

文献阅读笔记（六）

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 2, 2020

Contents

| | |
|------------------|----------|
| 1 文献简介 | 3 |
| 1.1 名称 | 3 |
| 1.2 作者 | 3 |
| 1.3 创新点 | 3 |
| 1.4 启发 | 3 |
| 2 量化因子 | 3 |
| 2.1 回归目标Y: PMAFE | 3 |
| 2.2 本地优势Local | 3 |
| 2.3 分析师经验Exp | 4 |
| 2.4 距离公司报告日Age | 4 |
| 2.5 跟踪公司个数Cover | 4 |
| 2.6 明星分析师Star | 4 |
| 2.7 其他控制变量 | 4 |
| 2.8 变量定义表 | 5 |
| 2.9 样本统计性描述 | 5 |
| 3 回归一 | 6 |
| 3.1 回归目标Y:PMAFE | 6 |
| 3.2 回归模型 | 6 |
| 3.3 回归区间 | 6 |
| 3.4 回归频率 | 6 |
| 3.5 回归结果 | 7 |
| 3.6 显著因子 | 7 |
| 3.7 回归逻辑/因子逻辑 | 7 |

| | | |
|----------|--|-----------|
| 3.7.1 | 本地优势Local | 7 |
| 3.7.2 | 企业信息披露质量Transparency | 7 |
| 3.7.3 | 分析师经验Exp | 8 |
| 3.7.4 | 分析师跟踪公司数Cover | 8 |
| 3.7.5 | 距离公司报告日Age | 8 |
| 3.7.6 | 公司规模Size | 8 |
| 3.7.7 | 财务杠杆Lev | 8 |
| 3.7.8 | 明星分析师Star | 8 |
| 3.7.9 | 市场化指数market | 8 |
| 3.7.10 | 审计意见Auditopinion | 8 |
| 4 | 回归二 | 8 |
| 4.1 | 回归目标Y:PMAFE | 8 |
| 4.2 | 回归模型 | 9 |
| 4.3 | 回归区间 | 9 |
| 4.4 | 回归频率 | 9 |
| 4.5 | 回归结果 | 9 |
| 4.6 | 显著因子 | 10 |
| 4.7 | 回归逻辑/因子逻辑 | 10 |
| 4.7.1 | 交叉项 $Local_{ijt} \times Transparency_{jt}$ | 10 |
| 5 | 存在问题 | 10 |
| 5.1 | 准确度定义 | 10 |
| 5.2 | 样本描述性统计 | 10 |
| 6 | 延伸 | 11 |

1 文献简介

1.1 名称

<<本地优势、信息披露质量和分析师预测准确性>>

1.2 作者

王会娟，北京大学光华管理学院；张路，北京大学光华管理学院；张然，北京大学光华管理学院副教授，博士生导师。

1.3 创新点

基于分析师信息/企业信息披露质量，从而判断分析师预测准确性。

1.4 启发

当因子中有利用到分析师报告时，判断分析师家乡及公司所在地，若一致则赋予该报告中给出的信息高权重。

另外，文章中提到信息披露质量，可以通过回归验证：信息披露质量对于股价是否有一定的预测作用。假如具有一定预测效果，此时可以自己定义信息披露质量，例如：在每年年报公布窗口期，挑选一份报告，将其披露质量指数作为基准1，以一定的评分方法，将其他年报的披露质量指数计算出来，从而作为因子。或者将该因子作为其他基本面因子的权重，因为披露质量指数是用来衡量一个财务报表的准确程度。

2 量化因子

2.1 回归目标Y：PMAFE

分析师年度盈余预测偏差比例 $PMAFE_{ijt}$ ，表示在第t年，分析师i对公司j的预测偏差比例，。

$$PMAFE_{ijt} = \frac{DAFE_{ijt}}{\overline{AFE_{jt}}}$$

定义 FE_{ijt} 表示第t年，分析师i对公司j的EPS的预测值； AE_{jt} 是在第t年，公司j的EPS的真实值。 AFE_{ijt} 是第t年，分析师i对公司j的EPS的预测偏差值。

$$AFE_{ijt} = |FE_{ijt} - AE_{jt}|$$

定义 $DAFE_{ijt}$ 为在第t年，分析师i对公司j的EPS的预测偏差的绝对值 AFE_{ijt} 与平均预测偏差绝对值 $\overline{AFE_{ijt}}$ 的差。

$$DAFE_{ijt} = AFE_{ijt} - \overline{AFE_{ijt}}$$

2.2 本地优势Local

$Local_{ijt}$ ，若第t年，分析师i所属的券商注册地和公司j注册地是同一省（或直辖市），则取1，否则取0

2.3 分析师经验Exp

$$Exp_{ijt} = \ln(1 + M)$$

其中，M为分析师在CSMAR数据库（国泰安数据库）出现的第一次预测公告日到在第t年分析师i对公司j做出预测公告日间隔的天数。

2.4 距离公司报告日Age

$$Age_{ijt} = \ln(1 + N)$$

其中，N为分析师预测公告日到公司报告期间隔的天数。

2.5 跟踪公司个数Cover

$Cover_{it}$ ，第t年，分析师i跟踪的公司个数。

2.6 明星分析师Star

$Star_{it}$ ，第t年，分析师i入选<<新财富>>则取1，否则为0。

2.7 其他控制变量

包括公司规模Size，资产负债率Lev，成长性Growth，公司所在地市场化进程Market，审计意见Auditopinion，信息披露质量Transparency。另外，文章还控制了年度和行业。

2.8 变量定义表

| 变量名称 | 变量符号 | 变量定义 |
|------------|--------------|--|
| 分析师盈余预测准确性 | PMAFE | 第 t 年, 分析师 i 对公司 j 的 EPS 的预测偏差的绝对值, 减去所有对公司 j 进行预测的分析师的预测偏差绝对值的平均, 然后除以平均值 |
| 本地优势 | Local | 第 t 年, 分析师 i 所属的券商注册地和公司 j 注册地是同一个省时 $Local_{ijt}$ 为 1, 否则为 0 |
| 信息披露质量 | Transparency | Transparency1 为虚拟变量, 当深交所的信息披露考核结果为不及格、及格、良好和优秀时, Transparency1 分别取 0、1、2 和 3 Transparency2: 当深交所的信息披露考核结果为不及格时取 0, 其他情况取 1 |
| 分析师经验 | Exp | 分析师 i 在 CSMAR 数据库出现的第一次预测公告日到在第 t 年分析师 i 对公司 j 做出预测公告日间隔的天数 M , 然后取 $\text{Log}(1+M)$ |
| 分析师跟踪的公司数 | Cover | 在 t 年分析师 i 跟踪的公司个数 |
| 预测的时间跨度 | Age | 分析师预测公告日到公司报告期的间隔天数 N , 然后再取对数 $\text{Log}(1+N)$ |
| 最佳分析师 | Star | 按照《新财富》每年入围的最佳分析师名单, 当分析师 i 在第 t 年入围最佳分析师时 $Star_{it}$ 为 1, 否则为 0 |
| 上市公司规模 | Size | 公司的总资产, 然后取自然对数 |
| 上市公司杠杆 | Lev | 公司的总负债/总资产的比率 |
| 上市公司成长性 | Growth | 公司的(年末主营业务收入-年初主营业务收入)/年初主营业务收入 |
| 市场化进程 | Market | 樊刚和王小鲁(2009)编制的《中国市场化指数——各地区市场化相对进程报告》中的市场化指数 |
| 审计意见 | Auditopinion | 公司上一年的审计意见为标准无保留意见时取 1, 否则取 0 |

Figure 1: 变量定义表

2.9 样本统计性描述

A股上市公司2003年-2008年, 信息披露质量数据来源深交所披露的“诚信档案”; 分析师预测数据和上市公司财务数据来自国泰安的CSMAR数据库; 券商注册地数据来源于Wind数据库; 公司注册地数据来自CCER中国经济金融数据库。文章将年报披露日之前的分析师预测报告作为当年样本(即上一年年报披露日至本年年报披露日之间分析师的所有报告)。其中一份报告有多个分析师签字, 则将多个分析师拆分为多个单人分析师报告(例如有两人签字, 则拆分为两份单人报告, 其中预测相同, 但预测人不同)

剔除了金融行业以及数据缺失样本, 并将资产负债率大于1的样本作为异常值剔除, 最后分析师预测准确性指标(PMAFE)进行1%和99%的截尾处理。未剔除前得到35228个观测值, 剔除披露质量缺失样本之后, 观测值减至15688个。

| 变量 | N | 均值 | 标准偏差 | 最小值 | 最大值 |
|---------------|--------|---------|---------|---------|----------|
| PMAFE | 35 228 | -0.0625 | 0.4972 | -0.9000 | 1.2198 |
| Local | 35 228 | 0.1205 | 0.3256 | 0.0000 | 1.0000 |
| Exp | 35 228 | 5.7114 | 1.8053 | 0.0000 | 7.9501 |
| Cover | 35 228 | 26.6945 | 31.8504 | 1.0000 | 226.0000 |
| Age | 35 228 | 4.8857 | 0.8036 | 0.0000 | 5.9026 |
| Star | 35 228 | 0.1470 | 0.3541 | 0.0000 | 1.0000 |
| Size | 15 688 | 22.0177 | 1.3115 | 18.6508 | 25.5044 |
| Growth | 15 688 | 0.2665 | 0.5990 | -0.9454 | 29.8578 |
| Lev | 15 688 | 0.4566 | 0.1870 | 0.0210 | 0.9994 |
| Market | 15 688 | 9.1182 | 2.0232 | 2.6000 | 12.3810 |
| Auditopinion | 15 688 | 0.9985 | 0.0381 | 0.0000 | 1.0000 |
| Transparency1 | 15 688 | 2.1376 | 0.6331 | 0.0000 | 3.0000 |

Figure 2: 样本统计性描述

3 回归一

3.1 回归目标Y:PMAFE

分析师预测准确度PMAFE。

3.2 回归模型

$$PMAFE_{ijt} = \beta_0 + \beta_1 Local_{ijt} + \beta_2 Exp_{ijt} + \beta_3 Cover_{it} + \beta_4 Age_{i,t} + \beta_5 Star_{it} + \beta_6 Size_{jt} + \beta_7 Lev_{jt} + \beta_8 Growth_{jt} + \beta_9 Market_{jt} + \beta_{10} Auditopinion_{jt} + \beta_{11} Transparency_{jt} + \varepsilon_{ijt} \quad (1)$$

3.3 回归区间

2003年至2008年

3.4 回归频率

年

3.5 回归结果

| 变量 | (1) | | (2) | | (3) | |
|-------------------|------------|-------|------------|-------|------------|-------|
| | 回归系数 | t 值 | 回归系数 | t 值 | 回归系数 | t 值 |
| Intercept | -0.9820*** | -4.04 | -1.4524*** | -4.01 | -1.4368*** | -4.18 |
| Local | -0.0159** | -2.43 | -0.0125** | -2.01 | -0.0169** | -2.20 |
| Exp | -0.0063*** | -3.91 | -0.0073** | -2.39 | -0.0078*** | -2.63 |
| Cover | -0.0004*** | -4.47 | -0.0047*** | -3.81 | -0.0065** | -2.14 |
| Age | 0.1983*** | 4.94 | 0.2968*** | 4.57 | 0.2873*** | 4.40 |
| Star | 0.0053 | 0.57 | 0.0385 | 0.97 | 0.0447 | 0.55 |
| Size | | | 0.0006** | 2.14 | 0.0003** | 2.08 |
| Growth | | | 0.0001 | 0.12 | 0.0001 | 0.14 |
| Lev | | | 0.0676** | 2.27 | 0.0701** | 2.42 |
| market | | | -0.0048* | -1.91 | -0.0049** | -2.03 |
| Auditopinion | | | -0.0350 | -0.31 | -0.0261 | -0.24 |
| Transparency1 | | | -0.0199*** | -2.68 | | |
| Transparency2 | | | | | -0.0264** | -2.56 |
| Industry Dummies | Included | | Included | | Included | |
| Fixed Year Effect | Included | | Included | | Included | |
| 调整 R ² | 10.34% | | 14.18% | | 14.37% | |
| 样本数 | 35 228 | | 15 688 | | 15 688 | |

注：t 值均为按公司和年度两个维度聚类调整后的结果；***、**、* 分别表示回归系数在 1%、5%、10%的水平上显著。

Figure 3: 回归结果

3.6 显著因子

未额外说明则代表1%显著性。

模型一：Local(5%)/Exp/Cover/Age

模型二：Local(5%)/Exp(5%)/Cover/Age/Size(5%)/Lev(5%)/market(10%)/Transparency1

模型三：Local(5%)/Exp/Cover(5%)/Age/Size(5%)/Lev(5%)/market(5%)/Transparency2(5%)

3.7 回归逻辑/因子逻辑

3.7.1 本地优势Local

在三个模型中，Local变量均在5%显著性上显著，且系数为负，故表明本地分析师预测更加准确。

3.7.2 企业信息披露质量Transparency

Transparency1和Transparency2在定义上略有不同，但都可以反映企业信息披露的质量程度。而在模型一中Transparency1以1%显著性显著，在模型二中Transparency2以5%显著性显著，且系数均为负数，故表明信息披露质量越高，则有助于分析师预测。

3.7.3 分析师经验Exp

在三个模型中，Exp变量均显著，且系数为负，故表明分析师经验有助于分析师预测。

3.7.4 分析师跟踪公司数Cover

在三个模型中，Cover变量均显著，且系数为负，故表明分析师跟踪的公司越多，则预测准确度越高。（一方面，跟踪数增加可以反映分析师的经验增加，另一方面，跟踪数增加分散了分析师的注意力，这个系数反应了两者的综合结果）。

3.7.5 距离公司报告日Age

在三个模型中，Age变量均在1%显著性上显著，且系数为正，故表明分析师的报告离年报公告日越近，预测越准确。（信息增加/能参考其他分析师报告）

3.7.6 公司规模Size

Size变量显著为正，表示公司规模越大，分析师预测越不准确。（公司的复杂度增加）

3.7.7 财务杠杆Lev

Lev变量显著为正，表示公司经营的财务杠杆越大，分析师预测越不准确。（公司经营风险增加）

3.7.8 明星分析师Star

Star在各个模型中均不显著，表明分析师是否是明星分析师，对盈余预测准确度没影响。（当然本文的预测准确度定义本身就不准确）

3.7.9 市场化指数market

market显著为负，故市场化有助于分析师预测（不明白）。

3.7.10 审计意见Auditopinion

审计意见对分析师预测无影响

4 回归二

4.1 回归目标Y:PMAFE

分析师预测准确度PMAFE。

4.2 回归模型

$$PMAFE_{ijt} = \beta_0 + \beta_1 Local_{ijt} + \beta_2 Exp_{ijt} + \beta_3 Cover_{it} + \beta_4 Age_{i,t} + \beta_5 Star_{it} + \beta_6 Size_{jt} + \beta_7 Lev_{jt} +$$
$$\beta_8 Growth_{jt} + \beta_9 Market_{jt} + \beta_{10} Auditopinion_{jt} + \beta_{11} Transparency_{jt} + \beta_{12} Local_{ijt} \times Transparency_{jt} + \varepsilon_{ijt}$$

(2)

回归二与回归一相比，增加了第12项变量，即本地优势Local与信息披露质量Transparency的交叉项。

4.3 回归区间

2003年至2008年

4.4 回归频率

年

4.5 回归结果

| 变量 | (1) | | (2) | |
|---------------------|------------|-------|------------|-------|
| | 回归系数 | t 值 | 回归系数 | t 值 |
| Intercept | -1.4533*** | -4.02 | -1.4374 | -4.19 |
| Local | -0.0080 | 1.25 | -0.0064* | 1.66 |
| Exp | -0.0073** | -2.40 | -0.0078*** | -2.64 |
| Cover | -0.0046*** | -2.78 | -0.0064** | -2.12 |
| Age | 0.2967*** | 4.56 | 0.2872*** | 4.39 |
| Star | 0.0387*** | 2.99 | 0.0448*** | 3.56 |
| Size | 0.0006** | 2.14 | 0.0003** | 2.06 |
| Growth | 0.0001 | 0.13 | 0.0001 | 0.15 |
| Lev | 0.0663** | 2.23 | 0.0694** | 2.40 |
| market | -0.0047* | -1.89 | -0.0049** | -2.01 |
| Auditopinion | -0.0368 | -0.33 | -0.0267 | -0.25 |
| Transparency1 | -0.0169** | -2.16 | | |
| Local×Transparency1 | -0.0240** | -2.12 | | |
| Transparency2 | | | -0.0225** | -2.07 |
| Local×Transparency2 | | | -0.0329** | -2.11 |
| Industry Dummies | Included | | Included | |
| Fixed Year Effect | Included | | Included | |
| 调整 R ² | 14.28% | | 14.42% | |
| 样本数 | 15 688 | | 15 688 | |

注：t 值均为按公司和年度两个维度聚类调整后的结果；***、**、* 分别表示回归系数在 1%、5%、10%的水平上显著。

Figure 4: 回归结果

4.6 显著因子

未额外说明则代表1%显著性。

模型一：Exp(5%)/Cover/Age/Star/Size(5%)/Lev(5%)/market(10%)/Transparency1(5%)/Local \times Transparency1

模型二：Local(10%)/Exp/Cover(5%)/Age/Star/Size(5%)/Lev(5%)/market(5%)/Transparency2(5%)/Local \times Transparency2

4.7 回归逻辑/因子逻辑

4.7.1 交叉项 $Local_{ijt} \times Transparency_{jt}$

由于该回归二与回归一相比，仅增加了交叉项，因此仅分析交叉项 $Local_{ijt} \times Transparency_{jt}$ 的影响。

首先，交叉项的作用是为了衡量，一个Predictor对Response的影响，是否受另一些变量的影响。举例：

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_1 \times X_2$$

则变量 X_1 对Y的影响（系数）为 $\alpha_1 + \alpha_3 X_2$ ，同理可得变量 X_2 的影响。

因此，增加交叉项，是为了检验上述所说的“影响”。

在这里，增加交叉项 $Local_{ijt} \times Transparency_{jt}$ ，则可以揭示当分析师具有本地优势时，信息披露质量对预测准确度的影响发生了怎样的变化；当信息披露质量较好时，分析师具有本地优势对预测准确度的影响发生了怎样的变化。

| | 不含交叉项 | 含交叉项 |
|------------------|---------|---------------|
| Local=L | -0.015 | T2=0: -0.0064 |
| | | T2=1: -0.0392 |
| Transparency2=T2 | -0.0264 | L=0: -0.0225 |
| | | L=1: -0.0554 |

Figure 5: 回归结果

从系数的比较当中可以看出，分析师本地优势与信息披露质量有相互促进的作用。

5 存在问题

5.1 准确度定义

该定义方法实际上是衡量分析师预测值与分析师预测均值之间的关系，而不能反应预测准确度。

5.2 样本描述性统计

从该文章对准确度的定义来看，PMAFE的均值应当为0，而实际为-0.0625。

6 延伸

对于文章中提到的“市场化指数”，主要衡量的方面有：政府与市场的关系/非国有经济的发展/产品市场的发育程度/要素市场的发育程度/市场中介组织发育和法律制度程度。具体内容可参考：

<http://www.doc88.com/p-9993581223649.html>