

# 文献阅读笔记（四）

郑骋, 香港中文大学（深圳）, 金融工程硕士  
电话: 15825675534, 邮箱: zhengzc@zju.edu.cn

April 1, 2020

## Contents

<b>1 文献简介</b>	<b>4</b>
1.1 名称	4
1.2 作者	4
1.3 创新点	4
<b>2 量化因子</b>	<b>4</b>
2.1 名词解释	4
2.1.1 SEC	4
2.1.2 OTCBB	4
2.1.3 文件S-1和SB-2	4
2.2 虚拟变量-是否收到意见信CL	4
2.3 第一封回复问题个数ISSUES	4
2.4 意见信数量ROUNDS	5
2.5 审核时间DAYSOLVE	5
2.6 提交版本数量FILEVERSION	5
2.7 市值SIZE	5
2.8 资产负债率LEV	5
2.9 利润率ROA	5
2.10 从事业务数量SEGMENT	5
2.11 虚拟变量-审计质量GOINGCONCERN	5
2.12 虚拟变量-是否解雇审计师DISMISS	5
2.13 虚拟变量-是否处于频繁诉讼行业LITIGATION	5
2.14 虚拟变量-审计师质量INSPECTION	5
2.15 虚拟变量-首席执行官是否持有股份CEO_OWN	6

2.16	虚拟变量-首席执行官是否是董事会主席CEO_CHAIR . . . . .	6
2.17	虚拟变量-首席财务官是否为CPA: CFO_CPA . . . . .	6
2.18	流通股比例AMT%OUT . . . . .	6
2.19	公司成立至提交文件的年数FILEAGE . . . . .	6
2.20	虚拟变量-提交文件时间顺序BF10K . . . . .	6
2.21	样本统计性描述 . . . . .	6
<b>3</b>	<b>回归一 . . . . .</b>	<b>7</b>
3.1	回归目标 (Y) . . . . .	7
3.2	回归模型 . . . . .	7
3.3	回归区间 . . . . .	7
3.4	回归样本 . . . . .	7
3.5	回归结果 . . . . .	8
3.6	显著因子 . . . . .	8
3.7	回归逻辑/因子逻辑 . . . . .	9
<b>4</b>	<b>回归二 . . . . .</b>	<b>9</b>
4.1	回归目标 (Y) . . . . .	9
4.2	回归模型 . . . . .	9
4.3	回归区间 . . . . .	9
4.4	回归样本 . . . . .	9
4.5	回归结果 . . . . .	10
4.6	显著因子 . . . . .	10
4.7	回归逻辑/因子逻辑 . . . . .	10
<b>5</b>	<b>分析、总结及问题 . . . . .</b>	<b>11</b>
5.1	意见信CL . . . . .	11
5.2	企业业务数量SEGMENT . . . . .	11
5.3	审计发表意见GOCONCERN . . . . .	11
5.4	解雇审计师DISMISS . . . . .	11
5.5	频繁诉讼行业LITIGATION . . . . .	11
5.6	审计师质量INSPECTION . . . . .	11
5.7	CEO是否持股CEO_OWN . . . . .	11
5.8	CEO是否是董事会主席CEO_CHAIR . . . . .	11
5.9	CFO是否通过CPA考试CFO_CPA . . . . .	12
5.10	企业上市至提交文件年数FILEAGE . . . . .	12
5.11	流通股比例AMT%OUT . . . . .	12

5.12 提交文件时间顺序BF10K . . . . .	12
5.13 问题个数ISSUE . . . . .	12
5.14 意见信个数ROUNDS . . . . .	12
5.15 审核时间DAYSOLVE . . . . .	12
5.16 版本数量FILEVERSION . . . . .	12
5.17 启发 . . . . .	12
5.18 问题 . . . . .	13
5.18.1 结合特征 . . . . .	13

# 1 文献简介

## 1.1 名称

<<SEC意见信与财务造假——基于中概股危机的实证分析>>

## 1.2 作者

张然，北京大学光华管理学院副教授，博士生导师；陈思，中央财经大学财政税务学院讲师；雷羽

## 1.3 创新点

基于SEC对企业开具的意见信识别企业财务造假：①是否受到意见信②第一封意见信中的问题个数③从提交第一份文件至文件审核通过过程中的意见信数量④审核时间⑤公司提交的文件版本数量⑥审计相关信息⑦CEO/CFO与公司关系⑧其他有关信息可在“量化因子”小节查看

# 2 量化因子

## 2.1 名词解释

### 2.1.1 SEC

美国证券交易委员会Securities and Exchange Commission

### 2.1.2 OTCBB

美国场外柜台交易系统Over the Counter Bulletin Board，是由NASDAQ的管理者全美证券商协会（NASD）所管理的一个交易中介系统。类似于国内创业板

### 2.1.3 文件S-1和SB-2

S-1和SB-2为股份转让前需要向SEC递交的文件，其内容为SEC要求的公司信息（如财务指标等），S-1和SB-2的区别在于两者所需提供的信息不同。根据公司规模不同可选择不同的文件递交。。

## 2.2 虚拟变量-是否收到意见信CL

如果公司收到S-1/SB-2的意见信，则 $CL_i$ 取1，否则取0（意见信CL，Comment Letter）

## 2.3 第一封回复问题个数ISSUES

$ISSUES_i$ 表示公司收到的第一封针对S-1/SB-2的意见信中SEC提出的问题个数。

## 2.4 意见信数量ROUNDS

$ROUNDS_i$ 表示公司提交第一封S-1/SB-2到SEC宣布生效的时间内，SEC发出的意见信数量。

## 2.5 审核时间DAYSOLVE

$DAYSOLVE_i$ 表示公司提交第一封S-1/SB-2到SEC宣布生效的天数。

## 2.6 提交版本数量FILEVERSION

$FILEVERSION_i$ 表示公司提交第一封S-1/SB-2到SEC宣布生效的时间内，公司一共提交的S-1/SB-2版本数量。

## 2.7 市值SIZE

$SIZE_i$ 为资产的自然对数

## 2.8 资产负债率LEV

$LEV_i$ 为总负债除以总资产

## 2.9 利润率ROA

$ROA_i$ 为净利润与总资产之比

## 2.10 从业务数量SEGMENT

$SEGMENT_i$ 为从业务数量

## 2.11 虚拟变量-审计质量GOINGCONCERN

如果公司在S-1/SB-2中披露了不确定性持续经营审计意见则为1，否则为0

## 2.12 虚拟变量-是否解雇审计师DISMISS

如果公司在S-1/SB-2前一年内解雇了审计师则为1，否则为0

## 2.13 虚拟变量-是否处于频繁诉讼行业LITIGATION

公司处于诉讼频繁的行业则为1，否则为0

## 2.14 虚拟变量-审计师质量INSPECTION

当S-1/SB-2的签字审计师接受美国公众公司会计监督委员会的检查时取1，否则取0

## 2.15 虚拟变量-首席执行官是否持有股份CEO\_OWN

如果公司首席执行官拥有公司股份则为1，否则为0

## 2.16 虚拟变量-首席执行官是否是董事会主席CEO\_CHAIR

如果公司首席执行官为董事会主席则为1，否则为0

## 2.17 虚拟变量-首席财务官是否为CPA：CFO\_CPA

如果公司首席财务官为CPA则为1，否则为0（此处的CPA含义是中国的注册会计师）

## 2.18 流通股比例AMT%OUT

$AMT\%OUT_i$ 为转售股份占目前流通在外的普通股的比例

## 2.19 公司成立至提交文件的年数FILEAGE

$FILEAGE_i$ 为公司成立至提交S-1/SB-2的年数

## 2.20 虚拟变量-提交文件时间顺序BF10K

如果公司在第一份年报之前提交了S-1/SB-2则 $BF10K_i$ 取1，否则取0。

## 2.21 样本统计性描述

文章搜索了2005年至2010年期间在美反向并购的中国公司，共353家。其中共130家公司提交了S-1/SB-2，因此最终样本由该130家公司构成。其中26家为问题公司，104家为非问题公司。

Panel A 样本提交 S-1/SB-2 年份分布					Panel B 样本提交 S-1/SB-2 相对于上市的年份分布				
	非问题公司	问题公司	总计	百分比		非问题公司	问题公司	总计	百分比
2005	3	1	4	3.08%	T	59	18	77	59.23%
2006	15	5	20	15.38%	T+1	27	4	31	23.85%
2007	31	7	38	29.23%	T+2	13	2	15	11.54%
2008	27	4	31	23.85%	T+3	4	1	5	3.85%
2009	7	4	11	8.46%	T+4	1	1	2	1.54%
2010	21	5	26	20.00%					
总计	104	26	130	100.00%	总计	104	26	130	100.00%

Figure 1: S-1/SB-2时间分布

其中PanelB是指，公司从（反向并购）上市开始计时，至递交文件的时间。

	均值	标准差	最小值	25%	中位数	75%	最大值
$CL_{it}$	0.746	0.437	0	0	1	1	1
$ISSUES_{it}$	31.408	28.361	0	2	30	55	153
$ROUNDS_{it}$	3.300	2.780	0	1	3	5	12
$FILEVERSION_{it}$	5.146	3.017	1	3	5	7	16
$DAYSOLVE_{it}$	195.608	185.549	5	82	148	260	1007

Figure 2: 意见信特征的描述性统计

### 3 回归一

#### 3.1 回归目标（Y）

企业是否造假FRAUD，此处FRAUD为虚拟变量，造假是取1，否则取0

#### 3.2 回归模型

$$\begin{aligned}
FRAUD_{it} = & \alpha_0 + \alpha_1 CL_i + \alpha_2 SIZE_i + \alpha_3 LEV_i + \alpha_4 ROA_i + \alpha_5 SEGMENT_i + \alpha_6 FILEAGE_i + \alpha_7 AMT\%OUT_i + \\
& \alpha_8 BF10K_i + \alpha_9 GOINGCONCERN_i + \alpha_{10} DISMISS_i + \alpha_{11} LITIGATION_i + \alpha_{12} CEO\_OWN_i + \\
& \alpha_{13} CEO\_CHAIR_i + \alpha_{14} CFO\_CPA_i + \alpha_{15} INSPECTION_i + (\sum YEAR) + \varepsilon_i
\end{aligned} \tag{1}$$

其中 $\sum YEAR$ 项文章并未提及是何含义，并且回归结果也无显示，或许是笔误？

#### 3.3 回归区间

2005年至2010年

#### 3.4 回归样本

文章搜索了2005年至2010年期间在美反向并购的中国公司，共353家。其中共130家公司提交了S-1/SB-2，因此最终样本由该130家公司构成。其中26家为问题公司，104家为非问题公司。

### 3.5 回归结果

	模型一	模型二	模型三	模型四
变量	FRAUD	FRAUD	FRAUD	FRAUD
$CL_{it}$	0.997 ** ( - 2.36)	1.094 ** ( - 2.47)	1.225 *** ( - 2.73)	1.884 *** ( - 3.37)
$SIZE_{it}$	0.533 ** ( - 2.51)	0.535 ** ( - 2.5)	0.497 ** ( - 2.24)	0.701 *** ( - 2.7)
$LEV_{it}$	-1.314 ( - 1.55)	-1.351 ( - 1.52)	-1.281 ( - 1.43)	-0.996 ( - 0.99)
$ROA_{it}$	0.455 ( - 0.47)	0.195 ( - 0.2)	0.752 ( - 0.72)	1.079 ( - 0.95)
$SEGMENT_{it}$	-0.365 ** ( - 2.19)	-0.375 ** ( - 2.19)	-0.432 ** ( - 2.44)	-0.475 ** ( - 2.32)
$GOINGCONCERN_{it}$		-0.423 ( - 0.54)		-0.371 ( - 0.44)
$DISMISS_{it}$		-0.353 ( - 0.77)		-0.531 ( - 0.98)
$LITIGATION_{it}$		-0.019 ( - 0.061)		0.268 -0.76
$INSPECTION_{it}$		0.652 ( - 1.39)		0.349 ( - 0.7)
$CEO\_OWN_{it}$			0.609 ( - 0.86)	0.911 ( - 1.13)
$CEO\_CHAIR_{it}$			-0.426 ( - 0.89)	-0.654 ( - 1.22)
$CFO\_CPA_{it}$			-0.635 * ( - 1.88)	-0.956 ** ( - 2.36)
$FILEAGE_{it}$				-0.014 ( - 0.41)
$AMT\%OUT_{it}$				0.11 ( - 0.35)
$BF10K_{it}$				-1.200 *** ( - 2.91)
Constant	-10.352 *** ( - 2.71)	-10.843 *** ( - 2.78)	-9.795 ** ( - 2.37)	-13.523 *** ( - 2.83)
Observations	130	130	130	130
Pseudo R <sup>2</sup>	0.214	0.2365	0.2493	0.3414

Figure 3: 回归结果

### 3.6 显著因子

模型一：CL/SIZE/SEGMENT

模型二：CL/SIZE/SEGMENT

模型三：CL/SIZE/SEGMENT

模型四：CL/SIZE/SEGMENT/CFO.CPA/BF10K



### 3.7 回归逻辑/因子逻辑

该回归通过CL变量来的显著性判断该特征是否对企业财务造假有预测作用。由于企业财务造假还受到其他变量的影响，因此提出模型一到模型四，四个模型在不同之处在于选择了不同的控制变量，通过控制变量，在一定程度上，来消除其他预测变量的影响。

## 4 回归二

### 4.1 回归目标（Y）

企业是否造假FRAUD，此处FRAUD为虚拟变量，造假是取1，否则取0

### 4.2 回归模型

$$\begin{aligned} FRAUD_{it} = & \alpha_0 + \alpha_1 CLCHARACTER_i + \alpha_2 SIZE_i + \alpha_3 LEV_i + \alpha_4 ROA_i + \alpha_5 SEGMENT_i + \alpha_6 FILEAGE_i + \alpha_7 AMT\%OUT_i + \\ & \alpha_8 BF10K_i + \alpha_9 GOINGCONCERN_i + \alpha_{10} DISMISS_i + \alpha_{11} LITIGATION_i + \alpha_{12} CEO\_OWN_i + \\ & \alpha_{13} CEO\_CHAIR_i + \alpha_{14} CFO\_CPA_i + \alpha_{15} INSPECTION_i + (\sum YEAR) + \varepsilon_i \end{aligned} \quad (2)$$

其中 $\sum YEAR$ 项文章并未提及是何含义，并且回归结果也无显示，或许是笔误？

另外 $CLCHARACTER_i$ 指意见信的各项特征，即ISSUE,ROUNDS,DAYSOLVE,FILEVERSION。

### 4.3 回归区间

2005年至2010年

### 4.4 回归样本

文章搜索了2005年至2010年期间在美反向并购的中国公司，共353家。其中共130家公司提交了S-1/SB-2，因此最终样本由该130家公司构成。其中26家为问题公司，104家为非问题公司。

## 4.5 回归结果

	模型五	模型六	模型七	模型八
变量	FRAUD	FRAUD	FRAUD	FRAUD
$ISSUES_i$	0.015 ** ( -2.35)			
$ROUNDS_i$		0.137 ** ( -2.35)		
$DAYSOLVE_i$			0.001 ( -1.62)	
$FILEVERSION_i$				0.137 ** ( -2.41)
$SIZE_i$	0.519 ** ( -2.31)	0.516 ** ( -2.3)	0.535 ** ( -2.38)	0.501 ** ( -2.26)
$LEV_i$	-0.776 ( -0.83)	-0.729 ( -0.78)	-0.688 ( -0.75)	-0.683 ( -0.74)
$ROA_i$	0.963 ( -0.89)	0.871 ( -0.83)	0.767 ( -0.68)	0.721 ( -0.68)
$SEGMENT_i$	-0.415 ** ( -2.21)	-0.363 ** ( -2.03)	-0.343 * ( -1.93)	-0.358 ** ( -1.98)
Constant	-9.390 ** ( -2.28)	-9.639 ** ( -2.36)	-9.592 ** ( -2.34)	-9.624 ** ( -2.40)
Observations	130	130	130	130
Pseudo R <sup>2</sup>	0.269	0.268	0.2444	0.2707

注：\*\*\*，\*\*和\* 分别代表在1%，5%和10%水平显著。省略部分控制变量回归结果。

Figure 4: 回归结果

## 4.6 显著因子

模型五：ISSUE/SIZE/SEGMENT

模型六：ROUNDS/SIZE/SEGMENT

模型七：SIZE/SEGMENT

模型八：FILEVERSION/SIZE/SEGMENT/CFO\_CPA/BF10K

## 4.7 回归逻辑/因子逻辑

该回归，在控制变量的基础上，分别加入不同的意见信特征，形成模型五至模型八，各自回归，来判断这些意见信特征是否有显著的预测作用。

这里猜想，可以把所有特征一起回归？假如这些特征都是有预测作用的话，理论上来说，能够提高识别财务造假的预测准确度。

## 5 分析、总结及问题

### 5.1 意见信CL

“有无意见信”是一个显著变量，即表示它对财务造假有显著预测作用，这在逻辑上行得通。显然，能够收到意见信说明企业信息披露方面存在问题，或者企业本身存在问题。如果是后者，显然其有大概率发生财务造假，若是披露方面存在问题，说明企业在法律合规模块也不尽人意，相对正规的企业来说，该企业也有较大可能进行财务造假。

从样本和特征的描述性统计来看，造假企业占样本企业数量的20%，而样本企业中收到意见信的企业占比75%，说明其实大部分企业都会收到意见信，而很有可能是由于未收到意见信的企业，财务造假概率较低，从而导致该变量显著。

### 5.2 企业业务数量SEGMENT

“企业业务数量”是一个显著变量。这里的原因可以这样想：如果企业业务数量较少，其难以实施财务造假（如通过母子公司关联交易操纵利润表，通过企业内部不同业务之间的造假合同操纵利润），因此业务数量少，则企业造假概率低。而企业业务数量多，企业更方便造假。因此这里可以将企业以一个业务数量阈值分组，观察两组的企业造假数量比例有显著区别。

### 5.3 审计发表意见GOCONCERN

“审计发表意见”并不是一个显著变量，这一点令人奇怪，如果审计师是专业的，则能够判断企业是否造假。（不知美国的审计市场如何）

### 5.4 解雇审计师DISMISS

“解雇审计师”并不是一个显著变量。（不明白该特征的作用）

### 5.5 频繁诉讼行业LITIGATION

“是否处于频繁诉讼行业”并不是一个显著变量，因此可以说明频繁诉讼与企业财务造假关系不大。

### 5.6 审计师质量INSPECTION

“审计师是否接受美国公众公司会计监督委员会的检查”并不是一个显著变量，这可能与“审计师发表意见GOCONCERN”有关。

### 5.7 CEO是否持股CEO\_OWN

“CEO是否持有本公司股票”并非显著变量。

### 5.8 CEO是否是董事会主席CEO\_CHAIR

“CEO是否是董事会主席”并非显著变量。

## 5.9 CFO是否通过CPA考试CFO\_CPA

“CFO是否通过CPA考试”，该变量在模型三中有10%显著性，在模型四中有5%显著性。该值说明CFO专业程度越高，企业财务造假概率越低，但是或许这不能用来判断企业是否经营良好（猜测企业更会“造假”，难以被发现？）

## 5.10 企业上市至提交文件年数FILEAGE

“企业从反向并购上市至提交文件年数”是一个无关变量。

## 5.11 流通股比例AMT%OUT

“流通股比例”是一个无关变量。这里可以从企业造假的目的出发去思考，企业财务造假是为了能够低成本的上市/融资（假设市盈率不变，如果能造假利润，则利润越高获得的融资越多，而实际上的利润并没有这么高，因此能够低成本融资），一些企业的目的是为了出售股票，因此流通股比例越高越好（使得融资出售股票收益大于造假成本），但另一方面，考虑到一些企业是出于政治利益造假上市，因此流通股比例对这种情况的财务造假没有预测作用。因此由于可能性太多，导致该变量并不显著。

## 5.12 提交文件时间顺序BF10K

“企业在第一份年报前提交文件BF10K”变量是一个显著变量，说明企业财务指标较优，无需造假，而造假需要时间，因此可以通过该变量反应。

## 5.13 问题个数ISSUE

“第一封意见信中的问题个数”是一个显著变量。

## 5.14 意见信个数ROUNDS

“收到的意见信数量”是一个显著变量。

## 5.15 审核时间DAYSOLVE

“从提交文件到通过审核的时间”并不是一个显著变量。原因可能是审核时间受到多方面因素影响。

## 5.16 版本数量FILEVERSION

“公司提交的版本数量”是一个5%显著变量，版本数量越多反应企业提交次数多，说明企业相关问题难以解决（存在造假现象，且掩盖手段困难）。

## 5.17 启发

对于企业财务造假问题，可以将企业根据有无财务造假分组，比较两组的描述性样本统计，观察哪些特征是有明显差异，从而判断显著性。

## 5.18 问题

### 5.18.1 结合特征

可以结合所有特征进行财务造假的预测，但文章并没有去做。