Predicting The Financial Failure Of Retail Companies In The United States
Mary Hilston Keener, The University of Tampa, USA

ABSTRACT

Predicting the financial failure of companies using financial ratios is a topic that has been explored in various ways for many years, and the current economic climate suggests that these models may still be more useful than ever. Various financial ratios and bankruptcy prediction methods have been used in order to try to find the most accurate prediction model possible. With historically successful retailers, like Sears, Kmart and JCPenney, struggling in recent years, predicting the future of retailers has become even more important.

Therefore, this paper focuses specifically on the application of a failure prediction model to companies from the retail industry. Logistic regressions are used in this study in order to attempt to predict which companies are likely to fail. The sample for this study includes publicly traded United States companies from the retail industry, and data is collected from the COMPUSTAT database for the period from 2005-2012. Based on prior studies, the author hypothesizes that companies are most likely to fail if they are unprofitable, highly leveraged, and having cash flow problems. As expected, the results demonstrate that smaller retail companies with fewer employees are more likely to fail. The results also provide strong evidence that firms with lower cash to current liability ratios, lower cash flow margins, and higher debt to equity ratios are more likely to file for bankruptcy.

Keywords: Bankruptcy Prediction; Retail; Financial Failure

INTRODUCTION

The headlines in the financial press over the last few years have been shocking. Several formerly successful retail companies, including Hollywood Video, Borders, and Fashion Bug, have unexpectedly gone bankrupt over the last few years. Predicting the financial failure of companies from financial ratios is a topic that has been explored in various ways for many years, and the current economic climate suggests that these models may still be more useful than ever. Various financial ratios and regression techniques have been used in order to try to find the most accurate prediction model possible. With historically successful retailers like Sears, Kmart and JCPenney struggling in recent years, predicting the future financial success of retailers has become even more important.

This paper focuses specifically on the application of a failure prediction model to companies from the retail industry. Logistic regressions are used in order to attempt to predict which companies are likely to fail. The sample for this study includes publicly traded United States companies from the retail industry, and data is collected from the COMPUSTAT database for the period from 2005-2012. Based on prior studies, including Lennox (1999), the author hypothesizes that companies are most likely to fail if they are unprofitable, highly leveraged, and having cash flow problems.
LITERATURE REVIEW

Many studies have previously addressed the issue of the predictive power of various financial ratios. In 1932, FitzPatrick published a study of 20 pairs of firms - one failed and one surviving - matched by factors including date, size and industry. Although he did not perform statistical analysis, he analyzed the ratios and trends in the ratios and conducted simple multivariate analysis. Beaver (1966) utilized t-tests to evaluate the predictive ability of various financial ratios using a pair-matched sample. In a landmark paper in the bankruptcy prediction area, Altman (1968) used multiple discriminant analysis on a pair-matched sample. Altman (1968) used five financial ratios to predict whether a company is going bankrupt. Altman’s study is very accurate in its prediction of which companies are going to go bankrupt, and the study has held up very well even after almost 35 years.

Ohlson’s (1980) paper has also been considered one of the landmark papers that has been published on the topic of using financial ratios to predict financial failure. Ohlson (1980) applied logistic regression in a much larger sample that did not involve pair-matching, and he found that the four statistically significant factors for identifying the probability of failure are the size of the company, measures of financial structure, measures of performance, and measures of current liquidity. Ohlson’s paper appears to have laid the foundation for a lot of later research on the predictive power of financial ratios, and therefore this paper is referred to in almost every later paper on this topic.

Lennox (1999) examined the causes of bankruptcy for a sample of UK listed companies for the period 1987 to 1994 and found that the most important determinants of bankruptcy are profitability, leverage, cash flow, company size, industry sector and the economic cycle. Also, Lennox (1999) argues that well-specified logit and probit models can be more accurate in identifying failing companies than discriminant analysis.

In recent years, many studies have developed and examined new methods for predicting bankruptcy. Barniv and McDonald (1999) discuss several alternative techniques to probit models and logit models for predicting company failure, including EGB2, lomit, and burrit models. Shirata et al. (2011), Divsalar et al. (2012), Kwak et al. (2012b), Olson et al. (2012), and Shiri et al. (2012) have attempted to use text mining or data mining to predict bankruptcy. Hwang et al. (2007) and Cheng et al. (2010) use semi-parametric methods to predict bankruptcy. Serrano-Cinca and Gutierrez-Nieto (2013) use partial lease square discriminant analysis in bankruptcy prediction. Future studies will be useful in determining which of these bankruptcy predictions methods will stand the test of time.

Some studies in this area have examined the impact of adding additional new variables as predictors of bankruptcy. Elam (1975) chose to look specifically at the effect that lease data has on the predictive power of financial ratios. Elam’s study uses 28 financial ratios that are commonly used in financial literature and textbooks. Elam ultimately found that the addition of capitalized lease data to financial statements did not increase the predictive power of financial ratios. Gentry et al. (1985) found that outflow components of financial statements, like dividends and investments, are better predictors of financial failure than inflow-related financial statement components. In other words, Gentry et al. (1985) found that cash-based funds flow components may provide superior results in the prediction of financial failure. Adding the cash components significantly improves predictive performance in the Gentry et al. paper. Gilbert et al. (1990) suggest that different variables should be used to distinguish between bankrupt and distressed firms than those variables that are typically used to discriminate between bankrupt and nonbankrupt firms. Also, Gilbert et al. (1990) and Giacomino and Mielke (1993) confirm Gentry et al.’s (1985) finding that cash flow variables add explanatory power to models for predicting bankruptcy.

Baldwin and Glezen (1992) compared the predictive power of quarterly financial data with previous studies’ results obtained using annual financial data. Baldwin and Glezen’s results confirm that more timely bankruptcy prediction models can be found using quarterly financial data without the loss of accuracy that is sometimes associated with the use of annual financial data.

Richard Morris (1997) looked at how useful bankruptcy prediction models actually are. He found that ultimately, most prediction models reflect the fact that listed companies that have low profits, large losses, and large debt burdens are more at risk than companies that are more profitable and that have less debt. Morris (1997) also points out the fact that corporate failure often occurs as a result of an unfortunate sequence of events or bad luck for
a company, and he suggests that it would be very difficult for a model to distinguish between companies that will go bankrupt and companies that are just in financial distress.

Because this study specifically looks at companies in the retail industry, a paper by Riordan (1998) is particularly relevant to this study. He looks at the surge in retail bankruptcies to identify the problems that caused the number of retail bankruptcies to increase suddenly the way they did. Riordan outlines recommendations for fixing the retail bankruptcy system and also offers stockholders with the challenge of finding capable management to try to turn around struggling companies in the retail industry. Many other studies have tried to predict bankruptcy in other industries, but very few papers in the extant literature examine bankruptcy prediction in the retail industry. Dietrich et al. (2005) attempt to predict financial failure in agricultural co-ops. Youn and Gu (2010) predict failure in the restaurant industry. Zaki et al. (2011) predict distress in banks. Ray (2011) predict corporate financial distress in the glass and glassware sector. Trewichayapong et al. (2011) predict bankruptcy for real estate firms, while Kim (2011) predicts hotel bankruptcy using various prediction methods.

Many studies have also attempted to predict bankruptcy in countries throughout the World. Ray (2011) and Bhunia et al. (2011) predict financial distress in India. Nam et al. (2000) and Kwak et al. (2012a) predict bankruptcy for Korean companies. Cheng-Ying (2004) and Chia-Liang and Kuan-Min (2011) attempt to predict bankruptcy in Taiwan. Studies including Alkhathib and Bzour (2011), Rashid and Abbas (2011), Zaki et al. (2011), and Karami et al. (2012) attempt to predict bankruptcy in various middle eastern countries. Stundžienė and Boguslauskas (2006), Dakovic et al. (2010), and Fijorek and Grotowski (2012) all predict bankruptcy for countries in Europe. One incremental contribution of this study is that it focuses on predicting bankruptcy for a specific industry segment in a specific country; namely, retail firms in the United States.

METHODOLOGY

As suggested by Altman (1968), companies are much more likely to fail if they are unprofitable, highly leveraged, and suffering from cash flow difficulties. For this study, companies that have been liquidated and those that have ultimately gone bankrupt are both included as failing companies. Although many early bankruptcy prediction models used discriminant analysis (DA) to predict financial failure, this technique has become less frequently used because of two assumptions made to use DA. First of all, to use DA, the variables used are assumed to have multivariate normal distributions, and it is well known that the variables normally used in bankruptcy studies are not normally distributed. Also, the samples of companies that are failing and non-failing are assumed to be randomly drawn, an assumption that does not allow for the commonly used technique of matched pairs. These problems, along with many others, have caused more researchers to turn to using probit and logit models instead of discriminant analysis.

This paper will focus, in particular, on using logistic regression modeling in order to predict bankruptcy. The unique incremental contribution of this paper is the focus on companies only from the retail industry for a recent sample period. Previous studies including Lennox (1999) have included a variable in the model for industry sector in order to rule out industry bias in the sample, but most studies have not examined the use of bankruptcy prediction models in the retail industry. The results of this study may provide valuable information to retail companies and investors in the retail industry.

The bankruptcy prediction model used in this study uses five variables that have been significant predictors of bankruptcy in previous studies. The variables to be used in this paper as predictors of bankruptcy are number of employees (EMP), return on assets (ROA), cash flow margin (CFLM), debt-to-equity ratio (DTEQ), and cash to current liabilities (CHELCT). Number of employees is expected to be a significant predictor of bankruptcy because of the increased likelihood that smaller companies with fewer employees will fail. Return on assets is included as a measure of profitability of the companies, and it is assumed that less profitable companies are more likely to fail. Cash flow margin and cash to current liabilities are included in the study as measures of company liquidity. The debt-to-equity ratio is a measure used to analyze the solvency of a company, and it is included in this study because it appears to be a good measure of company performance. The variables are best illustrated in Table 1.
Table 1: Variable Descriptions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Formula</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMP</td>
<td>Number of Employees</td>
<td>Employees</td>
</tr>
<tr>
<td>ROA</td>
<td>(Net Income + Net Interest Expense)/Average Total Assets</td>
<td>Return on Assets</td>
</tr>
<tr>
<td>CFLM</td>
<td>Cash Flow/Sales</td>
<td>Cash Flow Margin</td>
</tr>
<tr>
<td>DTEQ</td>
<td>(Total Debt/Total Stockholders' Equity)</td>
<td>Debt-to-Equity Ratio</td>
</tr>
<tr>
<td>CHELCT</td>
<td>Cash/Current Liabilities</td>
<td>Cash to Current Liabilities</td>
</tr>
</tbody>
</table>

A logit model is used in this study in order to predict failure of companies. The failure prediction model used in this paper is

\[ \text{FAILS}_i^* = \alpha + \beta_1 (\text{EMP}_i) + \beta_2 (\text{ROA}_i) + \beta_3 (\text{CFLM}_i) + \beta_4 (\text{DTEQ}_i) + \beta_5 (\text{CHELCT}_i), \]

where the prediction variables are as described in the paragraph above and the table and

\[ \text{FAILS}_i = 1 \quad \text{if the company fails} \]
\[ \text{FAILS}_i = 0 \quad \text{otherwise}. \]

Because the data has been collected as to which companies have actually failed, the actual results can be compared to these predicted from the model. The variable debt-to-equity ratio is the only variable expected to have a coefficient with a positive sign, while the other four variables should all have negative coefficients for this prediction model. Return on assets, number of employees, cash flow margin, and cash-to-current-liabilities are all expected to have negative signs because firms that have high numbers for these ratios are not as likely to go bankrupt.

DATA

The COMPUSTAT database was used to collect financial data about companies in the retail industry, and data was collected for each retail company for the period from 2005-2012. Data for both active and research companies was collected in order to be able to look at comparisons between companies that are still active and non-bankrupt and companies that are now bankrupt or liquidated. Research companies must be included in the study, for the research companies database contains companies that have already filed for bankruptcy.

The dates of deletion (code DLDTE) and the reasons for deletion (code DLRSN) for all of the research companies were obtained from COMPUSTAT for all of the research companies included in the study. Although there are eight reasons why companies are moved from COMPUSTAT’s active to research file, only companies that were deleted because of bankruptcy (reason code 2) and liquidation (reason code 3) are considered to have failed for the purposes of this study. Companies with codes other than two or three as reasons for deletion are simply left out of this study. One limitation of this study is the fact that the date given on COMPUSTAT as the date of deletion is, in many cases, not the actual date that a particular company failed. If this study is extended at some point, more research will have to be done to find out the actual date that all of the failed companies went bankrupt or were liquidated. Also, one of the reasons for deletion listed for many companies is “other” (code 10). More research would have to be done for each of these companies to see specifically why they are no longer listed on COMPUSTAT because in some cases, these companies may have gone bankrupt or been liquidated due to some special extenuating circumstance.

Data for all active and research companies in the retail industry was collected for the years 2005-2012, and the result is a listing of data for 3,595 companies. Many of the companies in this original list of data have missing data for some or all of the variables chosen for several reasons. First, retail companies that became unlisted at any time are all included in the list because of the inclusion of COMPUSTAT research companies. Secondly, some of the research companies are left out of this study because of the fact that they were deleted from COMPUSTAT for reasons other than bankruptcy and liquidation, and they are thus also missing data. Finally, some of the active companies are missing data because of the fact that many of the retail firms in the study did not start up until after 2005. Therefore, the final sample for this study only contains 1,203 firms. Because of the large quantity of missing data, firm-year observations are used in this study as opposed to firms.
Although the initial sample consists of 28,760 firm-years, the majority of these observations are missing data for the reasons discussed above. Data is therefore only available for 6,702 firm-years. After the top and bottom one percent of observations are trimmed to remove the impact of outliers on the results, 6,094 firm-year observations remain. Of these 6,094 observations in the study, 5,607 of the observations are for firms that are still active and 487 of the observations are for companies that have failed since 2005. Because of the fact that the exact dates of bankruptcy are unable to be obtained for this paper, the model will be using data for companies to predict whether they will go bankrupt at any time up until December 2012.

The descriptive statistics are given in Table 2 for the companies used in this study. The employees variable (EMP) is measured in thousands of employees, while the rest of the variables are ratios expressed as percentages (ratio * 100%). It is clear that failing companies are normally smaller, in terms of the number of employees that are not highly liquid or profitable but are leveraged. The means are significantly different for each of the variables for the non-failing and failing companies, as demonstrated by the t-test results in Table 2. On average, the non-failing retail companies have more employees, a higher return on assets average, a higher cash flow margin, a lower debt to equity ratio, and a higher cash-to-current liabilities ratio.

### Table 2: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Non-Failing Companies (5,607 Observations)</th>
<th>Failing Companies (487 Observations)</th>
<th>T-Test For Equality Of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Mean</td>
</tr>
<tr>
<td>ROA</td>
<td>-1.261</td>
<td>20.744</td>
<td>-16.381</td>
</tr>
<tr>
<td>CFLM</td>
<td>2.742</td>
<td>22.038</td>
<td>-12.652</td>
</tr>
<tr>
<td>DTEQ</td>
<td>250.972</td>
<td>5,317.894</td>
<td>4,371.928</td>
</tr>
<tr>
<td>CHELCT</td>
<td>68.881</td>
<td>90.095</td>
<td>39.196</td>
</tr>
</tbody>
</table>

*Indicates significance at the .01 level.

**RESULTS**

Logistic regressions are run on the model primarily because of previous empirical research, including Lennox (1999) that suggests the results from probit and logit models are similar to each other for this type of study. The results from the logistic regression are included in the Table 3.

### Table 3: Logistic Regressions To Predict Bankruptcy In Companies From 1992-1996

\[
\text{FAILS}^*_n = \alpha + \beta_1 (\text{EMP}_n) + \beta_2 (\text{ROA}_n) + \beta_3 (\text{CFLM}_n) + \beta_4 (\text{DTEQ}_n) + \beta_5 (\text{CHELCT}_n)
\]

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Expected Sign</th>
<th>Coefficients (Wald Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMP</td>
<td>-</td>
<td>-0.382 (125.651)*</td>
</tr>
<tr>
<td>ROA</td>
<td>-</td>
<td>0.033 (35.565)*</td>
</tr>
<tr>
<td>CFLM</td>
<td>-</td>
<td>-0.013 (9.236)*</td>
</tr>
<tr>
<td>DTEQ</td>
<td>+</td>
<td>0.000 (3.273)**</td>
</tr>
<tr>
<td>CHELCT</td>
<td>-</td>
<td>-0.098 (230.584)*</td>
</tr>
</tbody>
</table>

Nagelkerke $R^2 = .917$

*Indicates significance at the .01 level.

**Indicates significance at the .10 level

Several interesting results from Table 3 should be noted. First of all, all of the variables in the model are significant predictors of company failure at the 0.01 level of significance, except for the debt to equity variable, which is significant at the 0.10 level. The signs on the variables are as expected in all cases, except for the return on assets variable. The positive sign on the return on asset variable for this model indicates that an increase in ROA will cause the probability of a company failing to increase, which is an unexpected finding. This may be caused by...
the fact that the average value for ROA is very low for all the firms, as shown in Table 2. One last result that should be noted from Table 3 is the fact that the R-square value of 0.917 is very high, which means that the model developed in this study is very useful for predicting the financial failure of retail companies.

MODEL SPECIFICATION TESTS

Several tests were done to check for possible problems with the model that is used in this study, including tests for autocorrelation, heteroskedasticity, and model misspecification. Tests of autocorrelation were run first using the Durbin-Watson statistic, and the Durbin-Watson statistic of 1.938 indicates that autocorrelation is not a problem in this model. Therefore, no adjustments will need to be made to the model for problems with autocorrelation.

Next, tests for heteroskedasticity were conducted to rule out the changing variance problem. The Breusch-Pagan test was conducted on the OLS results from the model in this study and the results indicated that the model has a significant heteroskedasticity problem. The Likelihood Ratio was also calculated and the significance of this value is another indicator of the heteroskedasticity existing in the model. Unfortunately, the use of cross-sectional data in this study makes it almost impossible to avoid the heteroskedasticity problem. A weighted least square method was used in an attempt to reduce the heteroskedasticity, but the problem actually seemed to get worse when the weighted least square method was used. Therefore, the original model will be used for the remainder of the study.

The results were also examined carefully to see if there is a multicollinearity problem with the models. The most common indicator of multicollinearity - high \( R^2 \) but few significant t ratios - is not a problem for this study. The pair-wise sample correlation conducted between the regressors in the model indicates that only the two cash flow variables are significantly correlated, but the results in Table 3 show that including both of these variables did not negatively impact the results of this study. The low pair-wise correlations between most of the regressors is a good sign that there is not a significant problem with multicollinearity in this model. Finally, the low variance inflation factor values suggest that there is not a multicollinearity problem in this model.

Finally, several model specification tests were conducted to make sure that the model in this study has a good fit. Tests for underfitting were conducted first by running Ramsey’s RESET test on the data in the study. The significance of the F-value obtained from this test is indicative of a problem with underfitting. This means that additional variables that would increase the prediction power of this model have been omitted. Unfortunately, more extensive research would need to be done to find out exactly which financial indicator variables should be added to increase the predictive power of the model specifically for firms in the retail industry. This provides an excellent opportunity for future research. Tests for overfitting were done next to see if an unnecessary variable is included in the model. The results of the logistic regression models run in this study (Table 3) indicate that it is not necessary to include the ROA or debt to equity variables in this study. The Nagelkerke R-squared remains approximately the same when the logistic regression is run again with those variables are removed.

CONCLUSION

This model appears to be extremely accurate in predicting the failure of companies in the retail industry using several financial ratios. As expected, the results demonstrate that smaller retail companies with fewer employees are more likely to fail. The results also provide strong evidence that firms with lower cash to current liability ratios, lower cash flow margins, and higher debt to equity ratios are more likely to file for bankruptcy. Conversely, the results show that firms with higher return on assets are more likely to file for bankruptcy, and this unexpected result should be further examined in future studies.

Several limitations on the results of this study should be noted. First of all, the results of this study are only applicable to firms in the retail industry. Also, further research should be conducted in order to test which variables might add more additional predictive power to this model. By focusing more specifically on which ratios would be most important for retail companies, a better fitting model could probably be obtained. Next, the significant problems with heteroskedasticity and stability in this model should be further examined in order to possibly improve
the prediction model. Simply finding better predictors for the retail industry may decrease the heteroskedasticity and stability problems.

Another limitation on the results of this study is the fact that the dates that are used for the date of failure of the companies in this study may not be completely accurate. As discussed previously, COMPUSTAT provides only the date of deletion from COMPUSTAT and the reason for deletion, not the actual date of failure for the companies. Although the date of deletion is normally very close to the date of failure, the actual failure dates would need to be researched for each of the research companies included in this study. Another limitation that may have decreased the predictive power of this model is the fact that annual, as opposed to quarterly, financial data is used, and according to Baldwin et al. (1992), using annual data may contribute to a loss of accuracy compared to quarterly data. A final limitation that should be noted for this study is the fact that COMPUSTAT is missing a lot of data for many of the retail companies. The missing data and the fact that COMPUSTAT usually only includes large companies may bias the results obtained in this study.

AUTHOR INFORMATION

Dr. Mary Keener earned her PhD in Accounting from Kent State University in 2007. After becoming an Assistant Professor of Accounting at The University of Tampa in 2008, Dr. Keener became the faculty advisor for Beta Alpha Psi, an honorary organization for financial information students and professionals. Dr. Keener teaches courses in financial accounting at the introductory, intermediate and graduate levels. Dr. Keener’s research interests are in the areas of value relevance, restructuring charges, and analyst forecast accuracy, and she has recently published several papers in these areas. E-mail: mkeener@ut.edu

REFERENCES