

The Intricacies of College and University Closures: A Generalized Linear Model Perspective

Johnny Pang¹ & Ryan Sonn²

¹ Department of Mathematics, Monmouth University, West Long Branch, New Jersey, USA

² Department of Computer Science, Monmouth University, West Long Branch, New Jersey, USA

Correspondence: Johnny Pang, Department of Mathematics, Monmouth University, West Long Branch, New Jersey, USA. E-mail: wpang@monmouth.edu

Received: February 13, 2025

Accepted: March 30, 2025

Online Published: June 7, 2025

doi:10.5539/ijbm.v20n4p86

URL: <https://doi.org/10.5539/ijbm.v20n4p86>

Abstract

This study delves into the intricate and multifaceted factors driving state-level college and university closures, leveraging publicly accessible variables through a Generalized Linear Model (GLM) analysis, in contrast to Multiple Discriminant Analysis (MDA). The analysis identifies key determinants of closures, including institutional endowment, tuition, and the percentage of in-state students. A pivotal contribution of this research is the development of a predictive Z-score model and ranges, offering robust and efficient tools for classifying higher education institution closures with enhanced applicability.

Keywords: generalized linear model, GLM, bankruptcy, Z-score, university, endowment, institution, box-cox

1. Introduction

Pursuing higher education represents one of the most significant investments an individual or family can undertake, second only to the acquisition of a home. Many often overlook the importance of verifying an institution's operational status before making this substantial commitment. The allure of prestigious rankings and well-known university brands often overshadows critical considerations regarding an institution's long-term viability. While the opportunity to attend highly esteemed universities is limited, it is crucial to understand the financial stability of the institutions being considered. Unlike public companies that must provide detailed annual reports, colleges and universities that do not receive federal funding are not mandated to disclose their revenue and other financial metrics. This lack of transparency makes it challenging to assess the true health of an institution. From 2004 to 2021, roughly 861 US colleges and universities ceased operations, accounting for nearly fifteen percent of the 5,860 institutions that were operating at the time (Barshay, 2023). Closures arise from various causes. Financial difficulties can significantly impair an institution's operational capabilities. Some prominent financial challenges include endowment issues, declining enrollment, and accreditation problems. Public institutions also grapple with reduced state funding, straining budgets and necessitating cuts to programs and staff. Rising operational costs—such as faculty salaries, technology upgrades, and utilities—can become a substantial burden when coupled with diminishing tuition income. Without effective strategies to mitigate these financial pressures, colleges and universities may find themselves unable to continue operations, ultimately leading to closure.

This article aims to forecast the risk of college and university closures by leveraging widely accessible data from university annual reports and websites. While the precision of predictions would be significantly enhanced by incorporating technical financial metrics—such as total assets, working capital, debt-to-equity ratio, return on equity, and earnings before interest, taxes, depreciation, and amortization (EBITDA)—analogous to the data available from public companies, these metrics are generally not obtainable for colleges and universities. Although certain private universities may voluntarily provide financial information, it is not consistently available. This study endeavors to utilize readily accessible data and comprehensible terms to develop a robust score range for predicting the closures of higher education institutions. The intention is to equip individuals with a formula that is both straightforward and insightful.

1.1 Altman's Z-Score

The feasibility of predicting such closures using a Generalized Linear Model (GLM) and deriving an effective

score range is a central question of this research. Predicting institutional closures through mathematical formulas has a well-established precedent. In 1968, Edward Altman introduced the Altman Z-score formula in his seminal study, "Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy" (Altman, 1968). Altman's research involved sampling sixty-six firms, with an equal split between bankrupt and solvent manufacturing firms. These samples were used to construct a predictive function that generates a Z-score indicating the probability of default. Altman gathered balance sheet and income statement data for each firm and calculated twenty-two financial ratios, ultimately selecting the five most significant ratios as predictive variables. These variables included the net working capital to total assets ratio, earnings before interest and taxes to total assets ratio, retained earnings to total assets ratio, market value of equity to total liabilities ratio, and sales to assets ratio. Altman later revised his Z-score formulas for subsequent situations and industries in later publications (Altman, 2018). Other scholars, such as Begley, Ming, and Watts, have also developed their own Z-score formulas (Begley, Ming, and Watts, 1996).

1.2 Generalized Linear Model

Although Altman's original Z-score methodology is not utilized, this study designates the calculated score as a Z-score. This terminology is chosen intentionally, as the term Z-score is widely recognized in statistics for its role in identifying outliers. This study employs a GLM to compute the Z-score, as opposed to the traditional Multiple Discriminant Analysis (MDA) employed by Altman to compute the Altman's Z-score. The conventional MDA model's reliance on assumptions of normally distributed predictors and homoscedasticity often limits its practical application in real-world scenarios. Conversely, the GLM assumes that the predictors are linearly related to the log-odds of the response variable, offering a robust and adaptable framework that effectively manages mixed qualitative and quantitative predictors without enforcing stringent assumptions about data distribution or homoscedasticity. Particularly, the logistic regression component within the GLM is exceptionally suited for binary classification tasks, as demonstrated by the categorization of institutions in this study as either operational or bankrupt. All these factors endow the GLM Z-score formula with its own distinct merits.

The analysis will employ the following four qualitative and four quantitative variables:

1. **Status:** This qualitative response variable indicates whether the college or university is currently operational or bankrupt.
2. **Type:** This qualitative predictor specifies whether the educational institution is categorized as a college or a university.
3. **Religious:** This qualitative predictor specifies whether the educational institution is religiously affiliated.
4. **Public/Private:** This qualitative predictor specifies whether the educational institution is categorized as public or private.
5. **Endowment:** This quantitative predictor enumerates the endowment figures (in billions) for the latest academic year (2023-2024) for operational institutions, and the endowment at the time of closure for defunct institutions.
6. **Enrollment:** This quantitative predictor enumerates the annual enrollment figures for the latest academic year for operational institutions, and the annual enrollment figures at the time of closure for defunct institutions..
7. **Tuition:** This quantitative predictor enumerates the annual in-state tuition fees (in dollars) for the latest academic year (2023-2024) for operational institutions, and the annual in-state tuition at the time of closure for defunct institutions.
8. **Percentage_Instate:** This quantitative predictor, ranging between 0 and 1, details the proportion of in-state students relative to the total student population.

All predictors and the response variable should be comprehensible to a general audience and readily accessible on university websites. It is crucial to note that all qualitative predictors, including the response variable, exhibit only two distinct levels. For instance, Status is classified as either operational or bankrupt, Type distinguishes between college and university, and Religious indicates yes or no. These variables must be converted into binary (dummy) variables. Each dummy variable will represent one level of the categorical variable, with a value of 0 for the specific category and 1 for the others. The dummy variable for Status will assign 0 to operational and 1 to bankrupt, the dummy variable for Type will assign 0 to college and 1 to university, and the dummy variable for Religious will assign 0 to no and 1 to yes. It is important to emphasize that although dummy variables do not carry inherent quantitative meaning, they allow the model to contrast different categories against a reference group, which is typically omitted to avoid multicollinearity. In this context, the relationship between a binary

qualitative response variable and both qualitative and quantitative predictors is often curvilinear. A common curvilinear pattern is the S-shaped curve, which is effectively captured by the logistic GLM.

This dataset comprises a total of 403 institutions, comprising 282 universities and 121 colleges, 19 of which ceased operations between 2014 and 2023. Of these, 194 are private non-profit institutions, while 209 are public institutions. Additionally, 305 institutions are non-religious, and 98 are affiliated with religious organizations. These institutions are distributed across all 50 states, including Puerto Rico. Only private non-profit institutions are included, as most private for-profit institutions do not disclose endowment or other relevant information. Prior to the commencement of the research, it was assumed that endowment figures would be widely available and represent a key indicator of institutional strength and sustainability. The institutions within the dataset exhibit significant variation in key financial metrics—endowment, enrollment, and tuition—each of which plays a crucial role in shaping the financial health and operational priorities of these institutions. Endowments ranged from \$145,000 to an impressive \$49.5 billion, with a mean endowment of \$1.791 billion. Enrollment figures varied from a modest 145 students to a substantial 152,812 students. In-state tuition fees ranged from \$4,168 to \$65,000 annually, with a mean of \$28,730. Demographically, these colleges and universities show considerable regional variation. The proportion of in-state students at each institution ranged from as low as 5% to as high as 95%, reflecting the differing geographic dynamics and regional priorities in higher education across the United States.

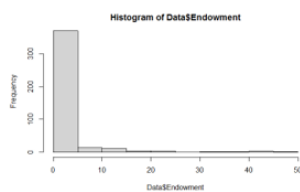


Figure 1. Histogram of Endowment

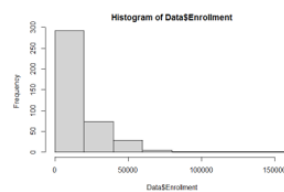


Figure 2. Histogram of Enrollment

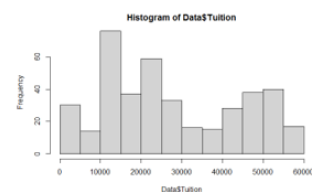


Figure 3. Histogram of Tuition

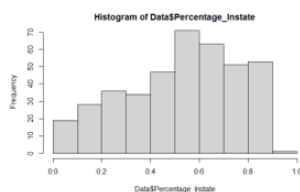


Figure 4. Histogram of Percentage_Instate

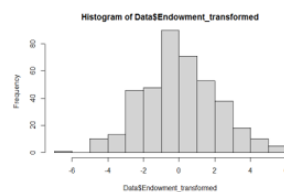


Figure 5. Histogram of Endowment_transformed

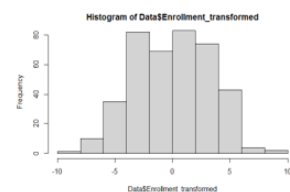


Figure 6. Histogram of Enrollment_transformed

1.3 Box-Cox Transformation

Figure 1 illustrates the histogram for Endowment, while Figure 2 illustrates the histogram for Enrollment. Both histograms exhibit a pronounced right skew, indicating a significant departure from normality. Unlike traditional linear models (LM), which require the normality of predictors, logistic regression within the GLM framework is robust to non-normality, as it assumes that the predictors are linearly related to the log-odds of the response variable. Thus, logistic regression focuses on estimating odds ratios rather than directly estimating the mean of the response variable. However, the skewness present in the two predictors may still pose challenges. These factors can affect the estimated coefficients and inflate standard errors, potentially influencing the model's performance and compromising the validity of the results. To address these issues, the skewness can be adjusted back to a normal distribution by applying the Box-Cox transformation method, which essentially raises each positive predictor to an optimal power and then normalizes the result by dividing by the corresponding optimal power value (Box, 1964). This approach tends to yield more satisfactory results compared to conventional log, square root, or reciprocal transformations. Like all transformations, the Box-Cox method has a key drawback: it complicates the interpretation of the coefficients for the transformed variables. However, model stability, interpretability, and performance are paramount. In this regard, the benefits of the Box-Cox transformation outweigh its drawbacks. Figures 5 and 6 present the histograms of the Box-Cox transformed and centered (without scaling) Endowment and Enrollment, respectively. Centering without scaling the two predictors offers the advantage of eliminating biases related to the respective means while preserving the original scales. Due to linear computations, the means can be easily adjusted at the end of the fitted model and incorporated as a constant. At this stage, GLM with the predictors using Box-Cox transformation makes the model quite complex, so the aim is to avoid complicating the model by introducing numerous calculations. Standardizing the variables may be done later if they become

necessary, though in most cases, they are not required. Figure 3, on the other hand, illustrates the histogram of Tuition, which appears to conform to the normality assumption, exhibiting a symmetric distribution. It is worth noting that the scale of Tuition is quite large, which may eventually lead to complications. Figure 4 presents the histogram for the percentage of in-state tuition, which also exhibits a roughly normal distribution.

$$E(\text{Status}) = \frac{\exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7)}{1 + \exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6 + \beta_7x_7)} \tag{1}$$

where Status = $\begin{cases} 0 & \text{if The institution is operational} \\ 1 & \text{if The institution is bankrupt} \end{cases}$
 $x_1 = \text{Endowment_transformed}$
 $x_2 = \text{Enrollment_transformed}$
 $x_3 = \begin{cases} 0 & \text{if Type = College} \\ 1 & \text{if Type = University} \end{cases}$
 $x_4 = \begin{cases} 0 & \text{if PublicPrivate = Private} \\ 1 & \text{if PublicPrivate = Public} \end{cases}$
 $x_5 = \begin{cases} 0 & \text{if Religious = No} \\ 1 & \text{if Religious = Yes} \end{cases}$
 $x_6 = \text{Tuition}$
 $x_7 = \text{Percentage_Instate}$

The seven identified predictors will be utilized to forecast institutional closures. The model presented in Equation 1 will be employed to classify institutions into three distinct zones: Operational (institutions in this zone exhibit strong and healthy operational performance), Gray Area (institutions within this zone demonstrate vulnerabilities in operational health and are deemed at potential risk of bankruptcy), and Bankrupt (institutions categorized in this zone are characterized by significant operational deficiencies and are assessed to face a high likelihood of bankruptcy).

Table 1. Correlation Matrix

	Endowment_transformed	Enrollment_transformed	Tuition	Percentage_Instate
Endowment_transformed	1.0000000	0.5751570	0.3919948	-0.3802254
Enrollment_transformed	0.5751570	1.0000000	-0.2760136	0.1976776
Tuition	0.3919948	-0.2760136	1.0000000	-0.8305723
Percentage_Instate	-0.3802254	0.1976776	-0.8305723	1.0000000

Table 1 presents the correlation matrix for all numerical predictors. The correlation coefficient between Tuition and Percentage_Instate is -0.8305723, indicating a strong negative linear relationship. This suggests potential multicollinearity between the two variables. While it may be beneficial to either standardize the variables or remove one from the model, it is important to note that correlation alone does not necessarily indicate an issue, provided that the Variance Inflation Factors (VIFs) for the predictors in the optimal model remain sufficiently low. Therefore, at this stage, we will retain both predictors and reassess them later, if necessary, based on VIFs.

```
Call: glm(formula = Status ~ (Enrollment_transformed + Endowment_transformed + Type + PublicPrivate + Religious + Tuition + Percentage_Instate), family = binomial(), data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
      -2.21010      -0.01708      -0.00001       0.00000       2.45806

Coefficients:
(Intercept)          -2.722e+00    3.012e+00   -0.904    0.366126
Enrollment_transformed  2.404e-02    1.442e-01    0.167    0.867599
Endowment_transformed -1.447e+00    4.070e-01   -3.556    0.000377 ***
TypeUniversity        -1.041e+00    9.788e-01   -1.063    0.287728
PublicPrivatePublic   -2.838e+01    2.248e+03   -0.013    0.989923
ReligiousYes          -1.049e+00    8.508e-01   -1.233    0.217591
Tuition              -1.302e-04    5.904e-05   -2.205    0.027477 *
Percentage_Instate     6.336e+00    3.841e+00    1.649    0.099078 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161 on 402 degrees of freedom
Residual deviance: 43.754 on 395 degrees of freedom
AIC: 59.754
Number of Fisher Scoring iterations: 21

Enrollment_transformed      Endowment_transformed      Type      PublicPrivate
      1.465481                1.057195                1.322918                1.000001

Religious      Tuition      Percentage_Instate
1.179653      1.264414      1.139249
```

Figure 7. R Printout of the GLM initial model

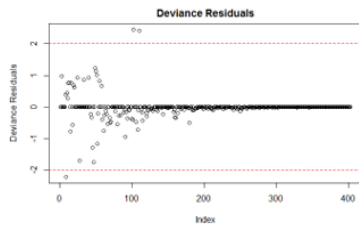


Figure 8. Deviance Residuals Plot

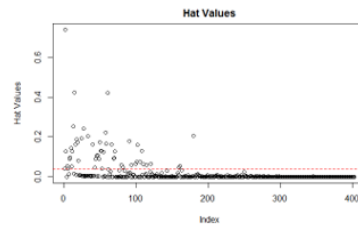


Figure 9. Hat Values Plot

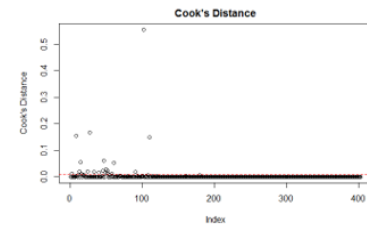


Figure 10. Cook's Distance Plot

2. Residual Analysis

Figure 7 illustrates the R printout for the initial model, employing a logistic GLM due to the binary nature of the response variable. The deviance residuals range from -2.21010 to 2.45806, suggesting the absence of extreme outliers within the model. This preliminary analysis highlights that *Endowment_transformed* and *Tuition* are statistically significant predictors of the outcome variable, *Status*, indicating their substantial influence on the model. It is noteworthy that none of the qualitative predictors were deemed significant in this initial evaluation. Furthermore, it is crucial to assess the VIFs of the model to identify potential multicollinearity among the predictors, thereby ensuring the stability and reliability of the model as it undergoes further refinement. Figure 7 also illustrates the VIF for each predictor in the initial model. None of the predictors exhibit signs of multicollinearity, as all VIFs fall within the acceptable threshold ($VIF < 10$). This indicates that multicollinearity is not a significant concern in this initial first-order model. While simpler models with fewer predictors generally encounter fewer issues related to multicollinearity, it is crucial to reassess VIFs as the model is refined and additional interaction terms are incorporated. This reassessment ensures the robustness and validity of the model. The deviance residual plot in Figure 8 indicates again that there are no extreme outliers requiring removal. In Figures 9 and 10, some data points exhibit higher hat values and Cook's distance values than the typical thresholds (with the hat value threshold at 0.0397 and Cook's distance threshold at 0.0099). While this suggests that these points are influential, a thorough examination of the dataset reveals that they do not qualify as outliers. Consequently, these influential non-outlier points will remain in the model.

3. Stepwise Regression

To enhance the initial model, the first step is to perform stepwise regression, followed by incorporating interaction terms of the remaining significant predictors. This sequential approach offers the advantage of first reducing the number of independent variables and identifying key predictors, ensuring that only the most relevant variables are retained. Once this step is completed, interaction terms can be added to the model, refining it into a more parsimonious and effective representation of the underlying relationships. This methodology preserves the model's predictive power while improving its efficiency and interpretability. Such an approach is particularly critical when working with limited sample sizes, as is the case with the current dataset of 403 institutions. One potential drawback is that significant predictors with interactions may be excluded during the stepwise selection, leading to the omission of important interaction effects. The benefits of this approach, particularly in terms of computational efficiency and minimizing the risk of having insufficient degrees of freedom, far outweigh this concern, given the constraints of the sample size.

Discerning readers will observe that the model is unweighted. Identifying bankrupt institutions correctly has been a significant challenge, even with the inclusion of all seven predictors, particularly the predictor *Endowment*. The limited number of bankrupt institutions—only 19 out of 403 total institutions—is especially concerning. Given that the proportion of bankrupt institutions is significantly smaller than that of operational institutions, applying weights to the sample is inadvisable. Assigning greater weight to the underrepresented group to ensure a more balanced contribution from both categories introduces the risk of instability to the model.

Beyond weighting in GLM, other methods could be considered. Synthetic techniques, such as SMOTE (Synthetic Minority Oversampling Technique) or ADASYN (Adaptive Synthetic Sampling), generate synthetic samples for the minority class. However, these methods may lead to overfitting if the synthetic samples do not accurately represent the minority class distribution. Alternatively, undersampling reduces the size of the majority class but risks losing valuable information and shrinking the dataset overall. Lastly, stratified resampling for validation ensures proportional representation during evaluation but does not balance the dataset during training.

Each method has its pros and cons. Stratified resampling for validation is preferred, as it ensures both classes are proportionally represented in each fold during cross-validation and preserves the original class distribution while

evaluating the model's performance. This approach will be implemented after finalizing the chosen model.

```
Call: glm(formula = Status ~ Endowment_transformed + PublicPrivate + Tuition + Percentage_Instate, family = binomial(),
data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-2.19435    -0.01923    -0.00002     0.00000     3.10504

Coefficients:
(Intercept)      -5.482e+00    2.644e+00   -2.073    0.038138 *
Endowment_transformed -1.479e+00    3.999e-01   -3.700    0.000216 ***
PublicPrivatePublic -2.742e+01    2.265e+03   -0.012    0.990341
Tuition          -1.080e-04    5.221e-05   -2.068    0.038670 *
Percentage_Instate  8.007e+00    3.800e+00    2.107    0.035108 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161  on 402  degrees of freedom
Residual deviance: 46.616  on 398  degrees of freedom
AIC: 56.616
Number of Fisher Scoring iterations: 21
      Endowment_transformed      PublicPrivate      Tuition      Percentage_Instate
      1.022226                1.000001                1.035062                1.013322
```

Figure 11. R Printout of the Stepwise Model

Figure 11 displays the printouts of the new GLM model and the VIFs of the significant predictors after stepwise regression. It is important to note that the VIFs do not need to be re-examined, as the printout reflects a parsimonious model derived from the initial model, assuming the predictors in the initial model did not exhibit high VIFs. Clearly, the parsimonious model of the initial model also does not show high VIFs. The printout indicates that the predictor PublicPrivate has non-significant p-value, suggesting it should be excluded from the model. While it is true that the predictor PublicPrivate could be excluded from the model, its inclusion may not necessarily lead to significant issues. Although it may result in less precise estimates or greater variability, these outcomes are not inherently problematic. The underlying reasons for the elevated p-value should be addressed. One potential explanation could be an imbalance in the proportions between Public and Private institutions. That said, this is not the case here, as the data is well-balanced with 194 private non-profit institutions and 209 public institutions.

```
glm(formula = Status ~ (Endowment_transformed + Tuition + Percentage_Instate), family = binomial(), data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-2.05076    -0.21450    -0.09376    -0.03105     2.67407

Coefficients:
(Intercept)      -6.860e+00    2.159e+00   -3.178    0.00148 **
Endowment_transformed -1.361e+00    2.378e-01   -5.724    1.04e-08 ***
Tuition          6.928e-05    3.110e-05    2.228    0.02591 *
Percentage_Instate -7.204e-01    2.407e+00   -0.299    0.76477
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161  on 402  degrees of freedom
Residual deviance: 88.674  on 399  degrees of freedom
AIC: 96.674
Number of Fisher Scoring iterations: 8
      Endowment_transformed      Tuition      Percentage_Instate
      1.476726                2.448803                2.417441
```

Figure 12. R Printout of the Stepwise Model with Predictor PublicPrivate Omitted

Figure 12 presents the model output excluding the predictor PublicPrivate. A more serious issue emerges as the Akaike Information Criterion (AIC) of the model escalates from 56.616 to 96.674, and the predictor Percentage_Instate now exhibits a non-significant p-value. It is crucial to recognize that stepwise regression in GLM does not rely on p-values to identify the most parsimonious model. Instead, it assesses the model's quality using AIC, which is calculated by considering the difference between the number of parameters in the model (including the intercept) and the natural log of the likelihood of the model.

A model with a lower AIC indicates a superior balance between predictive accuracy and simplicity, thereby making it the preferred choice for prediction. This indicates that the current model, after dropping the predictor PublicPrivate, performs worse in estimating the data than the previous model. Therefore, the new model should not be accepted. Another significant issue with the new model is that the p-value of the predictor Percentage_Instate is now 0.76477, indicating that it is not significant to the model. One possible explanation for these changes is the presence of multicollinearity and interactions between PublicPrivate and other predictors. However, the issue of multicollinearity can be ruled out, as it is not a concern. An examination of the VIFs in both Figures 11 and 12 confirms that no multicollinearity exists in either the previous or the present models. To address this, an interaction model will be examined to assess its effects. With all three predictors, a three-way

interaction model will be considered. Nevertheless, given the current sample size, it is important to be aware that multiple-way interaction models may not always be feasible, as they introduce too many parameters. This could potentially result in overfitting, where the model fits the noise of the data, loss of degrees of freedom leading to instability, or, worse, causing the model to fail to converge.

```

glm(formula = Status ~ (Endowment_transformed + Tuition + Percentage_Instate)^3, family = binomial(), data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-1.94504    -0.05276    -0.01054    -0.00339     2.60750

Coefficients:
(Intercept)                1.553e+01    6.402e+00    2.426    0.01527 *
Endowment_transformed      -2.253e+00    1.776e+00   -1.269    0.20457
Tuition                    -6.943e-04    2.316e-04   -2.998    0.00272 **
Percentage_Instate         -3.864e+01    1.437e+01   -2.688    0.00719 **
Endowment_transformed:Tuition  1.066e-04    5.130e-05    2.078    0.03770 *
Endowment_transformed:Percentage_Instate  1.119e+00    3.385e+00    0.330    0.74105
Tuition:Percentage_Instate   1.241e-03    4.895e-04    2.535    0.01126 *
Endowment_transformed:Tuition:Percentage_Instate -2.129e-04    1.126e-04   -1.892    0.05855 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161 on 402 degrees of freedom
Residual deviance: 43.723 on 395 degrees of freedom
AIC: 59.723
Number of Fisher Scoring iterations: 10
Endowment_transformed  Tuition  Percentage_Instate  Endowment_transformed:Tuition
26.68987      22.62660      24.32858      33.57402
Endowment_transformed:Percentage_Instate  Tuition:Percentage_Instate
29.65848
Endowment_transformed:Tuition:Percentage_Instate
40.87880
55.81698
    
```

Figure 13. R Printout of the Three-Way interaction model

3.1 Interaction Models

Figure 13 presents the printouts of the three-way interaction model. Notably, the AIC is significantly lower than that of the previous model in Figure 15 and only marginally lower than the initial model, which is a positive outcome. The p-value of the three-way interaction term is slightly larger, indicating that it may need to be considered for removal. Of particular concern is the interaction term between Endowment_transformed and Percentage_Instate, which has a p-value of 0.74105. Given this, it will undoubtedly and unquestionably be removed. All VIFs of the predictors in the new model have become quite high. With these high VIFs, it will be impossible to analyze the effect of each predictor on the response variable. This phenomenon indicates that multicollinearity is significantly present due to the interaction terms introduced. It is as though feedback loops have been created to collinear all predictors. Consequently, the next step will be to remove the predictor interaction term Endowment_transformed and Percentage_Instate. The next model will then be reassessed to determine if further actions are needed.

```

glm(formula = Status ~ (Endowment_transformed + Tuition + Percentage_Instate)^3, family = binomial(), data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-1.92807    -0.05312    -0.00932    -0.00265     2.61477

Coefficients:
(Intercept)                1.644e+01    6.027e+00    2.728    0.00636 **
Endowment_transformed      -1.742e+00    1.089e+00   -1.600    0.10965
Tuition                    -7.212e-04    2.418e-04   -2.983    0.00285 **
Percentage_Instate         -4.101e+01    1.313e+01   -3.125    0.00178 **
Endowment_transformed:Tuition  9.183e-05    4.571e-05    2.009    0.04456 *
Tuition:Percentage_Instate   1.314e-03    4.760e-04    2.760    0.00577 **
Endowment_transformed:Tuition:Percentage_Instate -1.794e-04    6.522e-05   -2.751    0.00595 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161 on 402 degrees of freedom
Residual deviance: 43.856 on 396 degrees of freedom
AIC: 57.856
Number of Fisher Scoring iterations: 11
Endowment_transformed  Tuition  Percentage_Instate  Endowment_transformed:Tuition
10.62705      25.36683      21.75538      23.57151
Tuition:Percentage_Instate  Endowment_transformed:Tuition:Percentage_Instate
55.05907
13.82886
    
```

Figure 14. R Printout of the Three-Way interaction model with endowment_transformed and percentage_instate omitted

Figure 14 illustrates the printouts with the interaction term Endowment_transformed:Percentage_Instate removed. Although the AIC is quite reasonable, this new model raises two concerns. Firstly, the predictor Endowment_transformed has a p-value of 0.10965, indicating that this predictor should be removed if all interactions related to Endowment_transformed exhibit high p-values. The three interaction

terms—Endowment_transformed:Tuition with a p-value of 0.04456, Tuition:Percentage_Instate with a p-value of 0.00577, and Endowment_transformed: Tuition:Percentage_Instate with a p-value of 0.00595—have significantly small p-values. These strongly suggest that the predictor Endowment_transformed should be retained. Secondly, despite a reduction from extremely high values, the VIFs remain elevated. It is important to emphasize that a few high VIFs do not necessarily indicate a serious issue. Predictions will remain reliable if the model coefficients are stable. However, when all VIFs are high, the stability of the model coefficients is compromised, leading to an unreliable prediction model. A new approach must be found to address this issue. It is noted that the predictor Tuition has a very low coefficient value, indicating a significant range from 4168 to 60000, compared to the other predictors: Endowment_transformed ranges from -6.6 to 5.5, and Percentage_Instate ranges from 0.05 to 0.95. With the substantial scale of the predictor Tuition, significant challenges arise in the GLM. The issue stems from the predictor Tuition exerting a disproportionately strong influence due to its measurement in larger units. Consequently, it appears highly significant, overshadowing other predictors measured on a smaller scale. To address this imbalance, one effective solution is to center it by subtracting the mean. As mentioned earlier, this centering process will scale down the range while preserving the original scale. This centering procedure should have been implemented from the beginning, despite adding complexity to the model. Initially, a Box-Cox transformation was applied to two of the other predictors. Given the significant impact of centering the predictor Tuition on the model, it is prudent to re-examine the three-way model including all interaction terms. This entails reintroducing Endowment_transformed:Percentage_Instate and conducting a thorough reassessment.

```

glm(formula = Status ~ (Endowment_transformed + Tuition_transformed + Percentage_Instate)^3, family = binomial(), data
= Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-1.94504    -0.05276    -0.01054    -0.00339     2.60750

Coefficients:
(Intercept)                Estimate Std. Error z value Pr(>|z|)
Endowment_transformed    -4.4206953    1.9263316  -2.295  0.02174 *
Tuition_transformed       0.8099931    0.8623838   0.939  0.34760
Percentage_Instate      -0.0006943    0.0002316  -2.998  0.00272 **
Endowment_transformed:Tuition_transformed  0.0001066    0.0000513   2.078  0.03770 *
Endowment_transformed:Percentage_Instate -4.9993616    1.8317170  -2.729  0.00635 **
Tuition_transformed:Percentage_Instate   0.0012407    0.0004895   2.535  0.01126 *
Endowment_transformed:Tuition_transformed:Percentage_Instate -0.0002129    0.0001126  -1.892  0.05855 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161  on 402 degrees of freedom
Residual deviance: 43.723  on 395 degrees of freedom
AIC: 59.723
Number of Fisher Scoring iterations: 10
Endowment_transformed      Tuition_transformed      Percentage_Instate      Endowment_transformed:Tuition_transformed
6.294294      22.626604      1.854094
10.519938
Endowment_transformed:Percentage_Instate      Tuition_transformed:Percentage_Instate
8.685596      27.473130
Endowment_transformed:Tuition_transformed:Percentage_Instate
14.222639

```

Figure 15. R Printout of Three-Way interaction model with centered tuition

Figure 15 illustrates the model's printouts subsequent to centering the predictor Tuition. The model demonstrates robust construction, as evidenced by a reasonable AIC. Notably, when comparing this model to the three-way interaction model without centering the predictor Tuition (as seen in Figure 14), substantial improvements in lowering the VIFs are observed. Despite this progress, the VIFs for the newly centered predictor Tuition_transformed and the interaction term Tuition_transformed:Percentage_Instate remain significantly elevated, which necessitates further attention. It may be advantageous to center the predictor Percentage_Instate from the outset. At this stage, it could be argued that Percentage_Instate should have been centered initially. Although this would have been beneficial, the pursuit of the optimal model inherently involves a process of trial and error, necessitating multiple iterations.

```

glm(formula = Status ~ (Endowment_transformed + Tuition_transformed + Percentage_Instate_transformed)^3, family =
binomial(), data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-1.94504   -0.05276   -0.01054   -0.00339    2.60750

Coefficients:
(Intercept)                Estimate Std. Error z value Pr(>|z|)
06 ***                    -6.081e+00  1.266e+00  -4.803   1.57e-
06 ***
Endowment_transformed      -1.968e+00  4.686e-01  -4.200   2.67e-
05 ***
Tuition_transformed        -4.854e-06  1.263e-04  -0.038   0.96935
Percentage_Instate_transformed -2.987e+00  3.968e+00  -0.753   0.45158
Endowment_transformed:Tuition_transformed -1.173e-05  4.047e-05  -0.290   0.77203
Endowment_transformed:Percentage_Instate_transformed -4.999e+00  1.832e+00  -2.729   0.00635
**
Tuition_transformed:Percentage_Instate_transformed 1.241e-03  4.895e-04  2.535   0.01126
*
Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed -2.129e-04  1.126e-04  -1.892
0.05855 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161 on 402 degrees of freedom
Residual deviance: 43.723 on 395 degrees of freedom
AIC: 59.723
Number of Fisher Scoring iterations: 10
Endowment_transformed Tuition_transformed Percentage_Instate_transformed Endowment_transformed:Tuition_transformed
1.858343 6.73034 1.854094 6.547689
Endowment_transformed:Percentage_Instate_transformed Tuition_transformed:Percentage_Instate_transformed
2.197335 2.154993
Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed
1.873673

```

Figure 16. R Printout of Three-Way interaction model with centered percentage_instate

Figure 16 illustrates the outputs of the three-way interaction model incorporating the centered predictor, Percentage_Instate_transformed. The acceptable VIF now indicates the multicollinearity has been successfully mitigated. This achievement is a testament to the iterative nature of model refinement, where multiple cycles of trial and error are essential for optimization. It's worth acknowledging that some may contend that earlier standardization could have preserved additional terms. Given model convergence, main effect predictors are generally retained in stepwise regression, whether employing forward or backward approaches, though exceptions do exist. The model still exhibits a reasonable AIC, and from a technical standpoint, it has improved. Despite these advancements, due to the non-significant p-values associated with one of the two-way interaction terms, Endowment_transformed:Tuition_transformed, and the three-way interaction term, Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed, it is prudent to omit these terms.

```

glm(formula = Status ~ (Endowment_transformed + Tuition_transformed + Percentage_Instate_transformed)^3 -
Endowment_transformed:Tuition_transformed, family = binomial(), data = Data)
Deviance Residuals:      Min       1Q   Median       3Q      Max
-1.98998   -0.05592   -0.01031   -0.0030    2.60996

Coefficients:
(Intercept)                Estimate Std. Error z value Pr(>|z|)
-6.058e+00  1.243e+00  -4.872   1.10e-06 ***
Endowment_transformed      -1.961e+00  4.615e-01  -4.249   2.15e-05 ***
Tuition_transformed        2.848e-05  5.260e-05  0.541   0.58821
Percentage_Instate_transformed -2.912e+00  3.925e+00  -0.742   0.45805
Endowment_transformed:Percentage_Instate_transformed -4.864e+00  1.764e+00  -2.757   0.00584 **
Tuition_transformed:Percentage_Instate_transformed 1.317e-03  4.258e-04  3.094   0.00198 **
Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed -1.974e-04  8.360e-05  -2.361   0.01824 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
Null deviance: 153.161 on 402 degrees of freedom
Residual deviance: 43.814 on 396 degrees of freedom
AIC: 57.814
Number of Fisher Scoring iterations: 10
Endowment_transformed Tuition_transformed Percentage_Instate_transformed
1.843173 1.151917 1.865934
Endowment_transformed:Percentage_Instate_transformed Tuition_transformed:Percentage_Instate_transformed
2.272072 1.819014
Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed
1.321161

```

Figure 17. R Printout of Three-Way Interaction Model with excluding the two-way interaction term Endowment_transformed:Tuition_transformed and the three-way interaction term Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed

3.2 Cross-Validation and Ridge Regression

Figure 17 illustrates the outputs of the interaction model, excluding the two-way interaction term Endowment_transformed:Tuition_transformed and the three-way interaction term

Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed. This model is deemed optimal with an AIC of 57.18, the second lowest after the model depicted in Figure 11, which has an AIC of 56.616. The VIFs are notably low, the deviance residuals are minimal (-1.98998, 2.60996), and all individual p-values are both significant and reasonable. The difference between the null deviance and the residual deviance is 109.314 (153.161 - 43.814) on 6 degrees of freedom, indicating a superior model fit. The Fisher scoring iterations total only 10, signifying rapid convergence. This model exhibits well-behaved characteristics and is therefore regarded as the best model identified thus far. To further enhance its performance, two additional steps will be undertaken. Firstly, cross-validation will be performed to obtain a more reliable estimate of the model's Mean Squared Error (MSE). This ensures the model's robustness and generalizability. Secondly, ridge regression will be applied to the current model. While ridge regression is typically employed to address multicollinearity among predictors or in models with numerous predictors and limited observations, the primary concern here is to avoid overfitting. There is a concern that this model performs excellently on the training data but poorly on the test data. By using ridge regression, the goal is to ensure that the model generalizes well to new, unseen data.

Table 2. Coefficients of ridge regression model

Best Lambda: 0.001 Cross-Validation MSE: 0.115295181314655	
8 x 1 sparse Matrix of class "dgCMatrix"	
(Intercept)	-4.978116e+00
Endowment_transformed	-1.496644e+00
Tuition_transformed	2.245556e-05
Percentage_Instate_transformed	3.285035e+00
Endowment_transformed:Percentage_Instate_transformed	-2.032878e+00
Tuition_transformed:Percentage_Instate_transformed	6.900580e-04
Endowment_transformed:Tuition_transformed:Percentage_Instate_transformed	-1.510567e-04

Table 2 presents the optimal coefficients of the model after applying Ridge regression. Notably, the tuning parameter lambda is set to 0.001, which implies that this Ridge regression model closely mirrors the model depicted in Figure 17. The 10-fold cross-validation MSE of 0.1153 indicates that the model's predictions are highly aligned with the actual values. The subsequent step involves making predictions based on this model.

$$\hat{Z} = \frac{\exp(-4.978116 - 1.496644x_1 - 0.00002245556x_8 + 3.285035x_9 - 2.032878x_1x_9 + 0.000690058x_8x_9 - 0.0001510567x_1x_8x_9)}{1 + \exp(-4.978116 - 1.496644x_1 - 0.00002245556x_8 + 3.285035x_9 - 2.032878x_1x_9 + 0.000690058x_8x_9 - 0.0001510567x_1x_8x_9)} \quad (2)$$

$$\text{where } x_1 = \text{Endowment_transformed} = \frac{\text{Endowment}^{0.02020202} - 1}{0.02020202} - 1.19425$$

$$x_8 = \text{Tuition_transformed} = \text{Tuition} - 28732.34$$

$$x_9 = \text{Percentage_Instate_transformed} = \text{Percentage_Instate} - 0.5557072$$

At this critical juncture, the fitted model illustrated in Equation 2 will be explicated for the purpose of predicting the response variable, Status. Although it encompasses only three quantitative predictors, the task has evolved in complexity due to the implementation of a Box-Cox transformation and centering applied to one predictor, while the remaining two predictors have undergone centering without scaling. Upon careful examination, it is imperative to ensure that the fitted model remains robustly applicable. With the formula clearly delineated, it becomes feasible to replicate all models presented in this article in a similarly precise manner, if necessary.

Any model can have overfitting issues, meaning the model fits perfectly on the original/training datasets but performs poorly on the new/test datasets. Even this final technically perfect model may have an overfitting issue. It is crucial to recognize that the most interpretable model does not always guarantee the highest predictive accuracy. Striking a balance between interpretability and predictive performance is essential for optimizing the model. Goals should be kept in mind, whether aiming to find a technically perfect model or the best model for predicting. Consequently, a comprehensive comparison between the previously mentioned models and the best model will be conducted, enabling the selection of the model that best suits preferences and requirements.

Table 3. Test Dataset

Name	Status	Type	Endowment	Enrollment	Public/Private	Religious	Tuition	Percentage_Instate
Antioch University	B	U	0.05	4000	Private	No	13704	0.5
Bard College	O	C	0.5	2000	Private	No	63612	0.5
Bates College	O	C	0.42	1790	Private	No	66590	0.4
Colby College	O	C	1.12	2200	Private	No	66600	0.35
College of WilliamsMary	O	C	1.2	8817	Public	No	65000	0.7
Corinthian College	B	C	0.1	72000	Private	No	15000	0.7
Davidson College	O	C	1.32	1927	Private	Yes	60300	0.6
U of Rochester	O	U	2.9	12000	Private	No	64348	0.5
U of Scranton	O	U	0.28	4825	Private	Yes	52309	0.5
U of San Francisco	O	U	0.51	9212	Private	Yes	58222	0.8
St. Gregory University	B	U	0.01	550	Private	Yes	22050	0.6
Mount Ida College	B	C	0.03	1500	Private	No	35720	0.55
U of Richmond	O	U	3.25	4000	Private	No	62600	0.45

Table 3 presents the test dataset utilized for validating the accuracy of the models. It is important to note that these 11 institutions were selected randomly, and the data may contain errors or inaccuracies, similar to the original dataset. This should have been highlighted at the outset. The purpose of this article is strictly academic and does not suggest or accuse any of these institutions of bankruptcy. For clarification, the original dataset comprises 384 operational institutions and 19 bankrupt institutions, while the test dataset contains 9 operational institutions and 4 bankrupt institutions.

3.3 Z-Score

Now defining the Z-score range. In this context, the Z-score range is determined by comparing the responses to the fitted values. 3 zones will be established: Operational (OP), Gray Area (GA), and Bankrupt (BK). The lower bound of GA is set at the lowest fitted value of the bankrupt institutions (Status=1). The upper bound of GA is established at the highest fitted value of the operational institutions (Status=0). Institutions will be classified as OP if their corresponding calculated Z-score is less than the lower bound of GA, classified in GA if their corresponding calculated Z-score falls within the range of the lower and upper bounds of GA, and classified as BK if their corresponding calculated Z-score exceeds the upper bound of GA. For the model accuracy, the counts and percentages will be categorized into true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP represents the correct identification of bankrupt institutions, whereas TN represents the correct identification of operational institutions. FP, also known as Type I errors, occur when operational institutions are incorrectly classified as bankrupt—a critical mistake to be avoided. FN, or Type II errors, occur when bankrupt institutions are incorrectly classified as operational. Although significant, this error is less severe than a Type I error, leading to potential inefficiencies.

Table 4. List of models for comparison

From Figure 7	From Figure 11	From Figure 12	From Figure 13	From Figure 14
OP: $Z < 0.0488$ GA: $0.0488 \leq Z \leq 0.9130$ BR: $Z > 0.9130$	OP: $Z < 0.0081$ GA: $0.0081 \leq Z \leq 0.9100$ BR: $Z > 0.9100$	OP: $Z < 0.0280$ GA: $0.0280 \leq Z \leq 0.8779$ BR: $Z > 0.8779$	OP: $Z < 0.0334$ GA: $0.0334 \leq Z \leq 0.8492$ BR: $Z > 0.8492$	OP: $Z < 0.0037$ GA: $0.0037 \leq Z \leq 0.9203$ BR: $Z > 0.9203$
AIC: 59.754	AIC: 56.616	AIC: 96.674	AIC: 59.723	AIC: 57.856
<u>OD Prediction</u> OP: 359 GA: 42 BR: 2	<u>OD Prediction</u> OP: 338 GA: 62 BR: 3	<u>OD Prediction</u> OP: 299 GA: 104 BR: 0	<u>OD Prediction</u> OP: 356 GA: 37 BR: 10	<u>OD Prediction</u> OP: 355 GA: BR: 6
<u>OD Accuracy</u> TP: 2 (10.53%) TN: 359 (93.49%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 3 (15.79%) TN: 338 (88.02%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 0 (0%) TN: 299 (77.86%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 10 (2.63%) TN: 356 (92.71%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 6 (31.58%) TN: 355 (93.45%) FP: 0 (0%) FN: 0 (0%)
<u>TD Prediction</u> OP: 9 GA: 4 BR: 0	<u>TD Prediction</u> OP: 9 GA: 4 BR: 0	<u>TD Prediction</u> OP: 8 GA: 5 BR: 0	<u>TD Prediction</u> OP: 9 GA: 4 BR: 0	<u>TD Prediction</u> OP: 10 GA: 3 BR: 0
<u>TD Accuracy</u> TP: 0 (0%) TN: 9 (100%) FP: 0 (0%) FN: 0 (0%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 9 (100%) FP: 0 (0%) FN: 0 (0%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 6 (66.67%) FP: 0 (0%) FN: 2 (50%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 8 (88.89%) FP: 0 (0%) FN: 1 (25%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 8 (88.89%) FP: 0 (0%) FN: 2 (50%)
From Figure 15	From Figure 16	From Figure 17	From Equation 2	Placebo
OP: $Z < 0.0334$ GA: $0.0334 \leq Z \leq 0.8492$ BR: $Z > 0.8492$	OP: $Z < 0.0334$ GA: $0.0334 \leq Z \leq 0.8492$ BR: $Z > 0.8492$	OP: $Z < 0.0334$ GA: $0.0334 \leq Z \leq 0.8492$ BR: $Z > 0.8492$	OP: $Z < 0.0588$ GA: $0.0588 \leq Z \leq 0.7416$ BR: $Z > 0.7416$	OP: $Z < 0.0535$ GA: $0.0535 \leq Z \leq 0.6680$ BR: $Z > 0.6680$
AIC: 59.723	AIC: 59.723	AIC: 57.814	AIC: 58.891	AIC: 58.891
<u>OD Prediction</u> OP: 356 GA: 37 BR: 10	<u>OD Prediction</u> OP: GA: 37 BR: 10	<u>OD Prediction</u> OP: 356 GA: 37 BR: 10	<u>OD Prediction</u> OP: 357 GA: 37 BR: 8	<u>OD Prediction</u> OP: 359 GA: 33 BR: 10
<u>OD Accuracy</u> TP: 10 (52.63%) TN: 356 (92.71%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 10 (52.63%) TN: 356 (92.71%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 10 (52.63%) TN: 356 (92.71%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 8 (44.44%) TN: 357 (92.97%) FP: 0 (0%) FN: 0 (0%)	<u>OD Accuracy</u> TP: 10 (55.56%) TN: 359 (93.45%) FP: 0 (0%) FN: 0 (0%)
<u>TD Prediction</u> OP: 11 GA: 2 BR: 0	<u>TD Prediction</u> OP: GA: 2 BR: 0	<u>TD Prediction</u> OP: 11 GA: 2 BR: 0	<u>TD Prediction</u> OP: 11 GA: 2 BR: 0	<u>TD Prediction</u> OP: 10 GA: 3 BR: 0
<u>TD Accuracy</u> TP: 0 (0%) TN: 9 (100%) FP: 0 (0%) FN: 2 (50%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 9 (100%) FP: 0 (0%) FN: 2 (50%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 9 (100%) FP: 0 (0%) FN: 2 (50%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 9 (100%) FP: 0 (0%) FN: 2 (50%)	<u>TD Accuracy</u> TP: 0 (0%) TN: 8 (88.89%) FP: 0 (0%) FN: 4 (100%)

4. Comparisons

Table 4 illustrates a comparative analysis of all the models developed in this article, arranged in ascending order. A total of ten models are included: nine created in this study, along with an additional placebo model. The placebo model is a straightforward GLM utilizing all original predictors, without any refinements such as variable transformations or the application of ridge regression, and serves primarily as a comparison placebo. For each model, the following details are provided: the Z-score ranges, the AIC, the prediction counts, and the model accuracy for both the original and test datasets. The following are key statistics and highlights: For the Z-score range in GA, the narrowest range is 0.6828, from Equation 2, while the widest range is 0.9166, from Figure 14. Regarding the AIC, most differences are marginal. The lowest AIC, 56.616, is from Figure 11, which includes four variables after stepwise regression. The highest AIC, 96.674, is from Figure 12. For OD prediction and accuracy, the lowest TP rate is 0% from Figure 12, whereas the highest TP rate is 5.56% from the placebo model. The lowest TN rate, 77.86%, is from Figure 12, while the highest TN rate, 93.49%, is from Figure 7 and the placebo model. All FP and false negative FN rates are 0% across all models in all figures. For TD prediction and accuracy, the TP rate is consistently 0% across all models. The lowest TN rate is 66.67% from Figure 12, whereas the highest TN rate is 100% from Figures 7, 11, 15, 16, 17, and Equation 2. FP rates for all models are 0%, but the FN rate varies, with the lowest being 0% from models in Figures 7 and 11, and the highest being 100% from the placebo model. When selecting the optimal model, one approach is to choose the most well-formulated model that satisfies all necessary assumptions and requirements and is well-defined. Alternatively, preference might be given to a model that excels in predicting TD. Ideally, a model would satisfy both criteria. The best-formulated model is Equation 2, but it does not offer the desired level of accuracy in TD prediction. Utilizing the counts in the three categories OP, GA, and BK to assess the percentages of TP, TN, FP, and FN. The

optimal model would be the one that achieves the highest percentages in both TP and TN, while simultaneously maintaining the lowest percentages in both FP and FN. Equation 2 reveals the following metrics in TD: OP: 11, GA: 2, BR: 0, TP: 0 (0%), TN: 9 (100%), FP: 0 (0%), and FN: 2 (50%). This model misidentified two bankrupt institutions, classifying them as operational, resulting in a 50% Type II error. While this is not as critical as a Type I error (any percentage in FP), it indicates that this model fails to detect financial distress and classifies the institution as financially stable. Consequently, it is not the best for predictions despite its robustness. Interestingly, there are two models with perfect TD predictions, both having no FN errors: the models from Figures 7 and 11. Remarkably, these two models were formulated from the outset. The model from Figure 7 exhibits the following metrics for TD: OP: 9, GA: 4, BR: 0, TP: 0 (0%), TN: 9 (100%), FP: 0 (0%), and FN: 0 (0%). Similarly, the model from Figure 11 reports OP: 9, GA: 4, BR: 0, TP: 0 (0%), TN: 9 (100%), FP: 0 (0%), and FN: 0 (0%). The model from Figure 7 incorporates all seven original predictors with Box-Cox transformations on both Endowment and Enrollment predictors. In contrast, the model in Figure 11 is derived from the model in Figure 7 after stepwise regression.

A discerning reader will recognize that while these two models have 0 FN, they are still not effective at predicting bankruptcy. This is because both models, as well as all the others, have 0 TP. Consequently, despite having 0 FN, the models' TP rate is 0%, meaning they fail to identify any bankrupt or financially distressed institutions. This indicates that the models do not accurately detect financial distress. However, upon closer examination, considering the prediction counts and FP, both models magically have perfect predictions. Analyzing further: both predict 9 in OP, 4 in GA, and 0 in BK. With 0 TP, 9 TN, 0 FP, and 0 FN, these metrics mean they correctly identify all nine operational institutions and the four bankrupt institutions, respectively, with no Type I and Type II errors. The four identified bankrupt institutions are incorrectly assigned to GA instead of BK, but the two models did correctly identify all four bankrupt institutions, which is precisely the most challenging task. Reassigning all 4 bankrupt institutions from GA to BK is not difficult. Z-score ranges can be redefined by eliminating GA and having only OP and BK. By establishing the original lower bound of GA as the new lower bound of BK, and including the endpoint, the classification process is effectively redefined.

Table 5. New Z-Score Ranges

Model	Previous Z-Score Ranges	New Z-Score Ranges
From Figure 7	OP: $Z < 0.0488$ GA: $0.0488 \leq Z \leq 0.9130$ BR: $Z > 0.9130$	OP: $Z < 0.0488$ BR: $Z \geq 0.0488$
From Figure 11	OP: $Z < 0.0081$ GA: $0.0081 \leq Z \leq 0.9100$ BR: $Z > 0.9100$	OP: $Z < 0.0081$ BR: $Z \geq 0.0081$

4.1 Optimal Z-score Model and Ranges

Table 5 illustrates the redefined Z-score ranges for the models from Figures 7 and 11. With these new Z-score ranges, both models exhibit TP and TN rates of 100%, and FP and FN rates of 0%, making them exemplary for predictions. The critical question now is which model to select between the two. Notably, the number of Fisher Scoring iterations is 21 for both models, indicating they converge at the same steps with similar stability and optimization behavior. The model in Figure 7 has an AIC of 59.754, whereas the model in Figure 11 has an AIC of 56.16. A smaller AIC suggests a superior balance between fit and complexity. The model in Figure 11, with four predictors compared to the seven predictors in the model in Figure 7, demonstrates this superior balance. The model in Figure 11 is the more robust choice in terms of statistical efficiency. As such, it is preferred for parsimonious predictions.

5. Conclusion

$$\hat{Z} = \frac{\exp(-5.482 - 1.479x_1 - 27.42x_4 - 0.000108x_6 + 8.007x_7)}{1 + \exp(-5.482 - 1.479x_1 - 27.42x_4 - 0.000108x_6 + 8.007x_7)} \quad (3)$$

$$\text{where } x_1 = \text{Endowment_transformed} = \frac{\text{Endowment}^{0.02020202} - 1}{0.02020202} - 1.19425$$

$$x_4 = \begin{cases} 0 & \text{if PublicPrivate} = \text{Private} \\ 1 & \text{if PublicPrivate} = \text{Public} \end{cases}$$

$$x_6 = \text{Tuition}$$

$$x_7 = \text{Percentage_Instate}$$

$$\text{Institution} \in \begin{cases} \text{OP} & \text{if } \hat{Z} < 0.0081 \\ \text{BK} & \text{if } \hat{Z} \geq 0.0081 \end{cases}$$

Figure 18. Optimal Z-Score Ranges

Equation 3 illustrates the model identified as the optimal choice for computing the Z-score, and Figure 18 presents the optimal Z-score ranges. When the calculated Z-score is below the threshold of 0.0081, the institution is considered operational. In contrast, when the calculated Z-score is at or above the threshold of 0.0081, the institution is considered at risk of bankruptcy. The equation for the Box-Cox transformation of the predictor Endowment_transformed is provided with the optimal lambda value of 0.02020202. It is important to note that the predictor Endowment is measured in billions of dollars. For the predictor Tuition, the in-state annual tuition (measured in dollars) is used. For the predictor Percentage_Instate, it ranges between the values of 0 and 1, not as a percentage. The coefficients reveal that all values, including the intercept, are negative except for the coefficient of Percentage_Instate. This suggests that public institutions with higher endowment and tuition values will experience a positive effect. Conversely, a higher percentage of in-state students will have an adverse effect, as out-of-state tuition is significantly higher than in-state tuition, which aligns with expectations. As a result, two zones have been established using TP, TN, FP, and FN instead of the initial three zones by verifying the fitted values and the true status. The current two zones provide more accurate predictions, as previously mentioned. This optimal model was developed immediately following stepwise regression. The decision to select this model was based not only on its derivation from stepwise regression but also on its low AIC for simplicity and rapid convergence, and, most importantly, its ability to provide perfect predictions for both OP and BK institutions in the test dataset. While this model was chosen, readers are strongly encouraged to explore other models presented in this article, based on their specific needs and preferences, to perform predictions accordingly. When utilizing alternative models besides the model from Equation 3, it is imperative that any institution's calculated Z-score falling within GA be regarded as questionable, high-risk, or bankrupt institutions to ensure caution.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal and publisher adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589-609. <https://doi.org/10.2307/2978933>
- Altman, E. I. (2018). A fifty-year retrospective on credit risk models, the Altman Z-score family of models and their applications to financial markets and managerial strategies. *Journal of Credit Risk*, 14(4), 1-34. <https://doi.org/10.3390/ijfs6030070>
- Barshay, J. (2023). Proof points: 861 colleges and 9,499 campuses have closed down since 2004. *The Hechinger Report*. Retrieved from <https://cappsonline.org/proof-points-861-colleges-and-9499-campus-have-closed-down-since-2004/>
- Begley, Ming, and Watts (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies*, 1(4), 267-284. <https://doi.org/10.1007/BF00570833>
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26(2), 211–252. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).