Market Variables and Financial Distress

Giovanni Fernandez
Stetson University

In this paper, I investigate the predictive ability of market variables in correctly predicting and distinguishing going concern opinion firms versus going concern firms. Following Fernandez et al. (2014), which demonstrated the benefit of adding market variables into a model to predict financial distress, the best distinctive market variables are employed in this paper. The bid-ask spread plays an important role the further one gets from the event, but a more subdued role as the event approaches than that found in the literature.

INTRODUCTION

Fernandez et al. (2014) argue that market variables are better predictors of corporate bankruptcy than financial ratios. The logic is as follows: accounting ratios are derived from the company's financial statements. These statements are reported quarterly and are a snapshot in time or a video of the previous quarter. Therefore, these are, at best, a current view of the firm, but more likely, stale numbers that are backward-looking. Market variables, on the other hand, are dependent on future expectations, those of corporate performance and future cash flows. This leads Fernandez et al. (2014) to conclude that, if the purpose is to predict a future event, the variables chosen should be those which are forward-looking, such as market variables. The authors used corporate bankruptcy to test this prediction. In this paper I employ similar methodology in order to test the ability of market variables to predict firms earning a non-going concern opinion.

Traditional models have been heavily employed in the bankruptcy prediction literature. Models such as multivariate discriminant analysis and logit regression analysis have been used and have demonstrated predictive power when discriminating between bankrupt and non-bankrupt firms. Receiving a non-going concern opinion can be a precursor for corporate bankruptcy; therefore, similar models are also suitable for predicting the receipt of a non-going concern opinion. The model employed in this paper is logit regression analysis.

Following the findings of Fernandez et al (2014), the predictability of a set of market variables is tested using the logit analysis technique. The variables are stock return, price, bid-ask spread, and standard deviation. These variables are shown in the literature, employing different models and techniques, to have the greatest predictive power of corporate financial distress. For the purposes of estimation, we find that the bid-ask spread plays a less important role in distinguishing between firms entering distress versus those that do not, when compared to the prior literature. However, it does play an important role further away in time from the opinion. This is important because it highlights the importance of including the information found through trading, specifically on the trading floor. Market microstructure is increasingly playing a bigger role in finance due to the use of high frequency trading (O'Hara (2014)) and data availability.

Lastly, as is desirable of any predictive test and model, out-of-sample tests must be performed. In order to test the model and variables' predictive ability, I use a holdout sample. Using data from five years before the non-going concern opinion to two years before the opinion, I fit the logit regression model. Following this procedure, the model's predictive ability is tested one year before the opinion. Because non-going concern opinions are less dire than all out bankruptcy, a model that correctly predicts this occurrence one year before the opinion can be very useful to many players in the market, especially lenders and investors. This would allow lenders to avoid bad business many years before glaring red flags are produced, and it would allow investors to better price assets.

The results, as often found with out of sample tests, suggest that more variables are needed, such as more market variables or including accounting ratios along with the market variables employed. Future work in this area may be best suited employing the different sets of variables, both market and accounting, found to be useful in Fernandez et al. (2014). Furthermore, more sophisticated models, such as neural networks, might prove to be better for predictive purposes.

The remainder of this paper is organized as follows: The next section reviews the previous literature associated with logit regression analysis as it pertains to financial distress. The following section discusses the data used in this study. Then the methodology is described. The next section discusses the results, and the final section summarizes and concludes.

LITERATURE REVIEW

Logit regression, other than multivariate discriminant analysis, is traditionally the most heavily used tool to predict bankruptcy. It is more commonly applied to predict failure through the use of financial ratios. However, recent research has started to widen the range of variables employed.

Lo (1986) tests the specification and application of discriminant analysis and logit regressions to corporate bankruptcy, and finds that even though logit regressions are more robust for parameter estimation than discriminant analysis, both methods result in consistent estimates and discriminant analysis estimators are asymptotically efficient.

Koh and Low (2004) test the classification of 165 going concern and 165 non-going concern firms by employing neural networks, decision trees, and logistic regressions. They find superior power in the decision tree models over the neural networks and logistic regressions. This is largely due to the rejection of the underlying assumptions of traditional models.

Mutchler (1985) applies multivariate discriminant analysis to test models of the non-going concern opinion decision with a sample of manufacturing companies that received a going concern opinion. Ohlson (1980) uses maximum likelihood estimation of the logit model to predict bankruptcy, employing financial ratios, which are strictly accounting numbers.

Fernandez et al. (2014) demonstrate that market variables, when used as the only predictive variables, are better at predicting corporate bankruptcy than accounting variables are when used alone. They employ multivariate discriminant analysis and find that the bid-ask spread, stock price, and stock price volatility are the best set of predictors.

DATA

The list of firms was obtained from bankruptcydata.com. The list includes companies that were issued a non-going concern opinion from January 2010 to December 2011, and the data for classification and predictive purposes goes back five years before the non-going concern opinion is issued. The subsample of going concern firms comes from the entire database of Bloomberg, consisting of firms that did not incur the non-going concern opinion during the sample period. The market data for all firms, non-going concern opinion and going concern, are obtained from Bloomberg.

METHODOLOGY

To begin, the going concern firms are selected in such a way that firm size does not severely affect the results. This is a common technique in the literature. Ex-ante, successful firms are to be larger than failing firms. Furthermore, firm success rates are highly influenced by their industry. For example, at the peak of the credit crisis, most financial firms performed well. To mitigate this issue, it is common to select firms in such a way that size and industry factors do not distort the results.

The procedure is two-fold. First, a non-going concern opinion firm is selected. Second, a going concern firm within the same industry with asset size closest to the non-going concern opinion firm is selected. These two firms (the one going concern opinion and the one going concern) are then stored and removed from the continuing procedure. A second non-going concern opinion firm is selected, and the selection process is continued with the remaining going concern firms. This is done for each non-going concern opinion firm.

The model applied for the analysis is logit regression analysis. The model estimates the probability of a discrete outcome given the values used to explain an occurrence, which in this case is a firm receiving a non-going concern opinion. The logit model is based on the logistic distribution and estimates the probability that the dependent variable equals one given the value of the explanatory variables, i.e. the probability that it received a non-going concern opinion given certain values of the market variables employed. Similar to linear regression, the logit regression uses one or more predictive variables that can be either continuous or discrete. However, the logit regression is employed to predict binary outcomes rather than continuous outcomes. The model is estimated using maximum likelihood ratios.

For the purpose of explaining which variables correctly classify the receipt of a non-going concern opinion, the model is estimated each year, from year one to year five before the opinion. The model is estimated employing all market variables previously used in the literature, along with specifications using the market variables found to distinguish firms by financial distress in Fernandez et al. (2014).

Lastly, since the true power of the model is in its predictive power, a pseudo-out-of-sample test is run. Estimating the model using data from five years before the opinion to two years before the opinion, I then test the models ability to correctly predict whether or not a firm will receive a non-going concern opinion one year before the actual opinion is rendered.

EMPIRICAL RESULTS

In order to estimate and test the ability of the logit regression to correctly classify and predict firms that earn a non-going concern opinion, an analysis of the underlying variables within and across groups is necessary. First, I review the characteristics of each group's stock price, return, bid-ask spread, and standard deviation.

The mean stock return for the going concern group is 13.57%, but as expected this value varies drastically. It is, however, surprising that firms deemed not on the precipice of financial distress produce such a high return. This contradicts the traditional, theoretical models. Investors require a higher return to invest in stocks with a higher systematic risk. Since this sample is produced from the market during a time when the market reached bottom and then bounced into positive territory, it is expected that these firms would be 'defensive' in nature. An explanation can be that traditional models and systematic risk are not robust explanatory variables of stock returns during market transitions from bear to bull markets.

Interestingly, the non-going concern opinion firms produce significantly lower average returns, 1.9% but with a substantially lower standard deviation. Investors seeking alpha during this period were likely to buy firms that performed poorly at the bottom of the credit crisis, trading heavily in to those firms in or about to be in financial distress, further requiring a higher rate of return for those investments.

The bid-ask spread is much higher for non-going concern opinion firms than for going concern firms. Traditionally, higher levels of uncertainty, which is demonstrated by the higher standard deviation of price for non-going concern opinion firms than that of going concern firms, lead to wider spreads, compensating the market makers and dealers for the extra risk involved in holding stock in such firms.

Following the findings of Fernandez et al. (2014), the logit regression is analyzed employing two sets of variables: stock price, bid-ask spread and standard deviation of price, and stock return, bid-ask spread, and standard deviation of return.

Table 1 displays the results of the first set of variables. Interestingly, price and standard deviation are significant at the 5% level one year prior to the non-going concern opinion, yet bid-ask spread is not. This differs from the findings in the literature. This is consistent up to three years before the non-going concern opinion. However, five years before the non-going concern opinion, only the bid-ask spread is significant. This is consistent with the theory that those on the floor, which receive much more information earlier than most of the market, can correctly distinguish between firms entering distress and those that are not. The reason it may not be significant as the time approaches the opinion is because the information is already priced in. For purposes of classification, the model is not robust and requires more variables, as suggested by the error rates.

Table 1

Results of Logistic Regression: Concern Modeling (Price, Bid Ask Spread and Standard Deviation)

YEAR 1

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	441
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	611.175	609.268
SC	615.264	625.624
-2 Log L	609.175	601.268

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.9072	3	0.0480
Score	6.9705	3	0.0728
Wald	5.4334	3	0.1427

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.0612	0.1026	0.3552	0.5512
M_PX_LAST	1	-0.00849	0.00380	4.9844	0.0256
M_Bid_Ask_Spread	1	-0.00188	0.00622	0.0911	0.7628
M_SD_PRICE	1	0.0153	0.00669	5.2116	0.0224

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_PX_LAST	0.992	0.984	0.999
M_Bid_Ask_Spread	0.998	0.986	1.010
M_SD_PRICE	1.015	1.002	1.029

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	2	231	5	203	52.8	1.0	97.9	71.4	46.8

YEAR 2

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	486
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	675.665	618.442
SC	679.851	635.187
-2 Log L	673.665	610.442

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	63.2233	3	<.0001
Score	19.7983	3	0.0002
Wald	27.3253	3	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.3404	0.1182	8.2944	0.0040
M_PX_LAST	1	-0.0634	0.0125	25.6798	<.0001
M_Bid_Ask_Spread	1	-0.6025	0.4198	2.0600	0.1512
M_SD_PRICE	1	0.1781	0.0383	21.6699	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_PX_LAST	0.939	0.916	0.962
M_Bid_Ask_Spread	0.547	0.240	1.246
M_SD_PRICE	1.195	1.109	1.288

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	206	95	145	40	61.9	83.7	39.6	41.3	29.6

YEAR 3

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	493
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	685.392	681.844
SC	689.593	698.646
-2 Log L	683.392	673.844

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.5481	3	0.0228
Score	1.6242	3	0.6539
Wald	3.3981	3	0.3342

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.00121	0.0926	0.0002	0.9896
M_PX_LAST	1	-0.00474	0.00278	2.9025	0.0884
M_Bid_Ask_Spread	1	-0.00210	0.00442	0.2257	0.6347
M_SD_PRICE	1	0.0106	0.00612	3.0215	0.0822

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_PX_LAST	0.995	0.990	1.001
M_Bid_Ask_Spread	0.998	0.989	1.007
M_SD_PRICE	1.011	0.999	1.023

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	60	128	121	184	38.1	24.6	51.4	66.9	59.0

YEAR 4

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITS ORT
Response Variable	C
Number of Response Levels	2
Number of Observations	518
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	716.359	719.476
SC	720.608	736.475
-2 Log L	714.359	711.476

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.8830	3	0.4100
Score	2.2058	3	0.5308
Wald	0.6758	3	0.8789

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1693	0.0886	3.6494	0.0561
M_PX_LAST	1	0.000084	0.000138	0.3743	0.5407
M_Bid_Ask_Spread	1	0.000163	0.000297	0.2991	0.5844
M_SD_PRICE	1	-0.00017	0.000275	0.3770	0.5392

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_PX_LAST	1.000	1.000	1.000
M_Bid_Ask_Spread	1.000	1.000	1.001
M_SD_PRICE	1.000	0.999	1.000

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	1	279	2	236	54.1	0.4	99.3	66.7	45.8

YEAR 5

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	504
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	696.114	692.561
SC	700.337	709.451
-2 Log L	694.114	684.561

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.5530	3	0.0228
Score	8.5180	3	0.0364
Wald	5.5779	3	0.1341

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.3040	0.1028	8.7448	0.0031
M_PX_LAST	1	-0.00004	0.000032	1.1923	0.2749
M_Bid_Ask_Spread	1	1.5792	0.7791	4.1082	0.0427
M_SD_PRICE	1	0.000170	0.000146	1.3469	0.2458

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_PX_LAST	1.000	1.000	1.000
M_Bid_Ask_Spread	4.851	1.054	22.338
M_SD_PRICE	1.000	1.000	1.000

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	26	257	19	202	56.2	11.4	93.1	42.2	44.0

Tables 2 displays the results of the second set of variables, which are stock return, bid-ask spread, and standard deviation. Overall, this model specification underperforms the first set of variables. All variables are mostly not statistically significant. The stock return variable is only significant three years prior to the opinion, while bid-ask spread is only significant at the 10% level two years before the opinion. Furthermore, the misclassifications are no better using this model specification when compared to the first model.

Lastly, while the classification and distinguishing power has been tested using the logit model, it is more important to practitioners to be able to properly and accurately predict a firms coming distress. Therefore, the model should be applied in a predictive environment. In order to do so, the logit model is again applied to the data using both sets of variables. However, to test the data out of sample, the data is parsed. Using the data from five years before the opinion to two years before the opinion, the model is specified. Following this procedure, the model is used to predict whether or not the firm will receive a non-going concern opinion during the one year before the opinion is actually handed out.

Since it was demonstrated above that the first set of variables performs better in classifying going-concern firms, the predictive model will employ the first set. While it is common for out-of-sample tests to underperform in-sample tests, the results here are still less than desirable. The model displays a bias toward predicting firms to be going concern firms. This can be due to one of two things. First, the model may be missing certain variables. There are more market variables that can be employed, outside of the three that are used here. Furthermore, as discussed in the prior literature, accounting ratios still include important information, so it can be beneficial to include the most discriminating ratios along with market variables. Second, more sophisticated and less restrictive models may outperform a traditional model like logit regression analysis for predictive purposes. Traditionally, logit analysis and discriminant analysis were employed, but with the improvement in technology and computing power, models like neural networks have shown to perform well without the restrictive assumptions needed in the traditional models.

Table 2

Results of Logistic Regression: Concern Modeling (Return, Bid Ask Spread and Standard Deviation)

YEAR 1

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	441
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	611.175	613.470
SC	615.264	629.826
-2 Log L	609.175	605.470

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.7052	3	0.2951
Score	3.4518	3	0.3271
Wald	2.7309	3	0.4350

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1543	0.0962	2.5714	0.1088
M_CUST_TRR_RETURN_HO	1	0.00882	0.00597	2.1852	0.1393
M_Bid_Ask_Spread	1	-0.00228	0.00625	0.1327	0.7157
M_SD_PRICE	1	0.000446	0.000676	0.4346	0.5098

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_CUST_TRR_RETURN_HO	1.009	0.997	1.021
M_Bid_Ask_Spread	0.998	0.986	1.010
M_SD_PRICE	1.000	0.999	1.002

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	14	226	10	191	54.4	6.8	95.8	41.7	45.8

YEAR 2

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	485
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	674.301	673.709
SC	678.485	690.445
-2 Log L	672.301	665.709

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.5926	3	0.0861
Score	5.1520	3	0.1610
Wald	4.1209	3	0.2487

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.0640	0.1053	0.3693	0.5434
M_CUST_TRR_RETURN_HO	1	0.00624	0.00587	1.1297	0.2878
M_Bid_Ask_Spread	1	-0.7041	0.4078	2.9814	0.0842
M_SD_PRICE	1	0.00146	0.00212	0.4756	0.4904

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_CUST_TRR_RETURN_HO	1.006	0.995	1.018
M_Bid_Ask_Spread	0.495	0.222	1.100
M_SD_PRICE	1.001	0.997	1.006

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	154	71	169	91	46.4	62.9	29.6	52.3	56.2

YEAR 3

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITS ORT
Response Variable	C
Number of Response Levels	2
Number of Observations	493
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	685.392	662.741
SC	689.593	679.543
-2 Log L	683.392	654.741

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	28.6516	3	<.0001
Score	24.4945	3	<.0001
Wald	21.3698	3	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.1077	0.0958	1.2652	0.2607
M_CUST_TRR_RETURN_HO	1	-0.0411	0.00893	21.2467	<.0001
M_Bid_Ask_Spread	1	0.00203	0.00412	0.2431	0.6220
M_SD_PRICE	1	0.000030	0.000085	0.1208	0.7281

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_CUST_TRR_RETURN_HO	0.960	0.943	0.977
M_Bid_Ask_Spread	1.002	0.994	1.010
M_SD_PRICE	1.000	1.000	1.000

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	165	158	91	79	65.5	67.6	63.5	35.5	33.3

YEAR 4

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITS ORT
Response Variable	C
Number of Response Levels	2
Number of Observations	517
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	715.134	717.825
SC	719.382	734.817
-2 Log L	713.134	709.825

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.3088	3	0.3464
Score	2.8810	3	0.4103
Wald	1.9181	3	0.5896

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1448	0.0917	2.4958	0.1141
M_CUST_TRR_RETURN_HO	1	0.00680	0.00668	1.0382	0.3082
M_Bid_Ask_Spread	1	0.000163	0.000294	0.3100	0.5777
M_SD_PRICE	1	7.457E-6	9.507E-6	0.6152	0.4328

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_CUST_TRR_RETURN_HO	1.007	0.994	1.020
M_Bid_Ask_Spread	1.000	1.000	1.001
M_SD_PRICE	1.000	1.000	1.000

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	5	269	11	232	53.0	2.1	96.1	68.8	46.3

YEAR 5

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGITSORT
Response Variable	C
Number of Response Levels	2
Number of Observations	504
Model	binary logit
Optimization Technique	Fisher's scoring

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	696.114	695.491
SC	700.337	712.381
-2 Log L	694.114	687.491

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.6229	3	0.0849
Score	5.8800	3	0.1176
Wald	4.6474	3	0.1995

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.3023	0.1036	8.5113	0.0035
M_CUST_TRR_RETURN_HO	1	-0.00064	0.00996	0.0041	0.9491
M_Bid_Ask_Spread	1	1.5851	0.7815	4.1141	0.0425
M_SD_PRICE	1	0.000018	0.000023	0.6191	0.4314

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
M_CUST_TRR_RETURN_HO	0.999	0.980	1.019
M_Bid_Ask_Spread	4.880	1.055	22.572
M_SD_PRICE	1.000	1.000	1.000

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	25	256	20	203	55.8	11.0	92.8	44.4	44.2

Table 3

Logit analysis results for Out of sample validation (Year 2 to 5): Concern Modeling (Return, Bid Ask Spread and Standard Deviation)

Model Information	
Data Set	CONCERN.C_NC_COMBINED_VARAVG_LOGIT_Y25
Response Variable	C
Number of Response Levels	2
Number of Observations	1999
Model	binary logit
Optimization Technique	Fisher's scoring

Response Profile		
Ordered Value	CONCERN	Total Frequency
1	1	954
2	0	1045

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	2769.058	2768.802
SC	2774.659	2791.204
-2 Log L	2767.058	2760.802

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.2559	3	0.0998
Score	4.8899	3	0.1800
Wald	2.3076	3	0.5111

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.0911	0.0451	4.0761	0.0435
M_CUST_TRR_RETURN_HO	1	-0.00252	0.00310	0.6620	0.4158
M_Bid_Ask_Spread	1	0.000175	0.000324	0.2903	0.5901
M_SD_PRICE	1	0.000011	9.771E-6	1.3723	0.2414

Classification Table									
	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.500	8	1038	7	946	52.3	0.8	99.3	46.7	47.7

CLASSIFICATION SUMMARY FOR TEST DATA (YEAR 1)

Actual	Predicted		Total
	Concern	Nonconcern	
Concern	4 1.19	332 98.81	336
Nonconcern	0 0.00	323 100.00	323
Total	4	655	659

CONCLUSION

The study of the causes and predictability of financial distress has been thoroughly investigated in the literature for decades. Traditionally, models employing accounting ratios have been used to predict corporate bankruptcy. Less has been done when it comes to the study of firms that receive a non-going concern opinion. Following Fernandez et al. (2014), which employs market variables to predict corporate bankruptcy, logit regression analysis is employed using market variables in order to predict whether or not a firm will receive a non-going concern opinion.

The reason for using market variables instead of, or along with, accounting ratios is as follows. Accounting ratios can be seen as stale numbers; the financial statements are produced after actual events occur. The statements are a snapshot in time, or at best a recent video. Market variables, however, are forward-looking, and therefore capture expectations along with the most recent available information. For purpose of prediction, any model should be better specified using such variables.

The variables studied are stock price, return, bid-ask spread, and standard deviation. The results are both consistent and inconsistent with the literature. Unlike that found in the literature, bid-ask spread plays a less important role in distinguishing between firms entering distress versus those that are not. Furthermore, the out-of-sample test which is estimated using these variables sharply underperforms other traditional models used, even when those models used accounting ratios. One more important finding, which strengthens the belief that variables that account for trading behavior and informed investors are important in a predictive framework, is that the bid-ask spread does play an important role the further one moves away from the event. Bid-ask spread plays a significant role five years prior to the receipt of the non-going concern opinion. Future research in this area should employ such variables as the bid-ask spread, along with the best discriminating accounting ratios. Lastly, more sophisticated and nonparametric models should be used.

REFERENCES

Fernandez, G., Prakash, A., and Mishra, S. (2014) With One Shot, Which Bullet Would You Use? Market versus Accounting Data in Bankruptcy Prediction Part I (The Univariate Case). *The International Journal of Finance*, 14(1).

Fernandez, G., Prakash, A., and Mishra, S. (2014) With One Shot, Which Bullet Would You Use? Market versus Accounting Data in Bankruptcy Prediction Part II (A Strong case of using Nonparametric MDA in Bankruptcy Prediction). *The International Journal of Finance*, 14(2).

Koh, H. C. and Low, C. K. (2004) Going concern prediction using data mining techniques. *Managerial Auditing Journal*. 19(3), 462 – 476.

Lo, A. W. (1986) Logit Versus Discriminant Analysis: A Specification Test and Application to Corporate Bankruptcies. *Journal of Econometrics,* 31(2), 151-178.

Mutchler, J. F. (1985) A Multivariate Analysis of the Auditor's Going-Concern Opinion Decision. *Journal of Accounting Research*, 23(2), 668-682.

O'Hara, M. (2014) High-Frequency Market Microstructure. *Journal of Financial Economics*, 116(2), 257-270.

Ohlson, J. A. (1980) Financial Ratios and Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109-131.