

附录

正文未报告部分

附录 1: 模型评估指标及计算方法

在评估模型性能时,本文从解释能力和预测精度两个维度综合考量。模型解释能力反映了模型在预测公司分红行为时的拟合程度,本文使用样本内拟合优度 R_{ls}^2 考察不同研究方法对训练集样本的拟合效果,样本内拟合优度越高,说明模型训练效果越好。本文采用样本外拟合优度 R_{oos}^2 和可解释方差 EVS_{oos} 以考察各研究方法对于测试集样本的预测效果,其值越大,说明模型对于公司分红行为的预测能力越强,样本外拟合效果也越好。模型预测精度则反映了模型在预测分红行为时的准确度,本文选取样本外均方误差 MSE_{oos} 衡量模型预测值与实际值之间的偏离程度。与此同时,考虑到均方误差可能受到离群值的影响,导致其估计值与实际情况偏离较大,本文还使用了平均绝对误差 MAE_{oos} 和绝对中位差 $MedAE_{oos}$ 来评估模型预测公司分红行为时的准确性。

附表 1: 模型评估指标及计算方法

评估指标	指标含义	计算方法
R_{ls}^2	样本内拟合优度,在训练集中,模型预测值对于实际观测值的拟合程度。	$R_{ls}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
R_{oos}^2	样本外拟合优度,在测试集中,模型预测值对于实际观测值的拟合程度。	R_{oos}^2 计算同 R_{ls}^2
EVS_{oos}	可解释方差,在测试集中,模型预测值变异程度对于实际观测值变异程度的拟合程度。	$EVS_{oos} = 1 - Var(y - \hat{y}) / Var(y),$ 其中 $Var(y - \hat{y}) = 1/n \sum_{i=1}^n (e_i - \bar{e})^2$, $e_i = y_i - \hat{y}_i$, $Var(y) = 1/n \sum_{i=1}^n (y_i - \bar{y})^2$ 。
MSE_{oos}	均方误差,样本外预测值与实际值之差平方的期望值。	$MSE_{oos} = 1/n \sum_{i=1}^n (y_i - \hat{y}_i)^2$
MAE_{oos}	平均绝对误差,样本外预测值与实际值之差绝对值的期望值。	$MAE_{oos} = 1/n \sum_{i=1}^n y_i - \hat{y}_i $
$MedAE_{oos}$	绝对中位差,样本外预测值与实际值之差绝对值的中位数。	$MedAE_{oos} = \text{median of } y_i - \hat{y}_i $

附录 2: 描述性统计结果

基于本文统计, 公司股利支付率的均值和中位数分别为 0.27 和 0.21, 标准差为 0.32, 这表明不同公司在股利分配决策上存在明显差异。其余变量的分布情况和数值也均在合理范围内。

附表 2: 主要变量描述性统计

变量名	样本量	均值	标准差	最小值	中位数	最大值
$Dividend_ratio_{i,t}$	31469	0.2696	0.3240	0.0000	0.2058	3.3247
$Managefee_ratio_{i,t}$	31469	0.0849	0.0694	0.0059	0.0677	0.7079
$Manageshare_{i,t}$	31469	11.3901	17.9620	0.0000	0.2434	68.7134
$Indep_ratio_{i,t}$	31469	0.3737	0.0532	0.2500	0.3333	0.6000
$Bgender_{i,t}$	31469	0.1462	0.1292	0.0000	0.1111	0.5714
$Bshare_{i,t}$	31469	0.0747	0.1306	0.0000	0.0001	0.5683
$Bage_{i,t}$	31469	3.9912	0.1368	3.5264	4.0073	4.3438
$Btenure_{i,t}$	31469	3.8758	0.9263	0.6931	4.0604	5.4467
$Bsalary_{i,t}$	31469	10.1526	5.7323	0.0000	13.0813	15.8477
$Equity_{i,t}$	31469	0.1060	0.3078	0.0000	0.0000	1.0000
$Da_abs_{i,t}$	31469	0.0551	0.0563	0.0003	0.0381	0.3799
$Tunneling_{i,t}$	31469	0.0162	0.0253	0.0001	0.0076	0.3275
$Top1_{i,t}$	31469	33.5536	14.6066	7.5200	31.2500	76.4400
$Sharebalance_{i,t}$	31469	0.7166	0.6025	0.0132	0.5486	2.9304
$Minorityrate_{i,t}$	31469	52.7822	16.1157	13.7600	52.5900	89.3500
$Institution_{i,t}$	31469	44.0131	24.0211	0.0688	45.7164	93.4019
$Pledge_{i,t}$	31469	6.0622	15.4204	0.0000	0.0000	100.0000
$Retainedearn_ratio_{i,t}$	31469	0.1625	0.1888	-1.1407	0.1662	0.6181
$Freecash_{i,t}$	31469	-0.0059	0.0939	-0.3728	-0.0012	0.3171
$Tax_avoid_{i,t}$	31469	0.0241	0.0582	-0.5221	0.0274	0.5550
$Tax_ratio_{i,t}$	31469	0.1540	0.1922	-1.3326	0.1483	1.4511
$Tax_volatility_{i,t}$	31469	0.1522	0.3660	0.0016	0.0546	4.1840
$Constraint_{i,t}$	31469	-3.8113	0.2510	-4.5576	-3.8074	-3.0793
$Refinance_{i,t}$	31469	0.0976	0.2967	0.0000	0.0000	1.0000
$Sentiment_inv_{i,t}$	31469	0.0726	1.0176	-1.8615	-0.1590	5.5564
$Dividend_lag_{i,t-1}$	31469	0.2755	0.3141	0.0000	0.2156	2.8662
$ROA_{i,t}$	31469	0.0333	0.0672	-0.5189	0.0343	0.2360
$Cashflow_{i,t}$	31469	0.4931	0.9325	-3.0564	0.3288	6.2000
$Tobinq_{i,t}$	31469	2.0251	1.3428	0.7954	1.6062	17.8865
$BM_{i,t}$	31469	0.6253	0.2504	0.0559	0.6226	1.2572
$Lev_{i,t}$	31469	0.4439	0.2002	0.0456	0.4424	0.9347
$Soe_{i,t}$	31469	0.4224	0.4940	0.0000	0.0000	1.0000
$Growth_{i,t}$	31469	0.1651	0.4080	-0.6762	0.1067	6.5667
$Lnsize_{i,t}$	31469	22.2331	1.2607	19.3834	22.0617	26.4557
$Analyst_num_{i,t}$	31469	1.3644	1.1851	0.0000	1.3863	3.9120
$Market_idx_{i,t}$	31469	9.5630	1.7656	3.3590	9.8060	12.8640

附录 3: 详细变量相对重要性排序

尽管已有研究发现再融资动机 ($Refinance_{i,t}$) 和控股股东股权质押比例 ($Pledge_{i,t}$) 等指标对公司股利分配具有一定的解释力, 但这些指标的相对重要性较低, 说明它们对公司股利分配行为的实际预测能力并不佳。

附表 3: 详细变量相对重要性排序

单位 (%)

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	17.8244	<i>Dividend_lag_{i,t-1}</i>	30.4594
2	<i>ROA_{i,t}</i>	7.9011	<i>ROA_{i,t}</i>	12.7493
3	<i>Retainedearn_ratio_{i,t}</i>	6.3092	<i>Retainedearn_ratio_{i,t}</i>	6.6660
4	<i>Tax_ratio_{i,t}</i>	4.4097	<i>Tax_ratio_{i,t}</i>	4.9492
5	<i>Tax_avoid_{i,t}</i>	3.9103	<i>Lev_{i,t}</i>	2.7796
6	<i>Tax_volatility_{i,t}</i>	3.2962	<i>Growth_{i,t}</i>	2.6070
7	<i>Growth_{i,t}</i>	2.9151	<i>Tax_volatility_{i,t}</i>	2.1892
8	<i>Lev_{i,t}</i>	2.8889	<i>Da_abs_{i,t}</i>	2.0731
9	<i>Tunneling_{i,t}</i>	2.6110	<i>Institution_{i,t}</i>	2.0386
10	<i>Freecash_{i,t}</i>	2.5321	<i>Tunneling_{i,t}</i>	1.9986
11	<i>Institution_{i,t}</i>	2.4066	<i>Tax_avoid_{i,t}</i>	1.9902
12	<i>Minorityrate_{i,t}</i>	2.3835	<i>Freecash_{i,t}</i>	1.8718
13	<i>Da_abs_{i,t}</i>	2.2948	<i>Minorityrate_{i,t}</i>	1.8667
14	<i>Cashflow_{i,t}</i>	2.2691	<i>Sharebalance_{i,t}</i>	1.7192
15	<i>Top1_{i,t}</i>	2.2330	<i>Cashflow_{i,t}</i>	1.5965
16	<i>Sharebalance_{i,t}</i>	2.2283	<i>Btenure_{i,t}</i>	1.5614
17	<i>Btenure_{i,t}</i>	2.1561	<i>Lsize_{i,t}</i>	1.5312
18	<i>Lsize_{i,t}</i>	2.1475	<i>Top1_{i,t}</i>	1.5019
19	<i>Managefee_ratio_{i,t}</i>	2.0838	<i>Bage_{i,t}</i>	1.4234
20	<i>Constraint_{i,t}</i>	2.0307	<i>Constraint_{i,t}</i>	1.3198
21	<i>Bage_{i,t}</i>	1.9923	<i>Managefee_ratio_{i,t}</i>	1.2640
22	<i>Sentiment_inv_{i,t}</i>	1.8801	<i>Manageshare_{i,t}</i>	1.1916
23	<i>BM_{i,t}</i>	1.8360	<i>Sentiment_inv_{i,t}</i>	1.1726
24	<i>Tobinq_{i,t}</i>	1.8220	<i>Bsalary_{i,t}</i>	1.1639
25	<i>Manageshare_{i,t}</i>	1.7761	<i>Market_idx_{i,t}</i>	1.0497
26	<i>Market_idx_{i,t}</i>	1.5941	<i>BM_{i,t}</i>	0.9386
27	<i>Bsalary_{i,t}</i>	1.5731	<i>Tobinq_{i,t}</i>	0.9004
28	<i>Analyst_num_{i,t}</i>	1.3351	<i>Bshare_{i,t}</i>	0.7406
29	<i>Bshare_{i,t}</i>	1.2393	<i>Bgender_{i,t}</i>	0.7295
30	<i>Bgender_{i,t}</i>	1.2127	<i>Indep_ratio_{i,t}</i>	0.7254
31	<i>Indep_ratio_{i,t}</i>	0.9492	<i>Analyst_num_{i,t}</i>	0.7063
32	Pledge_{i,t}	0.7455	Pledge_{i,t}	0.4564
33	<i>Soe_{i,t}</i>	0.2649	<i>Soe_{i,t}</i>	0.1612
34	Refinance_{i,t}	0.1984	Refinance_{i,t}	0.0959
35	<i>Equity_{i,t}</i>	0.1568	<i>Equity_{i,t}</i>	0.0489

附录 4: 控制各类别公司特征个数后的相对重要性比较

考虑到不同类别的特征数量的差异可能会影响本文的研究结论, 本文选取各自类别中相对重要性最大的一个特征加入到基本面特征中, 重新训练模型并计算此时各类别特征的相对重要性。未展示的结果表明, 尽管随着时间推移, 不同类别的公司特征累计相对重要性有所变动, 但整体而言, 生命周期特征和公司税率特征的相对重要性仍显著高于其他类别的公司特征。

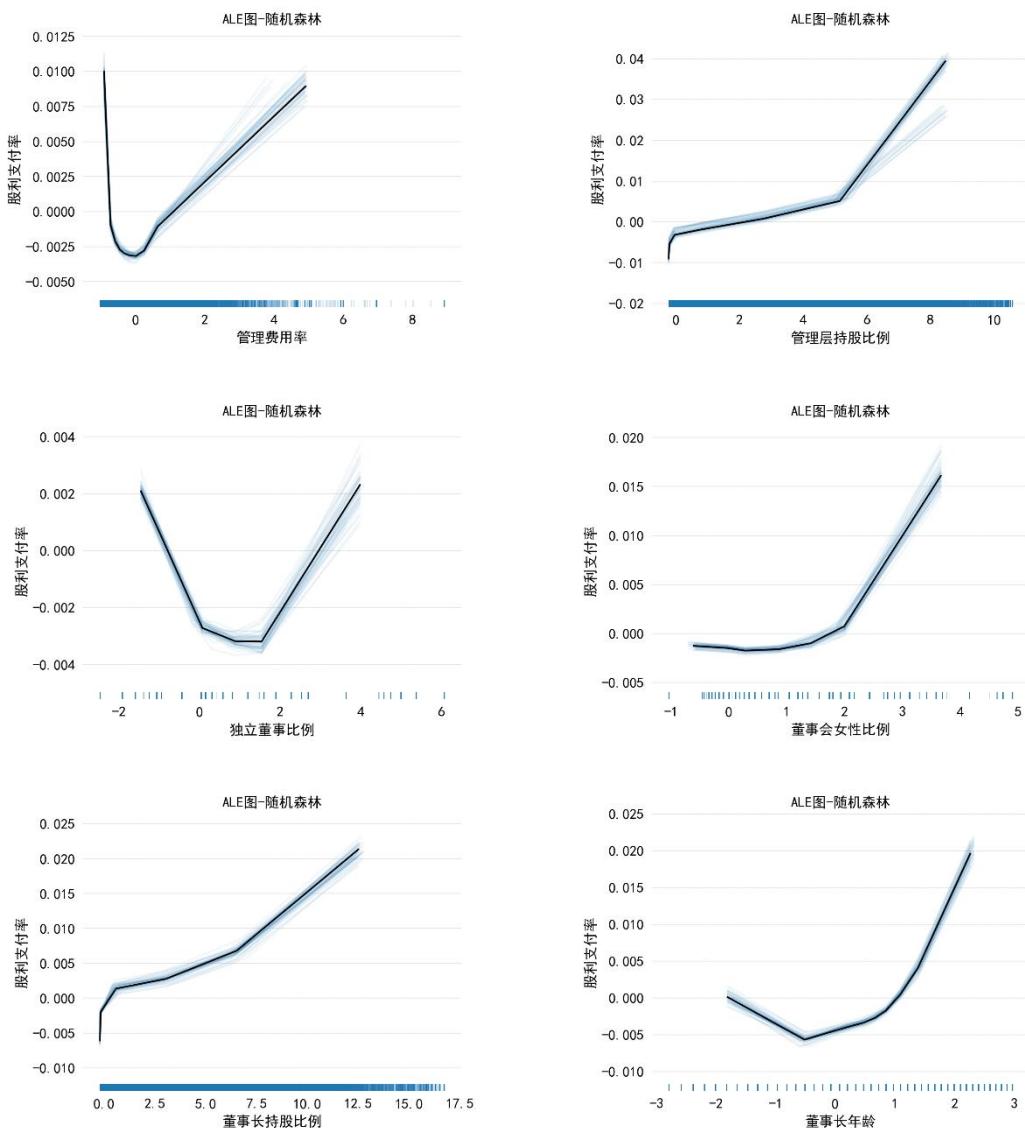
附表 4: 控制各类别公司特征个数后的相对重要性比较

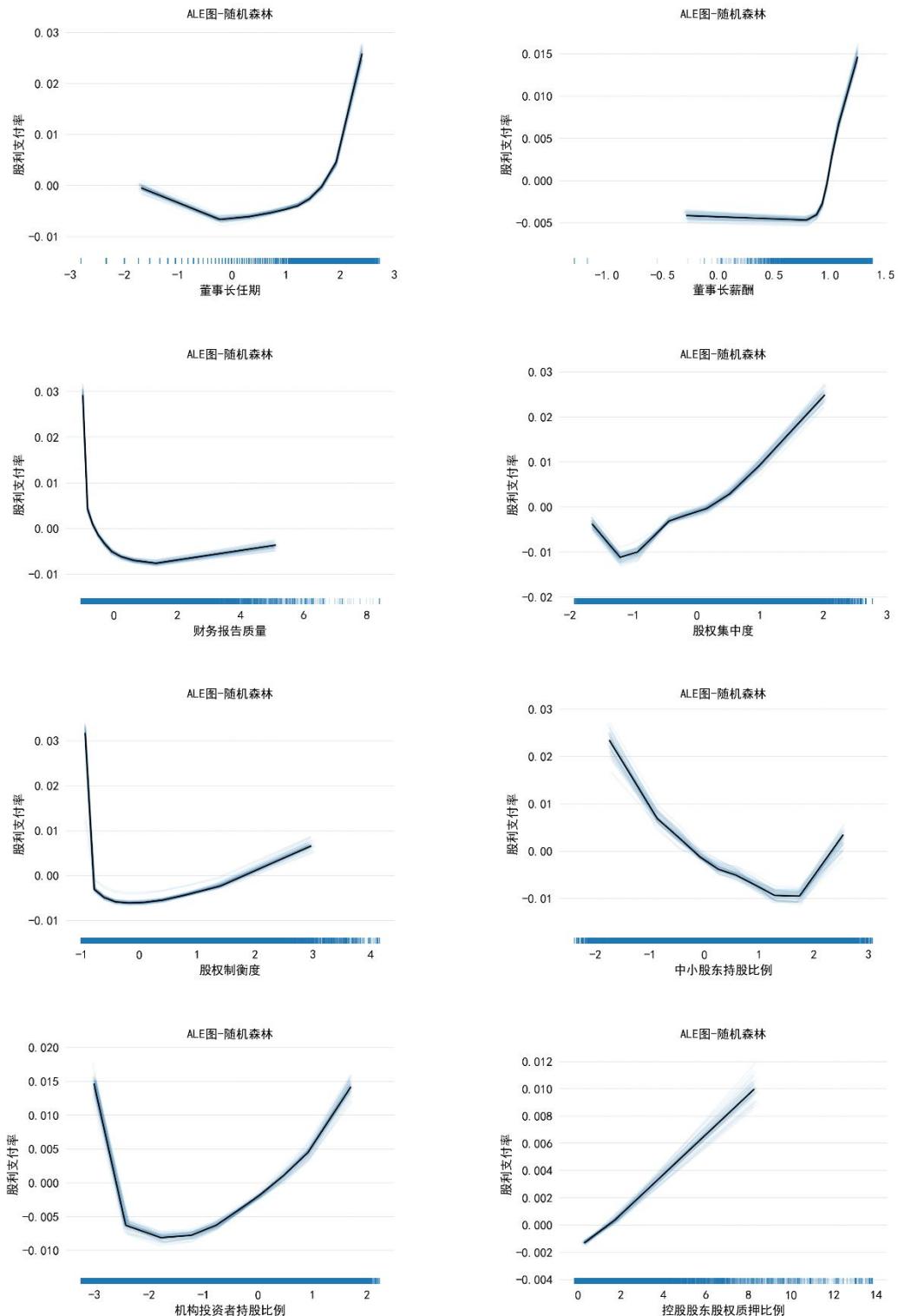
年份	第一类代理 问题特征	第二类代理 问题特征	生命周期 特征	公司税率 特征	融资需求 特征	投资者情 绪特征
2006	3.4108%	5.1152%	8.7787%	6.4549%	3.5323%	3.3967%
2007	3.9377%	6.1492%	7.7603%	6.0905%	4.2693%	3.9930%
2008	4.3943%	5.9929%	7.6843%	8.1217%	4.5186%	3.9783%
2009	4.6828%	4.4959%	9.3635%	6.6807%	4.2488%	3.9825%
2010	4.7071%	4.8163%	8.8057%	5.3736%	5.9102%	4.3455%
2011	4.0278%	4.6541%	7.5055%	6.1334%	4.9412%	4.2886%
2012	4.9226%	4.7172%	7.6061%	8.5181%	4.3473%	4.3062%
2013	4.4174%	4.1670%	9.0038%	6.4215%	3.9039%	3.7191%
2014	3.7395%	4.0106%	8.4078%	7.9104%	3.7373%	4.0331%
2015	4.0632%	3.9396%	7.6173%	7.3508%	4.7367%	3.9005%
2016	4.4954%	4.2500%	8.5038%	6.6587%	4.1013%	3.7604%
2017	4.1533%	4.7215%	8.7479%	6.5973%	3.6768%	3.7816%
2018	4.7633%	4.6092%	7.4618%	8.9603%	3.5764%	3.4547%
2019	4.3771%	4.2226%	8.9680%	7.3314%	3.6367%	3.5361%
2020	3.6329%	4.9818%	8.7452%	6.0949%	3.4277%	3.6209%
2021	3.9306%	4.1849%	9.7004%	6.6488%	3.3110%	3.8664%
2022	3.6509%	3.6604%	8.6212%	6.8364%	3.6754%	3.7346%
平均值	4.1945%	4.6287%	8.4283%	6.9520%	4.0912%	3.8646%

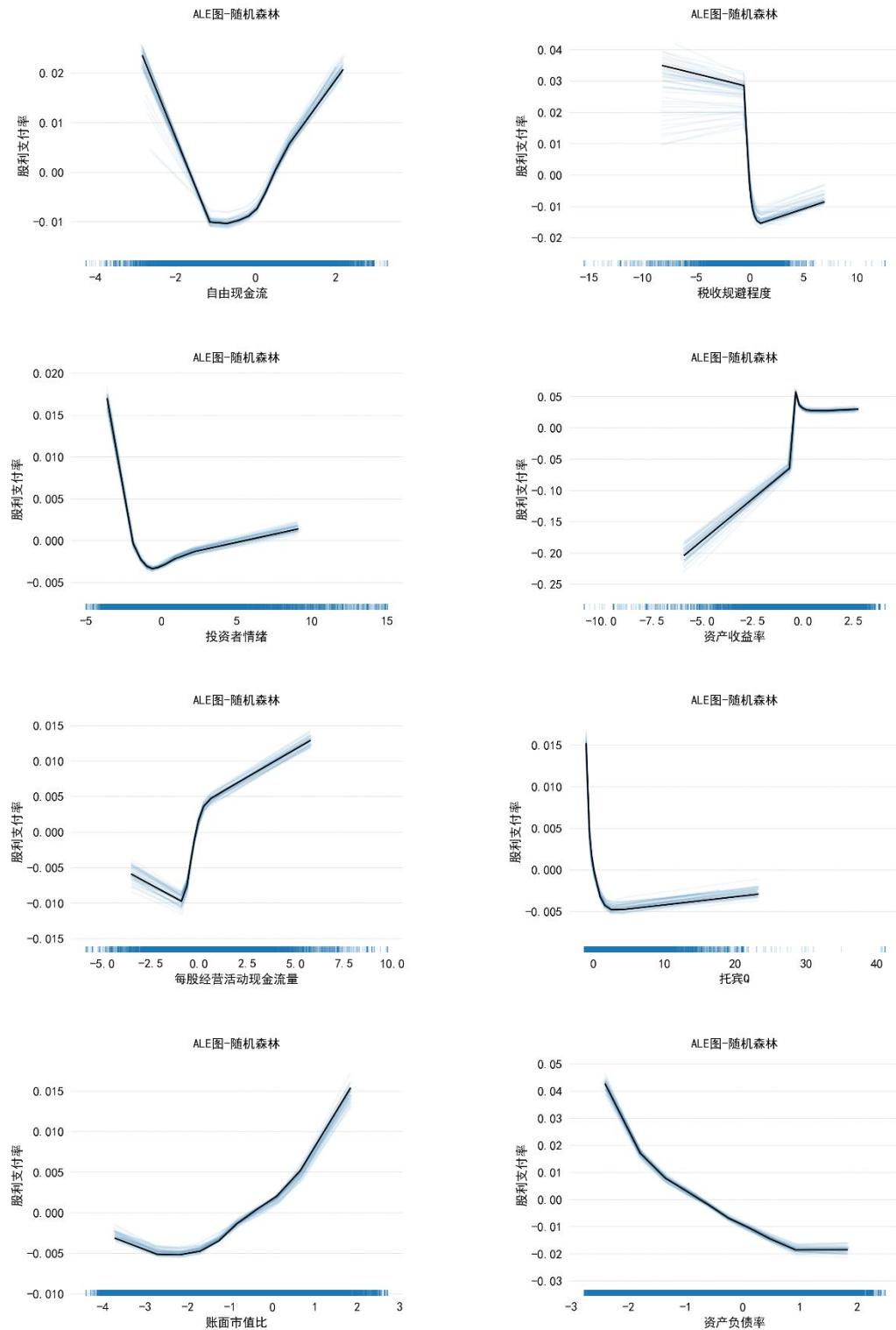
附录 5: 未列示的特征变量 ALE 图

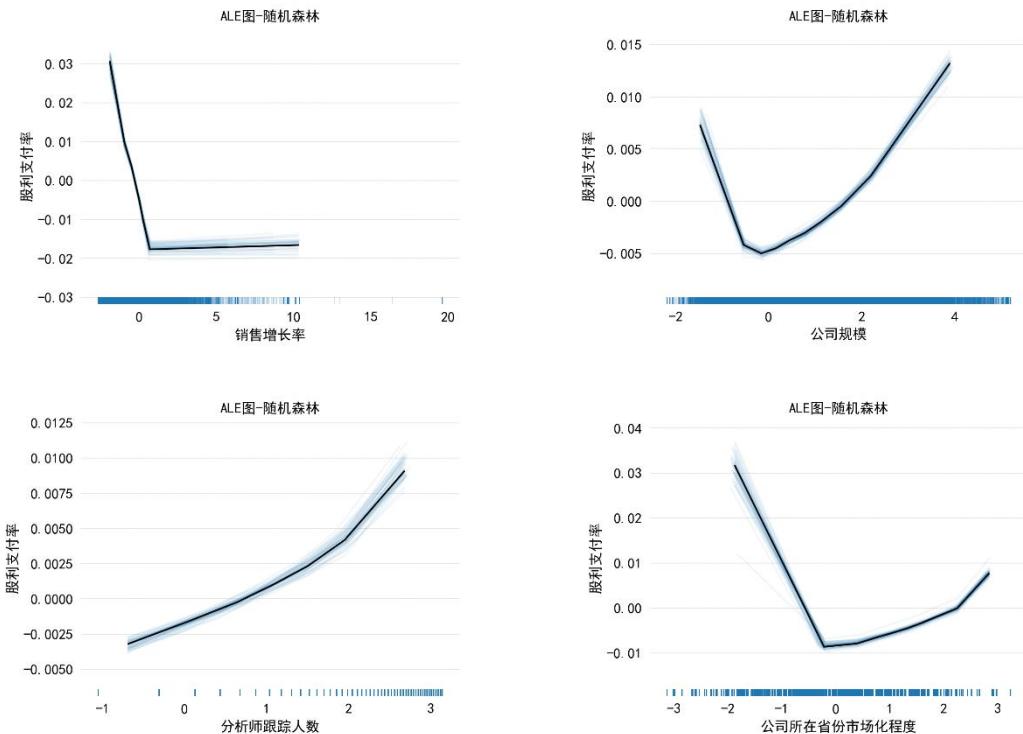
从结果来看, 除虚拟特征变量以及部分连续变量外, 其余特征与公司股利支付水平之间的关系呈现出非线性的特点, 这表明线性假定不能完全解释特征变量与公司股利水平之间的相关关系, 也在一定程度上解释了为何多元线性回归和 LASSO 等线性研究方法在公司分红行为的预测识别效果方面不如随机森林等非线性研究方法。

(一) 随机森林的 ALE 图:

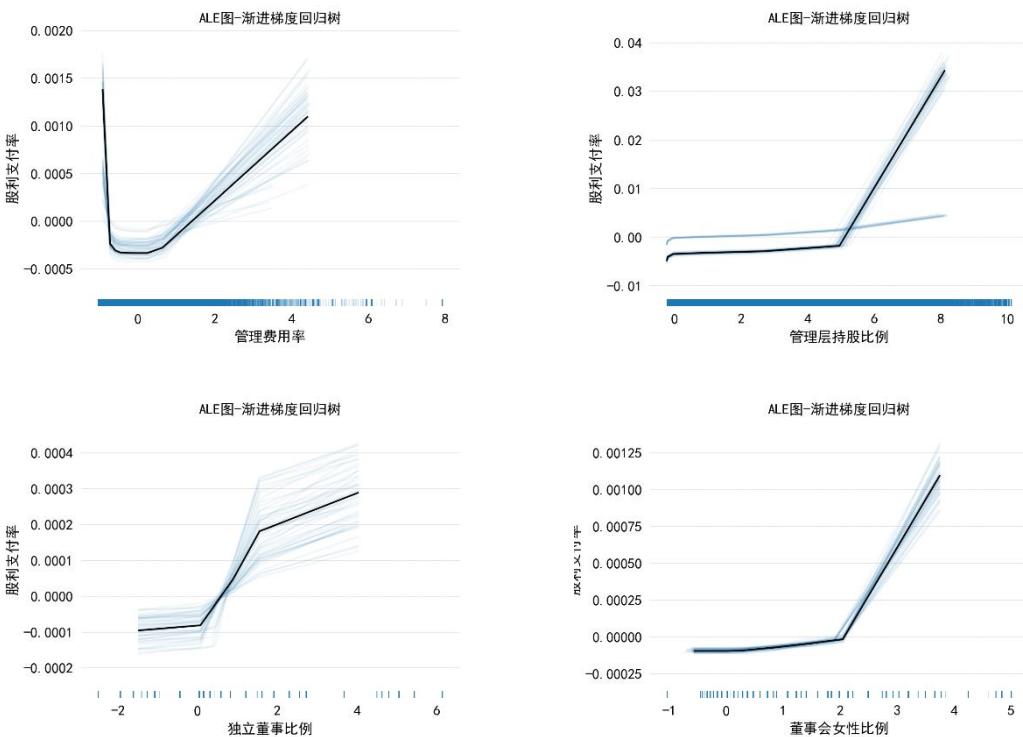


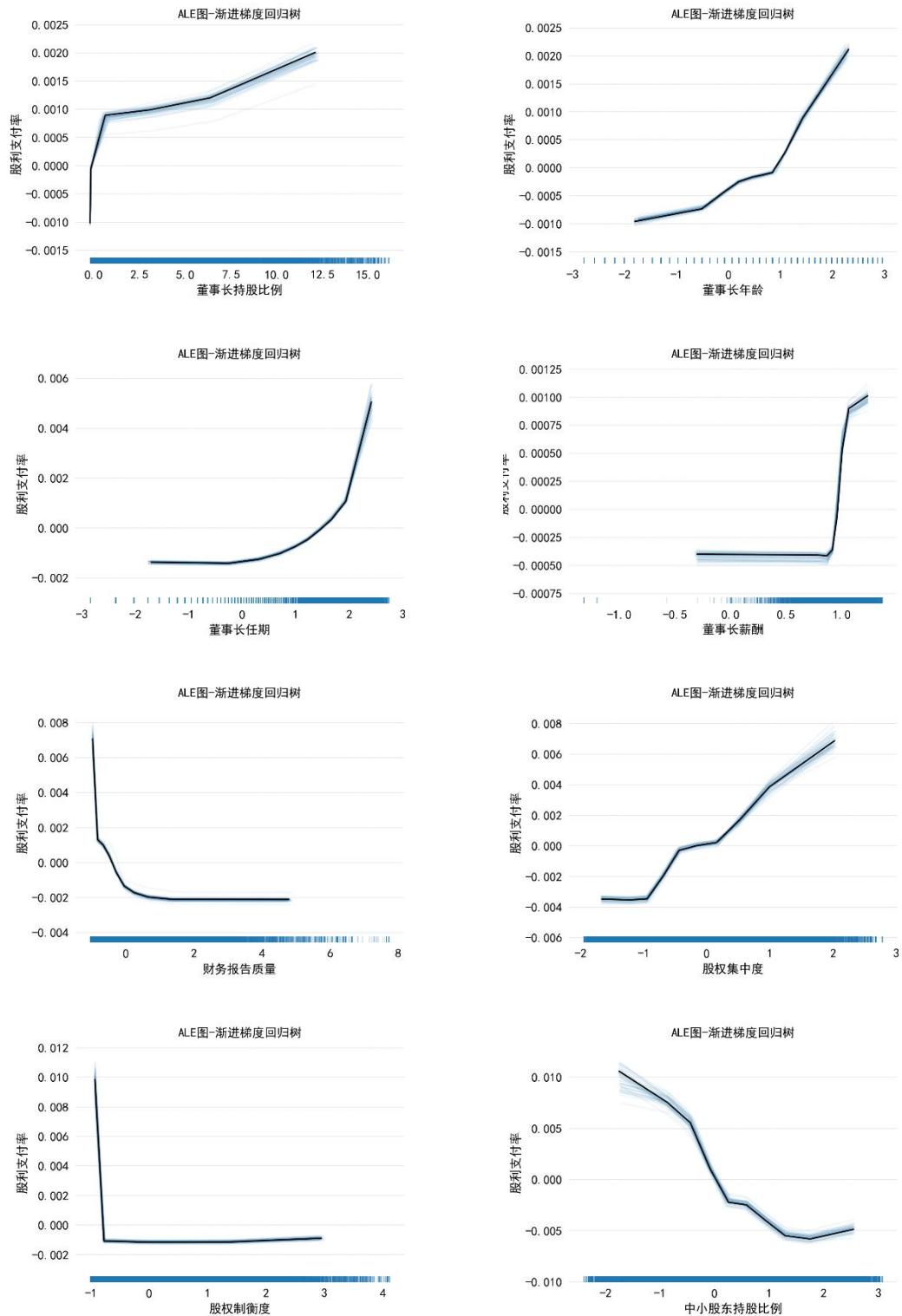


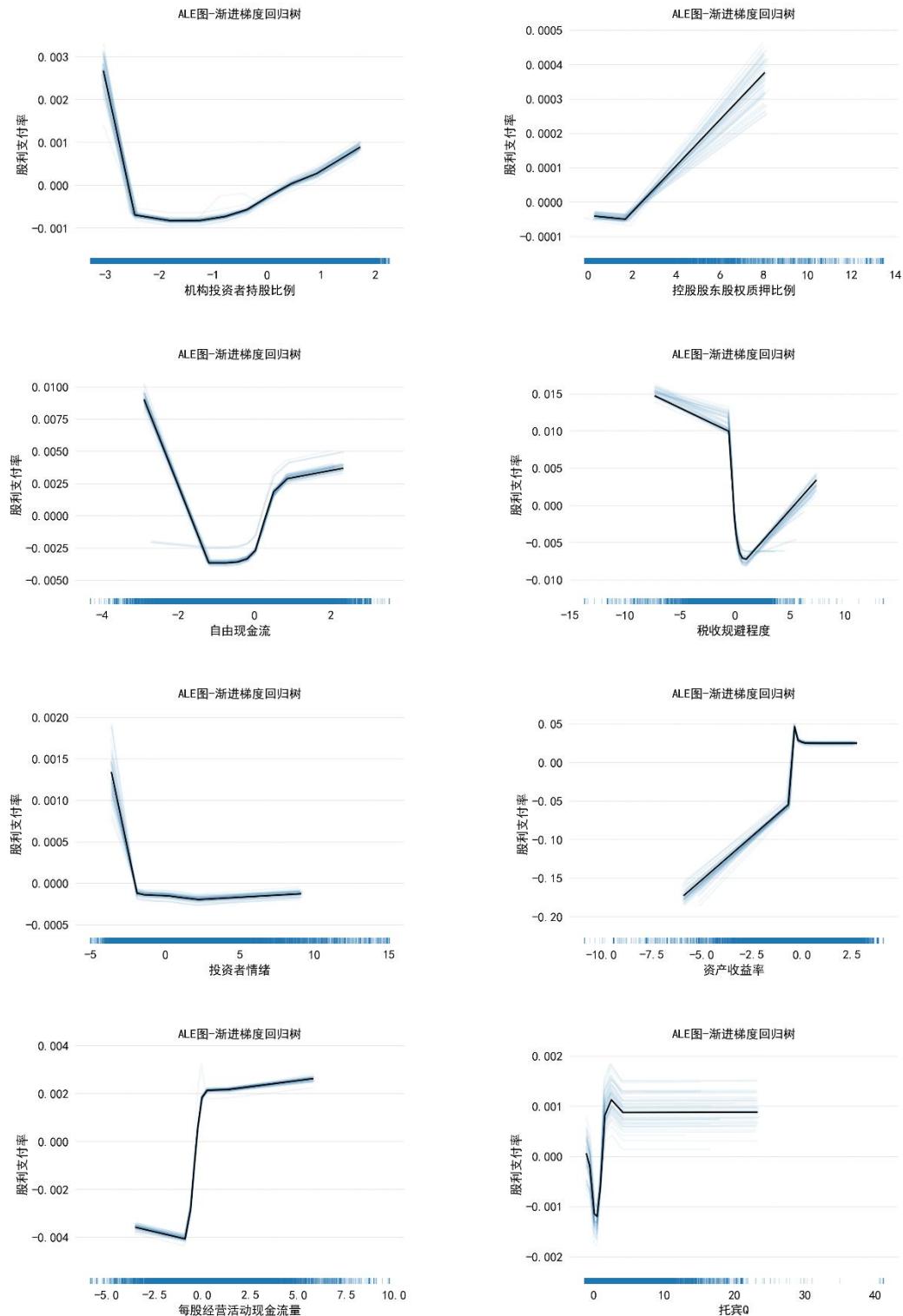


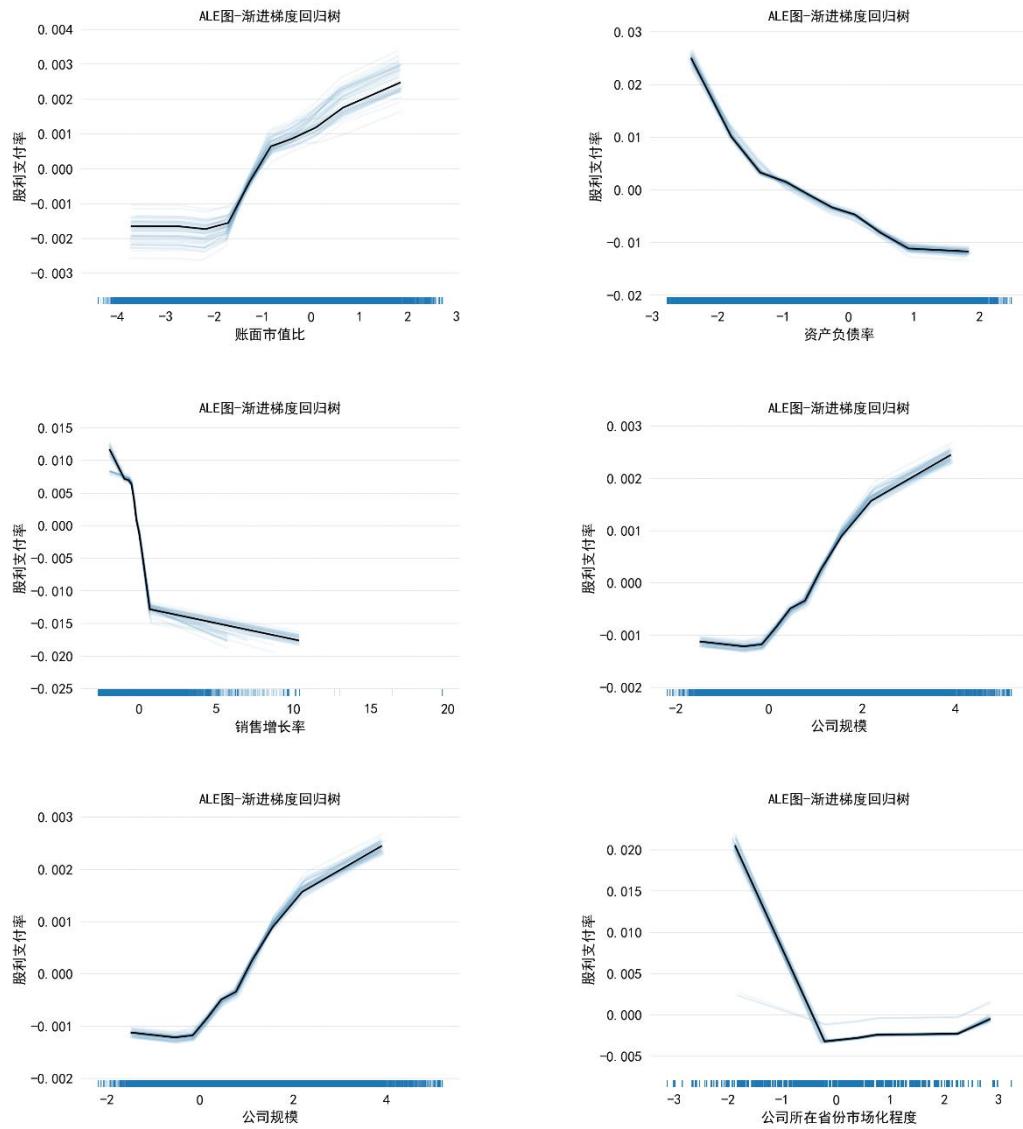


(二) 漐进梯度回归树的 ALE 图:









附录 6: 进一步分析部分实证结果

附表 5: 股利政策变更前后的变量相对重要性

Panel A 2007 年的变量相对重要性排序

排 序	随机森林 变量	相对重要性	渐进梯度回归树 变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	15.9569%	<i>Dividend_lag_{i,t-1}</i>	17.6938%
2	<i>ROA_{i,t}</i>	8.1345%	<i>ROA_{i,t}</i>	9.0593%
3	<i>Retainedearn_ratio_{i,t}</i>	6.1425%	<i>Retainedearn_ratio_{i,t}</i>	5.8490%
4	<i>Tax_avoid_{i,t}</i>	4.6026%	<i>Tax_ratio_{i,t}</i>	5.2753%
5	<i>Tax_ratio_{i,t}</i>	4.2559%	<i>Lev_{i,t}</i>	3.1939%
6	<i>Lev_{i,t}</i>	2.9513%	<i>Growth_{i,t}</i>	3.1399%
7	<i>Tax_volatility_{i,t}</i>	2.9430%	<i>Da_abs_{i,t}</i>	3.0456%
8	<i>Growth_{i,t}</i>	2.9066%	<i>Tax_avoid_{i,t}</i>	2.9018%
9	<i>Minorityrate_{i,t}</i>	2.5916%	<i>Freecash_{i,t}</i>	2.6974%
10	<i>Freecash_{i,t}</i>	2.5602%	<i>Tax_volatility_{i,t}</i>	2.6226%

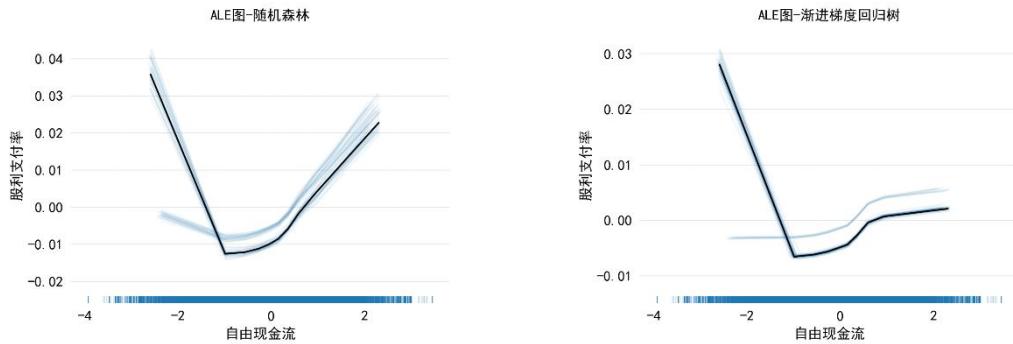
Panel B 2008-2011 年的变量相对重要性排序

排 序	随机森林 变量	相对重要性	渐进梯度回归树 变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	12.8220%	<i>Dividend_lag_{i,t-1}</i>	20.5659%
2	<i>ROA_{i,t}</i>	7.5592%	<i>ROA_{i,t}</i>	10.4965%
3	<i>Retainedearn_ratio_{i,t}</i>	6.7935%	<i>Retainedearn_ratio_{i,t}</i>	8.1032%
4	<i>Tax_ratio_{i,t}</i>	4.3135%	<i>Tax_ratio_{i,t}</i>	4.2858%
5	<i>Tax_avoid_{i,t}</i>	3.3657%	<i>Tunneling_{i,t}</i>	3.3504%
6	<i>Tunneling_{i,t}</i>	3.2147%	<i>Lev_{i,t}</i>	3.3375%
7	<i>Lev_{i,t}</i>	3.1810%	<i>Growth_{i,t}</i>	3.2320%
8	<i>Tax_volatility_{i,t}</i>	3.1186%	<i>Institution_{i,t}</i>	2.9281%
9	<i>Growth_{i,t}</i>	3.0448%	<i>Minorityrate_{i,t}</i>	2.8460%
10	<i>Institution_{i,t}</i>	2.7838%	<i>Da_abs_{i,t}</i>	2.3986%

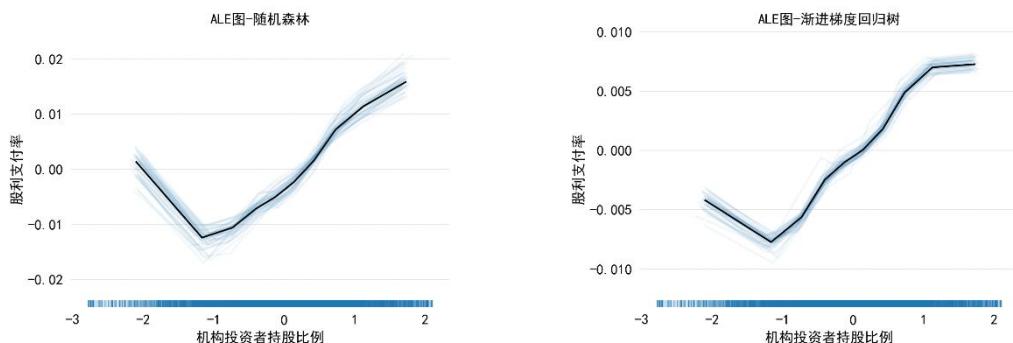
Panel C 2013 年之后的变量相对重要性排序

排序	随机森林 变量	相对重要性	渐进梯度回归树 变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	19.3600%	<i>Dividend_lag_{i,t-1}</i>	34.3328%
2	<i>ROA_{i,t}</i>	8.2599%	<i>ROA_{i,t}</i>	14.4949%
3	<i>Retainedearn_ratio_{i,t}</i>	6.1860%	<i>Retainedearn_ratio_{i,t}</i>	6.4858%

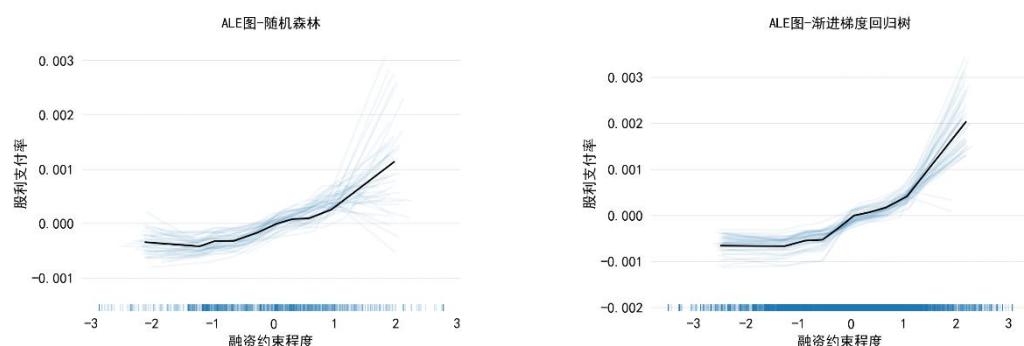
4	<i>Tax_avoid_{i,t}</i>	4.6413%	<i>Tax_ratio_{i,t}</i>	5.6211%
5	<i>Tax_ratio_{i,t}</i>	4.6197%	<i>Growth_{i,t}</i>	2.5649%
6	<i>Tax_volatility_{i,t}</i>	2.9894%	<i>Tax_avoid_{i,t}</i>	2.1349%
7	<i>Growth_{i,t}</i>	2.9541%	<i>Lev_{i,t}</i>	1.9634%
8	<i>Freeecash_{i,t}</i>	2.6948%	<i>Da_abs_{i,t}</i>	1.9279%
9	<i>Lev_{i,t}</i>	2.4506%	<i>Institution_{i,t}</i>	1.9016%
10	<i>Minorityrate_{i,t}</i>	2.4466%	<i>Top1_{i,t}</i>	1.8669%



附图 1：2007 年-自由现金流的累计局部效应图



附图 2：2008-2011 年-机构投资者持股比例的累计局部效应图



附图 3：2008-2011 年-融资约束的累计局部效应图

附表 6: 是否采取股利迎合策略企业的相对重要性差异

Panel A 采取股利迎合策略企业的变量相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	$ROA_{i,t}$	8.0376%	$ROA_{i,t}$	12.0315%
2	$Tax_ratio_{i,t}$	7.3299%	$Tax_ratio_{i,t}$	10.9309%
3	$Dividend_lag_{i,t-1}$	5.8604%	$Dividend_lag_{i,t-1}$	6.1581%
4	$Tax_avoid_{i,t}$	4.9604%	$Retainedearn_ratio_{i,t}$	5.6696%
5	$Retainedearn_ratio_{i,t}$	4.7089%	$Managefee_ratio_{i,t}$	4.9028%
6	$Growth_{i,t}$	3.7757%	$Tax_avoid_{i,t}$	3.7545%
7	$Da_abs_{i,t}$	3.7207%	$Da_abs_{i,t}$	3.6818%
8	$Tax_volatility_{i,t}$	3.5408%	$Top1_{i,t}$	3.5503%
9	$Lnsize_{i,t}$	3.1946%	$Tax_volatility_{i,t}$	3.4294%
10	$Managefee_ratio_{i,t}$	3.1872%	$Sentiment_inv_{i,t}$	3.3592%

Panel B 未采取股利迎合策略企业的相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	$Dividend_lag_{i,t-1}$	18.0421%	$Dividend_lag_{i,t-1}$	22.4632%
2	$ROA_{i,t}$	8.0493%	$ROA_{i,t}$	11.4067%
3	$Retainedearn_ratio_{i,t}$	6.3130%	$Retainedearn_ratio_{i,t}$	6.0259%
4	$Tax_ratio_{i,t}$	4.3981%	$Tax_ratio_{i,t}$	4.1652%
5	$Tax_avoid_{i,t}$	4.0452%	$Growth_{i,t}$	3.1591%
6	$Tax_volatility_{i,t}$	3.2317%	$Lev_{i,t}$	2.8150%
7	$Growth_{i,t}$	2.9092%	$Tax_volatility_{i,t}$	2.6754%
8	$Lev_{i,t}$	2.8506%	$Tax_avoid_{i,t}$	2.4885%
9	$Freecash_{i,t}$	2.5886%	$Freecash_{i,t}$	2.3959%
10	$Tunneling_{i,t}$	2.4688%	$Da_abs_{i,t}$	2.3823%

附表 7: 不同现金流状况企业的变量相对重要性

Panel A 现金流状况好的企业的变量相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	18.8571%	<i>Dividend_lag_{i,t-1}</i>	23.8517%
2	<i>Retainedearn_ratio_{i,t}</i>	6.7290%	<i>ROA_{i,t}</i>	6.7580%
3	<i>ROA_{i,t}</i>	6.2112%	<i>Retainedearn_ratio_{i,t}</i>	6.1411%
4	<i>Tax_ratio_{i,t}</i>	4.0652%	<i>Tax_ratio_{i,t}</i>	4.2792%
5	<i>Tax_volatility_{i,t}</i>	3.5926%	<i>Freecash_{i,t}</i>	3.2329%
6	<i>Tax_avoid_{i,t}</i>	3.5338%	<i>Growth_{i,t}</i>	3.1894%
7	<i>Freecash_{i,t}</i>	2.8857%	<i>Tax_volatility_{i,t}</i>	2.7922%
8	<i>Growth_{i,t}</i>	2.8779%	<i>Tax_avoid_{i,t}</i>	2.6656%
9	<i>Lev_{i,t}</i>	2.7612%	<i>Lev_{i,t}</i>	2.5341%
10	<i>Institution_{i,t}</i>	2.5752%	<i>Da_abs_{i,t}</i>	2.5045%

Panel B 现金流状况差的企业的变量相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	15.4083%	<i>Dividend_lag_{i,t-1}</i>	18.6686%
2	<i>ROA_{i,t}</i>	8.1167%	<i>ROA_{i,t}</i>	11.7504%
3	<i>Retainedearn_ratio_{i,t}</i>	5.5259%	<i>Retainedearn_ratio_{i,t}</i>	5.2033%
4	<i>Tax_ratio_{i,t}</i>	4.9388%	<i>Tax_ratio_{i,t}</i>	4.7696%
5	<i>Tax_avoid_{i,t}</i>	4.0503%	<i>Tax_volatility_{i,t}</i>	3.1415%
6	<i>Tax_volatility_{i,t}</i>	3.6093%	<i>Growth_{i,t}</i>	2.9703%
7	<i>Lev_{i,t}</i>	2.9967%	<i>Lev_{i,t}</i>	2.9620%
8	<i>Growth_{i,t}</i>	2.8223%	<i>Tax_avoid_{i,t}</i>	2.6469%
9	<i>Tunneling_{i,t}</i>	2.6450%	<i>Tunneling_{i,t}</i>	2.6392%
10	<i>Da_abs_{i,t}</i>	2.5229%	<i>Da_abs_{i,t}</i>	2.5331%

附表 8: 国有与非国有企业相对重要性差异

Panel A 国有企业变量相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	16.9430%	<i>Dividend_lag_{i,t-1}</i>	21.3160%
2	<i>ROA_{i,t}</i>	8.2218%	<i>ROA_{i,t}</i>	10.1093%
3	<i>Retainedearn_ratio_{i,t}</i>	7.5081%	<i>Retainedearn_ratio_{i,t}</i>	6.6845%
4	<i>Tax_ratio_{i,t}</i>	4.9543%	<i>Tax_ratio_{i,t}</i>	4.8458%
5	<i>Tax_avoid_{i,t}</i>	3.7043%	<i>Lev_{i,t}</i>	2.9075%
6	<i>Tax_volatility_{i,t}</i>	3.2410%	<i>Btenure_{i,t}</i>	2.4802%
7	<i>Lev_{i,t}</i>	2.9027%	<i>Tax_avoid_{i,t}</i>	2.4527%
8	<i>Institution_{i,t}</i>	2.6393%	<i>Tax_volatility_{i,t}</i>	2.4282%
9	<i>Growth_{i,t}</i>	2.5345%	<i>Growth_{i,t}</i>	2.4266%
10	<i>Freeecash_{i,t}</i>	2.3653%	<i>Sharebalance_{i,t}</i>	2.3548%

Panel B 非国有企业变量相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	17.7852%	<i>Dividend_lag_{i,t-1}</i>	22.1164%
2	<i>ROA_{i,t}</i>	6.4793%	<i>ROA_{i,t}</i>	8.3138%
3	<i>Retainedearn_ratio_{i,t}</i>	5.3597%	<i>Retainedearn_ratio_{i,t}</i>	4.8725%
4	<i>Tax_ratio_{i,t}</i>	4.2489%	<i>Tax_ratio_{i,t}</i>	4.2965%
5	<i>Tax_avoid_{i,t}</i>	3.8761%	<i>Tax_volatility_{i,t}</i>	3.4006%
6	<i>Tax_volatility_{i,t}</i>	3.5294%	<i>Growth_{i,t}</i>	3.3371%
7	<i>Growth_{i,t}</i>	2.9881%	<i>Tunneling_{i,t}</i>	2.8123%
8	<i>Tunneling_{i,t}</i>	2.8571%	<i>Tax_avoid_{i,t}</i>	2.5800%
9	<i>Lev_{i,t}</i>	2.7860%	<i>Lev_{i,t}</i>	2.5507%
10	<i>Freeecash_{i,t}</i>	2.7580%	<i>Minorityrate_{i,t}</i>	2.5401%

附表9: 投资者现金股利情绪不同年份下企业的相对重要性排序

Panel A 现金股利情绪高的年份变量相对重要性排序

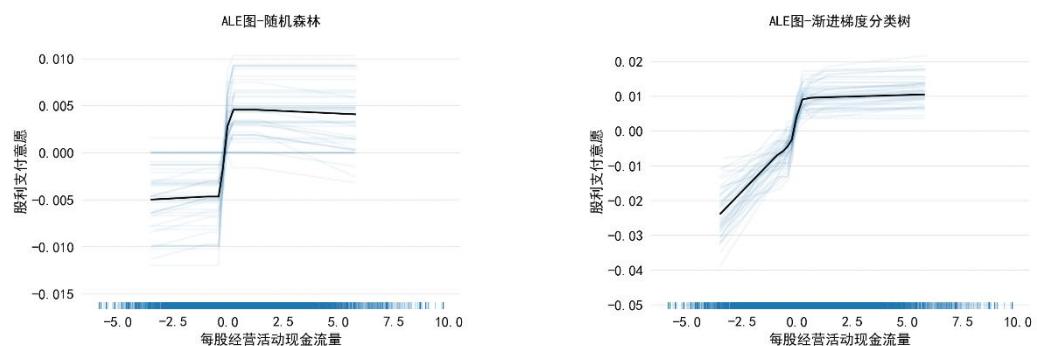
排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	14.9653%	<i>Dividend_lag_{i,t-1}</i>	15.8389%
2	<i>ROA_{i,t}</i>	7.4187%	<i>ROA_{i,t}</i>	9.6275%
3	<i>Retainedearn_ratio_{i,t}</i>	5.8655%	<i>Retainedearn_ratio_{i,t}</i>	5.5532%
4	<i>Tax_ratio_{i,t}</i>	4.1526%	<i>Tax_ratio_{i,t}</i>	4.8586%
5	<i>Tax_avoid_{i,t}</i>	3.9193%	<i>Lev_{i,t}</i>	3.3727%
6	<i>Lev_{i,t}</i>	3.3901%	<i>Da_abs_{i,t}</i>	3.3686%
7	<i>Tax_volatility_{i,t}</i>	3.1633%	<i>Growth_{i,t}</i>	3.1151%
8	<i>Growth_{i,t}</i>	3.1426%	<i>Sharebalance_{i,t}</i>	2.8984%
9	<i>Tunneling_{i,t}</i>	2.9034%	<i>Tunneling_{i,t}</i>	2.8466%
10	<i>Da_abs_{i,t}</i>	2.6636%	<i>Tax_avoid_{i,t}</i>	2.7285%

Panel B 现金股利情绪低的年份相对重要性排序

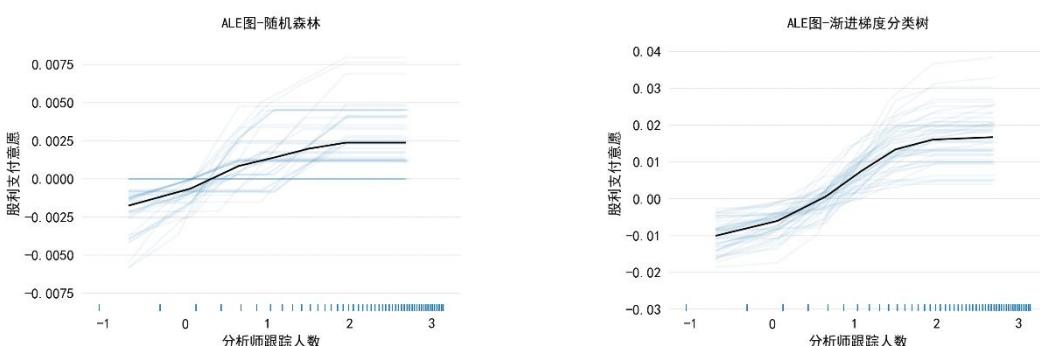
排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	<i>Dividend_lag_{i,t-1}</i>	16.0011%	<i>Dividend_lag_{i,t-1}</i>	17.6778%
2	<i>ROA_{i,t}</i>	8.5201%	<i>ROA_{i,t}</i>	10.0548%
3	<i>Retainedearn_ratio_{i,t}</i>	6.1090%	<i>Retainedearn_ratio_{i,t}</i>	5.3657%
4	<i>Tax_avoid_{i,t}</i>	4.7910%	<i>Tax_ratio_{i,t}</i>	5.0231%
5	<i>Tax_ratio_{i,t}</i>	4.5653%	<i>Growth_{i,t}</i>	3.0894%
6	<i>Tax_volatility_{i,t}</i>	3.1355%	<i>Da_abs_{i,t}</i>	3.0809%
7	<i>Growth_{i,t}</i>	2.9110%	<i>Tax_volatility_{i,t}</i>	2.9852%
8	<i>Lev_{i,t}</i>	2.8748%	<i>Lev_{i,t}</i>	2.9526%
9	<i>Freecash_{i,t}</i>	2.5282%	<i>Tax_avoid_{i,t}</i>	2.8994%
10	<i>Minorityrate_{i,t}</i>	2.4539%	<i>Freecash_{i,t}</i>	2.6910%

附表 10: 响应变量为是否发放现金股利的变量相对重要性排序

排序	随机森林		渐进梯度回归树	
	变量	相对重要性	变量	相对重要性
1	$ROA_{i,t}$	16.9873%	$Dividend_lag_{i,t-1}$	15.1441%
2	$Dividend_lag_{i,t-1}$	14.6575%	$ROA_{i,t}$	11.7178%
3	$Retainedearn_ratio_{i,t}$	12.2196%	$Retainedearn_ratio_{i,t}$	9.8536%
4	$Tax_avoid_{i,t}$	7.3019%	$Analyst_num_{i,t}$	4.6866%
5	$Tax_volatility_{i,t}$	5.2205%	$Growth_{i,t}$	4.1946%
6	$Tax_ratio_{i,t}$	4.0063%	$Cashflow_{i,t}$	4.0037%
7	$Cashflow_{i,t}$	2.2360%	$Freecash_{i,t}$	3.3591%
8	$Growth_{i,t}$	2.1712%	$Tax_ratio_{i,t}$	3.1654%
9	$Analyst_num_{i,t}$	2.0736%	$Tax_volatility_{i,t}$	3.0161%
10	$Lnsize_{i,t}$	1.9946%	$Bage_{i,t}$	2.9082%



附图 4: 经营活动现金流的部分依赖图



附图 5: 分析师跟踪人数的部分依赖图

附录 7: 稳健性检验结果

(1) 变更响应变量。在主检验中, 本文采用股利支付率作为响应变量。为进一步提高研究结论的普适性, 本文还采用股利收益率作为替代性响应变量。股利收益率是指每股现金股利与每股总资产的比值。结果表明, 替换响应变量后的研究结论与主检验一致。

附表 11: 变更响应变量后的模型拟合结果

	R^2_{ls} (1)	R^2_{oos} (2)	EVS_{oos} (3)	MSE_{oos} (4)	MAE_{oos} (5)	$MedAE_{oos}$ (6)
多元线性回归	0.5291	0.4532	0.4700	1.7539	0.8723	0.5719
LASSO	0.5305	0.4570	0.4722	1.7450	0.8606	0.5616
决策树	0.1439	0.1047	0.1084	2.8594	1.1077	0.8189
支持向量机	0.6005	0.4567	0.4763	1.7429	0.7454	0.3844
渐进梯度回归树	0.8349	0.6246	0.6313	1.1943	0.6359	0.3424
随机森林	0.9455	0.6014	0.6156	1.2637	0.6854	0.3999

(2) 变更样本划分方法。在主检验中, 本文采用了一年训练窗口期、一年测试窗口期的方式进行滚动预测。为避免因样本划分方法而导致的结果偏误, 本文还使用了以下样本划分方法重新拟合模型: ①按 7:3 的比例随机划分训练集和测试集; ②按 5:5 的比例随机划分训练集和测试集。变更样本划分方法后, 以渐进梯度回归树和随机森林为代表的集成学习方法对于股利支付行为的预测识别能力仍显著高于多元线性回归, 本文的研究结论保持不变。

附表 12: 变更样本划分后的模型拟合结果

Panel A 按 7:3 的比例随机划分训练集和测试集						
	R^2_{ls} (1)	R^2_{oos} (2)	EVS_{oos} (3)	MSE_{oos} (4)	MAE_{oos} (5)	$MedAE_{oos}$ (6)
多元线性回归	0.2300	0.2233	0.2234	0.0828	0.1743	0.1161
LASSO	0.2138	0.2108	0.2109	0.0841	0.1762	0.1226
决策树	0.3230	0.0779	0.0780	0.0983	0.1852	0.1237
支持向量机	0.3852	0.2375	0.2447	0.0813	0.1589	0.0934
渐进梯度回归树	0.9477	0.3108	0.3110	0.0735	0.1520	0.0818
随机森林	0.9097	0.3269	0.3274	0.0718	0.1499	0.0812
Panel B 按 5:5 的比例随机划分训练集和测试集						
	R^2_{ls} (1)	R^2_{oos} (2)	EVS_{oos} (3)	MSE_{oos} (4)	MAE_{oos} (5)	$MedAE_{oos}$ (6)
多元线性回归	0.2238	0.2297	0.2297	0.0808	0.1738	0.1172
LASSO	0.2077	0.2179	0.2179	0.0820	0.1756	0.1235
决策树	0.4212	0.0667	0.0667	0.0979	0.1761	0.1026
支持向量机	0.3924	0.2417	0.2485	0.0795	0.1598	0.0958

渐进梯度回归树	0.9727	0.3129	0.3133	0.0721	0.1526	0.0836
随机森林	0.9074	0.3276	0.3282	0.0705	0.1503	0.0840

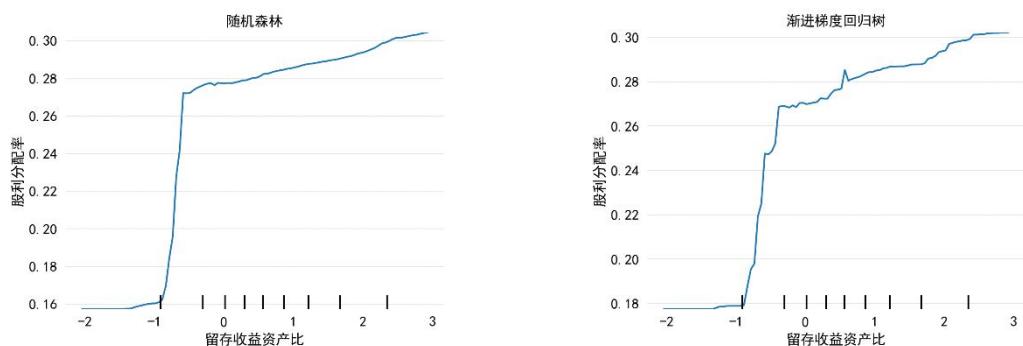
(3) 变更样本期间。鉴于 2007 年会计准则发生了变化, 为避免因准则变化导致结果发生改变, 本文还将样本起点调整至 2007 年。变更样本期间后的结果依旧稳健。

附表 13: 变更样本期间后的模型拟合结果

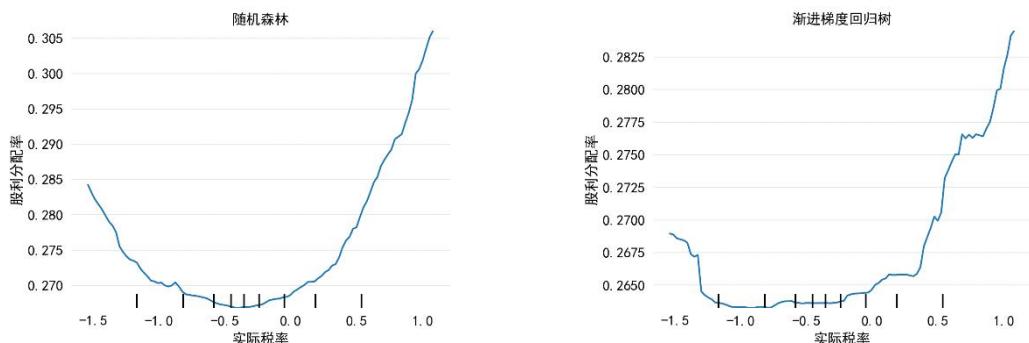
	R _{Is} (1)	R _{OOS} (2)	EVS _{OOS} (3)	MSE _{OOS} (4)	MAE _{OOS} (5)	MedAE _{OOS} (6)
多元线性回归	0.2407	0.1585	0.1807	0.0892	0.1843	0.1259
LASSO	0.2296	0.1794	0.1889	0.0872	0.1801	0.1225
决策树	0.0455	0.0132	0.0219	0.1039	0.2134	0.1791
支持向量机	0.5902	0.1649	0.1694	0.0888	0.1788	0.1162
渐进梯度回归树	0.6028	0.2371	0.2523	0.0790	0.1646	0.1010
随机森林	0.9002	0.2333	0.2537	0.0790	0.1620	0.0952

附录 8: 稳健性检验中部分依赖图的具体图像

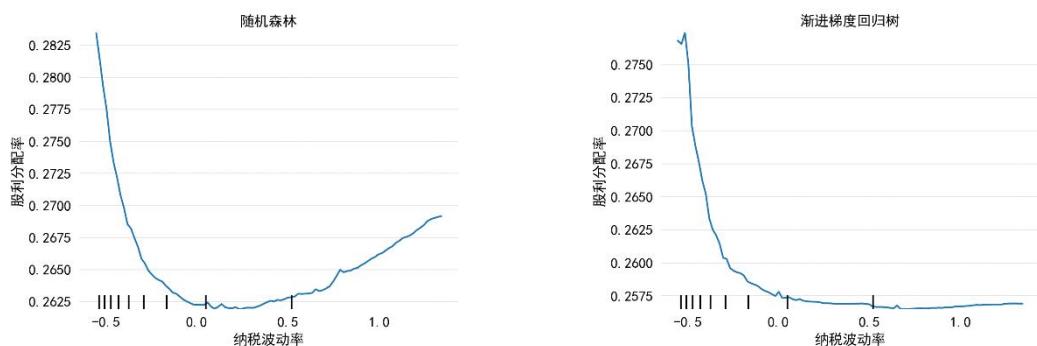
关于累计局部效应图的稳健性检验。除了利用累计局部效应图外，集成学习还常使用部分依赖图分析特征变量对于响应变量的影响模式。该方法的核心思想是，在其他特征保持恒定的条件下，观察某一特定变量的变化如何影响企业的股利支付水平，并将这种关系以直观的图像形式展示出来。但由于部分依赖图需要假设所考察的特征与其他特征不相关，而这一条件在现实中常常不符合，因而本文在主检验中采用累计局部效应图进行详细分析。为了进一步增强累计局部效应图的稳健性，本文将其与部分依赖图的图像模式进行了对比。结果显示，部分依赖图与累计局部效应图的图像模式基本一致，进一步增强了结果的稳健性。



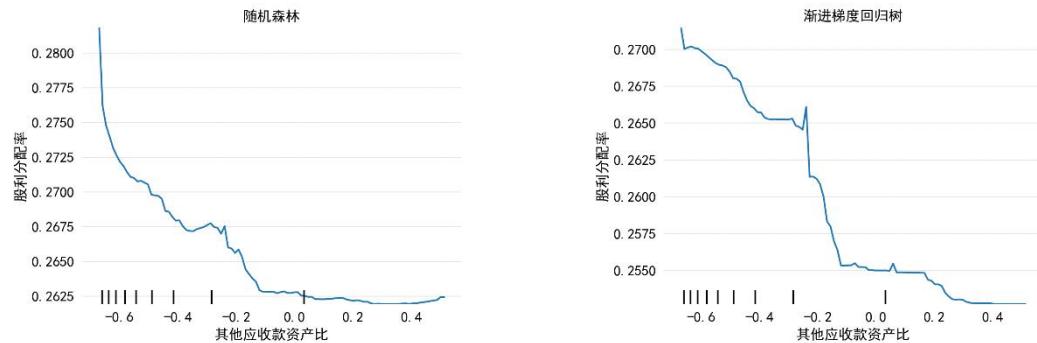
附图 6：正文图 1 对应的部分依赖图



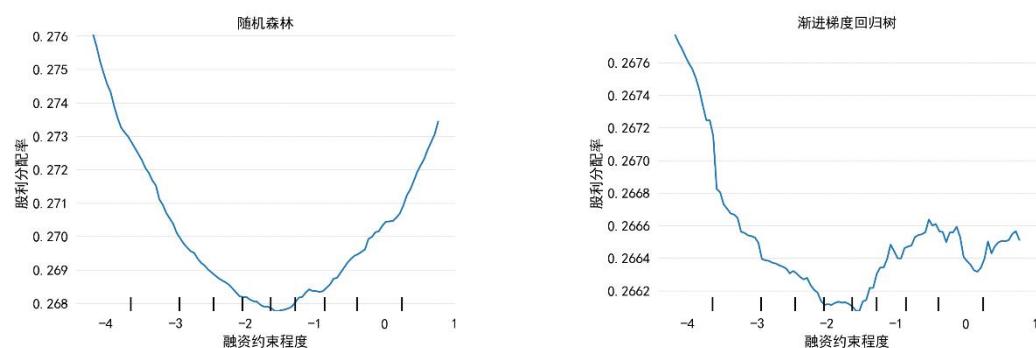
附图 7：正文图 2 对应的部分依赖图



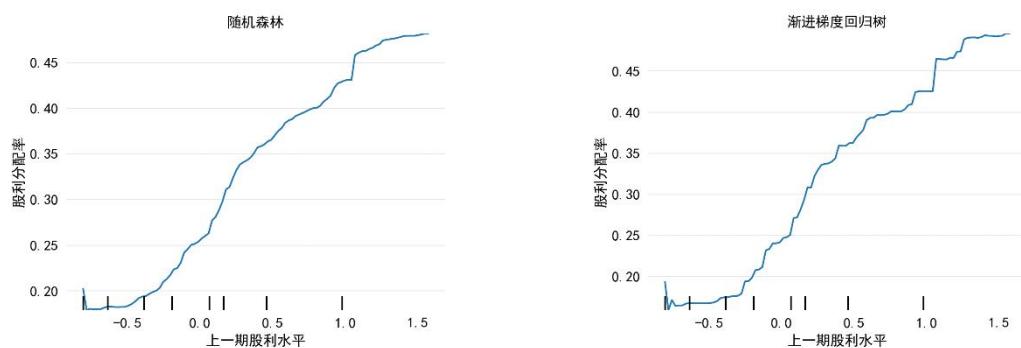
附图 8：正文图 3 对应的部分依赖图



附图 9：正文图 4 对应的部分依赖图



附图 10：正文图 5 对应的部分依赖图



附图 11：正文图 6 对应的部分依赖图

注: 该附录是期刊所发表论文的组成部分, 同样视为作者公开发表的内容。如研究中使用该附录中的内容, 请务必在研究成果上注明引文和下载附件出处。

引用示例:

参考文献引用范例:

- [1] 朱军. 技术吸收、政府推动与中国全要素生产率提升[J]. 中国工业经济. 2017, (1):5-24.

如果研究中使用了未在《中国工业经济》纸质版刊发、但在杂志网站上正式公开发表的数字内容(包括数据、程序、附录文件), 请务必在研究成果正文中注明:

数据(及程序等附件)来自朱军(2017), 参见在《中国工业经济》网站(<http://ciejournal.ajcass.org>)附件下载。