The Predictive Factors of Hospital Bankruptcy: A Longitudinal Analysis

Brad Beauvais, PhD [1]; Zo Ramamonjiarivelo, PhD [1]; C. Scott Kruse, PhD [1]; Lawrence Fulton, PhD [2]; Ram Shanmugam, PhD [1]; Arvind Sharma, PHD [2]; Aleksander Tomic, PhD [2]

[1] School of Health Administration, College of Health Professions, Texas State University; [2] Woods College of Advancing Studies, Boston College

Background

Regulation, reimbursement, complexity, and workforce costs in the US greatly challenge hospitals to remain financially viable, resulting in a 305% increase in bankruptcies since 2010. Despite earlier attempts by researchers to propose explanatory methods for predicting bankruptcy none of these models have undergone extensive testing in the health care industry.

Three Historical Bankruptcy Models

Altman Z Score = $(1.2 * A) + (1.4 * B) + (3.3 * C) + (0.6 * D) + (1.0 * E)$, where A=working capital, B=retained earnings/TA, C=EBITDA/TA, D=market value of equity/book value of debt, E= sales / TA

Ohlson O Score = O=-1.32 - 0.407Ln(O1) + 6.03 × O2 -1.43 × O3 + 0.0757 × O4 - 2.37 × O5 - 1.83 × O6 + 0.285 × O7 - 1.72 × O8 - 0.521 × O9, where O1=GNP Price index, O2=TL/TA, O3=WC/TA, O4=CL/CA, O5=1 if TL>TA, O6 = NI/TA, O7=Op Funds / TL, O8 = 1 if NI < 0 for two years, and O9 = Change in Net Income / |NI|

Zmijewski Score = $-4.336 - \left[4.513 * \left(\frac{NetIncome}{TotalAssets}\right)\right] + \left[5.679 * \left(\frac{TotalLiabilities}{TotalAssets}\right)\right] + \left[0.004 * \left(\frac{CurrentAssets}{CurrentLiabilities}\right)\right]$

Purpose

This study develops an explanatory and predictive logistic model for hospital bankruptcy utilizing only 8 financial and hospital-level variables (drawing from 3,091 hospitals spanning 2008-2021). This robust tool may prove useful to healthcare leaders to more accurately assess and predict financial distress and bankruptcy in their own institutions in the future.

Data Source

Data were acquired by custom query from *Definitive Healthcare* for years 2008-2021. Data sets were combined (joined) on Medicare Provider Number which is a unique identifier assigned by CMS. For those facilities entering bankruptcy in year y, data from years y-1 and y-2 served as potential predictors. Thus, bankrupt status at year y was modeled as a function of data gathered from year y-1 and year y-2.

Methods and Software

Python 3.x and R served as analytical tools for both explanatory and predictive models. The DV was dichotomous (Bankrupt = 1, not bankrupt = 0). Ohlson's O and Altman's Z were calculated as controls.

Machine learning was used to develop a new model for predictability, titled the "BRKFSST" model.

$$BRKFSST = \left(\widehat{\beta_0} + \widehat{\beta_1}X1 + \widehat{\beta_2}X2 + \widehat{\beta_3}X3 + \widehat{\beta_4}X4 + \widehat{\beta_5}X5 + \widehat{\beta_6}X6 + \widehat{\beta_7}X7 + \widehat{\beta_8}X8\right)$$

$$BRKFSST = (2.84 - .341X1 - .196X2 - .165X3 + .084X4 + 2.595X5 + .489X6 - .221X7 - 4.603X8)$$

Practical Application:

Sum	P (Bankruptcy)	Sum	P (Bankruptcy)
-2	11.92%	1	73.11%
-1	26.89%	2	88.08%
0	50.00%	3	95.26%

Model Estimation

Hospitals were split into a 50% training set and a 50% test set using a pseudo-random number seed and stratification based on bankruptcy status.

50% Training Set
(n=1,545)
1,513 Non-Bankrupt = 0
32 Bankrupt = 1

50% Test Set
(n=1,546)
1,513 Bankrupt = 1
33 Non-Bankrupt = 0

Balancing

The majority weighted minority oversampling (*mwmote*) was used in R, which manages "noisy data." This enabled us to leave the test set imbalanced.

Logistic Regression

LR models served as the most appropriate tool for estimating variable directionality and magnitude of the data in this study. This was appropriate for our data given the Bernoulli nature of each trial. Using this distribution provides the following formula for estimating a bankruptcy: $P(Y_i = 1) = \frac{exp(X\beta)}{1+exp(X\beta)}$. Its complement is therefore $P(Y_i = 0) = \frac{1}{1+exp(X\beta)}$. The odds ratio (OR) is then $\frac{P(Y_i=1)}{P(Y_i=0)}$ which simplifies to $OR = exp(X\beta)$.

By taking the log of the OR result, the equation becomes linear in parameters, $log(OR) = X\beta$ and can be estimated with maximum likelihood estimation. LR avoids homoskedasticity but relies on linearity of log odds for continuous variables, absence of collinearity, and extreme outliers, and independence of observations.

All models were built on the augmented training set. These models were then used to forecast the test set. Typical classification performance metrics including accuracy, precision, recall, specificity, and the F1-score were used to compare models

Coefficient Matrix of BRKFSST Model

	LogOdds	Std. Error	Pr(>|z|)	Odds	VIF
(Intercept)	2.840	0.142	<0.001	17.116	NA
(X1) AR by OpIncome y-1	(0.341)	0.031	<0.001	0.711	1.457
(X2) AR by OpIncome y-2	(0.196)	0.034	<0.001	0.822	1.454
(X3) Current Ratio y-1	(0.165)	0.025	<0.001	0.848	1.049
(X4) CL/CA y-1 (Ohlson O4)	0.084	0.033	0.010	1.087	1.019
(X5) Labor Comp Ratio Change	2.595	0.310	<0.001	13.398	1.097
(X6) Ohlson O4 Change	0.489	0.054	<0.001	1.631	1.045
(X7) Adj Px Days y-1	(0.221)	0.018	<0.001	0.802	1.049
(X8) Hospital Compare y-1	(4.603)	0.341	<0.001	0.010	1.062

Prevalence=0.213	BRKFSST	Altman Original	Altman New	Ohlson Orig	Ohlson New	Zm Orig	Zm New
Sensitivity/Recall	0.7576	0.6667	0.4848	0.0000	0.4848	0.5152	0.6061
Specificity	0.7951	0.5988	0.8050	0.9987	0.8130	0.1613	0.8268
Neg Pred Value	0.9934	0.9880	0.9862	0.9786	0.9864	0.9385	0.9897
PPV/Precision	0.0746	0.0350	0.0514	0.0000	0.0535	0.0132	0.0709
F1	0.1359	0.0665	0.0930	NaN	0.0964	0.0258	0.1270
Detection Rate	0.0162	0.0142	0.0103	0.0000	0.0103	0.0110	0.0129
Detection Prevalence	0.2167	0.4069	0.2012	0.0013	0.1934	0.8318	0.1824
Balanced Accuracy	0.7763	0.6327	0.6449	0.4993	0.6489	0.3382	0.7164

The BRKFSST model achieved an F1-score of 0.876 when predicting the test set and a recall of 0.758. While the model performed better than any of the other models, the positive predictive value (PPV) is somewhat problematic. While the recall was 0.758, the PPV was 0.078, but still the best of the models estimated.

Implications

Managerial

Organizations must monitor performance metrics. This hospital-specific model encompasses reliable predictive factors to enable insight that ensures long-term financial viability.

Policy Makers

Policy makers can use this model to scan the environment to examine geographical areas or specific hospital ownership characteristics that are struggling more than others and devise incentives or policy to ease this financial distress.

Based on our analysis, we contend both sound financial structure as well as supportive accreditation and quality performance all meaningfully insulate an organization against long-term economic underperformance.

Limitations

1. Our modeling falls short of providing hospital leaders with the exact values beyond which organizational solvency is impossible to sustain.
2. There may be other factors with a significant influence on bankruptcy that we did not consider in our study.
3. Our study does not capture those hospitals that are near bankruptcy or in other stages of financial distress. We used a dichotomous DV.
4. We are not able to verify the accuracy of the data beyond what is reported to the American Hospital Association, CMS, and other agencies.

Conclusions

Bankruptcy is the unfortunate result of many businesses throughout the US economy. Businesses start and fail daily in our country, but with the closure of a hospital comes the increased societal cost of not just lost jobs, but also poorer access to care, and other supportive clinical services. It is these implications that provided the impetus for our research into gaining a deeper understanding of what contributes to hospital bankruptcy.

Manuscript Under Consideration

This manuscript is currently under consideration for publication through *Healthcare Management Science*. Conceptualization, analysis, and paper draft development was performed in cooperation with colleagues at Boston College.

Future Research

Future study will leverage our model to evaluate all US short term acute care hospitals to determine the risk of bankruptcy and closure. This could be insightful to healthcare policy leaders.

TEXAS STATE UNIVERSITY®

MEMBER THE TEXAS STATE UNIVERSITY SYSTEM