第 50 卷第 12 期 中国科学技术大学学报 Vol. 50, No. 12
2020 年 12 月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Dec. 2020
Received: 2020-09-09; Revised: 2020-12-17 doi:10.3969/j.issn.0253-2778.2020.12.006

Heterogeneous causal inference of corporate profitability using two-scale distributional nearest neighbor estimation

HE Yu[1], WU Yue[2], ZHENG Zemin[1*], WU Jie[1*]
1. School of Management, University of Science and Technology of China, Hefei 230026;
2. School of Insurance, Central University of Finance and Economics, Beijing 100081
* Corresponding author. E-mail: zhengzm@ustc.edu.cn; wu12jie@mail.ustc.edu.cn

Abstract: The heterogeneous treatment effect is the center of gravity in modern causal inference and is also widely used in the financial field. Studying the heterogeneous effect of improving solvency on the profitability of listed companies with different operational capabilities can eliminate the impact of individual differences in listed companies and make the estimation results more accurate. The two-scale distributional nearest neighbor (DNN) is able to eliminate the first-order finite sample bias in the estimate, and it is asymptotically unbiased and asymptotically normal under regularization conditions. Therefore, an empirical analysis was conducted of the financial indicators of Chinese listed companies in the CSMAR financial statement database by using the two-scale distributional nearest neighbor estimation. And a study was carried out on that whether there is heterogeneity in the impact of corporate solvency reduction on corporate profitability under different operating capabilities. Corresponding suggestions were given for optimizing the financial structure and improving the profitability of listed companies.

Keywords: heterogeneous treatment effects; two-scale DNN; operating ability; profitability; solvency
CLC number: F830 **Document code**: A
2010 Mathematics Subject Classification: 62G86

1 Introduction

In almost all economic research, economists hope to infer the causal relationship. In the potential outcomes framework[1], we perceives results as the differences between two parallel universes. The event occurs in one universe and does not occur in the other. The usual interest is the mean difference between the two conceptual outcomes, which is the classic definition of average treatment effects (ATE). This unit of analysis is usually the average treatment effects conditional on an individual's fixed feature vector. In contrast, the unit has been under the name of conditional average treatment effects (CATE)[2] or heterogeneous treatment effects (HTE)[3]. The heterogeneous treatment effect can provide more information than the average treatment effect, and this information is invaluable in the big data era. In economic research, we often need to find causalities from observational research. However, observational experiments are very different from many laboratory and social experiments. It is difficult to obtain random treatment experiment data. Thus we assume another unconfoundedness condition[1], which implies that treatment assignments can be perceived as random by conditioning on a set of observed features. This assumption can greatly simplify the theoretical complexity.

Research on heterogeneity can enable economists to better explore the mechanism of some economic phenomena, but scholars currently have relatively few studies on the effect of heterogeneity, such as the research on the asset structure of listed companies. A listed company's profitability, operating ability, payment ability, debt-paying ability and development potential are indicators that can reflect the comprehensive strength[4]. By analyzing the company's asset structure, we can have an accurate assessment of its comprehensive strength. At present, domestic scholars have studied the correlation between the company's capital structure and its profitability, but they have not yet reached a unified opinion. Yang found that the total asset turnover rate and the profitability of listed companies in Hunan province are positively correlated, and the debt-to-asset ratio and the

Citation: HE Yu, WU Yue, ZHENG Zemin, et al. Heterogeneous causal inference of corporate profitability using two-scale distributional nearest neighbor estimation. J. Univ. Sci. Tech. China, 2020, 50(12): 1488-1497.

profitability are negatively correlated[5]. He found that there is an optimal capital structure interval in the asset structure of enterprises on Yunnan listed companies[6]. Zhang analyzed the performance of listed companies in China's coal industry and found that the operating capacity, the profitability, and the debt-paying change in the same direction[7]. Meng found that there is a negative correlation between the cash holdings of listed companies and their business performance[8]. Gu's research on the general machinery manufacturing sector of listed companies in Shanghai and Shenzhen found that the relation between the current asset ratio and the companys' business performance is not significant[9]. The above studies are all based on the ATE to study the relationship between financial indicators of listed companies, without considering the impact of individual differences between listed companies. Therefore, through an empirical analysis of listed companies, this article is to study the profitability's heterogeneous impact of listed companies with different operating capabilities as quick ratios are improved. This article is to study the impact of the company's solvency on its performance from a new perspective and enrich the current research.

The classic KNN estimation method is a very simple method to estimate heterogeneous treatment effects that well reflect distributional changes. However, the classic KNN estimation will lose precision quickly when the number of covariates increases[10]. The study of Demirkaya et al.[11] enhances and extends the classic KNN estimator to the distribution setting by a simple and effective method called Distributional Nearest Neighbor Estimation (DNN). It is to subsample data and average the first-order nearest neighbor estimators in all sub-sampling. The DNN is a statistic with a first-order finite sample bias and is asymptotically normal under the condition of regularization. Therefore, we obtain two DNN estimators by selecting different sub-sampling scales and then combine them to obtain the two-scale distributional nearest neighbor (two-scale DNN). By this method, the first-order finite sample bias of the DNN statistics can be eliminate to a certain extent. In previous experiments, it has been proved that the two-scale DNN estimation results are in line with the reality, and the accuracy of the experimental results is further improved[11]. The two-scale DNN estimation is an asymptotically unbiased, asymptotically normal statistic widely used in machine learning.

Based on the previous analysis of the relationship between financial indicators of listed companies, this article further analyzes the asset structure and the profitability of listed companies from a new perspective–heterogeneous treatment effects. We also analyze the mechanism of estimation results. In addition, when the heterogeneous treatment effect is significantly not zero, we can significantly improve the profitability of listed companies by changing the size of the solvency. Therefore, we use the bootstrap method[12] to estimate the variance of the two-scale DNN[13], and obtain the operating ability change interval where the heterogeneous treatment effect is significantly non-zero at a 90% point by point confidence level. And through the analysis, we give suggestions on the optimization of the asset structure of listed companies with the operating ability in this interval.

2 Problem description and related concepts

2.1 Problem description

The comprehensive strength of a listed company can be reflected by its profitability, solvency, operating ability, payment ability and development potential. Among them, as the three core financial indicators, the profitability, the operating capacity and the solvency are closely related. Previous studies have found that the operating ability represents the sales and operation of an enterprise. Therefore, the operating ability and the profitability of an enterprise are positively correlated[14]. However, domestic scholars have not reached a consensus on the impact of the solvency on the profitability. On the one hand, high solvency means that a large proportion of assets will be left idle for quick redemption. In this way, the asset utilization ratio of the enterprise is very low, which has a certain negative impact on the profitability of the enterprise. On the other hand, too low solvency means that there is a certain debt risk, which also has a certain negative impact on the profitability of enterprises[15]. The above impacts on the profitability of listed companies are proposed at the average level and targeted at the group of listed companies. However, there are differences among individuals of listed companies, which will lead to contradictions in the conclusion. By studying the heterogeneous treatment effect at the individual level and eliminating the influence of individual differences, we can better study the relationship among the solvency, the operating ability and the profitability, and explore their mechanism of action.

In this paper, we first take the average of various capabilities of listed companies as the empirical value of our estimated results, aiming to explore the impact of a listed company at an average level of development on the improvement of its solvency. After that, we change the indicators of our operating capacity within a certain range, aiming to explore heterogeneous impacts of improving the solvency on the profitability when the development level of listed companies is in an average state but at different operating levels.

2.2 Related concepts

For the estimation of heterogeneous models, we use the traditional potential outcomes framework[1] for processing. Although we only focus on the binary case in this paper, the conclusion can be extended to the multivariate case. Suppose we have a set of observations (X, W, Y), where Y is the response variable, X is the feature vector with dimensionality d and W is the binary treatment classification index ($W = 1$ represents the treated group and $W = 0$ represents the untreated group). Each treatment state corresponds to a potential outcome. $Y(0)$ and $Y(1)$ represent the results of binary treatment states. In empirical studies, there is often a dilemma that $Y(0)$ and $Y(1)$ cannot be observed at the same time. The observed result is actually $Y = Y(0)(1-W) + Y(1)W$.

Traditionally, we have always focused on the average treatment effect of the population, defined by

$$\tau = \mathbb{E}[Y(1) - Y(0)] \tag{1}$$

When a fixed eigenvector x is given, the heterogeneous treatment effect (HTE) of W on Y can be expressed as

$$\tau = \mathbb{E}[Y(1) - Y(0) \mid X = x] \tag{2}$$

Ideally, both $Y(0)$ and $Y(1)$ can be observed so that the problem can be reduced to a classical parametric regression problem. However, in general we cannot obtain such observations, so we assume an unconfoundedness hypothesis[1].

Condition 2.1 The treatment assignment is unconfounded in that it does not depend on the potential outcomes conditioning on X, that is,

$$Y(0), Y(1) \perp\!\!\!\perp W \mid X \tag{3}$$

The unconfoundedness condition entails that treatment assignments can be regarded as random for observations with the same feature vector. Under this condition, it holds that

$$\tau(x) = E[Y \mid X = x, W = 1] - E[Y \mid X = x, W = 0] \tag{4}$$

Therefore, the estimation of $\tau(x)$ can be divided into two problems: The first one is $\mathbb{E}[Y|X=x, W=1]$ in the treated group, and the second one is $\mathbb{E}[Y|X=x, W=0]$ in the untreated group, which is the classical nonparametric estimation problem.

In this paper, our estimates of $\mathbb{E}[Y|X=x, W=1]$ and $\mathbb{E}[Y|X=x, W=0]$ are first separated and then combined to estimate the heterogeneous treatment effect $\tau(x)$. To simplify the presentation process, we use the treated group $W=1$ to illustrate the estimation process. For the untreated group, we can use the same method of estimation. We assume that there is the following relationship between response variables and characteristic variables in the treated group.

$$Y = \mu(x) + \epsilon \tag{5}$$

where ϵ is an independent zero-mean white noise, and $\mu(\cdot)$ represents the unknown relationship between X and Y. In addition, an independent and identically distributed (i. i. d.) sample of size n, $\{(X_i, Y_i)_{i=1}^n\}$, is observed for the treated group. We assume (5) is our working model and our target is to estimate $\mu(x)$ for some given x, where x is a vector with fixed dimension d and probability density function $f(\cdot)$. Besides, there are some common regularization conditions as follow.

Condition 2.2 The density function $f(\cdot)$ is bounded away from 0 to ∞. $f(\cdot)$ and $\mu(\cdot)$ are twice continuously differentiable with bounded second derivatives in a neighborhood of x, and Y has finite second moment, $\mathbb{E}Y^2 < \infty$. In addition, random white noise has mean value zero and finite variance σ_ϵ^2.

Condition 2.3 There are independently identically distributed samples of size n,

$$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n).$$

In summary, Condition 2.1 is the unconfoundedness assumption, which is the basic assumption in the framework of potential causal inference. Under this assumption, the causal inference problem can be transformed into a nonparametric regression problem. Condition 2.2 is a general assumption in nonparametric regression problems. Condition 2.3 specifies the data we have. Although it may seem idealized at first glance, the assumption of such independent and identically distributed data is common in modern machine learning.

Before introducing the DNN estimation, let's first review the classical KNN estimation model. Given a fixed point x, the Euclidean distance between x and each point in the sample can be calculated, and then the sample can be reordered by using this distance,

$$\| X_{(1)} - x \| \leqslant \| X_{(2)} - x \| \leqslant \cdots \leqslant \| X_{(n)} - x \| \tag{6}$$

where $\|\cdot\|$ represents the Euclidean distance. $X_{(1)}$ represents the closest point to x, and $Y_{(1)}$ associated with $X_{(1)}$ is the 1-nearest neighbor estimate of $\mu(x)$. In general, the KNN estimator uses the first k nearest neighbors for estimation, which can be regarded as a special case of general weighted nearest neighbor estimation[16]

$$\widehat{\mu}_{\mathrm{KNN}} = \frac{1}{k}\sum_{i=1}^{k} Y_{(i)}.$$

However, the classical non-parametric estimation method is affected by the dimension. Some accuracy will be lost as dimension increases. Therefore, in this paper, we use a new estimator that can effectively reduce the estimation bias[17] and increase estimation accuracy.

3 Methods for estimating the heterogeneous treatment effects

3.1 Distributional nearest neighbor

Before we introduce the two-scale DNN algorithm, let's first introduce the DNN algorithm. The DNN algorithm makes the bias reduction practical and straightforward. Let $\{i_1,\cdots,i_s\}$, with $i_1<i_2<\cdots<i_s$ and s $(s\leqslant n)$ is a subset of the set of $\{1,\cdots,n\}$. With Z_{i_j} as a shorthand for (X_{i_j},Y_{i_j}), $\Phi(x;Z_{i_1},Z_{i_2},\cdots,Z_{i_s})$ is the first-order nearest neighboring estimator of $\mu(x)$ according to the subsample $\{(Z_{i_j})_{j=1}^{s}\}$,

$$\Phi(x;Z_{i_1},Z_{i_2},\cdots,Z_{i_s})=Y_{(1)}(Z_{i_1},Z_{i_2},\cdots,Z_{i_s}) \quad (7)$$

In other words, given x and the subsamples $\{(Z_{i_j})_{j=1}^{s}\}$, $\Phi(x;Z_{i_1},Z_{i_2},\cdots,Z_{i_s})$ is $Y_{(1)}$ associated with $X_{(1)}$ which is the closest point (defined in Eq. (6) to x in this subsample. The DNN estimation is the average of the 1-nearest estimators in all subsets of sample size s $(1\leqslant s\leqslant n)$. Thus, the DNN statistic is defined in the following form,

$$D_n(s)(x)=\binom{n}{s}^{-1}\sum_{1\leqslant i_1<i_2<\cdots<i_s\leqslant n}\Phi(x;Z_{i_1},Z_{i_2},\cdots,Z_{i_s}) \quad (8)$$

It has another equivalent form,

$$D_n(s)(x)=\binom{n}{s}^{-1}\left\{\binom{n-1}{s-1}Y_{(1)}+\binom{n-2}{s-1}Y_{(2)}+\cdots+\binom{s-1}{s-1}Y_{(n-s+1)}\right\}.$$

Lemma 3.1 Given $x\in\operatorname{supp}(x)$, under Conditions 2.2 and 2.3, we have

$$\mathbb{E}\,D_n(s)(x)=\mu(x)+B(s) \quad (9)$$

where

$$B(s)=\Gamma\left(\frac{2}{d}+1\right)\frac{f(x)\operatorname{tr}(\mu''(x))+2\mu'(x)^{\mathrm{T}}f'(x)}{2dV_d^{\frac{d}{2}}f(x)^{1+\frac{2}{d}}}s^{-\frac{2}{d}}+o(s^{-\frac{2}{d}}),$$

$$V_d=\frac{\pi^{d/2}}{\Gamma\left(1+\frac{d}{2}\right)} \quad (10)$$

$\Gamma(\cdot)$ denotes the Gamma function, $f'(x)$ and $\mu'(x)$ are the first order gradients at x for $f(x)$ and $\mu(x)$, $\mu''(x)$ is the Hessian matrix of $\mu(x)$ and $\operatorname{tr}(\cdot)$ denotes the trace.

Lemma 3.1 gives the form of the finite sample bias of the DNN estimator, which comes from Ref. [17]. When the sub-sampling scale $s\to\infty$, the DNN estimation is asymptotically unbiased. Besides, the coefficient in the leading order of bias term $B(s)$ does not depend on sub-sampling scale s.

Lemma 3.2 Given $x\in\operatorname{supp}(X)$, under Conditions 2.2 and 2.3, assuming that $s\to\infty$ with $s/n\to 0$, for some positive $\sigma_n^2=O\left(\frac{s}{n}\right)$, we have

$$\frac{D_n(s)(x)-\mu(x)-B(s)}{\sigma_n^2}\xrightarrow{\mathscr{D}}N(0,1) \quad (11)$$

Lemma 3.2 shows the asymptotic normal property of DNN, which comes from Ref. [11]. Because of these two properties, the DNN estimation can be weighted to asymptotically eliminate the bias term.

3.2 Two-scale distributional nearest neighbors

Suppose there are two DNN estimations composed of different subsamples s_1, s_2, and they have the following form of finite sample bias:

$$D_n(s_1)(x)=\mu(x)+cs_1^{-2/d}+o(s_1^{-\frac{2}{d}}) \quad (12)$$

$$D_n(s_2)(x)=\mu(x)+cs_2^{-2/d}+o(s_2^{-\frac{2}{d}}) \quad (13)$$

We then establish the following system of linear equations:

$$\omega_1+\omega_2=1 \quad (14)$$

$$\omega_1 s_1^{-2/d}+\omega_2 s_2^{-2/d}=0 \quad (15)$$

Thus, the weights can be calculated as follows:

$$\omega_1^*=\omega_1^*(s_1,s_2)=s_2^{-2/d}/(s_2^{-2/d}-s_1^{-\frac{2}{d}}),$$

$$\omega_2^*=\omega_2^*(s_1,s_2)=s_1^{-2/d}/(s_1^{-2/d}-s_2^{-\frac{2}{d}}).$$

Then we can get the two-scale DNN estimator,

$$D_n(s_1,s_2)(x)=\omega_1^*D_n(s_1)(x)+\omega_2^*D_n(s_2)(x).$$

Eq. (14) provides constraints to eliminate the finite sample bias, and Eq. (15) ensures that two-scale DNN is still unbiased for $\mu(x)$. Compared with the simple DNN estimation, the two-scale DNN estimation eliminates the finite sample bias.

Lemma 3.3 Given $x\in\operatorname{supp}(x)$, under the assumption of Conditions 2.2 and 2.3, assuming in addition, $s_1\to\infty$ with $s_1/n\to 0$, and $s_2\to\infty$ with $s_2/n\to 0$, for some positive $\sigma_n^2=O\left(\frac{s_1}{n}+\frac{s_2}{n}\right)$, we have

$$\left.\begin{aligned}&\frac{D_n(s_1,s_2)(x)-\mu(x)-\Lambda}{\sigma_n^2}\xrightarrow{\mathscr{D}}N(0,1),\\&\Lambda=o(s_1^{-\frac{2}{d}}+s_2^{-\frac{2}{d}})\end{aligned}\right\} \quad (16)$$

When exploring heterogeneous treatment effects, the data will be divided into treated group and untreated group. n_1 and n_0 denote the i. i. d. sample sizes, $s^{(1)}$ and $s^{(0)}$ denote the different subsampling scales.

Lemma 3.4 Given $x\in\operatorname{supp}(x)$, for the treated group and untreated group, under the assumption of Conditions 2.1–2.3, assuming that $s_i^{(1)}\to\infty$ with $s_i^{(1)}/n_1\to 0$, $(i=1,2)$, and $s_i^{(0)}\to\infty$ with $s_i^{(0)}/n_0\to 0$, $(i=1,2)$, we have

$$\left.\begin{aligned}&\frac{[D_n^{(1)}(s_1^{(1)},s_2^{(1)})(x)-D_n^{(0)}(s_1^{(0)},s_2^{(0)})(x)]-\tau(x)-\Lambda}{\sigma_n^2}\xrightarrow{\mathscr{D}}N(0,1),\\&\Lambda=o(s_1^{(1)-\frac{2}{d}}+s_2^{(1)-\frac{2}{d}}+s_1^{(0)-\frac{2}{d}}+s_2^{(1)-\frac{2}{d}})\end{aligned}\right\} \quad (17)$$

Lemma 3.3 and Lemma 3.4 demonstrate that the two-scale DNN is asymptotically unbiased and asymptotically normal, which come from Ref. [11].

4 Research design and result analysis

4.1 Research conjecture

The operating ability reflects a company's operating status and asset management efficiency. A strong operating ability means a company's asset management efficiency is high. On the contrary, a weak operating capacity means that the company's asset management efficiency is low and its operating condition is poor[18]. For listed companies with different operating abilities, improving their debt paying levels has two effects. On the one hand, the improvement of debt-paying ability increases the occupation of idle funds and the opportunity cost[19], which has a negative effect on improving the profitability of enterprises profitability (HTE^-). On the other hand, the improvement of the debt-paying ability reduces the financial risk of the enterprise operation[19], which has a positive effect to profitability (HTE^+) on improving the profitability of enterprises. With different operating conditions, the total profitability HTE of increasing solvency may have different effects on listed companies. For companies with poor operating conditions, it is necessary to improve the utilization ratio of assets, so keeping a low quick ratio can improve the profitability. For companies with good operating conditions, the management efficiency of assets is high, so the effect of keeping a low quick ratio on improving the company's profitability is not so obvious. In addition, for well-run listed companies, keeping a low quick ratio increases financial risks. Therefore, we propose the following model.

In this model, the values of the development potential indicator and the affordability indicator are fixed, and the profitability indicator is the response variable. As one of characteristic variables, the operating capacity index changes, while the solvency index is binary. Listed companies are divided into two categories. One category includes companies with quick ratios of less than 1. The other category includes companies with quick ratios greater than 1. Through the study of these two types of companies, we can explore the heterogeneous impacts of improving solvency on the corporate profitability.

Based on the above analysis, we put forward the following conjectures.

Conjecture 4.1 The heterogeneous treatment effects (HTE) of improving the quick ratio on the profitability of listed companies with different operating capacities are different.

Conjecture 4.2 With the improvement of the operating capacity, the HTE that increases the quick ratio to the profitability of listed companies changes from negative to positive. In other words, the dominant position of the profitability HTE^- is gradually replaced by the profitability HTE^+.

Conjecture 4.3 For listed companies with the poor operation ability, improving the solvency level has a negative heterogeneous treatment effect on improving their profitability, and the effect is significant.

4.2 Data sources and variable selection

The experimental data are from CSMAR Database of financial statements of listed companies in China, which are derived from the financial statements disclosed by listed companies and combined with the stock trading data. The sample selects the latest data in the database. In order to enhance the reliability of the empirical study, we screen the samples and delete them with vacant financial indicators. The final data contains a total of 1980 observation samples, among which 980 companies have quick ratios less than 1 and 1000 companies have quick ratios greater than 1. To simplify the model, we assume that the model satisfies the unconfoundedness assumption and the independent identically distributed hypothesis.

In the selection of variables, each ability to measure the comprehensive strength of listed companies is quantified by an indicator[20]. The profitability is expressed quantitatively by the return on equity (ROE), which reflects the efficiency of shareholders' own capital earnings. The higher the ROE is, the better the profitability of the enterprise will be. The paying ability is expressed by the dividend payout rate, which represents the ratio of dividends paid by a company. In general, mature companies distribute most of their earnings to shareholders. Growing companies do not distribute much of their earnings, but keep them. Therefore, the dividend payout rate also shows the maturity of an enterprise. The development potential is measured by the capital accumulation rate and the total assets growth rate. The capital accumulation rate indicates the increase and decrease of the enterprise's capital in the current year, which reflects the capital preservation and the growth of the enterprise. The total assets growth rate represents the growth of the enterprise's assets. The higher the index, the greater the growth of the enterprise's asset scale and the higher the stability of assets. The operation capacity is expressed by the total assets turnover rate. The higher the index is, the stronger the sales and the operation capacity of the enterprise will be. The solvency is measured by the quick ratio, which indicates the ability of an enterprise's current assets to be realized immediately and used to repay its current liabilities. When the quick ratio of a listed company is greater than 1, it indicates that the enterprise has a sufficient capacity to repay its liabilities and the liquidity of its assets is very high. When the quick ratio of a listed company is less than 1, we think that its solvency is not so good. Too low quick ratio indicates that an enterprise has a financial crisis. Too high quick ratio means that there is too much idle capital, which will increase the opportunity cost of the enterprise and affect its profit.

The above variables are summarized in Tab. 1.

Tab. 1 Variable definition.

Variable	Symbol	Rate	Meaning
Y	Profitablility	ROE	$\frac{\text{Net profit}}{\text{Balance of shareholders' equity}}$
W	Solvency	Quick ratio	$\frac{\text{Current assets}-\text{Inventories}}{\text{Current liabilities}}$
X_1	Operation capacity	Total assets turnover rate	$\frac{\text{Operating income}}{\text{Ending balance of total assets}}$
X_2	Paying ability	Dividend payout rate	$\frac{\text{Dividend per share before tax}}{\text{Earnings per share}}$
X_3	Development prospects	Capital accumulation rate	$\frac{\text{Equity at the ending of current period}}{\text{Equity at the beginning of current period}}$
X_4	Development potential	Total assets growth rate	$\frac{\text{Total assets change for current period}}{\text{Total assets for current period beginning}}$

Tab. 2 Descriptive statistics.

Variable	Mean	Median	MAX	MIN	SD
X_1	0.577772	0.470246	7.902779	0.000962	0.503358
X_2	0.335574	0.200182	79.27272	0	1.961881
X_3	1.39919	1.092457	57.25696	−5.49854	1.876856
X_4	0.36229	0.11728	47.92747	−0.61425	1.77953
Y	0.084339	0.067188	1.999521	0.000106	0.095217

4.3 Background and empirical analysis

4.3.1 Background

We divide the samples into two categories. If the quick ratio is lower than 1, it is the untreated group, that is, $W=0$; if the quick ratio is greater than 1, it is the treated group, that is, $W=1$. We calculate the average value of the four characteristic variables of 1980 companies and take the average as the final empirical value, that is to say, $x=(\overline{X}_1, \overline{X}_2, \overline{X}_3, \overline{X}_4)$. In order to verify the conjectures, we fix the empirical values of the other three characteristics, change the total asset turnover during the estimation, and get different x. Therefore, the estimated result of the profitability HTE on the improvement of the solvency of listed companies is

$$\tau(x) = E[Y \mid X = x, W = 1] - E[Y \mid X = x, W = 0] = D_n^{(1)}(s_1^{(1)}, s_2^{(1)})(x) - D_n^{(0)}(s_1^{(0)}, s_2^{(0)})(x) \quad (18)$$

When setting the change interval of the total asset turnover ratio, we take into account that when an enterprise has stable development and stable payment, the corresponding changes in operating capacity indicators will not be extreme. Therefore, the fluctuation interval of total asset turnover should be around the mean or median, which is the densest distribution interval.

Through the above analysis, we proposed the following model to verify our conjectures.

Explanatory variables: $x=(X_1, \overline{X}_2, \overline{X}_3, \overline{X}_4)$. $\overline{X}_2$, $\overline{X}_3$, $\overline{X}_4$ represent the average of dividend payout rates, capital accumulation rates and the total assets growth rates. X_1 represents the total asset turnover rate.

Binary index: W. $W = 0$ represents a listed company whose quick ratio is less than 1; $W = 1$ represents a listed company whose quick ratio is higher than 1.

Response variable: $D_n^{(0)}(s^{(0)}, 2s^{(0)})(x)$ represents the estimation of the return on equity at the empirical value x by using the two-scale DNN when $W = 0$. $D_n^{(1)}(s^{(1)}, 2s^{(1)})(x)$ represents the estimation of the return on equity at the empirical value x by using the two-scale DNN when $W=1$.

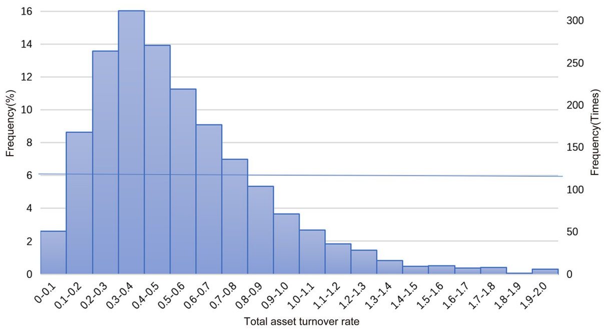

Fig. 1 Histogram of total asset turnover rates of sampled listed companies.

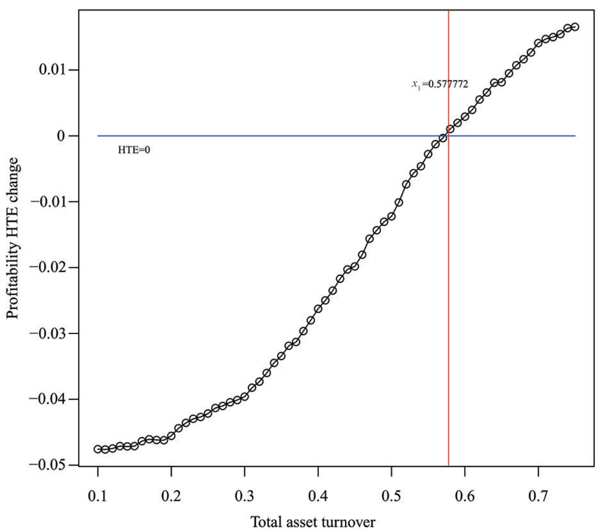

Fig. 2 Profitability HTEs under different operating capabilities.

HTE estimation results: $\widehat{\tau}(x)$ is the influence of the profitability by improving the quick ratio. $\widehat{\tau}(x) = D_n^{(1)}(s^{(1)},2s^{(1)})(x) - D_n^{(0)}(s^{(0)},2s^{(0)})(x)$.

4.3.2 Selection of the fluctuation interval of the total asset turnover rate

From Tab. 2, we can observe the distribution of several ratios of listed companies. The average value of the total asset turnover is 0.577772, the median is 0.470246 and the standard deviation SD is 0.503358, indicating that the distribution of the total asset turnover is relatively concentrated.

To set the reasonable range of the total asset turnover, we draw the frequency distribution histogram of total asset turnovers of 1980 listed companies in the sample.

From Fig. 1, we can see that the distribution of total asset turnovers of the sample listed companies is concentrated on the left side. As a result, the median of total asset turnovers can more accurately reflect the region with the most concentrated distribution. Therefore, [0.1, 0.8] is selected as the fluctuation interval of the total asset turnover rate. In addition, the total asset turnover of the sampled listed companies occurs about 78% of the time within this range.

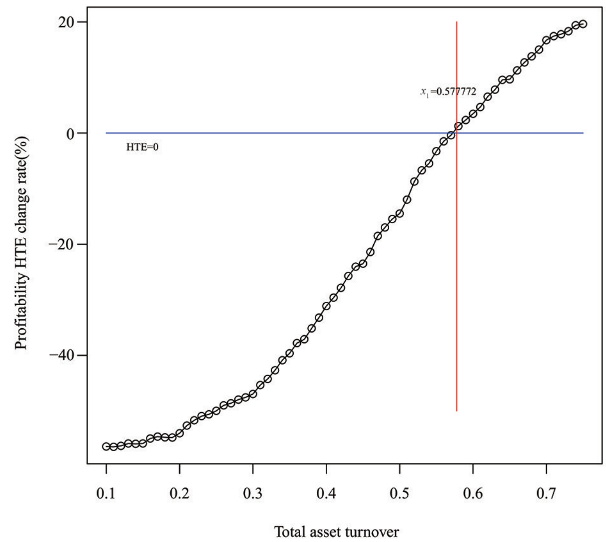

Fig. 3 Profitability HTE change rates under different operating capacities.

4.3.3 HTE estimation results

According to the construction method of two-scale DNN, we write the R program. The following is the estimation result of HTE. According to the empirical value of data, the subsample size $s=8$[11] is selected.

Fig. 2 shows the estimated results of the model. In the figure, the horizontal axis represents different operating capabilities of listed companies. Vertical axis represents estimated results

$$\widehat{\tau}(x) = D_n^{(1)}(s^{(1)},2s^{(1)})(x) - D_n^{(0)}(s^{(0)},2s^{(0)})(x).$$

In the figure, it can be observed that with the change of the total asset turnover, the heterogeneous treatment effect of profitability also changes, and the overall situation rises. Since the value of the profitability HTE is very small, we divide the profitability HTE by the mean ROE $(\overline{Y})$ to obtain the change rate of the profitability HTE as follows.

From Fig. 3, we can see that the greater the distance of total asset turnover rate from the mean value, the greater the profitability HTE change rate. When the total asset turnover is 0.1, the profitability HTE change rate even reaches about 50%. In Figures 2 and 3, the red lines represent the average of total asset turnover, that is, $X_1 = \overline{X}_1$. The blue line represents the estimate of the profitability HTE is 0, that is, $\widehat{\tau}(x) = 0$ and $\dfrac{\widehat{\tau}(x)}{\overline{Y}} = 0$. From the above two figures, it can be seen that the estimated results of the profitability HTE are two upward curves. With the increase of the total asset turnover, the estimated results of heterogeneous treatment effects change from negative to positive, from small to large. This verifies Conjecture 4.1 that the profitability HTE of improving solvency is different at different operating levels. In addition, from the two curves, we also find that the estimated results increase with the increase of the total asset turnover, which means that, with the growth of the total asset turnover, improving solvency increases the occupation of idle funds. The leading role of the negative effect (profitability HTE^-) is gradually replaced by the positive effect (profitability HTE^+) to reduce the financial risk of the enterprise operation. That is to say, when the total asset turnover rate of listed companies is relatively low and the asset utilization efficiency is low, under such circumstances, further improving the solvency of listed companies and increasing the occupation of idle funds can reduce financial risks, but seriously deteriorate the efficiency of asset utilization. Therefore, it has an obvious negative effect on the profitability, that is to say, HTE^- plays a dominant role. When the total capital turnover ratio is high, the capital utilization efficiency is also high. In this case, improving the solvency of listed companies increases the occupation of idle funds, but due to its high capital utilization efficiency, it has little negative effects on improving profitability. At the same time, the improvement of solvency reduces the financial risks in the operation of enterprises, which has a more obvious positive effect on profitability, that is to say, HTE^+ plays a dominant role. This confirms the Conjecture 4.2.

4.3.4 Listed companies with significantly non-zero HTE

When the total asset turnover of the listed company fluctuates within a certain range, the profitability HTE by improving the solvency of the enterprise is likely to be significantly non-zero. In this case, these listed companies can significantly improve their profitability by adjusting their debt repayment level and optimizing their asset structure. Therefore, we use bootstrap[12] to estimate the variance of experimental results[13], and calculate the confidence interval of the profitability HTE, so as to obtain the change interval of total asset turnover rates of listed companies with the significant non-zero heterogeneous treatment effect.

Experimental process:

① The existing samples in the treated group and the untreated group are sampled 1000 times with replacement, and the obtained samples form a new set of bootstrap samples$^{(0)}$ in the untreated group and a new set of bootstrap samples$^{(1)}$ in the treated group.

② Bootstrap samples$^{(0)}$ and bootstrap samples$^{(1)}$ are used to estimate a new $\widehat{\tau}(x)$ by the two-scale DNN.

③ By repeating the first two steps, we obtained 1000 estimates of the profitability HTE. Then, we can

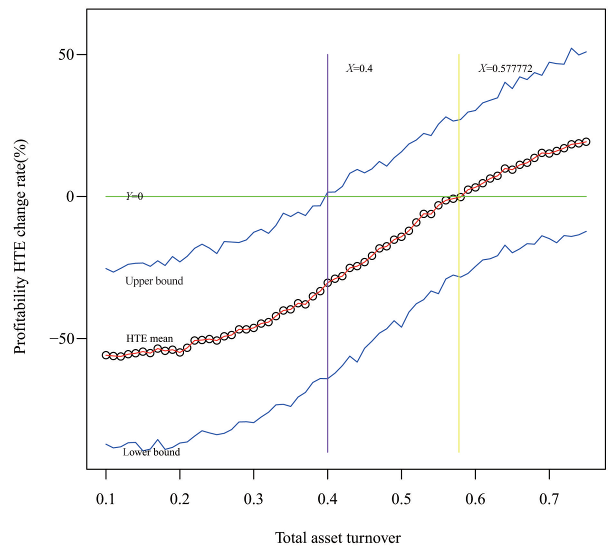

Fig.4 The confidence interval and mean estimations of the profitability HTE change rates under different operating capabilities.

get $\widehat{\tau}(x)_{(1)}$, $\widehat{\tau}(x)_{(2)}$, $\cdots$, $\widehat{\tau}(x)_{(1000)}$ by reordering the estimated results from small to large. The estimated distribution of the results of 1000 bootstrap samples was used to approximate the distribution of (x). Set $\alpha = 0.1$, so $[\widehat{\tau}(x)_{(50)}, \widehat{\tau}(x)_{(950)}]$ is the 90% point by point confidence interval of profitability HTE. The mean of $\tau(x)$ is estimated by averaging all the bootstrap estimated $\widehat{\tau}(x)$.

④ Then we change the first component of x so that the total asset turnover varies between $[0.1, 0.8]$. After repeating the first four steps, the mean value and 90% point by point confidence interval of the profitability HTE of improving solvency of listed companies at different operating levels can be obtained.

Fig. 4 shows the 90% point by point confidence interval estimate and mean estimate of profitability HTEs by improving solvency of listed companies under different operating capacities. $x_1 = 0.4$ represents the total asset turnover of 70% $\bar{x}_1$, and $x_1 = 0.577772$ is the average of total asset turnovers. It can be observed from the figure that when the total asset turnover rate is below 0.4, profitability HTE is significantly less than 0 as the upper bound of the confidence interval is less than 0. It shows that when the paying capacity and the development potential of listed companies are at the average level, and the total asset turnover rate fluctuates at $[0.1, 0.4]$, the profitability HTE of improving solvency is significantly less than 0 under 90% confidence. In this case, the listed company can optimize the asset structure by reducing its debt repayment level and reducing the occupation of idle funds, so as to achieve the purpose of significantly improving profitability. Thus, Conjecture 4.3 has been verified.

5 Conclusions

This paper explores the profitability HTE of improving the solvency for listed companies with the average payment ability and the development potential at different operating levels. The study of the heterogeneous treatment effects eliminates the influence of individual differences of listed companies, making the inconsistent conclusions from the level of population research no longer contradictory. Moreover, through the analysis of the heterogeneous treatment effects, we can explore the mechanism behind the economic phenomenon. Through the theoretical analysis and using the two-scale DNN estimation method, it is shown that the profitability HTE of improving solvency for the listed companies is different at different operating levels. With the improvement of operation level, the profitability HTE gradually changes from small to large and from negative to positive. It means that with the improvement of operating ability, when listed companies improve their debt paying ability, the positive effect (HTE^+) of reducing the financial risk of the enterprise operation on improving the profitability of the enterprise will gradually replace the dominant position of the negative effect (HTE^-) of idle capital occupation on improving the profitability of enterprises. The research results also show that when the turnover rate of total assets of listed companies fluctuates in [0.1, 0.4], reducing their debt repayment level can significantly improve their profitability. Therefore, for listed companies with the poor management ability, reducing the occupation of idle funds and improving the utilization efficiency of assets can significantly optimize the asset structure and improve the profitability.

In this article, we assume that the samples are independent of each other, but in reality, there are also cases where samples are not independent. In this case, if we still use the two-scale DNN to estimate, the asymptotic distribution is not valid, because the deviation of the estimate may overpower the asymptotic distribution term. However, when the sample size is large enough, the size of the two-scale DNN is still of certain reference value. At the same time, we can also use the bootstrap to get the distribution of the estimated results. In this paper, we use the financial data of listed companies. Listed companies must ensure their financial independence, and these listed companies are from different industries, so we can regard them as independent samples.

In reality, there may be also heteroscedasticity in the sample. In this case, the two-scale DNN estimator can be unbiased but the variance may be large. At this point, we can use bootstrap to get the estimated result distribution to make causal inference. In this paper, we assume that the error is homoscedasticity, but we also use the bootstrap to obtain 90% point-by-point confidence intervals of the estimated results to verify the validity of the results.

Acknowledgements

This work was supported by National Natural Science Foundation of China (72071187, 11901603, 11671374, 71731010, and 71921001), and Fundamental Research Funds for the Central Universities (WK347000 0017 and WK2040000027).

Conflict of interest

The authors declare no conflict of interest.

Author information

He Yu and Wu Yue are co-first authors.

HE Yu is a graduate student under the supervision of Prof. Zheng Zemin at University of Science and Technology of China (USTC). Her current research is mainly focused on statistical inference and applications.

WU Yue received her Ph. D. degree from State University of New York. She is currently an associate professor at Central University of Finance and Economics. Her research interests cover risk management, dynamic optimization and insurance.

ZHENG Zemin (corresponding author) is Full Professor at the department of management, University of Science and Technology of China. He received his B. S. degree from USTC in 2010 and Ph. D. degree from University of Southern California in 2015. Afterwards, he was working in Department of Management, University of Science and Technology of China. His research interests include high dimensional statistical inference and big data problems.

WU Jie (corresponding author) is currently a Ph. D. student from School of Management, University of Science and Technology of China. She received her B. S. degree from Anhui University of Technology in 2017. Her research mainly focuses on high dimensional variable selection, classification.

References

[1] Rubin D B. Estimating causal effect of treatments in randomized and nonrandomized studies. Journal of Educational Psychology, 1974, 66(5): 688-701.

[2] MaCurdy T, Chen X, Hong H. Flexible estimation of treatment effect parameters. American Economic Review, 2011, 101(3): 544-551.

[3] Heckman J J, Smith J, Clements N. Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. The Review of Economic Studies, 1997, 64(4): 487-535.

[4] Yang Chao. Analysis of the influencing factors of the comprehensive strength of my country's listed companies. Heilongjiang Foreign Economic Relations and Trade, 2011 (4): 94-96. (in Chinese)

[5] Yang Yuanxia. Research on the correlation between the

asset structure and company performance of listed companies in Hunan Province. Journal of Central South University (Social Science Edition), 2014,20(2): 30-34. (in Chinese)

[6] He Guozhong. Research on the impact of capital structure on corporate profitability and solvency: Taking Yunnan listed companies as an example. Friends of Accounting, 2019(3): 67-70. (in Chinese)

[7] Zhang Yi. Research on the performance evaluation system of listed companies in my country's coal industry. Xi'an: Xi'an University of Science and Technology, 2018. (in Chinese)

[8] Meng Shuangwu. Research on the impact of listed companies' cash holdings on investment behavior. Finance and Economics Theory and Practice, 2011(2): 53-58. (in Chinese)

[9] Gu Shuibin. An empirical study on the relationship between capital structure, asset structure, ownership structure and corporate performance -- Based on the research of the general machinery manufacturing sector of Shanghai and Shenzhen listed companies. Finance and Accounting News, 2009(21): 115-117. (in Chinese)

[10] Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 2018, 113: 1228-1242.

[11] Demirkaya E, Fan Y, Gao L, et al. Nonparametric inference of heterogeneous treatment effects with two-scale distributional nearest neighbors. https://arxiv. org/abs/1808.08469.

[12] Efron B. The Jackknife, the Bootstrap, and Other Resampling Plans. Philadelphia, PA: SIAM, 1982.

[13] Tu Dongsheng, Cheng Ping. Bootstrap approximation of non-censored L-statistic. Systems Science and Mathematics, 1989(1): 14-23. (in Chinese)

[14] Huang Yuan, Wu Yu. An empirical analysis of the relationship between the operating capability and profitability of retail enterprises: Based on the data analysis of 158 listed companies. Modern Business, 2017(14): 129-130. (in Chinese)

[15] Dan Zhen. Analysis of the financial capability of listed companies in Tibet. Chengdu: Southwestern University of Finance and Economics, 2011. (in Chinese)

[16] Mack Y. Local properties of k-NN regression estimates [J]. SIAM Journal on Algebraic Discrete Methods, 1980, 2(3): 311-323.

[17] Biau G, Devroye L. Lectures on the Nearest Neighbor Method. Cham, Switzerland: Springer, 2015.

[18] Niu Chunyan, Wang Baiyao. Analysis of financial statements of listed companies: A case study of Chongqing Fuling Mustard Group Co., Ltd.. Rural Economy and Technology, 2019, 30(15): 183-184. (in Chinese)

[19] Yan Wen. Application of financial index analysis method of listed companies: Taking Dongxu Optoelectronics as an example. Hebei Enterprise, 2019(9): 33-34. (in Chinese)

[20] Gao Han. Analysis of financial indicators of listed companies. Economist, 2018 (2): 99-100. (in Chinese)

运用双尺度分布最近邻估计对公司盈利能力作异质性因果推断

何雨[1],吴越[2],郑泽敏[1*],吴捷[1*]

1. 中国科学技术大学管理学院,安徽合肥 230026;2. 中央财经大学保险学院,北京 100081

摘要:异质性治疗效应是现代因果推断的一个重要内容,在金融领域也有非常广泛的应用. 研究提高偿债水平对不同营运能力的上市公司盈利能力的异质性效应,以消除上市公司个体差异带来的影响,使得估计结果更加准确. 在估计方法的选择上,选择了双尺度分布最近邻估计. 它能够消除估计中的一阶有限样本偏差,并且它在正则化条件下是渐近无偏和渐近正态的. 因此,应用双尺度分布最近邻估计对 CSMAR 中国上市公司财务报表数据库中的中国上市公司财务指标进行了实证分析,研究了不同经营能力下公司的偿债能力提高对于公司盈利能力是否存在异质性治疗效应. 根据其研究结果,对优化上市公司财务结构,提高上市公司的盈利能力给出了相应的建议.

关键词:异质性治疗效应;双尺度分布最近邻;营运能力;盈利能力;偿债水平