

毕设

* 解释变量

总之，本文采用金融化率公式如下：
$$\text{金融资产比例 (FAR)} = \frac{\text{金融资产}}{\text{总资产}} = \frac{(\text{货币资金} + \text{交易金融资产} + \text{衍生金融资产} + \text{应收利息} + \text{应收股利} + \text{债权投资} + \text{可供出售金融资产} + \text{持有至到期投资} + \text{长期股权投资} + \text{投资性房地产})}{\text{总资产}}$$

• 被解释变量

$$\text{经营收益率 (RO)} = \frac{\text{经营收入}}{\text{经营性资产}} = \frac{(\text{营业利润} - \text{金融收益})}{(\text{总资产} - \text{金融资产})}$$

• 控制变量

资本结构 (ALT) = 总负债/总资产
股权集中度(OC)：十大公司的股份占总股本的比重
企业规模(Size): 公司规模=ln 期末总资产
营业收入增长率(GOI)

表 4-1 模型变量表

变量类型	变量代码	变量名称	计量方法
因变量	RO	经营资产收益率	营业利润-金融收益/总资产-金融资产
自变量	FAR	金融资产比例	金融资产/总资产
控制变量	GOI	营业收入增长率	营业收入增长额/上年营业收入总额
	OC	股权集中度	前十大股东持股比例
	SIZE	公司规模	期末总资产的自然对数
	ALT	资本结构	总负债/总资产

额外要求:

4.2 样本选取与数据说明

本文选择 2010 年—2021 年为样本期，选取所有 A 股制造业上市公司作为研究样本，选择有关公司公开的年度数据作为研究样本，通过国泰安 CSMAR 数据库，搜集相关资料。本文从 2010 年开始选择数据，根据国泰安数据库检索到数据可查最新的 2021 年。并对样本中的公司进行了筛选，只选择了非 ST 的上市公司。同时采用 stata 16 的 Winsorize 命令，对变量进行 1%分位和 99%分位的缩尾处理(将小于百分位数 1%和大于百分位 99%的数替换为百分位 1%的数值)，规避了极端数值的干扰。对以下的公司进行了剔除：(1)数据缺失的公司；(2)被标注 ST、*ST 的公司；(3)财务报表标 B 的公司。最终得到 21855 个样本观测值，利用 stata 16 和 excel 2010 进行数据处理与分析。

2014.1-2023.12 时间 A股 非ST 制造业 合并报表 年度

1. 数据处理

```
gen year = substr(date,1,4)
destring year, replace
keep if typrep=="A"
drop if strmatch( shortname,"*ST*" )
sort stkcd year
```

数据相加

```
egen s=rowtotal(a1 a2 a3 a4)
```

收尾操作

```
winsor2 资本结构 公司规模 股权集中度 营业收入增长率 经营资产收益率 金融资产比例 cut(1 99)
```

2. 研究分析

2.1描述性统计分析

```
1. sum 经营资产收益率 资本结构 公司规模 股权集中度 营业收入增长率 金融资产比例
```

```
2. tabstat 经营资产收益率 资本结构 公司规模 股权集中度 营业收入增长率 金融资产比例, stats (N mean max sd p75 p50 p25 p1 min)
```

2.2 相关性分析：*显著相关 系数小于0.8不存在多重共线性，星号代表显著相关性

```
pwcorr_a 经营资产收益率 资本结构 公司规模 股权集中度 营业收入增长率 金融资产比例, star1(0.001) star5(0.01) star10(0.05)
```

相关性系数和显著性水平

在统计学中，皮尔逊相关系数(Pearson correlation coefficient) ， 又称皮尔逊积矩相关系数 (Pearson product-moment correlation coefficient, 简称 PPMCC或PCCs) <很多英文文献中的叫法>， 是用于度量两个变量X和Y之间的相关（线性相关）， 其值介于-1与1之间。

P值，也就是Sig值或显著性值。如果P值小于0.01即说明某件事情的发生至少有99%的把握，如果P值小于0.05（并且大于0.01）则说明某件事情的发生至少有95%的把握。当 $P < 0.01$ 或 $P < 0.05$ ，则为说明水平显著。

相关系数，是研究变量之间线性相关程度的量，用于说明两个变量之间是否存在相关关系，以及相关关系的紧密程度。分为pearson相关系数、Spearman相关系数。一般相关系数在0.7以上说明关系非常紧密；0.4-0.7之间说明关系紧密；0.2~0.4说明关系一般。

显著性回答的问题是它们之间是否有关系，说明得到的结果是不是偶然因素导致的（具有统计学意义）；相关系数回答的问题是相关程度强弱。

假如说我得到“ $P < 0.05$ ，相关系数 $R = 0.279$ ”，意味着二者之间确实（ $P < 0.05$ ）存在相关关系，而相关性为0.279。

而如果“ $P > 0.05$ 相关系数 $R = 0.799$ ”，则意味着二者之间相关性很强（ $R = 0.799$ ），而这个高相关的结果可能是偶然因素导致的，即不具有统计学意义。

2.3 多重共线性检验

```
reg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率  
estat vif
```

2.4 模型检验

- F检验：判断固定效应模型和OLS模型谁更优
- LM检验：判断随机效应模型和OLS模型谁更优
- Hausman检验：判断FE和RE模型谁更优

F检验

```
1. xtset stkcd year
```

```
2. xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率, fe
```

```
#固定效应
```

```
. xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率, fe
```

```
Fixed-effects (within) regression                Number of obs   =    20,273
Group variable: stkcd                          Number of groups =     3,407
```

```
R-squared:                                     Obs per group:
    Within = 0.2501                               min =          1
    Between = 0.3922                             avg =         6.0
    Overall = 0.3448                             max =        10
```

```
corr(u_i, Xb) = 0.0367                        F(5,16861)      =   1124.79
                                              Prob > F       =    0.0000
```

经营资产收益率	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
金融资产比例	.1672999	.0060442	27.68	0.000	.1554526	.1791472
资本结构	-.2435325	.006392	-38.10	0.000	-.2560614	-.2310035
公司规模	.0121611	.0013791	8.82	0.000	.009458	.0148643
股权集中度	.0008407	.0000828	10.15	0.000	.0006784	.001003
营业收入增长率	.0787163	.0016252	48.43	0.000	.0755307	.0819019
_cons	-.2119451	.0309867	-6.84	0.000	-.2726822	-.151208
sigma_u	.08254497					
sigma_e	.06964522					
rho	.58415616	(fraction of variance due to u_i)				

```
F test that all u_i=0: F(3406, 16861) = 5.91                Prob > F = 0.0000
```

```
.
```

下面的F能够看出FE模型是否显著与OLS, p (0.000) 小于0.001, 拒绝原假设

LM检验

```
1. xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率, re
```

```
2. xttest0
```

```
. xttest0
```

Breusch and Pagan Lagrangian multiplier test for random effects

经营资产收益率[stkcd,t] = $Xb + u[stkcd] + e[stkcd,t]$

Estimated results:

	Var	SD = sqrt(Var)
经营资~率	.0136891	.1170004
e	.0048505	.0696452
u	.0052586	.0725159

Test: $\text{Var}(u) = 0$

$\text{chibar2}(01) = 10076.32$
Prob > $\text{chibar2} = 0.0000$

同理P小于0.001,拒绝原假设, RE模型优于混合界面模型 (OLS)

Hausman检验

1. xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率, fe
2. est store fe
3. xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率, re
4. est store re
5. hausman fe re

```
. hausman fe re
```

	—— Coefficients ——		(b-B) Difference	sqrt(diag(V_b-V_B)) Std. err.
	(b) fe	(B) re		
金融资产比例	.1672999	.1766343	-.0093344	.0030486
资本结构	-.2435325	-.2368503	-.0066821	.0033319
公司规模	.0121611	.0125441	-.000383	.0009686
股权集中度	.0008407	.001204	-.0003632	.0000521
营业收入~率	.0787163	.078332	.0003843	.0004295

b = Consistent under H_0 and H_a ; obtained from xtreg.
B = Inconsistent under H_a , efficient under H_0 ; obtained from xtreg.

Test of H_0 : Difference in coefficients not systematic

$\text{chi2}(5) = (b-B)'[(V_b-V_B)^{-1}](b-B)$
= 57.35
Prob > $\text{chi2} = 0.0000$

同理P小于0.001,拒绝原假设, FE模型优于RE

2.4 异方差检验

- 异方差 (Heteroscedasticity) 是指随着自变量的变化, 因变量的方差也会发生变化的情况。在回归分析中, 异方差通常指因变量的方差不是恒定的, 而是随着自变量的不同取值而发生变化。这种情况下, 回归模型的残差 (预测值与观测值之间的差异) 的方差会随着自变量的变化而变大或变小。

为了判断是否需要稳健标准误robust

```
estat imtest,white  
logout, save(异方差检验) word replace: estat imtest, white
```

2.5 回归分析

双向固定效应模型: 个体和年份 (默认只固定个体)

```
xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率  
i.year, fe
```

##剧烈稳健标准误 robust 加上会消除异方差影响, 显著性降低

```
xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率  
i.year, fe robust
```

导出相应参数

```
xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率  
i.year, fe  
est store y1  
xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率  
, fe  
est store y2  
esttab y1 y2 using FE_both.rtf, nogap nocompress scalar(N F p) r2 ar2  
replace drop(*year*) b(4)
```

3. 异质性分析和稳步性检验

3.1 异质性分析

对不同解释变量, 分组变量进行固定效应回归, 比较模型结果之间的异质性

```
xtreg 经营资产收益率 金融资产比例 资本结构 公司规模 股权集中度 营业收入增长率  
i.year if 条件, fe
```

国企 1 民营 0 外资2 其他3

```
xtreg RO FAR ALT SIZE OC GOI i.year if SOE == 1, fe r
est store y1
xtreg RO FAR ALT SIZE OC GOI i.year if SOE == 0, fe r
est store y2
xtreg RO FAR ALT SIZE OC GOI i.year if SOE == 2, fe r
est store y3
xtreg RO FAR ALT SIZE OC GOI i.year if SOE == 3, fe r
est store y4
xtreg RO FAR ALT SIZE OC GOI i.year, fe r
est store y0
esttab y0 y1 y2 y3 y4 using 异质性分析.rtf, nogap nocompress scalar(N F
p) r2 ar2 replace drop(*year*) b(4)
```

3.2稳健性检验

变量替换法、样本调整法、控制变量法增减法、模型替换法、解释变量滞后期法(diff)
Utest检验

```
xtreg RO FAR FAR_2 ALT SIZE OC GOI i.year, fe r
utest FAR FAR_2, fieller level(99)
logout, save(utest_result) word replace: utest FAR FAR_2, fieller
level(99)
```

变量替换法：替换解释变量或者被解释变量，如果结果一致，具有稳健性

```
xtreg RO FAR ALT SIZE OC GOI i.year, fe
est store y1
xtreg ROA FAR ALT SIZE OC GOI i.year, fe
est store y2
esttab y1 y2 using 被解释变量替换.rtf, nogap nocompress scalar(N F p) r2
ar2 replace drop(*year*) b(4)
```

样本调整法：调整样本，改成2013-2014

控制变量法增减法：逐步加控制变量，观察被解释变量是否显著

模型替换法：替换成RE或者OLS、两阶段最小二乘法 xtivreg2

导出语句

```
outreg2 using *.doc:
logout, save () word:
```

```
est store y1  
esttab y* using *.doc,
```