Modern Management Based on Big Data IV
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA230171

A Probabilistic Guide for Domestic Stocks Delisting Risk from the Nature of Bayesian Matrix

Runnan Chen and Lei Zhang[1]
School of Economics and Management, Beijing Jiaotong University, Beijing 10044, China

Abstract. Finding an optimum way to identify stocks with less delisting risk is critical for every investor in the stock market. However, this procedure is often done based on personal experience, which doesn't fully utilize the historical delisting records. This convention of selecting stocks might result in a greater loss since it merely involves subjective judgment, especially for individual investors. Our research proposes a probabilistic approach for identifying the delisting risk associated with different industry sectors, given the P/B ratio level distribution. And this research offers a customized guide for individual investors to better choose the safer investment options related to the stocks' industry sectors. The completion of our conditional probability matrix is operated under the high-rank assumption, together with the features of Bayesian matrices. The experimental results for our domestic delisting stocks supports the validity and usefulness of our method.

Keywords. Matrix completion, delisting risk, Bayesian matrix, domestic stock data

1. Introduction

The delisting of stocks can bring detrimental impact on their current stockholders: decreased liquidity, less protections toward firms' frauds and increased risk of investing [1, 2]. However, with the little access to investigate the firm's internal operating status, it's hard for stockholders to learn their chances of delisting. Therefore, finding a proper criterion to evaluate the possibilities of delisting is crucial. And the price-to-book ratio (P/B ratio) has been a long-standing important indicator using in industry-wide stock risk evaluation [3, 4, 5]. Noticeably, the interpretation of this numeric ratio is highly dependent on the industry sector of the stock. Hence, considering the availability and usefulness of the historical information, a P/B ratio-based delisting evaluation offers practical methods for stockholders to select stock in different industrial sectors. As for individual investors, this selection process primarily relies on their personal analysis of the historical data of various indicators. The delisting risk evaluation can be highly biased among these subjective opinions. Focusing on China's domestic stock market, there's still no probabilistic guide about the delisting risk while delisting firms' information is avail. With a probabilistic guide regarding different delisting risk and related P/B ratio among different sectors, investors can make a more cautious choice.

[1] Corresponding Author, Zhang Lei, School of Economics and Managements, Beijing Jiaotong University, Beijing, 10044, China; E-mail: zhlei@bjtu.edu.cn.

Current research on delisting prediction majorly utilizes regression model and machine learning (ML) methods: Hwang et al. use a set of both nonfinancial and financial indicators to build logit regression model, and emphasizing an equal attention for both kinds of indicators [6]; Zhou applies Adaboost method to perform simple classification task on firms' delisting, which achieves a better accuracy than other existing classification methods [7]. Meanwhile, current investment-related research on matrix completion focuses on low-rank assumption or periodical behavior: Agrawal reviews existing methods for matrix completion, and shows the Singular Value Decomposition using the Order Method is the best to complete missing stock indices [8]; Athey et al. develop a series of matrix completion estimators to allow missing values have time-series pattern [9]. Previous techniques have been widely used in stock-related information prediction and completion. However, the previous research on prediction of the delisting risk cannot provide a stochastic reference for stock selection. Under the specific scenario, the P/B ratio and industrial sector distribution of the delisting stock form a high-rank matrix, which violates the low-rank assumption. Thus, our research proposes a novel approach to extract probabilistic information regarding the P/B ratio and industry-dependent delisting risk, aiming at better supporting domestic stockholders investing decisions.

2. The construction of the industry-PB ratio Bayesian matrix

From the data source, we can obtain the distribution of industry sectors of all delisted stocks: manufacturing, IT, real estate, and the other 13 sectors. For any individual stockholder, he/she might choose a portfolio of stocks, which is consisted of diverse industry sectors. It is worth noticing that, even the P/B ratio can be a generalized metric to evaluate the quality of stock, we can hardly interpret it directly by merely referring to its numeric value. So, we develop three predefined thresholds for benchmarking this ratio based on industry-wide convention [10]: low (P/B ratio less than 1), medium (P/B ratio falls between 1 and 3), high (P/B ratio falls between 3 and 10), extreme (P/B ratio greater than 10). Set the thresholds as $V = (v_1, v_2 \dots v_x)$, industry sectors as $H = (h_1, h_2 \dots h_y)$. And vector H should be initially given by the distribution of historic delisted data specified by different stock market.

The distribution of the delisted industry sectors is associated with the stocks' performance in the P/B ratio. Therefore, we can express their relationship in the form of conditional probabilities: $P(H|V)$ is the probability of the delisted stock belongs to the ith industry sector given its P/B ratio threshold. And $P(V|H)$ is the probability of the stock has jth P/B ratio threshold given its industry sector. We can write the two conditional probabilities information into a Bayesian matrix, M:

$$M = \begin{pmatrix} 0 & P(V|H) \\ p(H|V) & 0 \end{pmatrix}, V = \begin{pmatrix} v_1 \\ \cdot \\ \cdot \\ \cdot \\ v_x \end{pmatrix}, H = \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ \cdot \\ h_y \end{pmatrix} \quad (1)$$

We partition this matrix into four areas: the two areas that contain non-zero values are lower-left and upper-right: (1) the probabilities that the stock comes from this industry sector given its P/B ratio level. (2) the probabilities that the stock has this P/B ratio level given its industry sector. The two areas that are all zero-valued are upper-left and lower-right: since we set (3) there is no linkage between different stocks' industry sector and (4) there is no correlation between their P/B ratio level, the conditional probabilities will always be zero. Therefore, to solve this objective matrix M, we need to complement the lower-left and upper-right parts.

3. Complementing Bayesian matrix with high-rank settings

Because every element in our industry-PB matrix, M, represents for a certain probability, so they must fall in between the range from 0 to 1. Also, based on our previous settings that each column displays the all circumstances of conditional probabilities, so the column-wise sums of the industry-PB matrix will all equal to 1. After deriving these properties, the industry-PB matrix can be considered as a Markov matrix. And this industry-PB matrix is inherently possesses an eigenvector with eigenvalue equals to 1 because of the Markov matrix's feature. Now we need to find this eigenvector. Consider Equation (2) below:

$$\begin{pmatrix} 0 & P(V|H) \\ p(H|V) & 0 \end{pmatrix} \begin{pmatrix} P(V) \\ p(H) \end{pmatrix} = \begin{pmatrix} P(V) \\ p(H) \end{pmatrix} \tag{2}$$

This equation proves that, vector $r = \begin{pmatrix} p(V) \\ p(H) \end{pmatrix}$ is the corresponding eigenvector with the eigenvalue 1 of our industry-PB matrix, alternatively we can call it the principal eigenvector. Placing it in the scenario of delisted stock analysis, it consists of two components: the probability distribution of the P/B ratio levels and the probability distribution of the industry sectors among all delisted stocks. After getting this eigenvector, we need to reversely infer the original matrix elements. Consistent with the basic assumption that, there's great uncertainty and fluctuation around the coverage of every industry sector and every P/B ratio level, we discuss the finding process under high rank Markov matrix condition. So, our industry-PB matrix has a steady state after multiplying itself n times:

$$\lim_{n\to\infty} M^n = R = (r, r, \cdots, r), r = \begin{pmatrix} r_1 \\ \cdot \\ \cdot \\ \cdot \\ r_{x+y} \end{pmatrix} \tag{3}$$

With

$$MR = RM = R \tag{4}$$

Integrating these equations, we can express the eigenvalue λ by using matrix M:

$$\frac{r^T MRMr}{r^T r} = \frac{r^T Rr}{r^T r} = \lambda \tag{5}$$

So, the problem becomes an optimization problem. i.e., find the matrix M that maximizes the formula (5):

$$Maxmize \frac{r^T MRMr}{r^T r}$$

Given constraints (m_{ij} is the element of M at ith row, jth column; r_i is the ith element of vector r):

$$0 \leq m_{ij} \leq 1$$
$$\sum_{i=1}^{x+y} m_{ij} = 1$$
$$m_{ij}^3 r_j = m_{ji}^2 r_i$$

Eventually, we can solve a matrix M, which subordinates to all constraints above and conforms with the previous eigenvector r. Therefore, with the information drawing from a posterior probabilistic vector r, individual investors can be more well-informed about the delisting risk while knowing stocks' P/B ratio in particular industry.

4. Experiment

Using the delisted stocks data available on Shanghai Stock Exchange and Shenzhen Stock Exchange websites [11, 12], we derive the distribution of delisted stocks' industry sectors: we denote manufacturing as h1, IT as h2…and the rest of industry sectors following the same name convention, which showing with their corresponding percentages in the Figure 1. Also, among the total 124 delisted stock records, we can categorize their P/B ratio into four thresholds: v_1 (low): 29.8%, v_2 (medium): 27.4%, v_3 (high): 29%, v_4 (extreme): 13.7%. From this information, we can construct the eigenvector of our Bayesian matrix: (0.298387097, 0.274193548, 0.290322581, 0.137096774, 0.532258065, 0.088709677, 0.064516129, 0.056451613, 0.048387097, 0.040322581, 0.032258065, 0.032258065, 0.024193548, 0.024193548, 0.016129032, 0.008064516, 0.008064516, 0.008064516, 0.008064516, 0.008064516)^T.

Figure 1. Distribution of stocks' industry sector.

Utilizing the nonlinear optimization model developed in Section3, we obtain the complements of the upper right and lower left entries of the Bayesian matrix from our eigenvectors. The detailed information of the matrix entries and the complement results are displayed in Figure 2 and Figure 3. We use the value in row 5, column 1, to illustrate how an investor can utilize the results: it informs the investor that a delisted stock with a low P/B ratio level has a 58.79% probability that it comes from the manufacturing industry, h1. Compared with other entries located in column 1, this probability is significantly high. Now, the investor might become more cautious about selecting stocks from the manufacturing industry with a low P/B ratio for his/her portfolio.

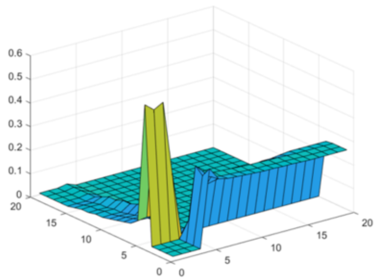

Figure 2. Values of the matrix entries.

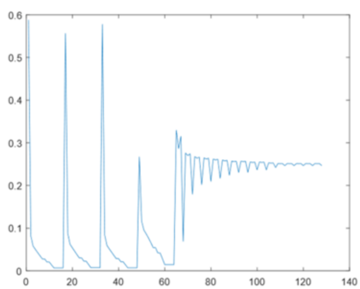

Figure 3. Complement results.

5. Conclusion

This paper provides a basic algorithm to extract the probabilistic relation between P/B ratio levels and delisting risk given the stock's industry sector using properties of Markov matrices by forming an eigenvector containing the distribution of industry sectors and the distribution of P/B ratio levels. It can serve as a probabilistic guide for individual investors to choose safer stocks regarding the delisting risk. The experiment implementation validates the viability of the approach we offered in this paper. The future research can be refined in different aspects: (1) the selection of more comprehensive delisting risk metrics; (2) the incorporation of dynamic risk prediction using time-series information.

Acknowledge

The work was supported by the National Natural Science Foundation of China (No.72271017).

References

[1] Macey J, O'Hara M, Pompilio D. Down and out in the stock market: the law and economics of the delisting process[J]. The Journal of Law and Economics, 2008, 51(4): 683-713.
[2] Loureiro G, Silva S. The impact of cross-delisting from the US On firms' financial constraints[J]. Journal of Business Research, 2020, 108: 132-146.
[3] Cheng C S A, McNamara R. The valuation accuracy of the price-earnings and price-book benchmark valuation methods[J]. Review of quantitative finance and accounting, 2000, 15: 349-370.
[4] Shittu I, Ahmad A C, Ishak Z. Price to book value, price to sales multiples and stock price; evidence from Nigerian listed firms[J]. Journal of Advanced Research in Business and Management Studies, 2016, 3(1): 85-93.
[5] Leonardo M P, Eduardo C M C, María T H A, et al. Formalization of a new stock trend prediction methodology based on the sector price book value for the Colombian market[J]. Heliyon, 2022: e09210.
[6] Hwang I T, Kang S M, Jin S J. A delisting prediction model based on nonfinancial information[J]. Asia-Pacific Journal of Accounting & Economics, 2014, 21(3): 328-347.
[7] Zhou L. Predicting the removal of special treatment or delisting risk warning for listed company in China with Adaboost[J]. Procedia Computer Science, 2013, 17: 633-640.
[8] Agrawal P. Predicting Stock Market Indices Using LRMC[J]. 2021.
[9] Athey S, Bayati M, Doudchenko N, et al. Matrix completion methods for causal panel data models[J]. Journal of the American Statistical Association, 2021, 116(536): 1716-1730.
[10] Pae J, Thornton D B, Welker M. The link between earnings conservatism and the price-to-book ratio[J]. Contemporary Accounting Research, 2005, 22(3): 693-717.
[11] Shanghai Stock Exchange. Suspension[EB/OL]. [2023-02-08]. http://www.sse.com.cn/disclosure/dealinstruc/suspension/.
[12] Shenzhen Stock Exchange. Market Data[EB/OL]. [2023-02-08]. https://www.szse.cn/market/stock/suspend/index.html.