

UC Irvine Papers

Title

Assessing the Risks of Small Business Borrowers

Permalink

<https://escholarship.org/uc/item/4dg4c1nb>

Journal

The Journal of Commercial Lending, 76(2)

Authors

McNamara, G
Bromiley, P

Publication Date

1993

Peer reviewed

Research Report: Assessing the Risks of Small Business Borrowers

by Gerry McNamara and Philip Bromiley

This article describes the development of a model to assess the risk of lending to small business borrowers. The model was developed using logistic regression analysis and a few common financial ratios from commercial borrowers at Norwest Bank Minnesota N.A., Minneapolis. The model predicts, with some accuracy, the future risk ratings assigned to loans by bankers.

The authors assess the model's predictive ability with Norwest data that was not used in developing the model. They also contrast their findings with other studies of financial distress and bankruptcy.

Over the past thirty years, many researchers have attempted to use financial data to evaluate the default risk of corporations. Although informative, these studies may be of limited use to commercial lenders.

Therefore, while small business loans are a major component of the loan portfolios of many banks, little information is available to assess the risk of small business lending quantitatively.

© 1993 by Robert Morris Associates. McNamara is a doctoral candidate and Bromiley is a professor of strategic management, Carlson School of Management, University of Minnesota. The authors are grateful to William Stansifer, Jerry Thompson, and the business banking managers and lending officers at Norwest Bank Minnesota N.A., Minneapolis, for their assistance with the project.

This article describes a model that we developed, using one bank's data, to predict the future risk of commercial loans. The model uses a few readily available financial ratios. Risk is evaluated based on a four-level risk-rating scale.

Previous Studies

While numerous studies have been undertaken to predict financial distress, our study overcomes at least three limitations of previous studies:¹

1. Our study focuses on predicting the risk of loans to small firms. Although the lessons learned from the collapse of Penn Central may be interesting, they are probably of little relevance when trying to understand the risk presented by a small commercial borrower. Also, while previous studies rely on financial ratios, just as our study does, the implications of some financial ratios may be different for small firms. And some types of data, such as stock prices, do not exist for the usual small business borrower.
2. Previous studies focused solely on corporate bankruptcy. A large corporate bankruptcy with reams of creditors differs from a small business bankruptcy in many ways. For example, lenders tend to play different roles in each. Also, understanding when a loan may be dete-

riorating is probably more relevant than determining possible bankruptcy since problem loans require higher loan loss reserves and can generate significant overhead costs regardless of whether or not actual bankruptcy occurs.

We use a *multilevel* risk-rating scale that provides more information on the relative risk of loans than previous models, which predicted only the probability of bankruptcy or nonbankruptcy. A multilevel risk-rating scale allows us to examine issues that are more commonly seen in small business borrowing situations.

3. Unlike previous studies, we analyze the data with logistic regression instead of multiple discriminant analysis. Logistic regression makes fewer questionable assumptions than other types of analysis.² Also, logistic regression not only predicts a loan's risk rating but can also indicate the likelihood of being at any given risk rating.

Data Collection

Working with Norwest Bank Minnesota N.A., we collected data on commercial loans to small corporations. The data were provided by the community banking division and came from five locations in the Minneapolis and St. Paul area. Both urban and suburban areas were represented.

The community banking division serves both small commercial customers and private individuals. As part of the credit review process, the division annually reviews all loans with balances greater than \$100,000.

Although the Norwest community banking division originates loans up to \$10 million, most customers have loan balances less than \$1 million.

The initial sample we collected included 377 corporations with loan balances greater than \$100,000 at year-end 1987.³ The data in the study were collected from loan reviews from 1985 to 1991.

At Norwest, information on problem loans is forwarded to a centralized loan workout office. The banking offices and loan workout area provided files on all active customers. For loans paid off or charged off before the collection of the data, we attempted to retrieve the loan files from the bank's storage facility. We were able to obtain loan files for 71% of the identified customers. An additional 12% of the total number of firms could not be included because their files included only one loan review or lacked other data. Information for 223 firms (59% of the original sample) was available for the research project.

Borrower Database

The bank's borrower database provided financial data for the firms

included in the study. Norwest records information from the borrower's annual financial statement on a standardized financial worksheet. The worksheet includes a balance sheet, cash flow statement, income statement, ratio report, and industry comparison report. Each worksheet can contain up to five years of financial data, based on the duration of the borrower's relationship with the bank. All borrowers in our study had multiple years of data, and a total of 784 usable observations were compiled.

Risk Ratings

Choosing the appropriate dependent variable for commercial risk-assessment modeling is an important step. We wanted to examine *gradations of risk*, so we tried to develop a model that could predict the risk ratings assigned by loan officers. This seems logical, as loans with higher risk ratings tend to result in increased loan loss reserves, higher overhead costs, and higher default risks.⁴

Norwest's risk-rating scale has seven levels: 1 (the lowest risk) to 7 (the highest risk). Because the bank assigns very few firms ratings of 1 or 2, we deleted observations with risk ratings 1 or 2 from the analysis. Also, the two observations with a 7 rating were rerated at 6 for the purposes of the study. Consequently, our study

¹ A bibliography of previous studies is included on page 31.

² Logistic regression analysis also does not require that the variance and covariance matrices of the independent variables be the same for the sets of firms in all risk-rating categories. By using logistic regression, we do not need to assume that the relations among variables used to predict loan risk remain consistent across all risk ratings. Also, logistic regression does not require all independent variables be normally distributed.

³ Sole proprietorships, partnerships, and subchapter S corporations were omitted because their financial data were not comparable to taxable corporation financial data.

⁴ William A. Charlton, "The Pricing Effects of the Risk-Based Capital Regulations," *Journal of Commercial Bank Lending*, April 1991, pp. 29-33.

uses a rating scale with four levels that nominally range from 3 to 6.⁵

Data Selection

We began our analysis by dividing the data into two parts. To develop the model, we used two-thirds of the database, and the remaining one-third was set aside to evaluate the predictive ability of the model.

We found we could not use the entire sample we had selected (a problem identified in previous studies) because the sample was overwhelmingly composed of loans with a risk rating of 3 or 4. These two categories accounted for 430 of the 506 observations. A majority of loans in a particular risk-rating category is a problem because the normal estimation process will minimize errors by almost always predicting the common rating. To correct this, we chose a random sample of 19 loans from each risk rating category for the database of the study. We chose 19 from each because it was the lowest number of observations in any given category. This adjustment winnowed the usable sample to 76, with equal representation in each rating category.

The sample reflects some selection problems: There are no data from customers who were initially rejected for loans, nor from customers who ended their relationships with the bank before the study period. These customers include borrowers who

changed banks and borrowers whose loans were paid or written off.

Model Development

We developed a model to analyze the data using stepwise logistic regression, which chooses the set of variables that best explains a criterion of interest—in our study, the risk rating. Logistic regression is based on the cumulative logistic probability function and can be used to estimate the probability of a loan receiving any risk rating. The logistic regression procedure developed three equations. Each equation gives the cumulative logistic probability that the loan will be rated at either the current rating or a lower risk rating. This logistic probability can be converted easily to a standard probability function.

Analysis

Initially, we attempted to develop the model by including 125 financial ratios that had been used in previous risk prediction studies. Although the analysis identified a six-variable model that "fit" the risk ratings of loans included in the database, we could not use the results because the model was "overfitted" to the data. In other words, because the model was developed using such a large number of ratios, it was accurate for predicting the risk of only the loans in the database used to develop it: When we applied the model to other data (the data that had been set aside for model

⁵ At Norwest, loans rated as 5 and 6 generally correspond to the Office of the Comptroller of the Currency designations of special mention and substandard loans. Loan loss reserves on these two categories typically range between 3% and 15%.

validation), the model performed poorly.

We realized a model needed to be developed using a much smaller set of ratios. After a review of the literature and discussions with bank managers, we identified six variables as the most likely to influence risk ratings:

1. Profitability, measured by profit before interest and taxes divided by total assets.
2. Cash flow, measured by cash flow after debt amortization divided by total assets.
3. Liquidity, measured by the current ratio.
4. Leverage, measured by net worth divided by total assets.
5. Collateral margin, which is related to net working capital divided by total assets.
6. Size, measured by the log of total assets.

Three of these ratios emerged as significant: net worth to total assets, profit before interest and taxes to total assets, and net working capital to total assets. The other three variables did not add to the explanatory value of the model.

Using the three financial ratios, logistic regression analysis generated three different sets of equations that can be used to predict the risk level of a loan. See Figure 1 for the equations and a brief example of how they can be used.

The Equations

The first set of equations calculates the logistic probability function for

each of the risk ratings. The second set translates the logistic function into a cumulative probability. The third set converts the cumulative probabilities into probabilities indicating the likelihood that a firm will receive any given risk rating.

Predictive Capabilities

The model correctly predicted the borrower risk ratings assigned by loan officers one year in the future 54% of the time. Our model is much more accurate than risk ratings selected at random. Given a risk-rating scale with four levels, a random rating system would rate a loan correctly only 25% of the time. In fact, there is less than a one in one hundred chance that a random model could rate 54% of the loans correctly.

The model misrated only 9% of the firms by more than one risk rating, compared with a 38% possibility of this occurring with a random rating system. Figure 2 presents these results.

Comparison with Past Studies

To compare the results of our research with previous bankruptcy prediction studies, we recoded the risk ratings to a two-level scale. That is, we combined risk ratings 3 and 4 to represent nonproblem loans and risk ratings 5 and 6 to represent problem loans. This model correctly identifies 80% of the loans as being either problem or nonproblem loans one year before the actual rating.⁶

⁶ These results are superior to those found by either Edmister or Hoeven in their earlier studies of small firms when they used one year of financial data.

Figure 1. Equations to Predict Risk Ratings**Set 1 - Logistic Probability Equations**

$$L(3) = -3.2170 + .0431 \times (\text{net worth} + \text{total assets}) \\ + .0264 \times (\text{net working capital} + \text{total assets}) \\ + .0637 \times (\text{profit before interest and taxes} + \text{total assets})$$

$$L(4) = -1.3243 + .0431 \times (\text{net worth} + \text{total assets}) \\ + .0264 \times (\text{net working capital} + \text{total assets}) \\ + .0637 \times (\text{profit before interest and taxes} + \text{total assets})$$

$$L(5) = 0.5124 + .0431 \times (\text{net worth} + \text{total assets}) \\ + .0264 \times (\text{net working capital} + \text{total assets}) \\ + .0637 \times (\text{profit before interest and taxes} + \text{total assets})$$

Set 2 - Cumulative Probability Equations

$$C(3) = 2.7183^{L(3)} + (1 + 2.7183^{L(3)}) \\ C(4) = 2.7183^{L(4)} + (1 + 2.7183^{L(4)}) \\ C(5) = 2.7183^{L(5)} + (1 + 2.7183^{L(5)})$$

Set 3 - Single Risk Rating Probability Equations

$$P(3) = C(3) \\ P(4) = C(4) - C(3) \\ P(5) = C(5) - C(4) \\ P(6) = 1 - C(5)$$

Example

Consider a borrower with the following characteristics:

Net worth + total assets = 29.4%

Working capital + assets = 6.5%

Profit before interest and taxes + total assets = 12.45%

Applying Set 1 Equations:

$$L(3) = -.9852 \\ L(4) = .9075 \\ L(5) = 2.7442$$

Applying Set 2 Equations:

$$C(3) = .2719 \\ C(4) = .7125 \\ C(5) = .9396$$

Applying Set 3 equations to produce the probability of the borrower being rated in a particular category:

$$P(3) = .2719 \\ P(4) = .4406 \\ P(5) = .2271 \\ P(6) = .0604$$

Figure 2. Risk Rating Results

Actual Risk Rating	Predicted Risk Rating				Total
	3	4	5	6	
3	12	6	1	0	19
4	5	9	3	2	19
5	1	5	8	5	19
6	0	3	4	12	19
Total	18	23	16	19	76

Model Validation

We also evaluated the predictive ability of the model on the one-third of the Norwest database not used to develop the model. Unlike the model development phase, the validation phase distribution of risk ratings should be representative of the total population of loan files. Tests of the predictive ability of the model used all 278 observations from the data that were not used for model development.

The results of this analysis showed that one year before the actual risk rating, the model correctly predicted 53% of the ratings. Only 7% of the predicted risk ratings differed from the actual risk ratings by more than one rating. When the four risk-rating levels were collapsed into a two-level rating scale, the model correctly classified 74% of the observations as either problem or nonproblem loans.

The accuracy of the model on the validation data almost equals that found in the data used to develop the

model. This indicates that the relationships between the three financial ratios and the firms' future risk ratings are not specific to the database used in the development of the model. In other words, the model is not "overfitted" to the database and can be used to predict the risk rating of other small business loans at Norwest Bank.

If we extend the analysis to two years before the actual risk rating, we find the model correctly predicts 47% of the risk ratings, and 9% of the predicted risk ratings differ from the actual risk ratings by more than one level. Using a two-level scale, the model correctly predicts 73% of the loans. The model's predictions are almost as accurate two years ahead as one year ahead. These results indicate that the model can be used as a leading indicator of the creditworthiness of the borrower. We believe the model can also be used by lenders to provide an early warning to loans that are likely to deteriorate into problem credits.

Prior Studies

Our last analysis compared the predictive ability of our model to models using the variables from previous bankruptcy studies. We identified the variables used in seven earlier studies of risk.⁷ Each set was entered into the stepwise logistic regression procedures, and the "best" set of variables and parameters was developed. We then used each of these seven estimations to predict the observations in the data that were set aside. These results indicated that although the equation that is derived using the variables from the Edmister study is slightly better at predicting loans on a two-level risk-rating scale, the equation from our study has the best overall predictive ability even though it relies on only a single year of financial data.

Although our model has the best overall predictive ability, several of the models perform almost as well. This is not surprising as all of the models use somewhat similar sets of variables, and these financial variables tend to be highly correlated. The results suggest that the specific variables used may be somewhat less critical (within a general set) than correct model development and prediction procedure.

Conclusion

This study contributes to the literature of risk prediction in at least three ways. First, it demonstrates that a useful model for predicting the risk of small business borrowers can be

developed using a few readily available financial ratios. Since the model has fairly accurate predictive abilities more than one year in advance, it can be used to understand both current and future borrower risk. Such a model may complement loan officer judgment in the assessment of corporate loans but, of course, cannot replace it. Second, since our model uses a multilevel risk-rating scale, it provides information that is more relevant to the commercial lender than the two-level rating scales used in prior studies. Third, unlike most prior risk prediction models, our model predicts not only what risk rating a borrower is most likely to be assigned in future periods but also the probability of the borrower being at any given risk level.

This constitutes the initial results of a research program on commercial lending risk. Although the study demonstrates the feasibility of predicting risk for small commercial borrowers, it has significant limitations:

1. The model omits some factors, such as firm management, which could significantly change the assessment of the borrower's risk.
2. The database represents only the experience of one regional bank over a five-year time period.

We are seeking additional banks to participate in future studies to examine how this model reacts to different lending practices, geographic areas, and time periods. From these efforts, we anticipate the production

of robust and well-tested models to assess the risk for small commercial borrowers.

Bibliography

- Altman, Edward, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, September 1968, pp. 589-609.
- Altman, Edward, "The Success of Business Failure Prediction Models: An International Study," *Journal of Banking and Finance*, June 1984, pp. 171-198.
- Beaver, William H., "Financial Ratios as Predictors of Failure," *Journal of Accounting Research*, 1966, pp. 71-127.
- Edmister, Robert O., "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction," *Journal of Financial and Quantitative Analysis*, March 1972, pp. 1477-1493.

- Gahlon, James M., and Robert L. Vigeland, "Early Warning Signs of Bankruptcy Using Cash Flow Analysis," *Journal of Commercial Bank Lending*, December 1988, pp. 4-15.
- Hoeven, James A., "Predicting Default of Small Business Loans," *Journal of Commercial Bank Lending*, April 1979, pp. 47-60.
- Keasey, Kevin, and Robert Watson, "Financial Distress Prediction Models: A Review of Their Usefulness," *British Journal of Management*, 1991, pp. 89-102.
- Ohlson, James A., "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, Spring 1980, pp. 109-131.
- Sharma, Subhash, and Vijay Mahajan, "Early Warning Indicators of Business Failure," *Journal of Marketing*, Fall 1980, pp. 80-89.
- Van Leeuwen, Peter H., "The Prediction of Business Failure at Rabobank," *Journal of Bank Research*, Summer 1985, pp. 91-98. ■

⁷ Altman, 1966; Edmister, 1972; Gahlon and Vigeland, 1988; Hoeven, 1979; Ohlson, 1980; Sharma and Mahajan, 1980; Van Leeuwen, 1985.