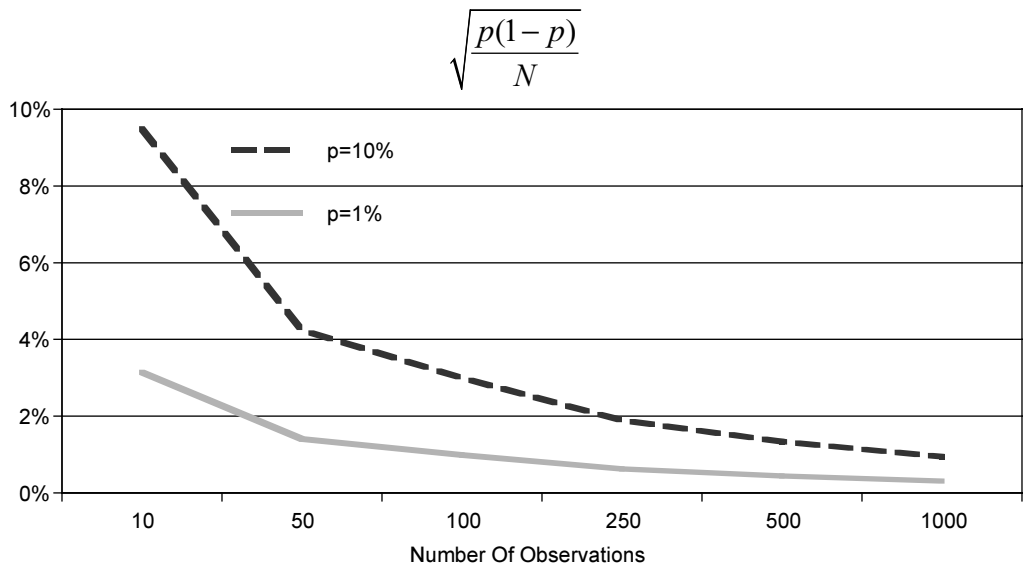
$$dr_t \sim N\left(p, \sqrt{\frac{p(1-p)}{n_t}}\right) \quad \text{and} \quad DR \sim N\left(p, \sqrt{\frac{p(1-p)}{N}}\right) \tag{1}$$

The standard deviation of the historical default rate is an extremely useful measure of empirical precision: the true underlying default probability is highly unlikely to be more than two or three standard deviations (if the standard deviation is measured correctly) from the underlying default rate.

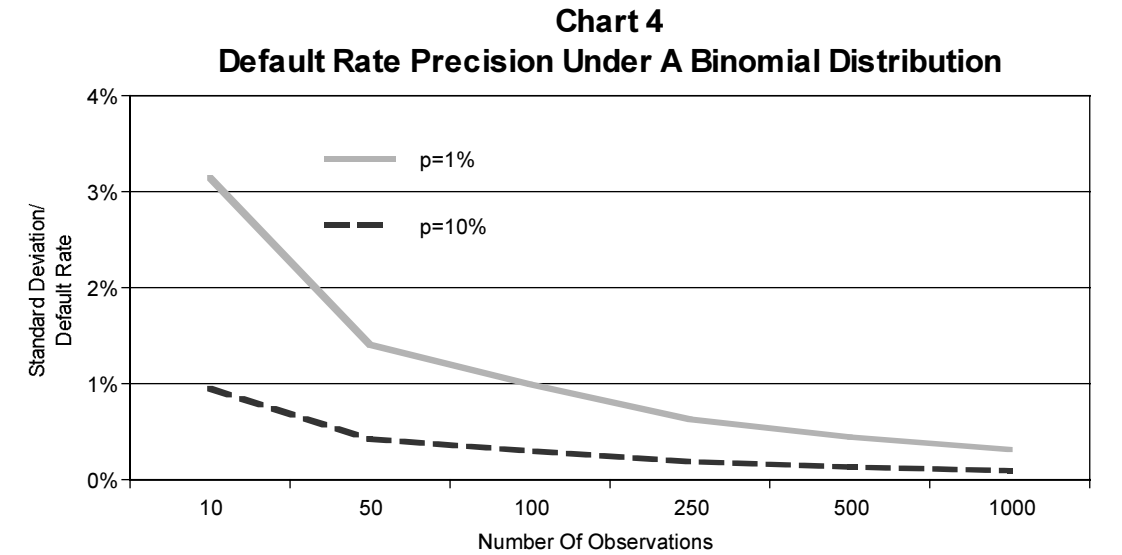
Findings

As shown in Chart 3, this framework implies that the standard deviation of the historical default rate declines rapidly as the sample size increases. The chart depicts two cases, one in which the underlying default probability is 1% and one in which it is 10%. Not surprisingly, default rate volatility is lower when the underlying expected default rate is lower.

Chart 3
Default Rate Standard Deviation Under A Binomial Distribution



However, as can be seen in Chart 4, the historical default rate is actually a less precise estimator when the underlying default rate is low if precision is measured as the ratio of the standard deviation to the mean, $\frac{\sigma}{\mu}$, which is often termed the ‘coefficient of variation.’



When the underlying default rate is 10%, only about 10 observations are needed to obtain a level of precision of about one, whereas over 100 observations are needed to obtain that level of precision if the underlying default rate is 1%. As a result, relatively more observations are necessary to evaluate consistency for investment-grade ratings than for speculative-grade ratings.

USING HISTORICAL DEFAULT RATES TO ESTIMATE UNDERLYING DEFAULT PROBABILITIES WHEN DEFAULT PROBABILITIES *FLUCTUATE* OVER TIME

Objective

This section extends the basic analysis of the previous section to the case in which the underlying probability of default varies over time, either due to fluctuations in the macroeconomic environment or conditions within a specific bond market sector.

Approach

To incorporate the concept of time-varying underlying default probabilities, we assume that each year’s default rate is normally distributed around a long-term underlying probability p with an annual shock of mean zero and standard deviation, σ .

$$\begin{aligned}
 p_t &= p + \varepsilon_t \\
 \varepsilon_t &\sim N(0, \sigma)
 \end{aligned}
 \tag{2}$$

Under these assumptions, the normal approximation to the probability distribution of the default rate for a single year is updated to include this new standard deviation term in an additive way:

$$dr_t \sim N\left(p, \sqrt{\frac{p(1-p)}{n_t} + \sigma^2}\right) \quad (3)$$

Findings

As might be expected, the precision of one year's annual default rate as an estimator of the long-term probability of default improves as n_t , the sample size within a given year, increases. But it can only improve so much. Multi-year observations are necessary in order to reduce the standard deviation below σ , which is likely to be about 1.5 times the default probability for a diversified portfolio of investment-grade credits and approximately equal to the default probability for a diversified portfolio of speculative-grade credits.⁸

Since each year's average default rate is approximately normally distributed, the average one-year default rate measured over multiple years is also approximately normally distributed. Furthermore, because the annual shocks to the default rate are assumed to be independent, and each year n_t companies are subject to independent time shocks, the multi-year default rate for the entire sample is distributed as follows,⁹

$$DR \sim N\left(p, \sqrt{\frac{p(1-p)}{N} + \frac{\sigma^2}{N^2} \sum_{t=1}^T n_t^2}\right) \quad (4)$$

The latter term under the square root sign comes from the fact that each time shock affects the total sample default rate by $\frac{n_t}{N}$, and assuming independence each time period's variance effect is reflected as $\sigma^2 \frac{n_t^2}{N^2}$. As the number of periods grows, $\frac{n_t}{N} \rightarrow 0$ and

$\sum_{t=1}^T \left(\frac{n_t}{N}\right)^2 \rightarrow 0$. Total volatility of the default rate estimate does go to zero as a function of *both* the number of observations (N) and number of time periods (T).

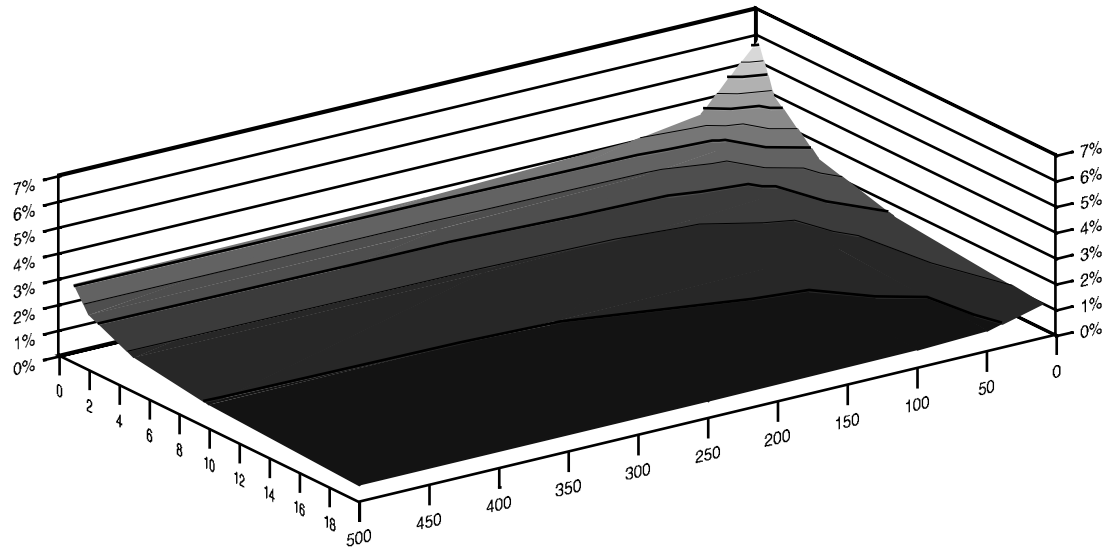
⁸ These rules of thumb were derived from a Monte Carlo simulation based on Moody's Corporate Bond Database, 1970-99. In each draw of the simulation, a random year was selected and portfolios of 1,000 random obligors within a specific rating category each were chosen with replacement. (With "replacement" means that firms could be counted more than once.) A default rate was then calculated for each draw. This draw was then repeated 10,000 times. Given the large sample size of each annual cohort, idiosyncratic risk was for all practical purposes eliminated from the estimated default rates. The rules of thumb reported above were derived from the estimated standard deviations of the 10,000 simulated default rates for each rating category.

⁹ As discussed in the following section entitled, "Accounting for Persistence in Default Rate Shocks," the formula needs to be modified if the annual shocks to the default rate are not independently distributed. If, as is quite likely, default rate shocks are positively serially correlated, the formula for the variance of long-term average default rate is more complicated but still solvable. From a practical perspective, simply assuming a higher volatility for the annual shocks can capture the main effect of serially correlated shocks. Positive serially correlation implies that annual shocks have larger impacts, extending over a number of years. Hence, like a higher variance assumption for annual shocks, positive serial correlation reduces the precision with which historical default rates can be used to infer underlying expected default rate parameters.

The coefficient on σ^2 , the time shock, is a function of the number of time periods observed and the heterogeneity of those periods. If all periods have equal numbers of firms the coefficient is minimized at $1/T$, and sample variance declines proportional to the size of the time sample. If periods have different numbers of firms, however, the coefficient is somewhat higher than $1/T$, but still strictly declining with the increasing size of the time sample.

Using equation (4), Chart 5 below illustrates how both the sample size (N) and the number of time periods (T) affects the standard deviation of the sample default rate as an estimate of the true default rate. For this illustration, we assumed an underlying default probability of 3.0% and an annual default volatility also of 3.0%, which approximately matches the speculative-grade default experience since 1970. We also assumed an approximately equal number of defaulting firms.¹⁰ The bottom line is that sample size is not enough, time is also important. This is particularly relevant to judging new areas. As a new sector may contain many obligors that been rated during only one to two credit cycles, the sector’s current size may give misleading signals on the reliability of its historical default rate as a guide to its long-term expected default rate.

Chart 5
Standard Deviation Of The Sample Default Rate
Example: $p = 3\%$ and $\sigma = 3\%$



ACCOUNTING FOR PERSISTENCE IN DEFAULT RATE SHOCKS

Objective

Thus far, we have assumed that the annual shocks to the default rate distribution are independent from one another, so that a high default rate year is equally likely to be followed by a low default year as another high default year. In practice, we know that the shocks to default rates are fairly persistent; that is to say, high default years are more likely to be followed by another high default year. In this section, we demonstrate that that the historical default rate becomes a much less precise estimator of the long-run underlying default rate if default rate shocks are highly persistent.

¹⁰ For a large sample such as the 1970-99 total speculative-grade universe, with over 17,000 firm-year observations, the annual default rate volatility, σ , is approximated well (i.e., within 0.1%) by the simple standard deviation of annual default rates.

Approach

For simplicity, assume that the persistence in default rate shocks can be modeled as a simple first-order autoregressive process,¹¹

$$p_t = p + \theta(p_{t-1} - p) + \varepsilon_t, \text{ where } 0 < \theta < 1. \quad (5)$$

For example, if the default rate is 10% above its long-run expected level (p), then equation (6) implies that next year's default rate is expected *on average* to be θ times 10% higher than its normal long-run value.

Findings

While this persistence has no impact on the unconditional distribution of any individual year's default rate, it can greatly increase the volatility of the average default rate measured over multiple years. Intuitively, when default rate shocks are independently distributed, it is quite reasonable to expect them to "average out" over, say, twenty years. If, however, default rate shocks are strongly persistent, then one shock can easily have the impact of three, four or more shocks on a particular twenty-year sample, so that it becomes less likely that annual shocks will average out over the same time period.

When shocks to the default rate are persistent, the formula that characterizes the volatility of a multi-year historical default rate can be calculated in a straightforward manner – although the result is a messy algebraic expression because each year's annual default rate is correlated with every other year's default rate. If the underlying long-run default rate is p , the number of observation years is T , the number of observations each year is n_t (with $t=1,2,\dots,T$), the default rate shock has a volatility of σ , and the shock has an autoregressive parameter θ , then the historical default rate, DR , is approximately normally distributed as follows:

$$DR \sim N \left(p, \sqrt{\frac{p(1-p)}{N} + \frac{\sigma^2}{N^2} X} \right), \text{ where } X \text{ is defined as } \dots \quad (6)$$

$$X = \left\{ \begin{aligned} & (n_T)^2 + (n_{T-1} + \theta n_T)^2 + (n_{T-2} + \theta n_{T-1} + \theta^2 n_T)^2 + (n_{T-3} + \theta n_{T-2} + \theta^2 n_{T-1} + \theta^3 n_T)^2 + \dots \\ & + (n_1 + \theta n_2 + \theta^2 n_3 + \dots + \theta^{T-1} n_T)^2 + \sum_{i=1}^{\infty} \theta^{2i} (n_1 + \theta n_2 + \theta^2 n_3 + \dots + \theta^{T-1} n_T)^2 \end{aligned} \right\} \quad (7)$$

Note that the infinite series in expression (7) converges as long as $\theta < 1$, and if θ equals zero, this complicated distribution collapses to the distribution given in equation (4).¹²

When default rate shocks are persistent, the explicit formula for the standard deviation of the average default rate is a complicated algebraic expression for a straightforward reason. Each

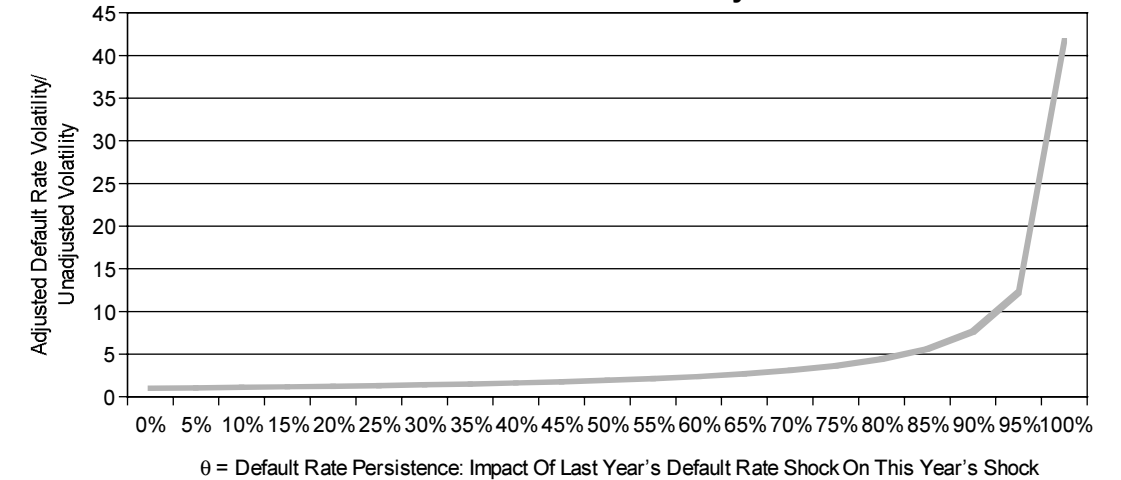
¹¹ While the assumption that default rates follow a first-order autoregressive process is reasonable, the approach presented here is quite general and can be applied to any autoregressive or moving average evolution of default rate shocks.

¹² The derivation of this formula is available upon request from the authors.

year's annual default rate is correlated with every other year's default rate, but the number of firms affected each year is different. The difference in the number of firms in each period creates complications. In many bond market sectors, historical data sets start with very small sample populations but then grow rapidly over time. To properly measure the expected variability of the long-term average default rate, it is imperative to capture the fact that early shocks affect only a handful of firms, while later shocks affect many more.

The effects of different values of θ on the expected volatility of multi-period default rates are presented in Chart 6. Here we have assumed that the sample consists of 20 years of data on 1000 firms in each year. Because of the large sample size, the effect of idiosyncratic risk on the historical default rate is nil because $p(1-p)/N$ is close to zero. As shown in the chart, the standard error of the historical default rate rises exponentially with increases in the default persistence parameter. Volatility rises by 25% if $\theta = 20\%$, by 50% if $\theta = 35\%$, by almost 100% if $\theta = 50\%$, by over 300% if $\theta = 70\%$, and so on. Of course, if $\theta = 100\%$, every default rate shock would last forever and default rate volatility would be unbounded.

Chart 6
Effect Of Persistence Of Default Rate Shocks On
Default Rate Volatility



SIGNIFICANCE TESTS FOR DIFFERENCES BETWEEN HISTORICAL DEFAULT RATES

Objective

Suppose we observe two different historical default rates, DR_1 and DR_2 , for two bond sectors, 1 and 2. How do we know whether (1) the sectors' long-term underlying default probabilities, p_1 and p_2 , are the same and the difference in default rates is due to pure chance, or (2) the underlying default probabilities are in fact the same (i.e., $p_1=p_2$)? The following discussion offers an approach for answering this question.

Generalized Approach

Statisticians generally address this question by estimating the probability of observing the difference in default rates ($DR_1 - DR_2$), under the assumption (the “null hypothesis”) that the true underlying default probabilities are the same and equal to the weighted average of their historical

default rates, $\hat{p} = \frac{N_1DR_1 + N_2DR_2}{N_1 + N_2}$. If the probability distribution of each sector’s default rate is

approximately normally distributed, the appropriate test statistic for this null hypothesis is the same one used whenever one tests for a difference in means between two normal distributions.

Without shocks to the underlying default probabilities, the test statistic is derived from the probability distribution of each sector's average default rate as summarized by equation (2), which is rewritten here for convenience.

$$DR_i \sim N\left(\hat{p}, \sqrt{\frac{\hat{p}(1-\hat{p})}{N_i}}\right), \text{ for sectors } i=1, 2, \quad (2)$$

where \hat{p} is the underlying default probability and $\sqrt{\frac{\hat{p}(1-\hat{p})}{N_i}}$ is the default rate's standard deviation.

Test Statistic When Default Probabilities Are Constant Over Time

Null Hypothesis H_0 : $E[DR_1 - DR_2] = 0$

$$\text{Test Statistic: } Z = \frac{DR_1 - DR_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} \quad (7)$$

This is just a familiar test statistic where the difference in the means of the two samples is divided by the standard deviation of that difference, which is simply the square root of the sum of the two variances when the distributions are independent. Under these assumptions, the Z statistic itself has a standard normal distribution, and the “two-sided test” likelihood (p-value) of observing this difference in default rates is given $2 \cdot (1-F(Z))$, where $F(\bullet)$ is the cumulative normal distribution function.

When the sectors' default probabilities themselves vary over time, the relevant distributions of the historical default rates are given by equation (6), again rewritten here for the sake of convenience,

$$DR_i \sim N\left(\hat{p}, \sqrt{\frac{\hat{p}(1-\hat{p})}{N_i} + \frac{\sigma_i^2}{N^2}} X_i\right), \text{ where } X_i \text{ is defined as } \dots \quad (6)$$

$$X_i = \left\{ \begin{aligned} & (n_{i,T})^2 + (n_{i,T-1} + \theta_i n_{i,T})^2 + (n_{i,T-2} + \theta_i n_{i,T-1} + \theta_i^2 n_{i,T})^2 + (n_{i,T-3} + \theta_i n_{i,T-2} + \theta_i^2 n_{i,T-1} + \theta_i^3 n_{i,T})^2 + \dots \\ & + (n_{i,1} + \theta_i n_{i,2} + \theta_i^2 n_{i,3} + \dots + \theta_i^{T-1} n_{i,T})^2 + \sum_{j=1}^{\infty} \theta_i^{2j} (n_{i,1} + \theta_i n_{i,2} + \theta_i^2 n_{i,3} + \dots + \theta_i^{T-j-1} n_{i,T})^2 \end{aligned} \right\}$$

for sectors $i=1,2$.¹³

¹³ The derivation of this formula is available from the authors upon request.

The appropriate test statistic, however, must not only account for the new volatility terms, σ_1 , and σ_2 , but also for the potential correlation of these shocks. In particular, the more correlated the shocks are with one another, the less likely they are to cause observed default rates to vary from one sector to another.

Specifically, assume samples, 1 and 2, which have the following default rate structure:

$$\begin{aligned}
 p_{1t} &= \hat{p} + \theta_1(p_{1,t-1} - p) + \varepsilon_{1t}, \\
 p_{2t} &= \hat{p} + \theta_2(p_{2,t-1} - p) + \varepsilon_{2t}, \\
 \varepsilon_{1t} &\sim N(0, \sigma_1^2) \\
 \varepsilon_{2t} &\sim N(0, \sigma_2^2) \\
 E(\varepsilon_{1t} \varepsilon_{2t}) &= \sigma_{12} \\
 \rho_{12} &= \sigma_{12} / \sigma_1 \sigma_2
 \end{aligned} \tag{7}$$

Again, the Z-statistic for the difference in means is simply the difference divided by the expected standard deviation of the difference. In this case, however, we need to take into account the potential covariance between the shocks to the two sectors. Using the fact that

$$Var(x - y) = \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}, \tag{8}$$

the appropriate test statistic is given by

$$Z = \frac{DR_1 - DR_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{N_1} + \frac{1}{N_2}\right) + \sum_{i=1}^2 \frac{\sigma_i^2}{N_i^2} X_i - \frac{2\sigma_{12}}{N_1 N_2} Q}}, \text{ where} \tag{9}$$

$$Q = \left\{ \begin{aligned} &n_{1,T}n_{2,T} + (n_{1,t-1} + \theta_1 n_{1,T})(n_{2,t-1} + \theta_2 n_{2,T}) + (n_{1,T-2} + \theta_1 n_{1,T-1} + \theta_1^2 n_{1,T})(n_{2,T-2} + \theta_2 n_{2,T-1} + \theta_2^2 n_{2,T}) + \dots \\ &+ (n_{1,1} + \theta_1 n_{1,2} + \theta_1^2 n_{1,3} + \dots + \theta_1^{T-1} n_{1,T})(n_{2,1} + \theta_2 n_{2,2} + \theta_2^2 n_{2,3} + \dots + \theta_2^{T-1} n_{2,T}) \\ &+ \sum_{j=1}^{\infty} \theta_1^j (n_{1,1} + \theta_1 n_{1,2} + \theta_1^2 n_{1,3} + \dots + \theta_1^{T-1} n_{1,T}) \theta_2^j (n_{2,1} + \theta_2 n_{2,2} + \theta_2^2 n_{2,3} + \dots + \theta_2^{T-1} n_{2,T}) \end{aligned} \right\}$$

Note that, if the two sectors have perfectly correlated default rate shocks, identical sample sizes, and the same persistence parameters, the effect of default rate shocks on the distribution of the mean difference in default rates disappears.¹⁴

¹⁴ The derivation of this formula is available from the authors upon request.

If we substitute the correlation coefficient for the covariance term equation (9), we have a complex but highly tractable test statistic for the difference in two mean default rates.

Test Statistic When Default Probabilities Fluctuate Over Time

Null Hypothesis H_0 : $E[DR_1 - DR_2] = 0$

$$Z = \frac{DR_1 - DR_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{N_1} + \frac{1}{N_2}\right) + \sum_{i=1}^2 \frac{\sigma_i^2}{N_i^2} X_i - \frac{2\rho_{12}\sigma_1\sigma_2}{N_1N_2} Q}} \quad (10)$$

Equation (10) puts the problem into a tractable form that accounts for all the variables need to test for the statistical significance of an observed difference in two sector's historical default rates:

- Mean default rates (DR_1, DR_2)
- Sample sizes (n_1, N_1, n_2, N_2)
- Persistence parameters (θ_1, θ_2)
- Correlation coefficient (ρ_{12})
- Default rate shock volatilities (σ_1, σ_2).

In general, the probability of an observed difference in default rates can be consistent with identical underlying default probabilities increases whenever the denominator of the equation (10) is large. Therefore, large differences in mean default rates can be expected when sample sizes are small, annual default rate volatilities are large, persistence parameters are large, and correlation coefficients are small

Three Case Studies: Bank v. Nonbanks, US v. Non-US Issuers, And Utilities v. Nonutilities

CASE STUDY I - COMPARING THE DEFAULT RATES OF US SPECULATIVE-GRADE BANKS AND NONBANKS

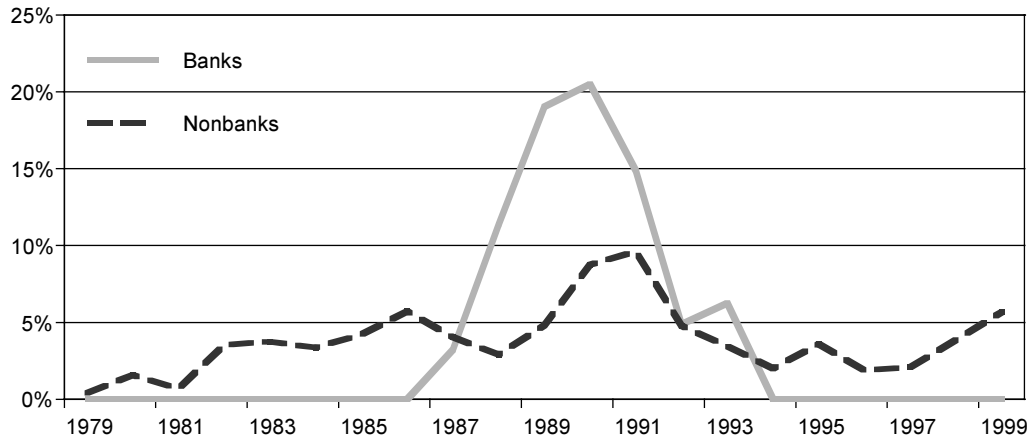
What the Data Shows

A comparison of average annual default rates of US bank vs. US nonbank speculative-grade issuers illustrates many of the concerns that arise when assessing this dimension of rating consistency. Moody's assigned its first (post-World War II) speculative-grade bank rating in 1979. Since then, 33 of the 434 US banks that began any given year with a speculative-grade rating, failed during that same year. By contrast, 559 of 13,401 speculative-grade nonbanks failed during the year they were rated speculative-grade. Using data such as these, some commentators have concluded that the annual historical default rate on speculative-grade US banks (7.61%) has been *significantly* higher than the corresponding default rate on US nonbanks (4.17%).¹⁵

¹⁵ See, for example, Ammer, John and Frank Packer, "How Consistent Are Credit Ratings? A Geographical and Sectoral Analysis of Default Risk," Board of Governors of the Federal Reserve System, International Finance Discussion Paper #668, June 2000. Following Ammer and Packer, we are using a broad definition of the banking sector that includes thrifts,

As depicted in Chart 7, the differences between the two default rate time series are also quite striking. Since 1979, the nonbank annual default rate has fluctuated fairly randomly between 2% and 6%, except that it reached 9% and 10% in 1990 and 1991, respectively. In contrast, the bank default rate was literally zero between 1979 and 1986, and again zero between 1994 and 1999. Between 1987 and 1993, however, the bank annual default rate averaged 11.5%.

Chart 7
Annual Default Rates Of Speculative-Grade Issuers

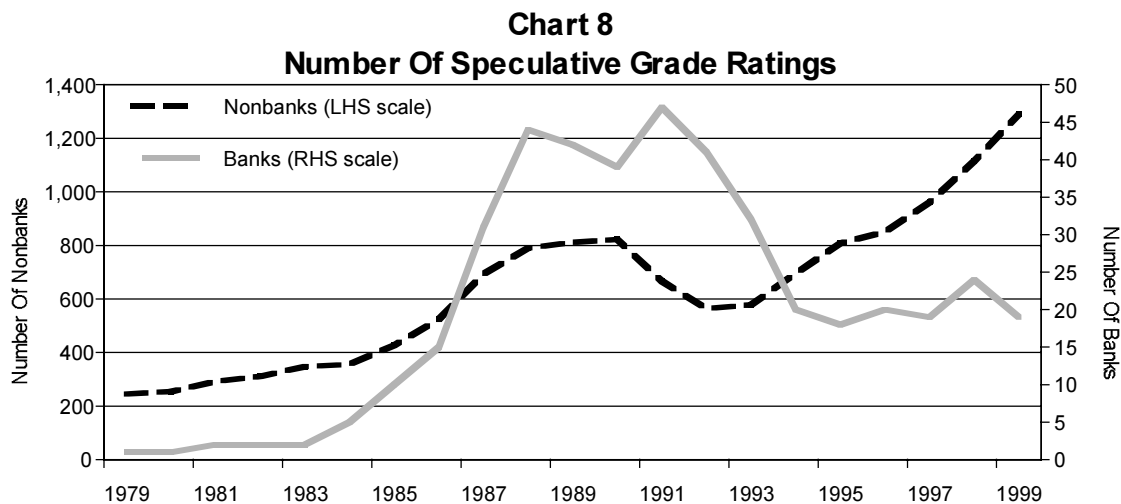


As depicted in Chart 8, the number of speculative-grade banks has historically been small compared to the number of speculative-grade nonbanks. Moreover, the time patterns of the two time series are quite different. The growth in speculative-grade nonbanks has a pronounced positive secular trend, whereas the number of speculative-grade banks peaked in the late 1980s and early 1990s.

How should these default statistics be interpreted? Does the higher default rate on banks imply that, in terms of annual default rates, speculative-grade banks are riskier than nonbanks?¹⁶ Or, does the banking industry’s greater default rate variability imply that the historical difference between bank and nonbank default rates is a statistical artifact, which is unlikely to be preserved over a longer sample period?

which account for many of the sector's historical defaults. Note that this example and the one that follows contrast the sectoral default rates of "speculative-grade" issuers, similar analysis could have been conducted on specific rating categories, such as Ba-rated issuers, because the percentage distributions of Ba-rated, B-rated and Caa-rated issuers were very similar across all the sectoral samples.

¹⁶ One might also challenge the consistency of the ratings over time. Were the ratings on banks too low between 1979 and 1986, and again between 1994 and 1999? Were they too high between 1987 and 1999?



Evaluating the Data Using Test Statistics

The test statistics discussed in the previous sections can be used to evaluate these hypotheses. The key data required for these tests are given in Table 1. The calculations of the *number of firm years*, *average default rates*, and *historical annual default rate standard deviations* are straightforward. The *idiosyncratic risk-based standard deviation* is the average theoretical volatility of annual default rate, assuming no shocks to the default rate and equals $\sqrt{\hat{p}(1-\hat{p})/(N_i/T)}$, for each sector. The *estimated standard deviation of annual shocks to the default rate* is the square root of the difference between the historical and the theoretical default rate variances. The estimated *autoregressive parameters* are obtained by regressing dr_t against dr_{t-1} for each sector, and the *correlation parameter* is sample correlation between the annual time series on default rates for each sector.¹⁷

Table 1: Significance Tests For Differences In Average Default Rates: Speculative-Grade US Banks vs. US Nonbanks (1979-1999 data)

	Banks	Nonbanks
Firm Years (N)	434	13,401
Average Default Rates (DR _t)	7.6%	4.2%
Average Combined Default Rate (\hat{p})		4.3%
Historical Standard Deviation of the Annual Default Rates	6.7%	2.3%
Idiosyncratic Risk-Based Standard Deviation of the Annual Default Rates Assuming No Shocks to the Default Rate ($\sqrt{\hat{p}(1-\hat{p})/(N_i/T)}$)	4.5%	0.8%
Estimated Standard Deviation of Annual Shocks to the Default Rate (σ_i)	5.0%	2.2%
Estimated Autoregressive Parameters (θ_i)	82%	59%
Estimated Correlation Between Sectoral Shocks (ρ_{12})		64%
	Z-Statistic	P-Value
Test for Difference in Means If Default Probabilities Are Constant	3.5	0.05%
Test for Difference in Means If Default Probabilities Fluctuate Over Time	0.6	54%

¹⁷ Our estimates of the historical autoregression and correlation parameters are biased downward slightly because the idiosyncratic risks make the historical default rate time series a “noisy” signal of the annual shocks to the underlying default rates. Given the amount of data generally available, other parameters are also likely to be measured imprecisely. In general, uncertainty around the true parameter values implies that it is theoretically more difficult to demonstrate that two default rates are significantly different than either the naive or the sophisticated testing models suggest.

If both the bank and nonbank samples were drawn from independently and identically distributed sample populations, the difference in historical default rates, 3.4% (=7.6% - 4.2%),

would have the standard deviation of. $\sqrt{0.043(1-0.043)\left(\frac{1}{13,401} + \frac{1}{434}\right)} \approx 1.0$. Hence, the

Z-statistic presented in equation (7) for the difference in historical default rates would be about 3.4 (or more precisely 3.48) standard deviations away from zero. *As a result, this model implies that the probability of observing this difference in default rates when the underlying default probabilities are the same for the two sectors is 0.05%.*¹⁸ That is to say, the likelihood of observing a 3.4% difference in average annual default rates for these two sectors is 1 out of 2000! Standard statistical measures of significance, therefore, seem to provide irrefutable evidence that the underlying default probability for speculative-grade bank issuers is lower than that of speculative-grade nonbank issuers.

However, adjustments for the presence of time-varying default rates, their persistence, and their correlation, implies that the data do not actually support that conclusion. Inserting the data from the table into equation (10), we calculate a revised Z-statistic of 0.6. That is, the likelihood of observing this difference in default rates is the same as the likelihood of observing a realization of any standard normal variable that is 0.6 standard deviations (plus or minus) away from zero. *Under this model, the probability of observing one sector with a default rate this much higher than another is 54%.* That is to say, the likelihood of observing this 3.4% difference in default rates is about 1 out of 2 – which is to say, this difference in default rates is not surprising at all.

Interpreting the Results

In conclusion, the naive test, which assumes there are no shocks to the annual default rates, suggests that speculative-grade banks are indeed riskier than nonbanks. However, the sophisticated test, which allows for time-varying shocks to annual default rates, implies that the observed difference in average default rates is not statistically significant.¹⁹

What is driving these sharply different conclusions? The results are really quite intuitive. The banking sector experienced a very large adverse and persistent default rate shock during the late 1980s and early 1990s. During the rest of the 1980s and 1990s, the banking sector experienced persistent, favorable shocks. In a sense, we have only three noisy observations on the long-term underlying banking sector default rate. Although the average default rate was a few percentage points higher for banks than nonbanks over the entire period, there is little reason to assume this historical difference will continue into the future.

18 The two-sided proposition – that the difference in average default rates would be either 3.48 standard deviations above or below the mean – occurs with 0.05% probability under the normal distribution.

19 One should, of course, keep in mind that the volatility and persistence parameters needed to implement the sophisticated model are measured very imprecisely, so that one should not assume that the resulting standard error calculations are themselves precise. However, in relying on the simple point estimates for these parameters, we have chosen the most reasonable, available estimates for these parameters. By contrast, the naïve model introduces false precision by implicitly setting all these parameters equal to zero.

CASE STUDY II - COMPARING THE DEFAULT RATES OF US AND NON-US SPECULATIVE-GRADE ISSUERS

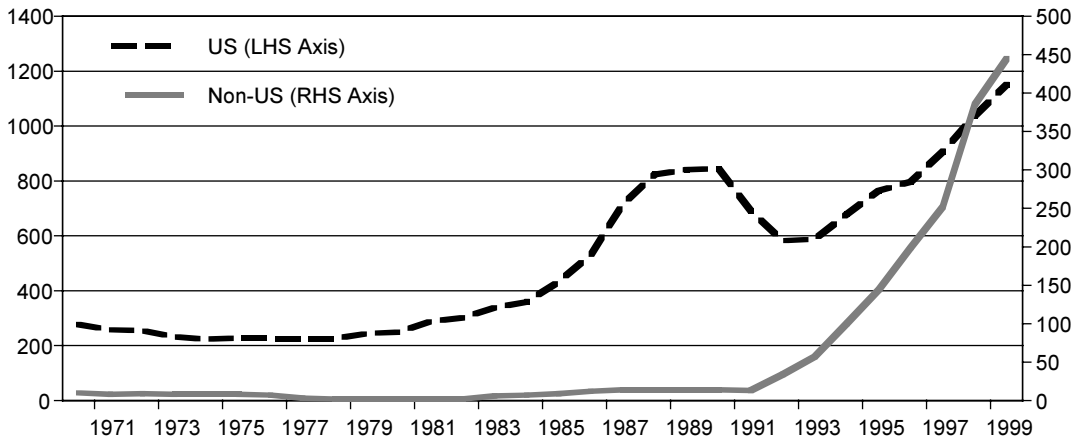
What the Data Shows

A case study analysis of the default rates of US and non-US speculative-grade issuers has many features in common with the bank v. nonbank example; however, there are some interesting differences. Like the previous case study, one sector (the US sector) has had a higher annual default than the other (non-US) sector, and the naive test suggests the difference is strongly significant.

In this case, however, a sophisticated test reaches a sharply different conclusion depends less on the presence of large, persistent shocks, and more on the fact that most of the non-US ratings are concentrated in a relatively small portion of the overall sample period. As a result, the historical data does not cover a sufficient amount of time for annual shocks to the default to “average out.”²⁰

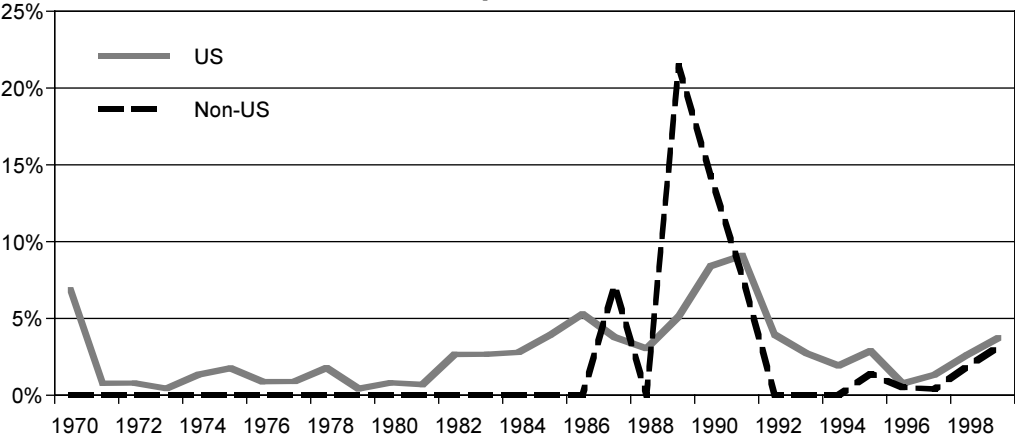
Consider Charts 9 and 10, which compare the number of speculative-grade issuers and annual default rates in the US and outside the US. Although there is a long history of speculative-grade issues outside the US, the bulk of all non-US speculative-grade ratings have been assigned in the last six or so years. This is not obvious simply by glancing at Chart 10, but the US one-year default experience has been higher (3.3%) than the non-US experience (1.8%) over the last 30 years. The non-US sector experienced extraordinarily high default rates for a few years in the early 1990s, but the size of the non-US sample was really quite small at that time.

Chart 9
Number Of Speculative-Grade Issuers



20 Our finds are similar to those reported by Ammer, John and Frank Packer, “How Consistent Are Credit Ratings? A Geographical and Sectoral Analysis of Default Risk,” Board of Governors of the Federal Reserve System, International Finance Discussion Paper #668, June 2000. These authors also conclude, using different methods, that the difference in default rates between US and non-US issuers is statistically insignificant.

Chart 10
Default Rates On Sepculative-Grade Issuers



Evaluating The Data Using Test Statistics

Table 2 presents the data needed to conduct tests of statistical significance for this case study. The naive test suggests that the difference in average annual default rates, 1.5% (= 3.3%-1.8%), is again highly significant. The Z-statistic is 3.5 and the p-value, which is the probability that such a difference in default rates could be observed if underlying default rates were actually the same, is again 0.05%. However, once again, the more sophisticated test reaches a sharply different conclusion. In particular, the adjusted Z-statistic becomes 0.31 and the p-value is now 75%.

Table 2: Significance Tests For Differences In Average Default Rates: Speculative-Grade US vs. Non-US Issuers (1970-1999 data)		
	US	Non-US
Firm Years (N _i)	15,314	1,790
Average Default Rates (DR _i)	3.3%	1.8%
Average Combined Default Rate (\hat{p})		3.2%
Historical Standard Deviation of the Annual Default Rates	2.3%	4.8%
Idiosyncratic Risk-Based Standard Deviation of the Annual Default Rates Assuming No Shocks to the Default Rate ($\sqrt{\hat{p}(1-\hat{p})/(N_i / T)}$)	0.8%	2.3%
Estimated Standard Deviation of Annual Shocks to the Default Rate (σ_i)	2.1%	4.3%
Estimated Autoregressive Parameters (θ_i)	60%	47%
Estimated Correlation Between Sectoral Shocks (ρ_{12})		58%
	Z-Statistic	P-Value
Test for Difference in Means If Default Probabilities Are Constant	3.5	0.05%
Test for Difference in Means If Default Probabilities Fluctuate Over Time	0.3	75%

Interpreting The Results

The basic intuition is that, given the historical volatility of default rate shocks, the expected difference between two sectors’ speculative-grade default rates is actually more likely than not going to at least 1.5%, given that the *effective* sample period for the non-US sector is only six years long.

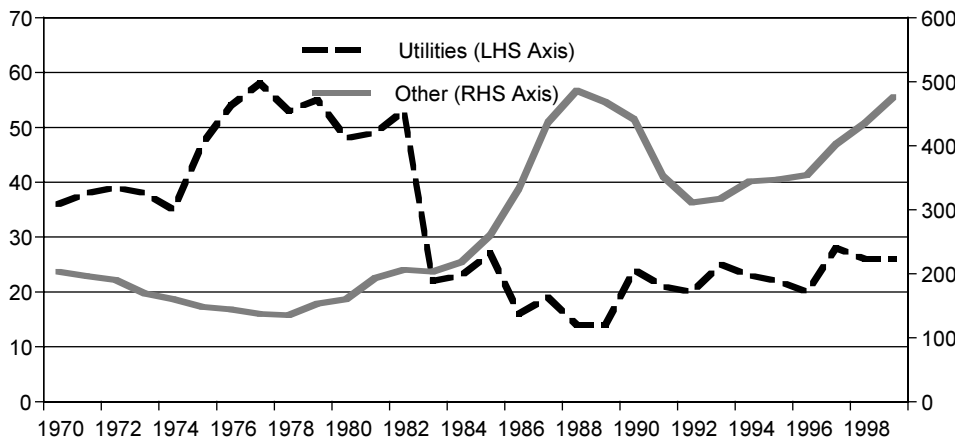
CASE STUDY III - COMPARING THE DEFAULT RATES OF BA-RATED US UTILITIES AND OTHER US BA-RATED COMPANIES

What The Data Shows

In this particular case, over the last thirty years, the default experience of Ba-rated US utilities has been very low (0.2%), despite the fact that many utilities have been rated in the Ba rating category during that time period.²¹ In fact, only two have Ba-rated holding companies defaulted. In contrast, the default experience of other Ba-rated US issuers has been higher (1.2%) during the same sample period.

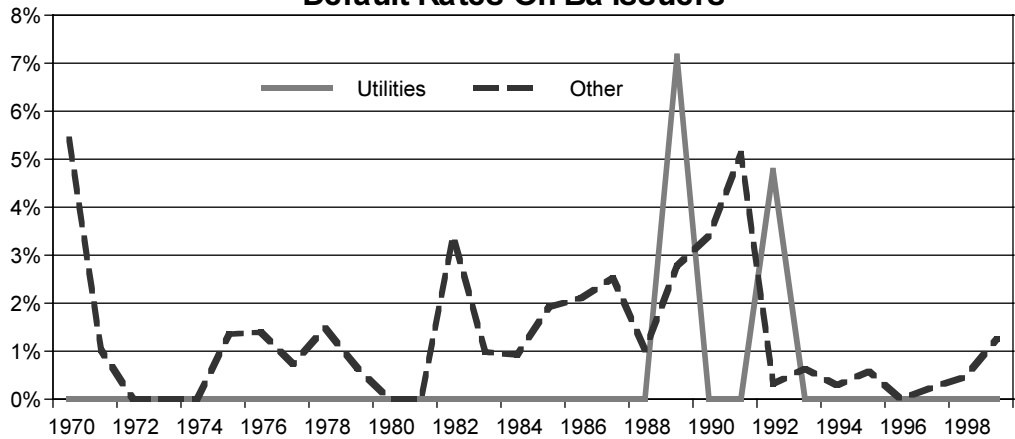
As shown in Chart 11, the sample populations of Ba-rated utilities and nonutility issuers have been fairly steady over the last thirty years, although Ba-rated utility population peaked in the mid- to late-1970s and the Ba-rated nonutility population peaked in the mid-1980s and late 1990s. The default rate experiences of the two sectors, however, have been quite different. In Chart 12, we see that the default rate pattern of the Ba-rated nonutilities was quite similar to the overall US speculative-grade default rate pattern that is shown in Chart 10. Ba-rated utilities, however, have experienced only two holding company defaults – one in 1989 and the other in 1992.

Chart 11
Number Of Ba-Rated Issuers



²¹ We have chosen to focus only on Ba-rated issuers rather than all speculative-grade issuers, because there have relatively few single-B or Caa-rated utilities compared to those rating categories' share of the overall speculative-grade universe. In contrast, it was appropriate to compare speculative-grade default rates in the other case studies because all the sectors being compared had very similar distributions within speculative-grade across the Ba, B and Caa rating categories.

Chart 12
Default Rates On Ba Issuers



Evaluating The Data Using Test Statistics

Table 3 presents the significance test statistics for this case study. Here we see that the 1.2% (1.4%-0.2%) difference in default rates is highly significant under the naive testing model. The Z-statistic is 3.2, with a p-value of 0.1%. That is, if the true underlying default rates were the same in the two sectors, the probability of seeing a difference in default rates this large would be only 1 out of 1000.

Table 3: Significance Tests For Differences In Average Default Rates: Ba-Rated US Utilities vs. US Nonutilities (1970-1999 data)		
	Utilities	Nonutilities
Firm Years (Ni)	973	8,377
Average Default Rates (DRi)	0.2%	1.4%
Average Combined Default Rate (\hat{p})		1.3%
Historical Standard Deviation of the Annual Default Rates	1.5%	1.5%
Idiosyncratic Risk-Based Standard Deviation of the Annual Default Rates Assuming No Shocks to the Default Rate ($\sqrt{\hat{p}(1-\hat{p})/(N_i/T)}$)	2.0%	0.7%
Estimated Standard Deviation of Annual Shocks to the Default Rate (σ_i)	0.0%	1.3%
Estimated Autoregressive Parameters (θ_i)	-	24%
Estimated Correlation Between Sectoral Shocks (ρ_{12})		0%
	Z-Statistic	P-Value
Test for Difference in Means If Default Probabilities Are Constant	3.2	0.13%
Test for Difference in Means If Default Probabilities Fluctuate Over Time	2.4	1.5%

In this case, the sophisticated test also suggests the difference in default rates is highly significant. The Z-statistic is 2.4 and the p-value is 1.5%, which implies that if the true underlying default rates were the same in the two sectors, the probability of seeing this difference in default rates would be only 15 out 1000. In this example, the volatility of shocks to the nonutility sector is indeed large (1.3%) and persistent (24%).²² However, the two utility sector defaults that were observed during the sample period can easily be explained by pure random variation from a constant default rate distribution. As a result, both the default rate volatility of the utility sector and its persistence are estimated to be zero.

Interpreting the Results

How should one use this sort of information to interpret the current ratings on US utilities? Does this mean that Ba-rated utilities are currently rated lower than they should be relative to other Ba-rated companies? Perhaps. However, there are a number of questions that need to be considered before drawing such a conclusion.

- 1) Are the experiences of Ba-rated utilities and nonutilities closer with respect to their long-term default rates than they are with respect to their annual default rates?
- 2) Are utilities subject to greater loss severity in the event of default than nonutilities?
- 3) Are utilities subject to greater transition risk than nonutilities?
- 4) Have the dynamics of the utility industry or Moody's ratings on the industry changed in a way that makes the historical default experience by rating category a poor guide to expected future experience?

While it seems quite likely that the answers to the first two questions are "no," the answers to other two questions may be yes. The US utility sector may be subject to higher than average transition risk. Many utilities have relatively weak intrinsic financial strength, and thus their long-term debt service capabilities are dependent on public support, which is highly likely to be forthcoming but nonetheless still uncertain.

Nevertheless, the low historical default incidence in the utility sector might have indeed led Moody's to upgrade many issuers had the industry not undergone dramatic change in recent years. The industry has recently been deregulated, and the market has become more competitive. This suggests that past experience may be a poor guide to future default experience in this sector.²³

This final case study is particularly interesting because it illustrates that the historical volatilities of annual default rates shocks do impose limits on the expected differences in average default rates. That is, the sophisticated testing model does find some differences in historical default rates to be significant.

²² Note that these parameter values for Ba-rated sector are, as expected, somewhat lower than their corresponding values for the overall speculative-grade universe, where average default rates are also higher.

²³ For more information, see "Changing Environment for Electric Utilities Prompts Additional Shifts in Analytic Focus," Moody's Special Comment, April 1998 and "Electric Utilities Industry Outlook," Moody's Special Comment, October 2000. Ironically, since this paper was first written, the holding companies of two electric utilities – Southern California Edison and Pacific Gas & Electric – have defaulted. From a purely technical perspective, our analysis would not be altered by the additional of these two defaults to our data set because these companies were never rated Ba. From a broader perspective, however, their defaults provides clear evidence that going forward there is little reason to expect that default incidence will remain lower in the utility sector compared to other sectors.

Appendix: Recent Academic Research on Rating Consistency

Public scrutiny of rating consistency has increased as ratings have come to play a larger role in portfolio risk management models and in financial market regulation. Academic research on rating consistency has recently been stimulated by the Basel Committee’s rating-sensitive bank capital model proposal, which was announced in June 1999. The academic literature on rating consistency has focused on three primary areas – potential inconsistencies across rating agencies, across industry segments, and across geographical regions.

INCONSISTENCIES ACROSS RATING AGENCIES

Differences in the meaning of ratings across rating agencies have been widely documented (Beattie and Searle,²⁴ Cantor and Packer,²⁵ Jewell and Livingston,²⁶ and the Basle Committee on Banking Supervision²⁷). The absence of uniformity in the meanings of the rating scales used by different agencies is hardly surprising. Each rating agency employs a different verbal set of meanings for its ratings. Moreover, even for Moody’s and S&P – the only rating agencies that publish their historical record – the default rate experiences of their ratings are quite different if one accepts the standard mapping that is usually applied between their ratings. Furthermore, there are strong competitive forces that are likely to drive differences in the meanings of ratings across agencies.

INCONSISTENCIES ACROSS INDUSTRIES

Academic researchers have also examined differences between the performance of Moody’s bank and nonbank ratings. In particular, Nickell, Perraudin, and Varotto²⁸ note that, between 1970 and 1997, Moody’s bank ratings were more volatile than Moody’s industrial ratings. These authors did not, however, find evidence that the default rate experiences of these two sectors were significantly different.

In contrast, Ammer and Packer²⁹ conclude that Moody’s ratings on speculative-grade banks were associated with significantly higher default rates than similarly rated US industrials during the 1983-1998 period. Further analysis by these authors, however, reveals that the differences in default experiences are attributable entirely to US thrift industries’ unusually high default rates between 1989 and 1991.

INCONSISTENCIES ACROSS GEOGRAPHICAL REGIONS

A second topic generating academic interest is that of the potential geographical inconsistencies. In particular, Beattie and Searle investigate whether there might be a so-called “home-country” bias. They find some evidence that local rating agencies rate their domestic companies higher than foreign rating agencies. However, Cantor and Packer demonstrate that, at least for ratings of international banks, differences in ratings assigned by home and foreign agencies result principally from differences in the rating scales applied by individual rating agencies, rather than a home-

24 Beattie, Vivien and Susan Searle, “Bond Ratings and Inter-Rater Agreement,” *Journal of International Securities Markets*, Summer 1992.
25 Cantor, Richard and Frank Packer, “The Credit Rating Industry,” *Federal Reserve Bank of New York Economic Policy Review*, Summer-Fall, 1994.
26 Jewell, Jeff and Miles Livingston, “A Comparison of Bond Ratings from Moody’s, S&P and Fitch IBCA,” *Financial Markets, Institutions, & Instruments*, New York University Salomon Center, Volume 8, Number 4, August 1999.
27 “Credit Ratings and Complementary Sources of Credit Quality Information,” *Basel Committee on Banking Supervision*, Working Paper #3, August 2000.
28 Nickell, Pamela, William Perraudin, and Simone Varotto, “Stability of Rating Transition Matrices,” *Journal of Banking and Finance*, 24 (1&2), 2000.
29 Ammer, John and Frank Packer, “How Consistent Are Credit Ratings? A Geographical and Sectoral Analysis of Default Risk,” *Board of Governors of the Federal Reserve System, International Finance Discussion Paper #668*, June 2000.

country bias. That is, the local rating agencies uniformly rate higher than the international rating agencies.

Ammer and Packer further examine the potential for a home-country bias, focusing specifically on the historical database that Moody's has made publicly available. They observe that, between 1983 and 1998, the average one-year default rate for speculative-grade non-US issuers was lower than that for US issuers. However, after controlling for the year of issuance, they find that the differences are not statistically significant. In particular, they note that 78% of their non-US observations (compared to 48% of their US observations) come from the last seven years of their sample, when default rates were particularly low both in the US and abroad.

IMPLICATIONS FOR THE CONSISTENCY OF MOODY'S RATINGS

The academic literature has identified two cases where the average default rates associated with Moody's corporate ratings were different across sectors. In each case, however, the literature has concluded that the observed differences can be attributed special circumstances that support the view that the ratings had been assigned in a consistent manner. As discussed in this *Special Comment*, these factors can be introduced systematically into the analysis of rating consistency by explicitly recognizing the role of global/macroeconomic and sectoral shocks into the estimated statistical distributions of expected default rates.

Going forward, however, the sector may in fact become more exposed to sectoral shocks and, perhaps, higher average default experience by rating category due to the ongoing structural and competitive changes currently underway in that industry.

To order reprints of this report (100 copies minimum), please call 800.811.6980 toll free in the USA. Outside the US, please cal 1.212.553.1658

Report Number: 63945