Journal of High School Science

Evaluation of the Altman Z-score framework for financial distress prediction in Korean firms

Seung E

Submitted: April 2, 2026, Revised: version 1, May 11, 2026, version 2, May 11, 2026, version 3, June 16, 2026, version 4, July 12, 2026
Accepted: July 17, 2026

Abstract
The Altman Z-score is one of the most widely used models for assessing corporate financial distress, but its coefficients were originally derived from U.S. manufacturing firms and may not generalize to other markets. This study evaluates the performance of the original Altman framework in Korean manufacturing firms using 12,651 Korean firm-year observations (2011–2024) constructed from publicly available DART financial statements and compares it with a Korean-specific logistic regression model that re-estimates the coefficients of the five original Altman variables. Model performance was evaluated using a chronological train-test split (2011–2020 training; 2021–2024 testing), confusion matrices, ROC curves, AUC, and bootstrap confidence intervals. The re-estimated Korean model outperformed the original Altman framework on the held-out test set (AUC 0.947 vs. 0.859). To assess generalizability beyond the accounting-based distress definition used for model estimation, both models were additionally validated using an independent dataset of 94 firms subsequently delisted from KOSPI or KOSDAQ. The Korean logistic model consistently achieved higher AUC across all pre-delisting horizons (pooled AUC 0.849 vs. 0.775), indicating improved discrimination of real-world corporate failure events. The study also releases the datasets and source code to support reproducibility. These findings suggest that market-specific statistical re-estimation of the Altman framework can improve financial distress prediction in Korean firms while retaining the interpretability of the original variables.

Keywords
Altman Z-score, Bankruptcy prediction, Financial distress prediction, Corporate failure prediction, Logistic regression, Financial ratio analysis, Credit risk assessment, Manufacturing firms, Korean manufacturing firms, Predictive modeling

---

Eunsu Seung, Light House Academy, 3621 Garnet St, Unit 6, Torrance, CA 90503, USA. harryseungjr@gmail.com

1. Introduction

1.1 The historical significance of the Altman Z-score

Professor Edward Altman of the NYU Stern School of Business introduced the Altman Z-score model in 1968 as a quantitative framework for predicting corporate bankruptcy. By combining five financial ratios through multivariate discriminant analysis, the model enabled corporate financial health to be summarized as a single numerical score that provides an early indication of financial distress. Its strong predictive performance and interpretability established it as one of the most influential models in financial risk assessment, and it continues to be widely used by researchers, financial institutions, and practitioners more than five decades after its introduction.

1.2 The Altman Z-score formula

The Altman Z-score model is based on five financial ratios that capture complementary aspects of a firm's financial condition, including liquidity, accumulated profitability, operating performance, solvency, and asset utilization. The first variable, Working Capital/Total Assets (X_1), measures short-term liquidity by assessing the firm's ability to meet its current obligations. The second, Retained Earnings/Total Assets (X_2), reflects cumulative profitability and the extent to which the firm's assets have been financed through retained earnings rather than external debt. The third, Earnings Before Interest and Taxes (EBIT)/Total Assets (X_3), measures operating profitability by evaluating how efficiently the firm's assets generate earnings independent of its financing structure. The fourth, Market Value of Equity/Total Liabilities (X_4), assesses solvency by comparing the market value of shareholders' equity with the firm's total liabilities, thereby reflecting the firm's capacity to absorb financial losses. Finally, Sales/Total Assets (X_5) measures asset turnover, indicating how efficiently the firm's assets are utilized to generate revenue.

The Altman Z-Score model is evaluated by assigning statistical weights to the five significant financial ratios above and by summing them. The formula is represented by eq1. Based on the calculated Z-Score, a firm's financial condition is classified into three zones as shown in Table 1.

$$Z = 1.2\,X_1 + 1.4\,X_2 + 3.3\,X_3 + 0.6\,X_4 + 1.0\,X_5 \qquad \text{(eq1)}$$

Table 1. A firm’s financial condition correlated with the calculated Z-score

Zones	Score Range	Descriptions
Safe Zone	$Z > 2.99$	The company exhibits strong financial health with minimal bankruptcy risk
Grey Zone	$1.81 \le Z \le 2.99$	The company is in a borderline state with uncertain financial stability
Distress Zone	$Z < 1.81$	The company faces a high risk of bankruptcy

This study uses these five variables as the foundation for developing a Korean financial distress prediction framework by examining whether structural characteristics of Korean manufacturing firms—particularly their relatively high leverage and lower asset turnover resulting from capital-intensive operations—reduce the effectiveness of the original Altman Z-score model and therefore warrant market-specific coefficient re-estimation.

1.3 Large-scale empirical validation of the Altman model in the U.S. market

To establish a benchmark for comparison, this study evaluated the original Altman Z-score using a large U.S. manufacturing firm-year dataset constructed directly from SEC EDGAR financial statements through an automated Python-based extraction pipeline. The dataset comprised 4,575 firm-year observations spanning 2009–2024, including 117 distress-related observations identified independently of the Altman score using accounting-based indicators of severe financial weakness.

Figures 1 and 2 show that the original Altman framework clearly distinguishes distressed from non-distressed U.S. firms. Most observations fall within the traditional safe range, whereas distress-related observations exhibit substantially lower Z-scores and a marked downward shift in their distribution. Because the distress labels were defined independently of the Altman score, this comparison avoids circular evaluation and provides an empirical benchmark demonstrating that the original model retains strong discriminatory capability within the U.S. market.

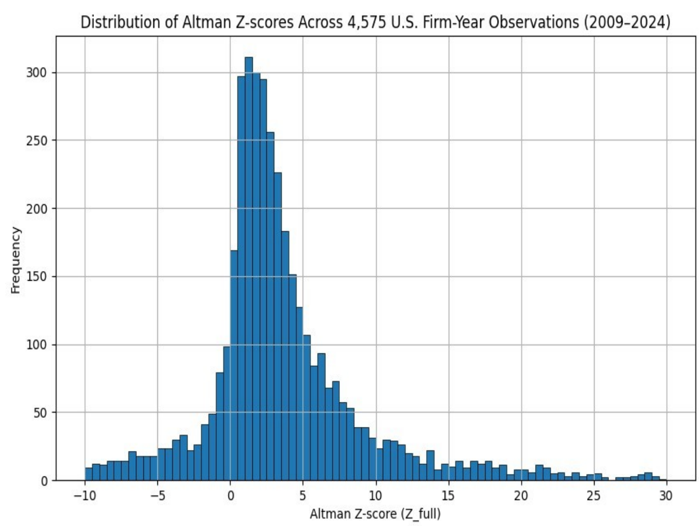

Figure 1. Distribution of Altman Z-scores Across 4,575 U.S. Firm-Year Observations (2009–2024). Histogram showing the overall distribution of Altman Z-scores computed from SEC EDGAR firm-year observations between 2009 and 2024.

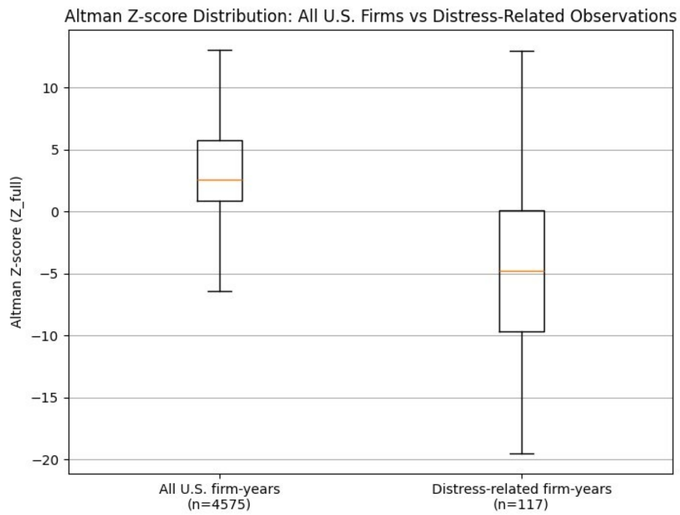

Figure 2. Comparison of Altman Z-score Distributions Between All U.S. Firms and Distress-Related Observations. Boxplot comparison between the full U.S. firm-year sample and distress-related firm-year observations identified through severe accounting based financial distress criteria.

1.4 Empirical evaluation of the Altman Z-score in Korean firms

Having established that the original Altman Z-score retains strong discriminatory capability within the U.S. market, this study next examines whether that performance generalizes to Korean manufacturing firms. Unlike the previous section, which evaluated the model using the market for which it was originally developed, the present analysis investigates the transferability of the Altman framework to a substantially different financial and industrial environment.

To enable a direct comparison with the U.S. analysis, a large-scale Korean firm-year dataset was constructed using publicly available financial statements from the Financial Supervisory Service's DART disclosure system. Using the same automated Python-based preprocessing pipeline, the five Altman financial ratios (X_1–X_5) and corresponding Z-scores were calculated for 12,651 Korean firm-year observations spanning 2011–2024. The identical analytical procedure used for the U.S. dataset was then applied to the Korean firms, allowing any differences in performance to be attributed primarily to differences in market characteristics rather than changes in methodology.

Figure 3 presents the overall distribution of Altman Z-scores across the Korean dataset. Similar to the U.S. distribution shown previously, the majority of Korean firms are concentrated within positive Z-score ranges, while a right-skewed tail captures firms with exceptionally strong financial indicators.

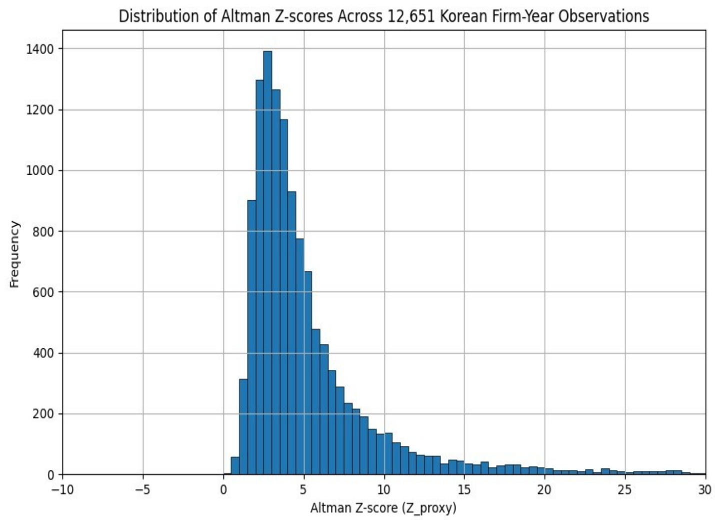

Figure 3. Distribution of Altman Z-scores Across 12,651 Korean Firm-Year Observations (2011–2024). Histogram showing the distribution of Altman Z-scores computed from Korean firm-year observations obtained through the DART disclosure system.

To further evaluate the discriminatory capability of the Altman framework in Korea, this study separately identified financially distressed observations using accounting-based distress criteria. Figure 4 compares the Z-score distribution between the overall Korean sample and the distress-related observations.

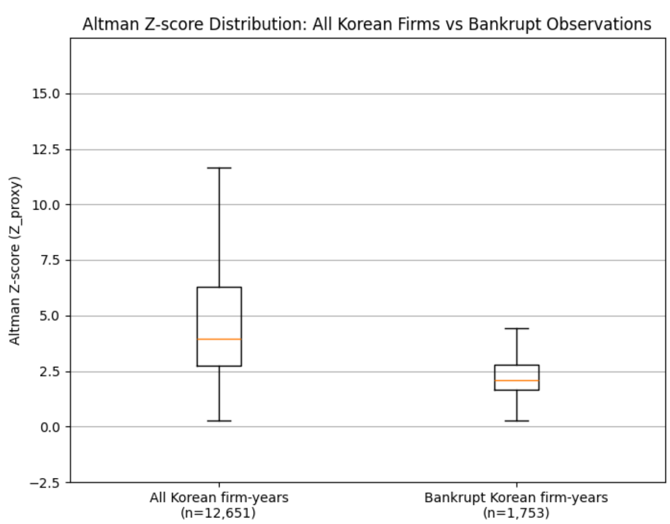

Figure 4. Comparison of Altman Z-score Distributions Between All Korean Firms and Distress-Related Observations. Boxplot comparison between the full Korean firm-year sample and distress-related observations identified through accounting-based financial distress screening.

Unlike the U.S. results presented in Section 1.3, the Korean distress-related observations exhibit considerably less separation from the overall market distribution. Although the median Z-score of the distressed sample remains lower than that of the full Korean sample, the overlap between the two distributions is substantially greater than that observed for U.S. firms. This finding suggests that the original Altman Z-score discriminates less effectively between distressed and non-distressed firms in the Korean market.

One possible explanation is that Korean manufacturing firms differ structurally from the historical U.S. firms on which the original Altman model was developed. In particular, Korean firms often operate with relatively high leverage while maintaining substantial tangible assets and making long-term capital investments. Consequently, variables such as X_4 (Market Value of Equity/Total Liabilities) and X_5 (Sales/Total Assets) may contribute differently to financial distress prediction than assumed by the original U.S.-derived coefficients.

These results do not imply that the original Altman model is ineffective. Rather, they suggest that coefficients estimated from historical U.S. manufacturing firms may not fully capture the financial characteristics of Korean firms. This observation provides the motivation for the market-specific coefficient re-estimation presented in the following section.

2. Methods

2.1 Construction of the Korean firm-year dataset

This study used publicly available corporate financial statements obtained from the Financial Supervisory Service's DART disclosure system (https://dart.fss.or.kr). The dataset comprises Korean manufacturing firm-year observations spanning 2011–2024.

An automated Python-based preprocessing pipeline was developed to extract, clean, and organize the financial information required for analysis. For each firm-year observation, the five financial ratios comprising the original Altman Z-score (X_1–X_5), as defined in Section 1.2, were computed directly from balance sheets, income statements, and market-related financial data.

Observations with missing values for any of the required variables were excluded. After data filtering and quality control, the final dataset consisted of approximately 12,651 Korean firm-year observations.

2.2 Definition of financial distress

For model development, financially distressed observations were identified using accounting-based indicators rather than formal bankruptcy filings, which are relatively infrequent and difficult to collect consistently over long time periods. A firm-year observation was classified as financially distressed if it satisfied at least two of the following four criteria: (i) negative net income, (ii) negative return on assets, (iii) a leverage ratio above the sample median, and (iv) a current ratio below one. The leverage threshold was defined relative to the overall sample distribution rather than by a fixed cutoff. Observations meeting at least two criteria were assigned a distress label of 1, while all remaining observations were assigned 0. Because these criteria are independent of the Altman Z-score, this approach avoids defining the outcome using the model being evaluated and thereby reduces circularity in the analysis. Applying these criteria identified approximately 1,755 financially distressed observations in the Korean dataset.

2.3 Logistic Regression framework

To adapt the Altman framework to Korean firms, this study re-estimated the relationship between the five financial ratios (X_1–X_5) and financial distress using logistic regression. Whereas the original Altman Z-score employs coefficients derived from historical U.S. manufacturing firms, the logistic model estimates market-specific coefficients directly from the Korean training dataset. The logistic regression model is defined as,

$$P(Bankruptcy) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5)}} \qquad (eq\,2)$$

where P(Bankruptcy) represents the estimated probability of financial distress and X_1–X_5 correspond to the five Altman financial variables. Model coefficients were estimated from the training data using maximum likelihood estimation (MLE), which determines the coefficient values that maximize the likelihood of the observed distress outcomes under the logistic regression model. In practice, this procedure identifies the set of coefficients that best fits the relationship between the financial ratios and the observed distress status in the training data. By retaining the original Altman variables while re-estimating their coefficients, the model preserves the interpretability of the underlying financial ratios while adapting their relative importance to the characteristics of Korean firms.

2.4 Train/test split and Out-of-Sample validation

To evaluate the generalizability of the re-estimated logistic model, a chronological train/test split was used rather than random sampling. Firm-year observations from 2011–2020 were used to estimate the logistic regression coefficients, while observations from 2021–2024 were reserved as an independent out-of-sample test set.

This temporal validation strategy evaluates model performance on future observations that were unavailable during training, providing a more realistic assessment of predictive performance and reducing the risk of overfitting. Applying the same train/test partition to the original Altman Z-score model ensured a fair comparison between the original and re-estimated frameworks.

2.5 Performance evaluation metrics

Model performance was evaluated using Accuracy, Precision, Recall, and F1-score to compare the predictive performance of the re-estimated logistic regression model with that of the original Altman Z-score. Together, these metrics assess overall classification performance, the reliability of distress predictions, the ability to identify financially distressed firms, and the balance between precision and recall.

To evaluate discrimination across all possible classification thresholds, Receiver Operating Characteristic (ROC) curves and the corresponding Area Under the Curve (AUC) were computed for both models. Because AUC is independent of any single decision threshold, it provides a robust measure of overall discriminative performance.

Confusion matrices were also examined to characterize the distribution of true positives, true negatives, false positives, and false negatives. To quantify uncertainty in model performance, 95% confidence intervals for the AUC values were estimated using 1,000 bootstrap resamples.

2.6 External validation using delisted firms

To evaluate model performance using an independent failure outcome, this study identified Korean firms formally delisted from the KOSPI or KOSDAQ exchanges between 2016 and 2024 using publicly available Korea Exchange (KRX) disclosure records. Because KRX does not provide a dedicated historical delisting API, the delisting records were compiled manually from publicly accessible disclosures. A total of 94 delisted firms were identified. Since delisting is determined according to KRX listing and regulatory requirements rather than the accounting-based distress criteria used to train the models, it provides an independent benchmark that avoids circularity. None of the delisted firms overlapped with the 12,651 firms used for model development, ensuring complete independence of the external validation dataset.

For each delisted firm, financial statement data were obtained from the DART disclosure system using the preprocessing pipeline described previously. Financial information was collected for each of the five fiscal years preceding delisting (t−1 through t−5, where t−1 denotes the fiscal year immediately before delisting). Because complete DART records are consistently available only from 2015 onward, fewer firms contributed observations at longer pre-delisting horizons. In total, approximately 400 positive (delisted) firm-year observations were available across the five prediction horizons.

The comparison group consisted of non-distressed observations drawn exclusively from the 2021–2024 out-of-sample test period. After excluding 527 observations classified as financially distressed under the accounting-based criteria described in Section 3.2, the final comparison pool contained 3,269 non-distressed firm-year observations. This comparison set was held constant across all prediction horizons.

At each pre-delisting horizon, predictive performance of both the original Altman Z-score and the re-estimated logistic model was evaluated by comparing the delisted observations with the fixed non-distressed comparison pool. ROC curves, AUC, confusion matrices, accuracy, precision, recall, and F1-score were computed using the same evaluation procedures described in Section 3.5. To quantify statistical uncertainty, 95% confidence intervals for the AUC values and the difference in AUC between models were estimated using 1,000 bootstrap resamples at the firm-year observation level.

2.7 Reproducibility and open-source release

To promote transparency and reproducibility, the large-scale U.S. and Korean financial datasets and the Python source code used for data preprocessing, Altman Z-score calculation, financial distress classification, logistic regression modeling, and visualization have been made publicly available. The repository is available at https://github.com/owenseung/K-Altman-Score

3. Results

3.1 Logistic Regression re-estimation based on Korean firms

The results presented in Section 1 indicate that the original Altman Z-Score model does not adequately predict financial distress when applied directly to Korean firms. One possible explanation is that the original model coefficients were estimated using historical financial data from U.S. companies and therefore reflect the capital structure, accounting practices, and financial characteristics of U.S. firms rather than those of Korean firms. Motivated by this observation, this study re-estimates the Altman coefficients using a large-scale dataset of Korean firms.

A total of 12,651 Korean firm-year observations obtained from the DART disclosure system between 2011 and 2024 were used to estimate a logistic regression model that predicts financial distress directly from the five Altman financial ratios (X_1–X_5).

To evaluate predictive performance on previously unseen data, the dataset was divided chronologically into training and testing periods. Firm-year observations from 2011 to 2020 were used to estimate the logistic regression coefficients, while observations from 2021 to 2024 were reserved as an out-of-sample test set. This temporal separation avoids using future information during model estimation and provides a more rigorous evaluation of whether the relationships learned from the 2011–2020 data generalize to subsequent years. The resulting estimated coefficients are shown in Table 2.

Table 2. Estimated coefficients using a logistic regression model on the Korean firms DART disclosure system

Variable	X_1	X_2	X_3	X_4	X_5
Estimated Coefficient	11.0396	4.5027	1.5118	0.0526	0.0455

The re-estimated coefficients indicate that the liquidity ratio (X_1) and retained earnings ratio (X_2) have substantially stronger associations with financial distress among Korean firms than in the original Altman model. In contrast, the coefficients for the market value of equity to total liabilities ratio (X_4) and the sales to total assets ratio (X_5) are considerably smaller, suggesting that these variables contribute less to distress prediction in the Korean context.

These findings are broadly consistent with the structural characteristics of Korean manufacturing firms discussed in the previous section. In particular, Korean firms often maintain relatively high leverage ratios and substantial tangible asset bases even when they are financially healthy, reducing the discriminatory power of leverage- and asset-related variables. Conversely, measures of liquidity and accumulated profitability appear to provide stronger signals of financial distress.

Finally, the model maintained its predictive performance on the independent 2021–2024 out-of-sample test set, indicating that the re-estimated coefficients generalize beyond the estimation period and capture stable relationships between financial ratios and financial distress in Korean firms.

3.2 Out-of-Sample performance evaluation

Both models were evaluated to determine whether the statistically re-estimated Korean logistic framework improved predictive performance compared to the original Altman Z-score model based on Korean firms. The two models were assessed by the identical 2021-2024 Korean out-of-sample test data set. The evaluation was based on predictive performance for previously unseen Korean firm-year observations. In the course of the evaluation, various classification performance metrics were calculated, such as accuracy, precision, recall, F1-score, confusion matrix analysis, Receiver Operating Characteristic (ROC) curves, and Area Under the Curve(AUC) (Table 3).

Table 3. Out-of-Sample metrics for Korean firms evaluated using the two models

Model	Accuracy	Precision	Recall	F1-Score	AUC
Original Altman	0.760	0.380	0.620	0.470	0.859
Korean Logistic	0.833	0.452	0.953	0.613	0.947

The Korean logistic regression model outperformed the original Altman Z-score model across all classification metrics on the out-of-sample test set. Most notably, recall increased from 0.620 to 0.953, indicating that the re-estimated model identified substantially more financially distressed Korean firms while missing far fewer distress cases than the original Altman model. Precision also improved from 0.380 to 0.452, indicating that a greater proportion of firms classified as distressed were correctly identified (Table 4).

Table 4. Confusion matrix from the Korean logistic regression model

	Predicted Non-Distress	Predicted Distress
Actual Non-Distress	2661	608
Actual Distress	25	502

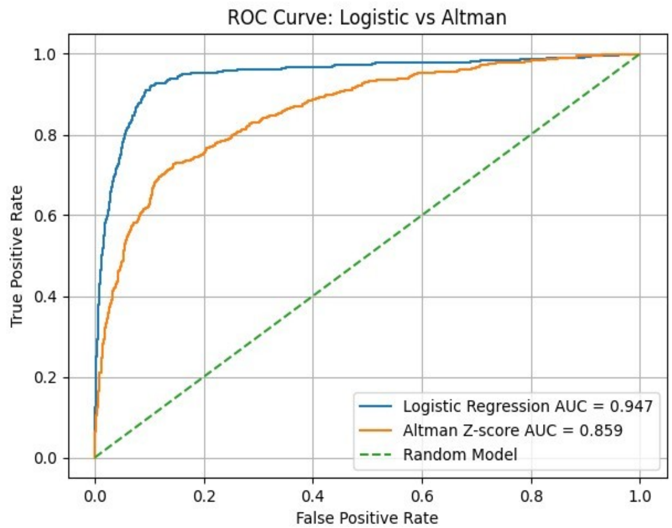

Figure 5. ROC Curve Comparison Between Logistic Regression and the Original Altman Model. ROC curve comparison between the logistic regression framework trained on Korean firm-year observations and the original Altman Z-score model.

The ROC analysis also supports the improved discriminatory capability of the re-estimated Korean framework. The Korean logistic regression model achieved an Area Under the Curve (AUC) of 0.947 (95% CI: 0.935–0.960), compared with 0.859 (95% CI: 0.841–0.877) for the original Altman model. Importantly, the non-overlapping 95% confidence intervals between the two frameworks indicate that the observed improvement in discriminatory performance is statistically robust and cannot be attributed solely to sampling variability. In most classification thresholds, the ROC curve of the Korean logistic regression framework stayed above that of the original Altman model, suggesting that the Korean framework more effectively distinguishes financially distressed Korean firm-year observations.

Overall, the results indicate that statistically re-estimating the relationship between the Altman financial ratios and financial distress conditions using Korean firm-year data substantially improved predictive performance relative to directly applying the original U.S.-based Altman coefficients to Korean firms.

3.3 Distributional changes under the re-estimated Korean logistic score

Using coefficients estimated from the Korean firm-year logistic regression model, this study developed a re-estimated Korean financial distress score (K-score). The original Altman Z-score model was developed from historical U.S. corporate data, whereas the K-score was built using Korean financial observations collected between 2011 and 2024.

Figure 6 compares the distribution of the original Altman Z-score and the re-estimated Korean logistic K-score across 12,651 Korean firm-year observations. Both distributions are generally right-skewed, but the Korean logistic K-score spreads the firms more widely across the financial risk range, creating a stronger separation between financially stable firms and financially distressed firms.

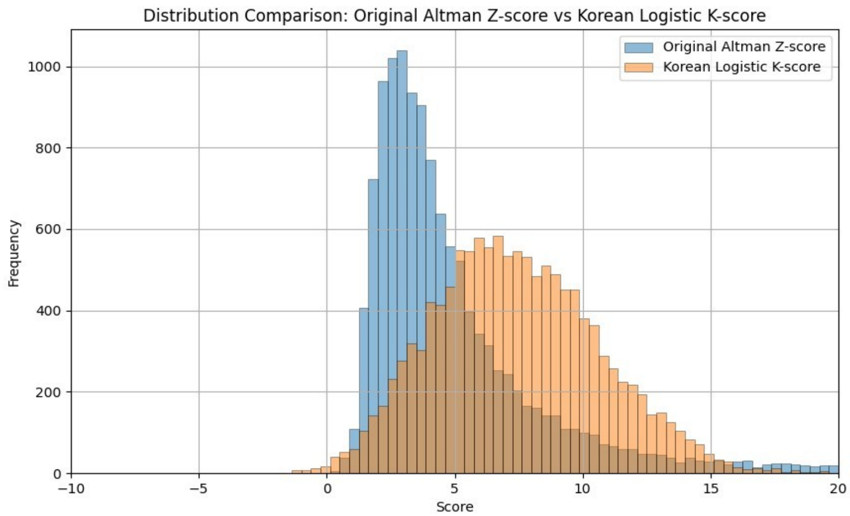

Figure 6. Distribution Comparison Between the Original Altman Z-score and the Re-estimated Korean Logistic K-score. Histogram comparing the distributions of the original Altman Z-score and the re-estimated Korean logistic K-score across 12,651 Korean firm-year observations.

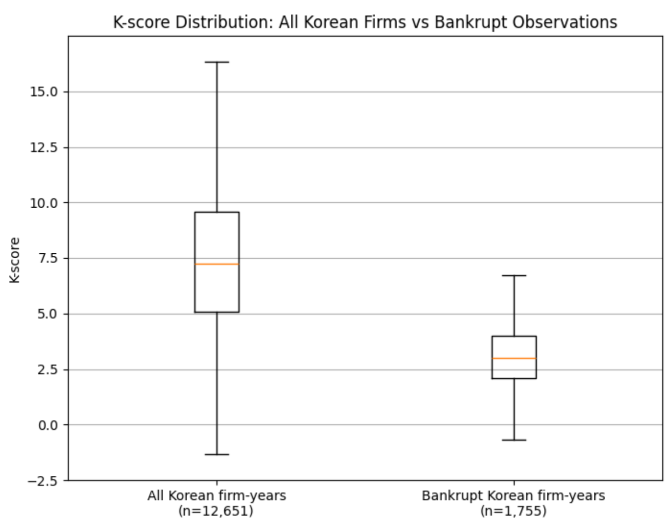

Figure 7. Distribution of the Re-estimated Korean Logistic K-score Between All Korean Firms and Financially Distressed Observations. Boxplot comparison of the re-estimated Korean logistic K-score between all Korean firm-year observations and financially distressed Korean observations.

Specifically, the re-estimated Korean K-score shifts financially stable firms toward higher score ranges while keeping financially distressed firms concentrated in lower score ranges. As a result, the separation between the two groups is more pronounced than under the original Altman framework. Figure 7 further demonstrates this improved discrimination by comparing the distribution of the re-estimated Korean K-score between the overall Korean firm-year sample and the financially distressed observations.

Compared with the weaker separation observed under the original Altman framework (Figure 4), the re-estimated Korean logistic K-score demonstrates substantially better discrimination between financially distressed and non-distressed firms. The median K-score of the financially distressed observations remains considerably lower than that of the overall Korean sample, and the reduced overlap between the interquartile ranges further indicates improved separation between the two groups.

These findings are consistent with the ROC and AUC results presented earlier, which showed that the Korean logistic regression model outperformed the original Altman model in distinguishing financially distressed firms from non-distressed firms. Collectively, the results suggest that estimating the model coefficients using Korean firm-year data more accurately captures the financial characteristics and distress patterns of Korean firms than directly applying the original U.S.-derived Altman coefficients.

Rather than relying on manually adjusted coefficients, the proposed framework estimates the variable weights directly from large-scale Korean financial data using logistic regression. Consequently, the re-estimated Korean logistic K-score provides a more empirically grounded, reproducible, and market-specific framework for evaluating financial distress in Korean firms.

3.4 External validation on KRX delisted firms

To assess whether the improved discrimination observed during internal validation translated into real-world prediction of corporate failure, both models were evaluated using an external dataset of firms that were subsequently delisted from the Korea Exchange (KRX). A total of 94 firms delisted from KOSPI or KOSDAQ between 2016 and 2024 were identified from publicly available KRX records. Because DART financial statements are consistently available beginning in 2015, financial data from the five fiscal years preceding delisting (t−1 through t−5) were collected whenever available, resulting in approximately 400 firm-year observations. As firms delisted earlier in the study period (e.g., 2016–2020) did not always have five complete years of pre-delisting financial data, the composition of firms differed slightly across the five prediction horizons, with more recent delisting cohorts contributing more observations at t−4 and t−5.

As a benchmark (non-distress) comparison group, we used the held-out 30% test split from the original 12,651-firm dataset (n = 3,796). Of these, 527 observations previously classified as financially distressed according to the study's risk classification criteria were excluded, leaving 3,269 firm-year observations classified as non-distressed. Because these observations were drawn exclusively from the held-out test set, they were not used during estimation of the logistic regression model and therefore remained out-of-sample with respect to model training. The same non-distressed comparison pool was used for each prediction horizon (t−1 through t−5), allowing consistent comparison across time horizons.

This validation design should nevertheless be interpreted with an important limitation. Although the delisted firms constitute an independent external failure cohort, the non-distressed comparison group was drawn from the original study population rather than from an entirely independent external dataset. Furthermore, exclusion of firms previously classified as financially distressed produced a benchmark consisting of relatively healthy firms, which may enhance the apparent discrimination between the failure and non-failure groups. Consequently, this analysis represents a partially external validation rather than a fully independent external validation, and the reported performance should be interpreted accordingly.

Table 5. Comparison of classification performance at five pre-delisting horizons (t−1 through t−5) for the Altman Z-score and the logistic regression model, based on 94 KRX-delisted firms evaluated against the held-out non-distress comparison pool

	Altman AUC (95% CI)	Logistic AUC (95% CI)	Diff (Altman – Logistic)
t-1	0.781 (0.716–0.842)	0.854 (0.798–0.903)	−0.072 (−0.125, −0.022)
t-2	0.806 (0.745–0.863)	0.877 (0.829–0.921)	−0.071 (−0.104, −0.042)
t-3	0.783 (0.717–0.846)	0.864 (0.808–0.914)	−0.080 (−0.117, −0.046)
t-4	0.744 (0.673–0.811)	0.819 (0.758–0.876)	−0.075 (−0.116, −0.035)
t-5	0.743 (0.663–0.818)	0.816 (0.744–0.882)	−0.073 (−0.118, −0.029)
pooled	0.775 (0.744–0.805)	0.849 (0.824–0.874)	−0.074 (−0.093, −0.056)

For each prediction horizon, receiver operating characteristic (ROC) curves were constructed and the area under the ROC curve (AUC) was calculated for both the original Altman Z-score and the Korean logistic regression model. Ninety-five percent confidence intervals were estimated using 1,000 bootstrap resamples at the individual firm-year observation level. As shown in Figure 8 and Table 5, the logistic regression model AUC consistently outperformed the original Altman Z-score across all prediction horizons, as well as in the pooled analysis. Moreover, the 95% confidence intervals for the AUC differences excluded zero at every horizon (pooled difference: −0.074; 95% CI: −0.093 to −0.056), demonstrating that the superior discrimination of the Korean logistic regression model was statistically significant and consistently maintained across all pre-delisting horizons.

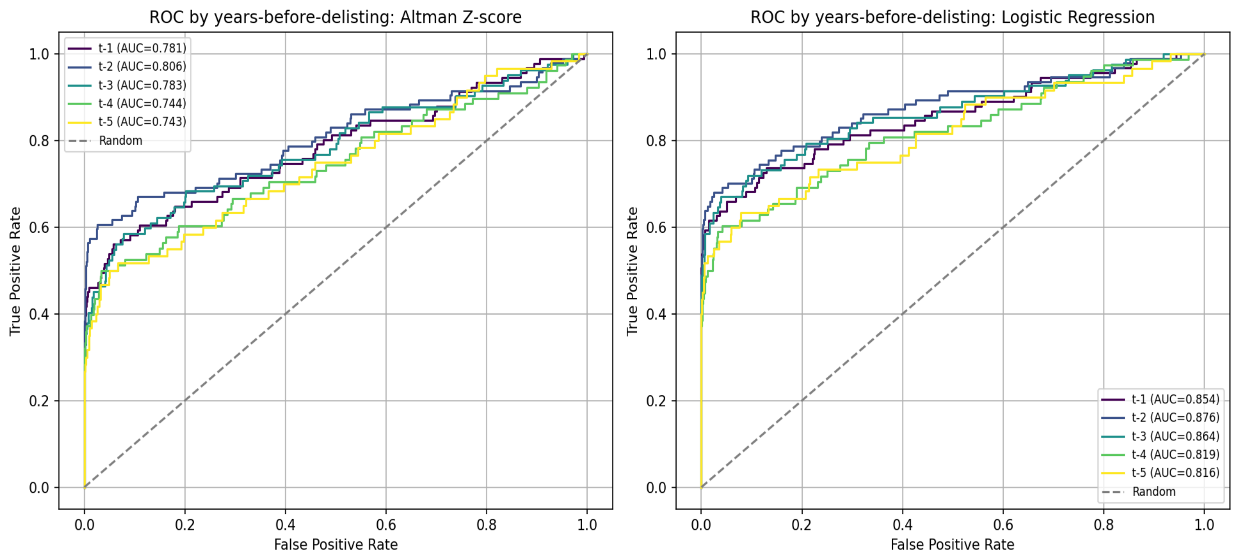

Figure 8. ROC Curves by Years Before Delisting for the Altman Z-score and Logistic Regression Model. Comparison of classification performance at five pre-delisting horizons (t−1 through t−5) for the Altman Z-score (left) and the logistic regression model (right), based on 94 KRX-delisted firms evaluated against the held-out non-distress comparison pool. AUC values for each horizon are shown in the legend.

To complement the threshold-independent comparison based on the area under the receiver operating characteristic curve (AUC), classification performance was also evaluated using pre-specified decision thresholds: $Z < 1.81$ for the original Altman model and a predicted probability threshold of 0.5 for the logistic regression model. Evaluation was performed on the pooled external validation sample (n = 3,674; 405 distressed and 3,269 non-distressed firm-year observations). As shown in Tables 6–8, the two models exhibited distinct classification characteristics. The Altman model identified a greater proportion of distressed firms (recall = 0.630) but also generated substantially more false-positive classifications (608 of 3,269 non-distressed observations; false-positive rate = 18.6%), resulting in relatively low precision (0.296). In contrast, the logistic regression model markedly reduced false-positive classifications (65 of 3,269; false-positive rate = 2.0%), improving precision (0.786) and overall accuracy (0.937). This improvement was accompanied by a modest reduction in recall (0.590), indicating that more distressed firms were classified as non-distressed than under the Altman model. Consequently, the logistic model achieved a substantially higher F1 score (0.674 vs. 0.402), reflecting a more balanced trade-off between precision and recall.

Because both decision thresholds were specified *a priori*—the conventional Altman cutoff of 1.81 and the standard logistic probability threshold of 0.5—rather than optimized using the external validation data, the comparison is not affected by post hoc threshold selection. Combined with the consistently higher AUC values reported earlier, these findings indicate that the Korean logistic regression model provides superior overall discrimination while substantially reducing false-positive classifications. The preferred model may nevertheless depend on the intended application: contexts in which missing a distressed firm carries a high cost may favor a higher-recall decision threshold, whereas applications emphasizing fewer false alarms may benefit from the higher precision achieved by the logistic model.

Table 6. Confusion matrix for the original Altman Z-score model (decision threshold: $Z < 1.81$) on the pooled external validation dataset

	Predicted Non-Distress	Predicted Distress
Actual Non-Distress	2661	608
Actual Distress	150	255

Table 7. Confusion matrix for the Korean logistic regression model (decision threshold: predicted probability ≥ 0.50) on the pooled external validation dataset

	Predicted Non-Distress	Predicted Distress
Actual Non-Distress	3204	65
Actual Distress	166	239

Table 8. Performance comparison of the Altman Z-score and Korean logistic regression models on the pooled external validation dataset

Model	Accuracy	Precision	Recall	F1	AUC
Altman	0.794	0.296	0.630	0.402	0.775
Logistic	0.937	0.786	0.590	0.674	0.849

Several limitations of this external validation should be noted. First, the number of unique delisted firms (94) is modest, and the resulting confidence intervals widen at longer horizons ($t-4$, $t-5$), reflecting reduced statistical power. Second, because the non-distress comparison pool was reused across all five horizons, the horizon-specific results are correlated rather than fully independent, and the pooled estimate should be regarded as the principal basis for inference. Third, the changing cohort composition across horizons — driven by the requirement of five years of pre-delisting data — means that comparisons across horizons partly reflect differences in which delisting cohorts are represented, rather than purely temporal effects. Fourth, bootstrap resampling was performed at the individual firm-year observation level rather than clustered by firm; because multiple fiscal years from the same firm are not statistically independent, this approach may understate the true sampling variability, and the reported confidence intervals should be interpreted as a lower bound on the actual uncertainty. Fifth, because the non-distress comparison pool was drawn from the 2021–2024 out-of-sample test period, while positive (delisted) observations at longer horizons (t−4, t−5) draw predominantly on earlier fiscal years (approximately 2016–2019, given the requirement of five years of pre-delisting data), the horizon-specific comparisons may partly reflect period-specific differences in the macroeconomic or financial reporting environment rather than purely the temporal proximity to failure. This concern applies less to the pooled estimate and to the shorter horizons (t−1, t−2), which overlap more closely with the negative pool's observation period.

4. Discussion

The empirical results of this study suggest that the original Altman Z-score framework exhibits substantially different discriminative behavior when applied to Korean firms compared to the U.S. financial environment in which the model was originally developed. While the original Altman model continues to separate financially distressed firms effectively within the U.S. dataset, the Korean firm-year analysis revealed noticeably weaker separation between financially distressed and non-distressed observations.

One possible explanation is that the statistical relationships embedded in the original Altman coefficients primarily reflect the historical capital structures and industrial characteristics of U.S. manufacturing firms during the period in which the model was originally developed. In contrast, Korean manufacturing firms frequently operate under relatively high leverage structures while simultaneously maintaining substantial tangible asset bases and long-term industrial investment strategies. These structural differences appear to be reflected in the re-estimated coefficient structure obtained from the Korean dataset. In particular, the lower importance of X_4 and X_5 suggests that market-value-based leverage and asset turnover may have weaker discriminatory power in the Korean manufacturing environment than in the original U.S.-based framework. This occurs since many Korean firms remain operationally stable while maintaining relatively high leverage ratios and large asset bases. At the same time, the increased importance of X_1 and X_2 suggests that liquidity conditions and accumulated retained earnings may provide stronger signals of financial distress within Korean firms. This implies that the short-term liquidity pressure and declining internal reserves may become visible before significant market-based deterioration occurs.

The logistic regression re-estimation performed in this study supports this interpretation. After re-estimating the relationship between the five Altman variables and financial distress using 12,651 Korean firm-year observations, the resulting Korean logistic framework showed substantially stronger predictive performance than the direct application of the original Altman coefficients. In particular, the ROC analysis showed that the Korean logistic framework achieved an AUC of approximately 0.947, compared to 0.859 for the original Altman model.

The confusion matrix results exemplify the predictive characteristic of the Korean logistic framework. Among non-distressed observations, 2,661 were correctly classified, while 608 were incorrectly classified as distressed. First and foremost, the model correctly identified 502 out of 527 observations, missing only 25 distress cases among the actual financially distressed observations. This implies a recall value of approximately 0.953, indicating that the model successfully detects the vast majority of financially distressed firms within the Korean dataset.

From a practical financial risk-management perspective, this result is important because missing financially distressed firms may create larger economic losses than incorrectly classifying healthy firms as distressed. In this sense, the model places greater emphasis on detecting distress, even though some false positive classifications still occur.

At the same time, several limitations remain. First, the study primarily focuses on manufacturing related financial structures and therefore may not generalize equally well to industries where intangible assets and platform-based business models dominate, such as biotechnology, software, and internet services. Second, because the framework relies partially on market-value-related variables, the approach is more naturally suited for publicly listed firms than for private companies or small unlisted businesses.

Third, there is an inherent potential for overlap between the predictor variables and the criteria used to define the financial distress label. The proposed logistic framework utilizes financial ratios that share common underlying accounting characteristics with the target distress definition. Although the distress definition was intentionally separated from the Altman Z-score itself to mitigate direct circularity, a portion of the exceptionally strong in-sample predictive performance—such as the AUC of 0.947—may still reflect the model's proficiency in identifying internally defined accounting characteristics rather than predicting an entirely independent financial outcome. To address this concern directly, Section 3.4 evaluated both models against an external, independently determined failure event—formal delisting from KOSPI or KOSDAQ—using firms and observation periods not used in model estimation. The logistic model retained a statistically significant advantage over the Altman Z-score in this external setting, suggesting that its superior performance is not solely an artifact of shared accounting structure with the distress label used for estimation.

Fourth, because the accounting-based distress definition used to estimate and initially evaluate the models does not, by itself, confirm actual corporate failure, the corresponding classification metrics should be interpreted primarily as measures of financial distress identification. This concern is directly addressed by the external validation presented in Section 3.4, in which both models were evaluated against KRX delisting—a regulatory outcome independent of the accounting ratios used in either model. While this validation strengthens confidence that the logistic framework's advantage extends to real-world failure events, the number of delisted firms [at 94] remains modest relative to the full Korean dataset, and further validation using a larger sample of confirmed bankruptcy filings or legal restructuring cases would help establish the robustness of this finding across a broader range of failure definitions.

Notably, the AUC values obtained in the external validation (0.849 for the logistic model and 0.775 for the Altman model, pooled across horizons) were lower than the corresponding in-sample test-set values reported earlier (0.947 and 0.859, respectively). This decline is expected rather than concerning: predicting an externally determined, regulator-confirmed failure event such as delisting is inherently more difficult than reproducing an accounting-based distress label that shares statistical structure with the predictor variables themselves. The fact that the logistic model's advantage over the Altman Z-score persisted—and remained statistically significant—under this more demanding external test provides stronger evidence for the practical value of the re-estimated framework than the in-sample results alone.

5. Conclusion

This study re-estimated the original Altman Z-score using logistic regression and a large sample of Korean firm-year financial data to develop a Korean-specific financial distress prediction model. The results demonstrate that statistically estimating the model coefficients from Korean firms substantially improves discrimination between financially distressed and non-distressed firms compared with directly applying the original Altman coefficients derived from U.S. firms. During internal validation, the Korean logistic regression model consistently outperformed the original Altman model across all major classification metrics, including accuracy, precision, recall, F1 score, and AUC.

External validation using firms subsequently delisted from the Korea Exchange further confirmed these findings. Across all pre-delisting horizons from one to five years before delisting, the Korean logistic regression model achieved consistently higher AUC values than the original Altman model, indicating superior ranking of financial distress risk. Evaluation at conventional decision thresholds further showed that the logistic model substantially reduced false-positive classifications while maintaining comparable overall detection performance, resulting in considerably higher precision, accuracy, and F1 score. These findings suggest that the re-estimated model provides a more reliable framework for assessing financial distress among Korean firms.

The improved performance is likely attributable to estimating the model coefficients directly from Korean financial data, thereby capturing country-specific financial characteristics and distress patterns that are not reflected in the original U.S.-based Altman coefficients. At the same time, the proposed approach preserves the simplicity and interpretability of the original Altman framework while providing a statistically derived and reproducible method for adapting it to a different economic environment.

This study has several limitations. The external validation employed an independent cohort of delisted firms but used non-distressed observations drawn from the held-out study sample; therefore, the validation should be regarded as partially external rather than fully independent. In addition, the model relies solely on accounting-based financial ratios and does not incorporate market-based, macroeconomic, or industry-specific variables that may further improve predictive performance.

Future research may extend this framework by incorporating additional financial and market indicators, evaluating time-varying or survival-based prediction models, and validating the approach using independent datasets from other countries. Nevertheless, the present findings demonstrate that country-specific statistical re-estimation of the Altman framework offers a practical, interpretable, and empirically robust approach for improving financial distress prediction in the Korean market.

6. References

1. Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

2. Altman, E. I. (2000). *Predicting financial distress of companies: Revisiting the Z-score and ZETA models*. NYU Stern School of Business. https://pages.stern.nyu.edu/~ealtman/Zscores.pdf

3. Bretscher, O. (2013). *Linear algebra with applications* (5th ed.). Pearson.

4. Financial Supervisory Service. (2026). *Data Analysis, Retrieval and Transfer System (DART).* https://dart.fss.or.kr/

5.Johansson, T., & Kumbaro, J. (2011). *Predicting corporate default: An assessment of the Z-score model on the U.S. market 2007–2010* (Master's thesis, Lund University). Lund University Student Papers. http://lup.lub.lu.se/student-papers/record/2061526

6. MacCarthy, J. (2017). Using Altman Z-score and Beneish M-score models to detect financial fraud and corporate failure. *International Journal of Finance and Accounting, 6*(3), 81–92. https://doi.org/10.5923/j.ijfa.20170603.01

7. Macrotrends LLC. (2025). *Financial data and stock statistics*. https://www.macrotrends.net/

8. Naver. (2026). *Naver finance: Financial market data*. https://finance.naver.com/

9. Ofori, E. (2016). Detecting corporate financial fraud using modified Altman Z-score and Beneish M-score: The case of Enron Corp. *Research Journal of Finance and Accounting, 7*(4), 1–10.

10. U.S. Securities and Exchange Commission. (2025). *EDGAR full text search*. https://www.sec.gov/edgar/search/

11. Korea Exchange. (2026). *KIND (Korea Investor's Network for Disclosure System)*. https://data.krx.co.kr/

12. Yahoo. (2025). *Yahoo Finance: Financial market data*. https://finance.yahoo.com/