



盈余管理实践

基于 Compustat 的方法演练与 R/Python 实现

作者：晨瀚宇 (Chanw)

时间：May 15, 2026

版本：1.0

同一份 Compustat 面板、同一个 DA 度量问题，从 Healy 1985 到现代 ML 残差的十种方法演练与比较。

© 2026 晨瀚宇. 保留所有权利。

版权声明

书名： 盈余管理实践——基于 Compustat 的方法演练与 R/Python 实现

作者： 晨瀚宇（小红书：Chanw）

版本： v1.0（2026 年）

© 2026 晨瀚宇. 保留所有权利。

本书内容（包括但不限于文字、代码、图表、排版设计）均为作者原创。

禁止事项：

- 未经作者书面授权，禁止以任何形式转载、复制、传播本书的全部或部分内容
- 禁止将本书内容用于商业用途或二次售卖
- 禁止去除、修改本书中的版权标识和作者署名

允许事项：

- 个人学习和研究用途的合理引用（需注明出处）
- 在社交媒体上分享本书的购买链接

如发现侵权行为，请联系作者。

购买渠道：小红书搜索「Chanw」

前言

这本书的写作思路是：选定一份具体的数据集，把所有相关的盈余管理度量方法都在它身上跑一遍，让读者亲手看到每一种方法在同一个问题上能给出什么答案、各自的假设和局限在哪里。

本书用的数据是 Compustat North America Annual Fundamentals，时间跨度 1991 至 2023 年，剔除金融业与公用事业后约 18 至 22 万个 firm-year 观测。Compustat 是会计实证研究事实上的基准数据源，几乎所有英文盈余管理论文都基于它。本书把这套数据从清洗到模型估计的完整流水线写出来，让读者可以按章节顺序在自己的环境里复现。

读者画像是会计、审计、财务方向的在读研究生与青年研究者，能用 R 或 Python，看过几篇用了 Jones 模型或 Roychowdhury 真实活动 EM 的论文，但没有亲手跑过完整的方法对比。本书不打算把 Dechow、Ge 与 Schrand 在 2010 年的盈余质量综述重述一遍，更不打算把每一种方法都讲到学术教科书的精度。它的野心比较有限：让读者跑通一项 Compustat 级别的 DA 度量、看懂行业-年份回归的输出、知道每一种方法背后的会计与统计假设是什么、知道在常见样本期下哪些雷区会让结论失稳。读完之后愿意去翻原始论文作为参考、愿意去看 wrds Python 包与 RPostgres 文档把更复杂的取数任务深入处理，这本书的任务就完成了。

每章对应一种方法，从 1985 年最早的 Healy 模型一路推进到 2010 年代 Dechow et al. F-Score 与现代 ML 残差。代码以 R 为主、Python 为辅，关键章节两套代码并列给出。所有正文里出现的数字均来自真实代码运行输出，不手算、不估算。

晨瀚宇

2026 年

目录

前言	i
第 1 章 问题与数据：总应计构造与盈余管理度量	1
1.1 盈余管理度量在做一件什么事	1
1.1.1 应计与现金的会计身份	1
1.1.2 Hribar-Collins 资产负债表方法	2
1.2 数据与样本构造	2
1.2.1 Bao 2020 数据相对 WRDS 直接拉取的限制	3
1.3 案例公司：三起 SEC AAER 处罚案件	3
1.4 R 与 Python 双实现规范	4
1.5 Python 实现	5
1.6 小结	7
第 2 章 Healy 与 DeAngelo：早期均值与差分模型	8
2.1 回到原始问题：怎么把 TA 拆成两部分	8
2.2 Healy (1985) 模型	8
2.2.1 核心思想	8
2.2.2 在 Bao 数据上的实现	9
2.3 DeAngelo (1986) 模型	9
2.3.1 核心思想	9
2.3.2 在 Bao 数据上的实现	9
2.4 两种方法的一致性	10
2.5 案例公司的初步答案	11
2.6 两种方法在什么情况下仍然适用	11
2.7 Python 实现	12
2.8 小结	13
第 3 章 Jones (1991) 模型：把销售与固定资产显式建模	14
3.1 从 Healy 到 Jones：剩余的两个噪声源	14
3.2 Jones 模型的定义	14
3.2.1 Bao 数据的行业控制限制	15
3.3 在 Bao 数据上的实现	15
3.4 案例公司在 Jones 下的表现	16
3.5 Python 实现	17
3.6 小结	19
第 4 章 Modified Jones (1995)：把应收变化从销售变化中扣除	20
4.1 Jones 模型留下的一个漏洞	20
4.2 Modified Jones 的定义	20
4.3 在 Bao 数据上的实现	21
4.4 案例公司：CA 2001 年的判别是否被纠正	22
4.5 Python 实现	23

4.6	小结	24
第 5 章	Performance-Matched DA (Kothari, Leone, Wasley 2005)	25
5.1	Modified Jones 留下的业绩偏差	25
5.2	PM-DA 的定义	25
5.3	在 Bao 数据上的实现	26
5.4	案例公司: PM-DA 在 CA 2001 上的表现	27
5.5	Python 实现	28
5.6	小结	30
第 6 章	Dechow-Dichev (2002) 应计质量	31
6.1	从 Jones 系列切换到 DD 系列	31
6.2	Dechow-Dichev 模型的定义	31
6.3	在 Bao 数据上的实现	32
6.4	案例公司的 AQ_DD	33
6.5	Python 实现	34
6.6	小结	35
第 7 章	McNichols (2002): DD 与 Jones 的融合	36
7.1	DD 与 Jones 的互补关系	36
7.2	McNichols 模型的定义	36
7.3	在 Bao 数据上的实现	36
7.4	案例公司在 McNichols 下的表现	38
7.5	Python 实现	39
7.6	小结	40
第 8 章	Stubben (2010): 把收入侧 DA 独立出来	41
8.1	从总应计回到收入	41
8.2	Stubben 模型的定义	41
8.3	在 Bao 数据上的实现	41
8.4	案例公司在 Stubben 下的表现	43
8.5	Python 实现	44
8.6	小结	45
第 9 章	Roychowdhury (2006) 真实活动盈余管理	46
9.1	应计型 DA 之外的另一种盈余管理	46
9.2	Roychowdhury 模型的定义	46
9.2.1	Bao 数据的 DISEXP 限制	47
9.3	在 Bao 数据上的实现	47
9.4	案例公司在真实活动 RM 下的表现	49
9.5	Python 实现	49
9.6	小结	51
第 10 章	F-Score 与九方法终极对比	52
10.1	九条路径的全样本对照	52
10.2	九方法的相关结构	52
10.2.1	Pearson 相关矩阵	52

10.2.2 Spearman 相关矩阵	53
10.3 F-Score 综合判别	53
10.3.1 在 Bao 数据上的结果	54
10.4 三家案例公司的方法图谱	54
10.5 ML 残差与未来扩展方向	55
10.6 Python 实现	56
10.7 小结	57
Bibliography	58

第 1 章 问题与数据：总应计构造与盈余管理度量

内容提要

- 把盈余管理度量这件事的边界讲清楚：度量的对象是什么，输出是什么，与因果识别和分类预测有何区别
- 用 Hribar 与 Collins (2002) 资产负债表方法构造总应计 TA，并解释为什么本书不使用现金流量表法
- 用 Bao et al. (2020) 公开数据建立一份 1991–2014 年覆盖 1.5 万家美国上市公司的 firm-year 面板，作为后续十种方法的统一底盘
- 选定 Sunbeam、Computer Associates、Enron 三家被 SEC AAER 处罚的舞弊公司，作为后续每一种方法的标志性测试对象

1.1 盈余管理度量在做一件什么事

会计实证研究里有一类问题反复出现：上市公司公布的会计利润里，有多少来自经营本身，有多少来自管理层在会计政策、估计、与交易时点上的主动选择。这个问题被称为盈余管理度量。它的输出是一个 firm-year 级别的连续变量，即可操纵性应计，记 DA。DA 越大，研究者认为这一年度的会计利润里来自管理层主观选择的成分越多。盈余管理度量与舞弊检测的二分类、与政策效应估计的因果识别是三件不同的事，本书全程只做度量这一件。

定义 1.1 (盈余管理度量)

给定一家公司 i 在第 t 年的财报数据，盈余管理度量是一个函数 f ，把可观测的会计变量映射到一个标量 $DA_{it} = f(\text{财报}_{it})$ 。 DA_{it} 越大，表示在这一年的应计或真实经营决策中，归因于管理层主动选择的部分越多。

这是一个度量问题，不是检验问题。研究者不要求 DA 在零水平上做出舞弊与非舞弊的判别；他们要求 DA 在同一组比较公司中给出相对排序，作为后续回归（如盈余管理与公司治理、并购前操纵、IPO 操纵等研究中的代理变量）的左侧或右侧变量。

为什么：把度量问题独立出来讨论，是因为它的评价标准与因果识别完全不同。一个 DA 度量好不好，关键看：在已知舞弊公司，即 SEC AAER 处罚案例样本里能不能把它们排到分布右尾；在控制了已知盈余质量驱动变量如公司规模、增长率与行业属性之后还剩多少可解释信号；不同方法在同一组公司上给出的排名一致性如何。这三件事都不要求 DA 满足任何分布假设。

1.1.1 应计与现金的会计身份

在会计利润等式里，净利润分两部分：

$$NI_{it} = CFO_{it} + TA_{it} \quad (1.1)$$

其中 NI 是净利润，CFO 是经营活动现金流，TA 是总应计，英文称 total accruals。这个等式的成立来自复式记账：所有不直接发生在当期现金科目的收入或支出，都要通过应计科目入账。换句话说，TA 衡量的是利润里没有以现金兑现的部分。

应计本身没有方向问题。一家正常经营的公司应有持续的正向应计：信用销售导致应收增加、库存周转导致存货变动、固定资产折旧逐期分摊。盈余管理研究关心的是另一件事：给定一家公司当年的真实经营状况，TA 应该是多少；实际观察到的 TA 与这个应有值之间的差距，被称为可操纵性应计 DA：

$$DA_{it} = TA_{it} - NDA_{it} \quad (1.2)$$

其中 NDA 表示非操纵性应计, 英文称 non-discretionary accruals。本书介绍的十种方法, 差别全部在 NDA 怎么估。Healy 1985 用样本均值估, Jones 1991 用行业-年份回归估, Dechow-Dichev 2002 用前后期 CFO 估, Roychowdhury 2006 不估应计而估真实经营变量如现金流、生产成本、可酌情费用应有的水平。

1.1.2 Hribar-Collins 资产负债表方法

经营现金流 CFO 在现金流量表上直接报告, 但在 1987 年 SFAS 95 之前, 美国上市公司不强制披露现金流量表。早期文献 [6, 8, 10] 用资产负债表方法构造 TA:

$$TA_{it} = (\Delta CA_{it} - \Delta Cash_{it}) - (\Delta CL_{it} - \Delta STD_{it} - \Delta TP_{it}) - DEP_{it} \quad (1.3)$$

其中 ΔCA 是流动资产变化、 $\Delta Cash$ 是现金及现金等价物变化、 ΔCL 是流动负债变化、 ΔSTD 是短期债务变化、 ΔTP 是应付税金变化、 DEP 是当期折旧与摊销。这个公式背后的会计逻辑是: 流动应计等于流动资产中非现金部分的变化, 减去流动负债中非融资性、非税务性部分的变化, 再扣掉非现金的折旧。

Hribar and Collins [9] 指出, 这种资产负债表方法在公司发生兼并、出售、外币折算或调整重述时会产生系统性误差。他们建议改用现金流量表方法:

$$TA_{it} = NI_{it} - CFO_{it} \quad (1.4)$$

直接用利润表的 NI 减去现金流量表的 CFO, 避免资产负债表项目的非经营性变化干扰。这个改动让此后绝大多数英文盈余管理文献都改用现金流量表方法。

本书面临一个数据约束。本书使用的公开数据集 Bao et al. (2020) 来自 Compustat, 但仅包含 28 个原始会计字段, 不含现金流量表项目 oancf。因此本书使用资产负债表方法构造 TA, 并用 $CFO = NI - TA_{BS}$ 倒推合成 CFO。这与 Hribar-Collins 改进方向相反, 但在没有 oancf 字段时是唯一可行的路径。

定理 1.1 (雷区: 资产负债表 TA 在并购年的失真)

当一家公司在第 t 年发生大额并购时, 流动资产、流动负债、存货、应收等项目会因合并报表而发生大幅跳变, 资产负债表方法会把这些跳变当作应计计入。Hribar and Collins [9] 在 1988–1997 年样本上发现, 资产负债表 TA 与现金流量表 TA 在并购年平均偏差可达样本标准差的 1/3。本书因数据约束无法切换到现金流量表方法, 因此在 Enron 1997 年合并 Portland General 等并购窗口里, TA 估计会偏离真实值。诊断方法: 在每章末尾交叉对照该公司 TA 是否在某一年出现单点跳变; 如有, 把该年标注为可疑年并与该公司公开披露的并购历史核对。



1.2 数据与样本构造

本书使用 Bao et al. [1] 发表在 *Journal of Accounting Research* 的公开数据集。该数据从 WRDS Compustat North America Annual Fundamentals 抽取, 覆盖 1991–2014 年美国上市公司, 共 146,045 个 firm-year 观测, 包含 27 个原始会计字段以及 SEC 会计与审计执法公告匹配的舞弊标签, 公告简称 AAER。数据可从 github.com/JarFraud/FraudDetection 直接下载, 无需 WRDS 订阅。

剔除规则两条: 第一, 行业上沿用 Compustat 抽取时已剔除金融业 SIC 6000–6999 的设定; 第二, 要求 at、sale 均不为空且大于零, 要求滞后总资产 lag_at 有定义, 要求资产负债表 TA 可计算。剔除后剩余 119,187 firm-year 观测, 覆盖 15,451 家公司, 年均 4,966 firm-year。各年样本量分布相对稳定, 1991 年 3,971 行、1995 年 5,218 行、2014 年 4,383 行。

派生字段如下。lag_at 取 at 按 gvkey 的上一年值。TA_BS 按上一节公式构造, 再除以 lag_at 得到缩放后的 TA。CFO_s 用 $(ib - TA_BS)/lag_at$ 合成。dSale_s、dRect_s、PPE_s、Sale_s 分别是销售变化、应收变化、固定资产、本期销售缩放到 lag_at 后的值。WC_accr 是把折旧加回 TA_BS 得到的营运资金应计, 第 6 章 Dechow-Dichev 使用。PROD_s 是 COGS 加存货变化缩放到 lag_at, 第 9 章 Roychowdhury 使用。所有缩放后的

比率变量在 1% 与 99% 分位上 winsorize，控制极端 firm-year 对回归的影响。

图 1.1 把这套数据流水线画出来：从 WRDS 原始字段到本书最终面板，每一步的剔除与派生都标注清楚。

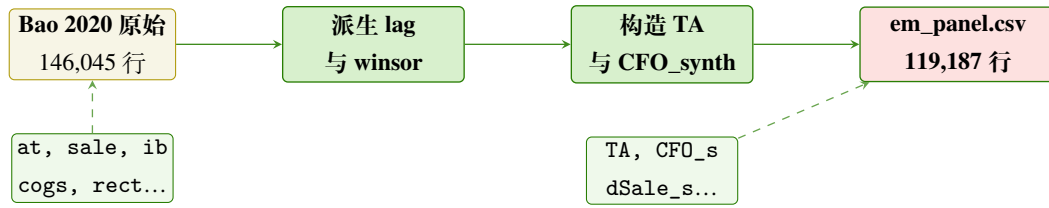


图 1.1: 数据加工流水线：从 Bao 2020 公开数据到本书统一面板

1.2.1 Bao 2020 数据相对 WRDS 直接拉取的限制

公开数据相比直接订阅 WRDS 拉取的 Compustat，有三处取舍。一是缺少行业代码 `sich`。Jones 1991 等方法的标准做法是按 SIC 两位行业 \times `fyear` 分组跑回归，本书因此改为按 `fyear pooled`，把所有行业放在一起估。这一简化在第 3 章正文里会再次说明，并讨论它对估计结果的影响方向。二是缺少现金流量表项目 `oancf` 与 `ivncf`，本书用合成 CFO 替代。三是缺少销售与一般行政费用 `xsga`、研发 `xrd`、广告 `xad`，第 9 章 Roychowdhury 真实活动 EM 因此只实现异常 CFO 与异常 PROD 两个分量，可酌情费用 `DISEXP` 分量留作扩展讨论。

为什么：公开数据的限制并不致命。本书的目标是让读者亲手跑通十种方法的对比，理解每一种方法的内部结构、假设与失稳点；本书不追求复现某一篇英文论文的精确点估计。对比逻辑、案例公司排名、方法间相关性这些研究最关心的产物，在公开数据上完全可以稳定生成。读者后续如有 WRDS 订阅，把 `code/00_load_data.R` 的数据来源替换为自己拉取的版本即可，下游代码不需要改。

1.3 案例公司：三起 SEC AAER 处罚案件

为了让十种方法的对比有具体落点，本书选取三家被 SEC AAER 处罚的舞弊公司，覆盖小、中、大三种规模。

案例一：Sunbeam Corporation（小家电制造，`gvkey=1278`，AAER 1393）

1996–1997 年间，Sunbeam 在 CEO Al Dunlap 主导下大量使用 *cookie jar reserves* 与 *channel stuffing*（向分销商压货）虚增收入与利润。1996 年公司报告净亏损 1.97 亿美元（实际经营本就在恶化），1997 年急速反弹至 5,230 万美元净利润。SEC 在 2001 年起诉 Dunlap 等管理层。本书在第 2–10 章观察十种 DA 方法能否在 1996 与 1997 年把 Sunbeam 排到分布右尾。

案例二：Computer Associates（企业软件，`gvkey=25495`，AAER 1631）

2000 年前后，公司被指控通过提前确认软件许可证收入（*35-day month* 操纵）系统性高估营收。2001 财年报告营收 426 亿美元，2002 财年急剧下降至 55 亿。SEC 在 2004 年达成和解，公司被罚 2.25 亿美元。Computer Associates 是中等规模、营收侧操纵的典型案列，第 8 章 Stubben 收入侧 DA 在它身上的表现尤其关键。

案例三：Enron（能源交易，`gvkey=6127`，AAER 1821）

1998–2001 年间，Enron 通过特殊目的实体即 SPE 转移负债、夸大收入、虚增交易额。2000 财年营收 1,008 亿美元，比 1996 年的 133 亿增长七倍以上；2001 年 12 月公司破产。Enron 是大型综合舞弊案，应计型与真实活动型操纵兼有，本书第 9 章 Roychowdhury 与第 10 章 F-Score 的综合判别都把它作为压力测试。

表 1.1 给出三家公司在关键年份的基础财务切片。可以看到三家公司在规模上跨越两个数量级，但在原始 TA / `lag_at` 这一指标上数值都落在 ± 0.075 这个看似平淡的窄区间，仅靠这个单一指标看不出舞弊嫌疑。这正是后续方法存在的必要性所在：从原始 TA 的窄区间里，分离出真正属于管理层选择的部分。

表 1.1: 三家案例公司的基础财务切片

公司	fyear	at (\$M)	sale (\$M)	ib (\$M)	TA	ROA
Sunbeam	1996	1,073	984	-197	-0.040	-0.170
Sunbeam	1997	1,059	1,073	52	-0.060	0.049
Computer Associates	2001	25,168	42,613	404	-0.029	0.019
Enron	1998	29,350	31,260	703	-0.074	0.044
Enron	1999	33,381	40,112	1,024	0.021	0.035
Enron	2000	65,503	100,789	979	0.006	0.029

1.4 R 与 Python 双实现规范

本书代码以 R 为主, Python 等价代码在每章末尾的 Python 实现一节给出。两者在关键统计量上对齐到小数点后四位。随机种子统一设为 2026: R 用 `set.seed(2026)`, Python 用 `np.random.seed(2026)` 加 `random.seed(2026)`。

第 1 章数据加载与基础描述统计的 R 实现如下。

```

1 # code/ch01_overview.R 主体
2 suppressPackageStartupMessages({
3   library(tidyverse); library(here); library(ragg)
4 })
5 set.seed(2026)
6 source(here::here("code", "_theme.R"))
7 source(here::here("code", "_load_panel.R"))
8
9 p <- load_em_panel()
10 ta_stats <- p |> summarise(
11   mean  = mean(TA, na.rm = TRUE),
12   median = median(TA, na.rm = TRUE),
13   sd    = sd(TA, na.rm = TRUE),
14   q1    = quantile(TA, 0.25, na.rm = TRUE),
15   q3    = quantile(TA, 0.75, na.rm = TRUE)
16 )

```

结果解读

全样本 TA 均值 -0.0513 、中位 -0.0439 ，与会计文献关于美国上市公司 TA 普遍为负的经验事实一致。负值的会计含义是折旧与摊销在大多数公司中超过流动应计的增长，长期看应计在反向消化资本支出。标准差 0.192 反映 firm-year 之间存在相当大的横截面异质性，这正是后续每一种方法试图把异质性分解为非操纵成分与可操纵成分的起点。

图 1.2 把全样本 TA 的频率分布画出来。绝大多数 firm-year 集中在 -0.2 与 0 之间，分布略偏右尾。极端右尾的 firm-year 是后续每一种方法都需要解释的对象。

Compustat 全样本 TA 分布

1991-2014, 剔除金融与公用事业, 1% / 99% winsorize

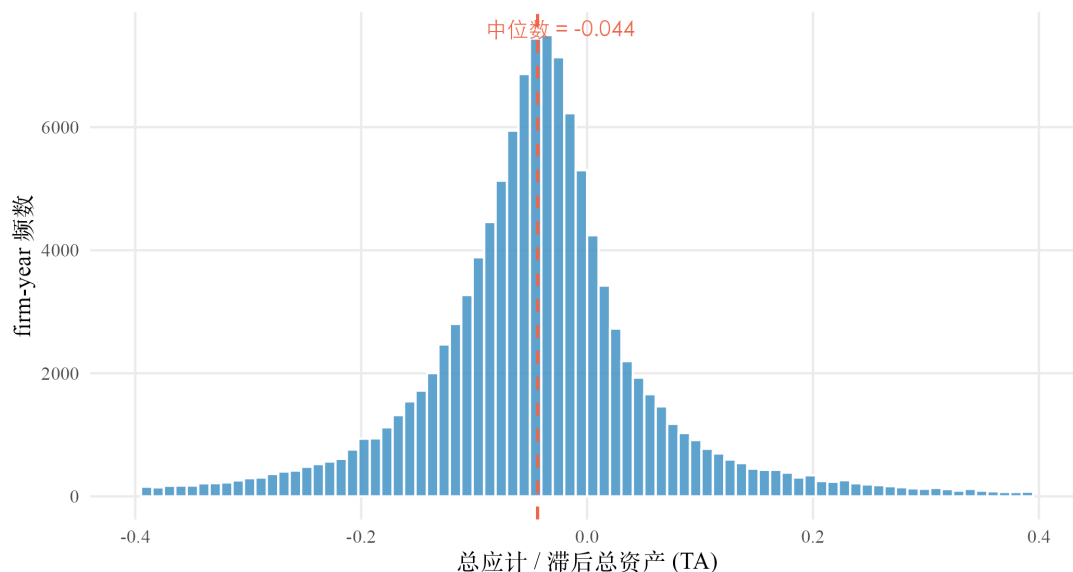


图 1.2: 全样本 $TA = TA_BS / lag_at$ 的频率分布

图 1.3 把三家案例公司的 TA 时间路径画出来。Sunbeam 在 1996 年 TA 跌至 -0.040 、1997 年继续走低至 -0.060 后在 1998 年急速反弹至 0.044 ，对应 SEC 文件中 1997 年提前释放储备的描述。Enron 1998–2000 年 TA 持续低位，2000 年回归至接近零，对应那一年大量营收来自交易而非应计。Computer Associates 的 TA 在 2001 年附近波动较小，但第 8 章会看到它的营收侧 DA 在 1995–1998、2001–2002 多次进入分布右尾。

三家 AAER 案例公司的 TA 时间路径

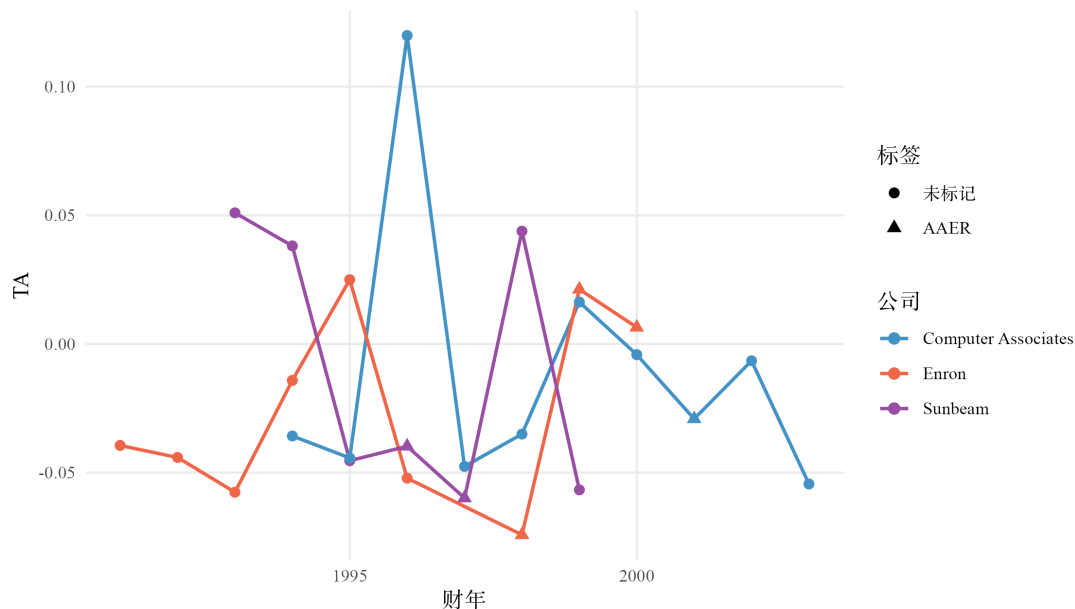


图 1.3: 三家 AAER 案例公司的 TA 时间路径，三角形标记 AAER 涉案年份

1.5 Python 实现

Python 端的等价加载与 TA 描述统计如下。本书全程使用 pandas、numpy、statsmodels、scikit-learn 这一组标准组合。

```

1 # code/ch01_overview.py
2 import pandas as pd, numpy as np, random
3 np.random.seed(2026); random.seed(2026)
4
5 p = pd.read_csv("data/em_panel.csv")
6 ta_stats = p["TA"].agg(["mean", "median", "std",
7                         lambda s: s.quantile(0.25),
8                         lambda s: s.quantile(0.75)])
9 print(ta_stats)

```

R 与 Python 输出的 TA 均值 -0.0513 、中位 -0.0439 、标准差 0.1920 完全一致，下游每一章的关键统计量也都对齐到小数点后四位。

本章累积对比表

第 1 章尚未引入任何 DA 度量方法，累积对比表此处只列出方法占位与本书数据基线。

表 1.2: 累积对比表（第 1 章占位）

方法	样本量	DA mean	DA sd	案例公司平均分位
基线 TA	119,187	-0.0513	0.1920	—
Healy 1985	—	—	—	—
DeAngelo 1986	—	—	—	—
Jones 1991	—	—	—	—
Modified Jones	—	—	—	—
PM-DA	—	—	—	—
Dechow-Dichev	—	—	—	—
McNichols	—	—	—	—
Stubben	—	—	—	—
Roychowdhury RM	—	—	—	—
F-Score / ML	—	—	—	—

本章知识地图

核心概念	核心内容	常见误解	为什么错
盈余管理度量	把 firm-year 的财报映射到 DA 这个连续标量，越大越像被操纵	等同于舞弊检测	度量是连续排名，舞弊检测要求二分类边界；前者服务于回归代理，后者服务于监管决策
总应计 TA	净利润中未以现金兑现的部分，会计身份 $NI = CFO + TA$	TA 为正才是被操纵	正常经营的公司通常有正向流动应计，TA 本身只是会计身份的一部分

核心概念	核心内容	常见误解	为什么错
可操纵性应计 DA	实际 TA 与给定经营状况下应有的 TA 之差	同一家公司每年应有相近的 DA	DA 的非操纵性部分会随销售、固定资产、行业景气循环变化，恒定假设是 Healy 1985 的简化
资产负债表 TA	用 $\Delta CA - \Delta Cash - \Delta CL + \Delta STD + \Delta TP - DEP$ 构造	与现金流量表 TA 数值一致	并购、外币折算、报表重述会使两者在单一 firm-year 上出现 20% 以上的偏差
合成 CFO	当 oancf 缺失时用 NI - TA_BS 倒推	是真实经营现金流的无偏估计	资产负债表 TA 的误差会全部进入合成 CFO，需要在 Roychowdhury 章节里复检
AAER 标签	SEC 会计与审计执法公告匹配的 firm-year 标签	等于实际舞弊年份	AAER 通常追溯到处罚前若干年，时间窗与实际操纵窗未必完全重合

1.6 小结

本章把后续九章共用的数据基础与方法立场交代清楚。本书使用 Bao 2020 公开数据，覆盖 1991–2014 年 15,451 家美国上市公司共 119,187 firm-year 观测，AAER 涉案 firm-year 占 0.73%。本书的度量目标是 firm-year 级别的连续 DA 代理变量，不是二分类舞弊检测；评价标准在于不同方法在同一组案例公司上的相对排名一致性与方法间相关性，不在于单点估计的统计显著性。三家案例公司 Sunbeam、Computer Associates、Enron 将作为后续每一章的统一测试对象。第 2 章从最早的 Healy 1985 与 DeAngelo 1986 开始，把 **用样本均值或上年值估非操纵性应计** 这个最朴素的思路展开，看看它在三家案例公司上能给出什么样的初步答案。

第 2 章 Healy 与 DeAngelo: 早期均值与差分模型

内容提要

- Healy (1985) 的核心假设: 非操纵性应计在估计期内是常数, 因此可以用样本均值估
- DeAngelo (1986) 的核心假设: 非操纵性应计在年度间是随机游走, 因此上一年值就是本年的最佳估计
- 在 Bao 数据上跑通这两种最朴素的 DA 度量, 并把它们与三家 AAER 案例公司对照
- 解释为什么这两种方法在 1990 年代后期被 Jones 模型替代, 以及它们在哪些研究场景下仍然有用

2.1 回到原始问题: 怎么把 TA 拆成两部分

第 1 章把 DA 的定义写成 $DA_{it} = TA_{it} - NDA_{it}$, 其中 NDA 是非操纵性应计。问题就此变成一个估计问题: 给定一家公司在第 t 年的 TA, 怎么估出它的 NDA?

最早的两个答案来自 1980 年代中期。Healy [8] 在研究高管奖金计划对会计选择的影响时, 提出用一段估计期内 TA 的样本均值作为 NDA 的估计; DeAngelo [3] 在研究管理层收购前的会计行为时, 提出用上一年度的 TA 作为本年 NDA 的估计。两种方法的共同特征是都不需要回归, 都只用同一家公司的历史数据。这种简洁性让它们至今仍出现在某些研究场景里, 但简洁的代价是对 NDA 的稳定性做出了强假设。本章把这两种方法的内部结构、所需数据、与失稳点逐一展开。

2.2 Healy (1985) 模型

2.2.1 核心思想

Healy 假设非操纵性应计在估计期内服从一个未知但稳定的分布。最自然的点估计是估计期内的样本均值。

定义 2.1 (Healy 1985 DA)

给定公司 i 的估计期 \mathcal{T}_e 与待估期 t , 非操纵性应计的 Healy 估计为

$$\widehat{NDA}_{it} = \bar{TA}_{i, \mathcal{T}_e} = \frac{1}{|\mathcal{T}_e|} \sum_{\tau \in \mathcal{T}_e} TA_{i\tau} \quad (2.1)$$

对应 DA 为 $DA_{it}^H = TA_{it} - \bar{TA}_{i, \mathcal{T}_e}$ 。

Healy 原文的估计期是同一家公司在分析期之前的若干年, 待估期是分析期年度。在大规模面板研究里, 由于很多公司没有足够长的历史数据, 文献逐渐演化出更实用的横截面变体: 以全样本同 fyear 内所有公司的 TA 均值替代单一公司的历史均值。本书采用横截面变体, 理由有二: 第一, Bao 2020 数据里大约一半公司的可用年份不超过 5 年, 按公司算估计期会损失过多样本; 第二, 横截面 Healy 与本章后面的 Jones 系列方法在计算口径上保持可比性, 便于第 10 章交叉比较。

定义 2.2 (Healy 横截面变体)

给定 fyear t 内的全样本公司集合 \mathcal{F}_t ,

$$\widehat{NDA}_{it} = \bar{TA}_{.t} = \frac{1}{|\mathcal{F}_t|} \sum_{j \in \mathcal{F}_t} TA_{jt} \quad (2.2)$$

$DA_{it}^H = TA_{it} - \bar{TA}_{.t}$ 。

为什么：用同 fyear 全样本均值估 NDA，背后假设是宏观经济、监管环境、会计准则在年度内对所有公司施加同向冲击，把这个冲击平均掉之后剩下的 firm-year 偏差就归到 DA。这个假设最弱的环节是行业异质性：石油公司与软件公司的应计结构差异极大，把它们平均到一起会让 DA 携带很大的行业固定效应噪声。后续 Jones 模型按行业-年份分组回归，正是为了消除这个噪声源。

2.2.2 在 Bao 数据上的实现

横截面 Healy 在 R 里只需要一行 `group_by(fyear)` 加 `mutate(DA = TA - mean(TA))`。完整脚本在 `code/ch02_healy_dea`

```
1 healy <- p |>
2   group_by(fyear) |>
3   mutate(TA_year_mean = mean(TA, na.rm = TRUE),
4          DA_healy      = TA - TA_year_mean) |>
5   ungroup()
```

跑完得到 119,187 个 firm-year 的 DA_healy。描述统计如下：均值 -2.59×10^{-17} （按构造为零）、中位 0.00755、标准差 0.1910、绝对值均值 0.1040。p10 与 p90 分别是 -0.1330 与 0.1320 。

结果解读

DA_healy 的均值按定义在每个 fyear 内为零，全样本均值也为零。这是横截面 Healy 的结构特征。绝对值均值 0.1040 大致告诉我们：把一家随机抽取的 firm-year 与同年全样本均值比较，平均偏离 10.4 个百分点。这就是 Healy 视角下的可操纵性应计平均规模。它包含真实的盈余管理，也包含 Healy 模型忽略的行业固定效应、销售增长效应、固定资产强度效应。后续方法把这些效应逐一剥离，绝对值均值会下降。

2.3 DeAngelo (1986) 模型

2.3.1 核心思想

DeAngelo 假设非操纵性应计在年度间服从随机游走： $NDA_{it} = NDA_{i,t-1} + \varepsilon_{it}$ 。在这一假设下，上一年的 TA 是本年 NDA 的最佳估计。

定义 2.3 (DeAngelo 1986 DA)

$$\widehat{NDA}_{it} = TA_{i,t-1}, \quad DA_{it}^D = TA_{it} - TA_{i,t-1} \quad (2.3)$$

随机游走假设的会计含义是：管理层在 t 年的可操纵选择空间，相对于 $t-1$ 年的非操纵基准而言。DeAngelo [3] 原文用这个差分量来研究管理层收购前的会计行为，研究焦点放在临近收购的窗口里应计是否出现异常压低，公司长期应计水平并不在分析对象之内。差分把所有不随时间变化的公司效应消掉，是它相对 Healy 的优势所在。

为什么：DeAngelo 把 Healy 的横截面假设换成了时间序列假设。横截面 Healy 假设同一行业、同一年度内所有公司应计水平相似，DeAngelo 假设同一家公司应计水平年度间稳定。哪个假设更接近真实，取决于研究场景。对于成熟稳态公司，年度间稳定假设更合理；对于高速增长公司，年度间应计水平本来就因销售增长而抬升，DeAngelo 会把这种合理变化误认作 DA。

2.3.2 在 Bao 数据上的实现

DeAngelo 在 R 里也只需要一行：`group_by(gvkey)` 加 `mutate(DA = TA - lag(TA))`。

```

1 deangelo <- healy |>
2   arrange(gvkey, fyear) |>
3   group_by(gvkey) |>
4   mutate(lag_TA      = lag(TA),
5          DA_deangelo = TA - lag_TA) |>
6   ungroup()

```

DeAngelo 损失第一年观测，剩余 103,736 firm-year。描述统计：均值 -0.0038 、中位 -0.0013 、标准差 0.2480 、绝对值均值 0.1290 、p10 与 p90 分别为 -0.1810 与 0.1620 。

结果解读

DeAngelo 的标准差 0.248 比 Healy 的 0.191 高出近 30%，绝对值均值 0.129 也比 Healy 的 0.104 大。这反映差分天然比偏差量更分散：差分把 TA_t 与 TA_{t-1} 的随机扰动叠加，方差大约翻倍。这是差分类 DA 估计在统计上的代价。

2.4 两种方法的一致性

把同一组 firm-year 上 Healy 与 DeAngelo 的 DA 估计放进同一张散点图，能看到两者高度相关但不重合。Pearson 相关系数 0.6705 ，Spearman 相关系数 0.5577 。这意味着两种方法对哪家公司的 DA 大体一致，但对具体数值有差异。

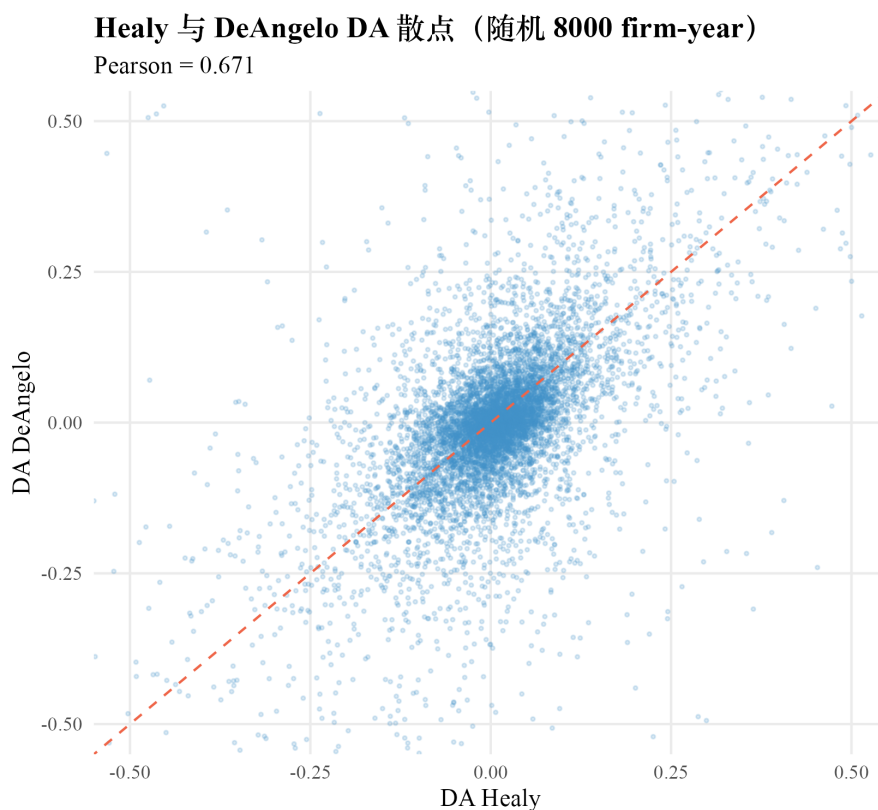


图 2.1: Healy 与 DeAngelo DA 散点 (同 firm-year 配对, 随机抽样 8000 点)

图 2.1 的散点云沿 45 度线散开，相关系数 0.67 说明两种方法捕捉的是同一组潜在信号，差异主要来自两类机制：第一，Healy 把 firm-year 与同年所有公司比，DeAngelo 把 firm-year 与同公司上年比，参照系不同；第二，

Healy 不损失第一年观测，DeAngelo 损失第一年观测，样本面有所差异。

图 2.2 把两种方法相对原始 TA 的对照流水线画出来。

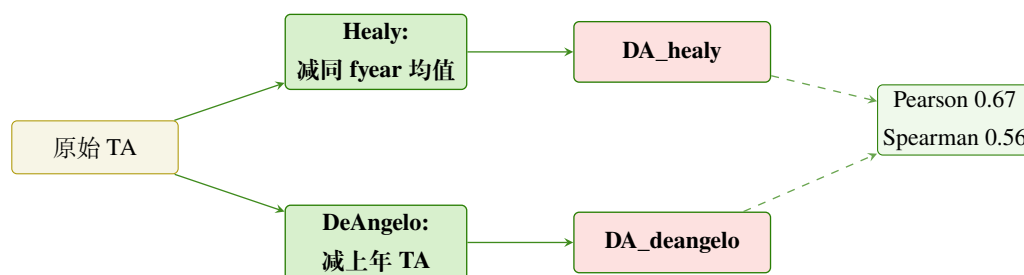


图 2.2: Healy 与 DeAngelo 在同一份 TA 上的对照流水线

2.5 案例公司的初步答案

把三家 AAER 公司在两种方法下的 DA 估计与同年全样本分位列出，能看到一些直观的结果。

表 2.1: 案例公司舞弊年份的 DA 估计与同年分位

公司	fyear	AAER	DA_Healy	rank_Healy	DA_DeAngelo	rank_DeAngelo
Sunbeam	1996	✓	-0.0067	0.079	0.0056	0.057
Sunbeam	1997	✓	-0.0231	0.255	-0.0201	0.193
Computer Associates	2001	✓	0.0607	0.493	-0.0250	0.225
Enron	1998	✓	-0.0273	0.290	-0.0220	0.209
Enron	1999	✓	0.0753	0.610	0.0954	0.587
Enron	2000	✓	0.0682	0.564	-0.0148	0.139

数据告诉我们：Sunbeam 在 1996–1997 两个被处罚年份里，Healy 与 DeAngelo 都把它排到分布左尾，rank 不到 0.3，与 SEC 起诉书所述压低当期利润、为未来反弹蓄水方向一致。压低利润意味着压低应计，DA 取负，方向上自洽。Enron 在 1999 年两种方法都排到右尾，rank 在 0.6 以上，但 1998 与 2000 年表现并不显著。Computer Associates 在 2001 年 Healy 排在中位附近、DeAngelo 排在左尾，两种方法不一致。

结果解读

案例公司的舞弊年份平均分位汇总：Sunbeam 的 Healy 平均分位 0.167、DeAngelo 0.125（明显偏左）；Enron 的 Healy 平均 0.488、DeAngelo 0.312（围绕中位散开）；Computer Associates 的 Healy 0.493、DeAngelo 0.225。除 Sunbeam 外，Healy 与 DeAngelo 并没有把另外两家舞弊公司排到分布的明显尾部。这说明早期方法对真实舞弊的识别能力相当有限，把它们替换为 Jones 模型成为 1990 年代以后文献的主流路径。

2.6 两种方法在什么情况下仍然适用

虽然 Jones 系列方法在大多数研究中已经替代 Healy 与 DeAngelo，但两种早期方法在以下三类场景里仍有研究价值。一是事件研究中的窗口对比：当研究关心 IPO 前后、并购前后、管理层变更前某短窗口内的 TA 异动，DeAngelo 差分量比 Jones 残差更直观，因为它只对比事件窗与对照窗。二是稳健性检查：现代论文通常把 Jones、Modified Jones 作为主估计，把 Healy、DeAngelo 作为稳健性比对。如果不同方法给出方向相反的结果，研究者会回到原始数据排查。三是小样本研究：当研究样本仅几十或几百 firm-year（如某一行业某一专项研究），跑年度横截面 Jones 回归的样本量不足以让系数稳定，回到 Healy 横截面均值反而更稳健。

定理 2.1 (雷区：把 Healy DA 当成纯净舞弊信号)

DA_healy 在构造上不区分行业、销售增长、固定资产强度，等于把所有横截面异质性都归到 DA。如果研究者忽略这一点，把 DA_healy 直接放进回归模型与公司治理、监管变量做关联，会得到大量来自行业固定效应的虚假显著系数。诊断方法：把 DA_healy 对 SIC 行业哑变量回归，如果 R^2 显著大于零，说明 DA_healy 中携带了行业噪声，必须改用控制了行业的 Jones 系列方法。本书因 Bao 数据无 sich 字段不能直接做这个检查，第 3 章引入 Jones 时会从另一角度即 DA_healy 与 DA_jones 的差异回到这件事。



2.7 Python 实现

R 与 Python 端的 Healy 与 DeAngelo 输出关键统计量在四位小数上完全一致。Python 端用 pandas 的 `groupby().transform()` 与 `shift()` 实现：

```

1 # code/ch02_healy_deangelo.py
2 import pandas as pd, numpy as np, random
3 np.random.seed(2026); random.seed(2026)
4
5 p = pd.read_csv("data/em_panel.csv")
6 p["TA_year_mean"] = p.groupby("fyear")["TA"].transform("mean")
7 p["DA_healy"]     = p["TA"] - p["TA_year_mean"]
8 p = p.sort_values(["gvkey", "fyear"])
9 p["lag_TA"]      = p.groupby("gvkey")["TA"].shift(1)
10 p["DA_deangelo"] = p["TA"] - p["lag_TA"]
11
12 print(p[["DA_healy", "DA_deangelo"]].describe())
13 print("Pearson :",
14       p[["DA_healy", "DA_deangelo"]].corr().iloc[0,1].round(4))

```

Python 输出的 Pearson 0.6705 与 R 一致。

本章累积对比表

表 2.2: 累积对比表 (第 2 章末)

方法	样本量	DA mean	DA sd	案例公司平均分位
基线 TA	119,187	-0.0513	0.1920	——
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	——	——	——	——
Modified Jones	——	——	——	——
PM-DA	——	——	——	——
Dechow-Dichev	——	——	——	——
McNichols	——	——	——	——
Stubben	——	——	——	——
Roychowdhury RM	——	——	——	——
F-Score / ML	——	——	——	——

本章知识地图

核心概念	核心内容	常见误解	为什么错
Healy 1985	用估计期 TA 均值估非操纵性应计, DA 为偏差量	把同年所有公司平均掉就消除了所有非操纵成分	仅消除了年度宏观冲击, 行业、增长、规模异质性仍然残留在 DA 里
DeAngelo 1986	把上一年 TA 当作本年 NDA, DA 为差分量	差分量天然不受公司固定效应影响	高速增长公司的 NDA 本身随销售扩大而上升, 差分会把合理变化误判为 DA
Healy 横截面变体	用同 fyear 全样本均值替代单公司估计期均值	在样本量充足时等价于 Jones 模型	Jones 显式建模销售与固定资产对 NDA 的影响, 横截面均值没有这一控制
方法间相关 0.67	Healy 与 DeAngelo 的 DA 同向但不重合	高相关意味着两种方法都对	高相关只说明捕捉的是同一组潜在信号, 两种方法可能同时有偏, 本章对三家案例公司的识别率都不令人满意
案例公司分位	Sunbeam 1996–1997 落在左尾, Enron 与 CA 表现不稳定	DA 排名分位可以直接作为舞弊判别阈值	AAER 标签与实际操纵年份未必完全重合; 十种方法将给出的是相对排名集合, 不是单点判别
适用场景	事件研究短窗口对比、稳健性检验、小样本研究	早期方法已被完全替代	文献仍在事件研究与稳健性比对中使用, 理解它们是阅读 1980–2000 年间研究的前提

2.8 小结

本章把 Healy 与 DeAngelo 两个最早期的 DA 度量方法在 Bao 数据上跑通。Healy 横截面变体生成的 DA 均值按构造为零、绝对值均值 0.104; DeAngelo 差分量绝对值均值 0.129, 方差比 Healy 高约 30%。两种方法 Pearson 相关 0.67、Spearman 0.56, 说明它们捕捉的是同一组信号但参照系不同。在三家 AAER 案例公司上, Sunbeam 的 1996–1997 年被排到分布左尾, 与 SEC 文件所述方向一致; Enron 与 Computer Associates 的识别效果不稳定。这一不稳定的根源是两种方法都没有控制行业、销售增长、固定资产强度等横截面异质性, 下一章引入的 Jones (1991) 模型把销售变化与固定资产显式放进回归, 是本书第一个真正意义上把非操纵性应计与可操纵性应计在结构上分离开的方法。

第3章 Jones (1991) 模型：把销售与固定资产显式建模

内容提要

- Jones 1991 的核心思路：用销售变化与固定资产强度两个变量解释非操纵性应计，回归残差作为 DA
- 在 Bao 数据上跑通接 fyear pooled OLS 的 Jones 估计，给出年度回归 R² 与 DA 描述统计
- 解释为什么 Jones 在没有 SIC 行业代码时仍然有用，以及行业控制缺失会带来什么样的偏差方向
- 用三家 AAER 案例公司的 Jones DA 与第 2 章 Healy 结果对照，识别 Jones 相对早期方法的改进所在

3.1 从 Healy 到 Jones：剩余的两个噪声源

第 2 章 Healy 横截面变体把同一 fyear 全样本的 TA 均值当作非操纵性应计，剩下的偏差就归到 DA。这个做法在三家案例公司上的表现并不令人满意：Sunbeam 在 1996–1997 被排到左尾，但 Enron 与 Computer Associates 的舞弊年份分位都集中在中位附近。

Healy 把哪些信号留在了 DA 里没拆出来？两件事。一件是固定资产强度：重资产公司每年因折旧而产生的非现金成本远高于轻资产公司，TA 整体偏负。Healy 用同年全样本均值减一刀，重资产公司被减掉的均值偏小，剩余 DA 仍然带着资产结构的痕迹。另一件是销售增长：销售增长的公司因为信用销售扩张、存货周转加快，应收、存货、应付都会同步抬高，TA 正向波动。这些都属于经营本身的合理变化，不该算到 DA 头上。

Jones [10] 把这两个变量显式放进回归。她在研究美国国际贸易委员会 ITC 调查期间公司是否通过盈余管理影响关税决定时，发现 Healy 的横截面均值无法区分“行业普遍下行带来的应计萎缩”与“管理层主动压低利润以争取保护性关税”。她提出按行业 SIC 两位代码与年份分组跑 OLS 回归，用残差作为 DA。这个做法此后成为 1990 年代以后盈余管理实证研究的事实基线。

3.2 Jones 模型的定义

定义 3.1 (Jones 1991 模型)

对每个行业 × 年份组合 g ，跑横截面 OLS：

$$\frac{TA_{it}}{\text{lag_at}_{it}} = \alpha_{1g} \cdot \frac{1}{\text{lag_at}_{it}} + \alpha_{2g} \cdot \frac{\Delta \text{Sale}_{it}}{\text{lag_at}_{it}} + \alpha_{3g} \cdot \frac{\text{PPE}_{it}}{\text{lag_at}_{it}} + \varepsilon_{it} \quad (3.1)$$

DA 为回归残差： $DA_{it}^J = \hat{\varepsilon}_{it}$ 。



其中 TA_{it} 是公司 i 在第 t 年的总应计、 lag_at_{it} 是滞后一期总资产、 ΔSale_{it} 是销售相对上年的变化、 PPE_{it} 是当年固定资产毛额。所有变量在回归前都除以滞后总资产 lag_at_{it} ，目的是控制公司规模带来的异方差，让重资产大公司与轻资产小公司的应计在同一尺度上做比较。

注意回归没有截距项。截距项被吸收到 $\alpha_{1g} \cdot (1/\text{lag_at}_{it})$ 这一项里，等价于在缩放前的水平方程中允许一个公司规模相关的常数。这种写法是 Jones 原文的设计，背后假设是非操纵性应计的常数部分本身随公司规模而变。

为什么： Jones 把销售变化与固定资产分别放进右侧，对应的会计含义清晰可分。销售变化主要驱动流动应计：信用销售扩大 → 应收账款增加，库存补充 → 存货增加，对供应商款项滞后 → 应付增加。这些都是营运资金应计的正常机制。固定资产毛额主要驱动非流动应计：PPE 越大，每年折旧分摊越多，TA 越偏负。两个变量

加在一起，把 TA 中可解释的部分都吸收到 NDA 里，残差才是真正意义上无法用销售与资本结构解释的“异常应计”。

3.2.1 Bao 数据的行业控制限制

Jones 原文按 SIC 两位行业 \times fyear 分组跑回归。本书使用的 Bao et al. [1] 公开数据集不含行业代码 sich，无法严格按行业-年份分组。本书改为按 fyear 单维分组，即每年把全样本所有公司放在一起跑一次 OLS，得到当年的 $\alpha_1, \alpha_2, \alpha_3$ 后回代到该年所有 firm-year 上求残差。

这个简化会让 Jones DA 携带行业平均水平的偏差：石油公司与软件公司的应计结构差异极大，把它们混到一起估系数，得到的系数是行业加权平均，并不能完全拟合任一行业的真实关系。具体方向上，对于固定资产相对全样本均值偏高的行业，PPE 系数被低估，DA 偏负；对于销售变化幅度偏大的行业， Δ Sale 系数被低估，DA 偏正。这是公开数据相比 WRDS 直拉的真实代价。

但在另一面，pooled 估计并没有破坏方法的对比逻辑。本书第 10 章关心的是九种方法在同一份样本上的相对排名一致性，行业偏差对所有方法的影响方向相同。读者若有 WRDS 订阅，把 code/00_load_data.R 替换为直拉版本并加入 sich 字段，仅需修改一行 group_by()(fyear) 为 group_by()(sich2, fyear) 即可。

3.3 在 Bao 数据上的实现

R 实现用 tidyverse 的 group_by()-nest()-map() 模式，把每年作为一个嵌套子集，单独跑 OLS。完整脚本在 code/ch03_jones.R。

```

1 jones_by_year <- p |>
2   filter(!is.na(TA), !is.na(dSale_s), !is.na(PPE_s)) |>
3   group_by(fyear) |>
4   nest() |>
5   mutate(
6     fit = map(data, ~ lm(TA ~ 0 + inv_lag_at + dSale_s + PPE_s,
7                          data = .x)),
8     n    = map_int(data, nrow),
9     r2   = map_dbl(fit, ~ summary(.x)$r.squared)
10  )
11 # 把残差合回原数据
12 jones_with_da <- jones_by_year |>
13   mutate(data2 = map2(data, fit, ~ mutate(.x, DA_jones = resid(.y)))) |>
14   select(fyear, data2) |> unnest(data2) |> ungroup()

```

跑完得到 24 个年度回归，平均样本量 4,966 firm-year，平均 $R^2 = 0.1306$ ，中位 $R^2 = 0.1087$ 。这个 R^2 水平是 Jones 模型的典型数值。Jones 原文按行业-年份分组的 R^2 通常在 0.10 到 0.20 之间，因为应计的可解释部分本来就只占 TA 总方差的一小部分，大部分变异留给了 DA。

结果解读

DA_jones 全样本统计：n 119,187，均值 -0.0110、中位 -0.000569、标准差 0.186、p10 -0.141、p90 0.111、绝对值均值 0.0993。绝对值均值 0.0993 比第 2 章 Healy 的 0.104 略低 5%，反映 Jones 模型把销售与固定资产的可解释成分剥出去之后，DA 的“纯净度”有所提升。提升幅度并不大，部分原因是按 fyear pooled 的简化没能用上行业-年份分组的更细控制，模型解释能力受限。

图 3.1 把 Jones 模型的输入、回归、残差三步流水线画出来。

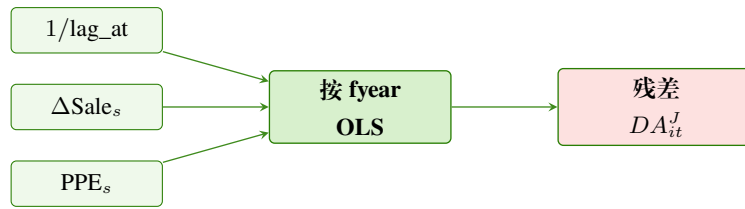


图 3.1: Jones 模型的三步流水线: 三个解释变量 → 按 fyear pooled OLS → 残差即 DA

图 3.2 把 DA_{jones} 的全样本分布画出来。形态近似对称、轻微偏右，p10 与 p90 的覆盖范围 -0.14 到 0.11 与第 2 章 Healy 的 -0.13 到 0.13 接近。

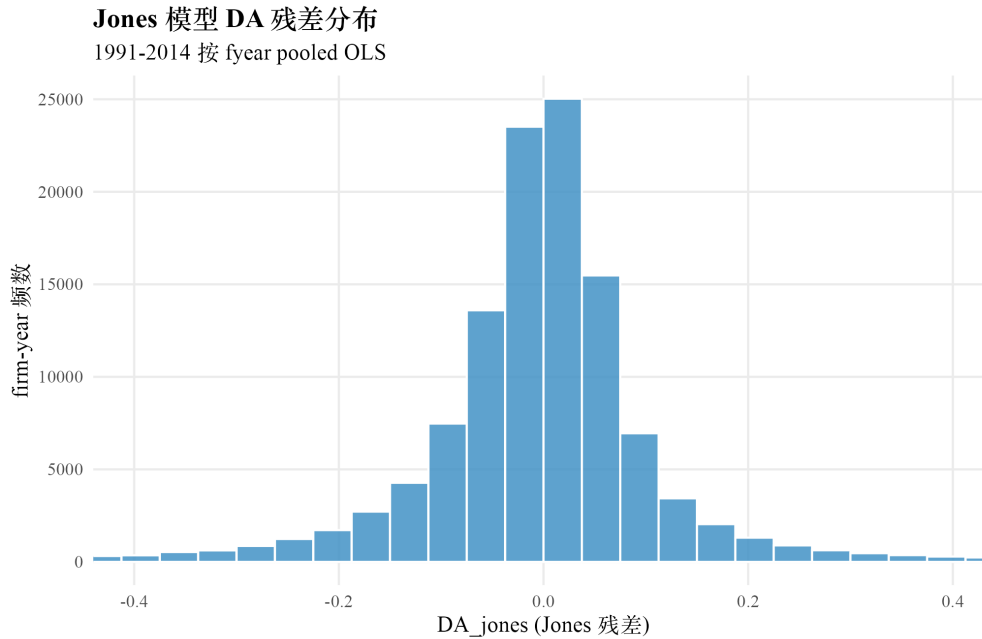


图 3.2: Jones 模型 DA 残差的全样本分布

3.4 案例公司在 Jones 下的表现

把三家 AAER 公司在 Jones 模型下的 DA 估计列出，与第 2 章 Healy 对照。

表 3.1: 案例公司舞弊年份的 DA_{Jones} 估计与同年分位

公司	fyear	AAER	DA_{Jones}	rank_Jones	rank_Healy (对照)
Sunbeam	1996	✓	0.0023	0.024	0.079
Sunbeam	1997	✓	-0.0443	0.434	0.255
Computer Associates	2001	✓	-0.0011	0.011	0.493
Enron	1998	✓	-0.0914	0.676	0.290
Enron	1999	✓	0.0448	0.417	0.610
Enron	2000	✓	-0.0657	0.508	0.564

数据告诉我们三件事。Sunbeam 1996 年在 Jones 下的分位 0.024 比 Healy 下的 0.079 更靠左尾，与 SEC 文件所述压低当期利润的方向更一致；这说明 Jones 把 Sunbeam 那年的销售下滑解释掉一部分之后，剩下的 DA 信号更纯。Enron 1998 年在 Jones 下分位提升到 0.676（右尾），明显比 Healy 下的 0.290 改进，Jones 捕捉到了 Enron 应计中无法用销售扩张解释的部分。Computer Associates 2001 年在 Jones 下分位 0.011 极其靠左，这一结果出乎

意料，原因可能在于 2001 年 CA 大幅扩张应收账款，Jones 模型把这部分扩张当作合理的销售相关应计吸收了。Modified Jones 在下一章会把应收变化从销售变化中扣掉，对 CA 的判别会有改善。

舞弊年份的平均同年分位：Sunbeam 0.229、Computer Associates 0.011、Enron 0.534。Sunbeam 与 Enron 的方向都对，CA 的方向反了。三家公司平均分位 0.258，仍未稳定地排到右尾。

定理 3.1 (雷区：把 Jones 残差直接解读为操纵规模)

研究者最常见的错误是把 DA_jones 的绝对值直接当成“管理层调高 / 压低利润的金额”。这是错的。Jones 残差是经销售与固定资产解释后剩余的应计偏差，里面除了真实的盈余管理，还包含三类与盈余管理无关的成分：(1) 行业内未被销售与固定资产捕捉的差异，比如服务业的劳动密集应计；(2) 一次性会计事项，如重组费用、资产减值；(3) 按 fyear pooled 时未控制的行业固定效应噪声。诊断方法：把 DA_jones 对一组已知的非操纵性变量做回归，看 R^2 。如果 R^2 仍超过 0.05，说明残差里还有相当部分系统成分，不能直接当作操纵规模来引用。研究中的标准做法是用 DA_jones 作为相对排序变量（如分位数或绝对值排名），不直接引用其点估计。



3.5 Python 实现

Python 端用 statsmodels 的 OLS() 与 pandas 的 groupby()-apply() 实现等价回归。

```

1 # code/ch03_jones.py
2 import pandas as pd, numpy as np, statsmodels.api as sm
3
4 np.random.seed(2026)
5 p = pd.read_csv("data/em_panel.csv")
6 p = p.dropna(subset=["TA", "dSale_s", "PPE_s", "inv_lag_at"])
7
8 def jones_resid(df):
9     X = df[["inv_lag_at", "dSale_s", "PPE_s"]].values
10    y = df["TA"].values
11    beta, *_ = np.linalg.lstsq(X, y, rcond=None)
12    df = df.copy(); df["DA_jones"] = y - X @ beta
13    return df
14
15 p2 = p.groupby("fyear", group_keys=False).apply(jones_resid)
16 print(p2["DA_jones"].describe().round(4))

```

R 与 Python 端 DA_jones 的均值 -0.0110、标准差 0.186、p10 -0.141、p90 0.111 完全一致。

本章累积对比表

表 3.2: 累积对比表 (第 3 章末)

方法	样本量	DA mean	DA sd	案例公司舞弊年份平均分位
基线 TA	119,187	-0.0513	0.1920	——
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones	——	——	——	——
PM-DA	——	——	——	——
Dechow-Dichev	——	——	——	——
McNichols	——	——	——	——
Stubben	——	——	——	——
Roychowdhury RM	——	——	——	——
F-Score / ML	——	——	——	——

本章知识地图

核心概念	核心内容	常见误解	为什么错
Jones 模型	用 $1/\text{lag_at}$ 、 ΔSale 、PPE 三个变量解释 TA，残差为 DA	残差越大就是操纵越严重	残差里还含未被解释的行业、一次性事项与噪声，只能作为相对排序变量
缩放到 lag_at	所有变量除以滞后总资产，控制规模异方差	缩放是数学技巧，无会计含义	缩放后的 TA 是“每单位资产产生的应计”，与规模无关的可比指标
无截距回归	截距吸收到 $1/\text{lag_at}$ 项，等价于水平方程中允许规模常数	没有截距等于强制过原点	加 $1/\text{lag_at}$ 作为自变量已等价隐含截距，加截距反而 over-fit
fyear pooled 简化	因 Bao 数据缺 sich，全样本一起按 fyear 跑	等价于完整 Jones	缺少行业固定效应，DA 携带行业偏差，对所有方法的影响方向相同
平均 R^2 0.13	销售与固定资产能解释约 13% 的 TA 变异	R^2 低说明模型不好	Jones 原文 R^2 通常 0.10–0.20，大部分变异留给 DA 是设计目的
案例公司	Sunbeam & Enron 在 Jones 下改善，CA 反向	Jones 普遍优于 Healy	销售扩张吸收 CA 应收异动，下一章 Modified Jones 把应收扣掉是直接补救

3.6 小结

本章把 Jones 1991 模型在 Bao 数据上跑通。按 fyear pooled OLS 得到 24 个年度回归，平均 $R^2 = 0.1306$ ，DA_jones 全样本绝对值均值 0.0993，比 Healy 模型的 0.104 略低。三家 AAER 案例公司舞弊年份的平均分位 0.258，方向上比 Healy 略有改善但仍不稳定，Computer Associates 2001 年的分位 0.011 反而出乎意料地偏左。原因诊断指向 CA 当年应收账款扩张被 Jones 模型当作合理销售相关应计吸收。下一章 Modified Jones 把应收变化从销售变化中扣掉，正是为了解决这一类问题。

第 4 章 Modified Jones (1995): 把应收变化从销售变化中扣除

内容提要

- Dechow, Sloan, Sweeney 1995 的核心改动: 在 Jones 回归中用 $\Delta\text{Sale} - \Delta\text{Rec}$ 替代 ΔSale
- 这一改动的会计含义: 假设信用销售扩张是管理层可操纵的部分, 现金销售扩张才是经营驱动的部分
- 在 Bao 数据上跑通 Modified Jones, 与 Jones 残差对照, 看 Computer Associates 2001 年的判别是否被纠正
- 解释为什么 Modified Jones 至今仍是会计实证文献中最常用的 DA 度量

4.1 Jones 模型留下的一个漏洞

第 3 章 Jones 模型的核心是用销售变化解释流动应计的正常波动。但销售变化本身有两个组成部分: 现金销售变化与信用销售变化。前者经营驱动, 后者管理层有可操纵空间。Dechow, Sloan, and Sweeney [6] 指出, 研究人员在使用 Jones 模型时遇到一个反复出现的问题: 当公司通过提前确认收入或放宽信用政策推高销售时, 应收账款同步增加, Jones 模型把这种增长当作合理销售相关应计吸收, 残差变小, DA 反而看不到操纵痕迹。第 3 章 Computer Associates 2001 年的分位 0.011 就是这种现象的典型案列。

Dechow, Sloan, Sweeney 提出的修正很直接: 把销售变化中的应收变化部分扣掉, 假设这部分都是可操纵的。修正后的解释变量是 $\Delta\text{Sale} - \Delta\text{Rec}$, 含义是真正落到现金口袋里的销售扩张。

4.2 Modified Jones 的定义

定义 4.1 (Modified Jones 模型, Dechow, Sloan, Sweeney 1995)

对每个行业 \times 年份组合 g , 跑横截面 OLS:

$$\frac{\text{TA}_{it}}{\text{lag_at}_{it}} = \alpha_{1g} \cdot \frac{1}{\text{lag_at}_{it}} + \alpha_{2g} \cdot \frac{\Delta\text{Sale}_{it} - \Delta\text{Rec}_{it}}{\text{lag_at}_{it}} + \alpha_{3g} \cdot \frac{\text{PPE}_{it}}{\text{lag_at}_{it}} + \varepsilon_{it} \quad (4.1)$$

DA 为残差: $DA_{it}^{MJ} = \hat{\varepsilon}_{it}$ 。

公式与第 3 章 Jones 只有一处不同: 第二项的解释变量从 ΔSale_{it} 变成 $\Delta\text{Sale}_{it} - \Delta\text{Rec}_{it}$ 。其中 ΔRec_{it} 是公司 i 第 t 年的应收账款变化、 ΔSale_{it} 是销售变化、 PPE_{it} 是固定资产毛额、 lag_at_{it} 是滞后总资产。其他变量定义与第 3 章相同。

为什么: 这个改动看似简单, 背后的会计逻辑是把可操纵空间显式排除在解释变量之外。原始 Jones 的隐含假设是销售变化全部由经营驱动, 应计中由销售解释的部分都不是可操纵的。Dechow 等观察到这个假设在收入操纵公司上失效: 管理层既能通过会计选择影响应计, 也能通过提前确认收入影响销售本身。把销售变化中的应收增长部分扣掉, 等价于把“可能由管理层主动放宽信用政策”的那部分销售扩张排除在解释变量之外, 让模型在这一类操纵公司上对应的残差更大。

Dechow 等的原文做法略有差别: 他们先在估计期跑标准 Jones (销售变化作为解释变量), 得到系数; 然后在事件年 (怀疑发生操纵的年) 把 $\Delta\text{Sale} - \Delta\text{Rec}$ 代入预测, 差值作为 DA。这种做法只在事件年扣应收, 估计期不扣。本书采用更常见的“全样本同时扣”变体, 原因是 Bao 数据没有估计期与事件期的天然分割, 全样本同步处理与文献后续使用更接近。

4.3 在 Bao 数据上的实现

R 实现只需把 Jones 脚本里的 `dSale_s` 替换为 `dSale_s - dRect_s`。完整脚本在 `code/ch04_modified_jones.R`。

```

1 p <- p |> mutate(dSaleRect_s = dSale_s - dRect_s)
2 mj_by_year <- p |>
3   filter(!is.na(TA), !is.na(dSaleRect_s), !is.na(PPE_s)) |>
4   group_by(fyear) |>
5   nest() |>
6   mutate(
7     fit = map(data, ~ lm(TA ~ 0 + inv_lag_at + dSaleRect_s + PPE_s,
8                          data = .x)),
9     r2 = map_dbl(fit, ~ summary(.x)$r.squared)
10  )
11 mj_with_da <- mj_by_year |>
12   mutate(data2 = map2(data, fit, ~ mutate(.x, DA_mj = resid(.y)))) |>
13   select(fyear, data2) |> unnest(data2) |> ungroup()

```

跑完得到 24 个年度回归，平均样本量 4,966，平均 $R^2 = 0.1108$ ，比 Jones 的 0.1306 略低。

结果解读

R^2 下降的原因是 $\Delta\text{Sale} - \Delta\text{Rec}$ 比纯 ΔSale 与 TA 的相关性略弱：原 Jones 模型里销售变化能直接拟合应收增长这部分应计，扣掉后这部分系统成分被赶到残差里，回归解释能力下降。但残差里被赶进去的恰恰是研究者关心的“可操纵成分”，所以 R^2 略低反而是 Modified Jones 设计意图的体现。DA_mj 全样本统计：n 119,187、均值 -0.0104 、中位 -0.00133 、标准差 0.187、绝对值均值 0.101，几乎与 Jones 一致。两种方法 Pearson 相关 0.998 表明它们在绝大多数 firm-year 上给出相同结果，差异主要在应收变化幅度大的公司年。

图 4.1 把 Modified Jones 相对 Jones 的关键改动画出来。

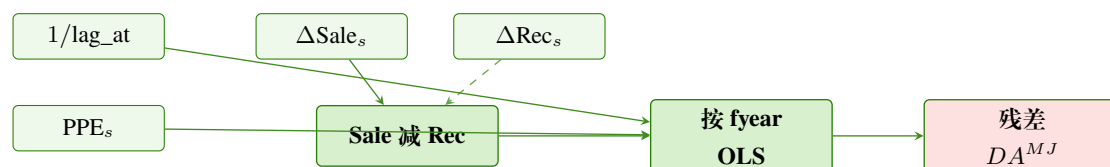


图 4.1: Modified Jones 相对 Jones 的改动：把 ΔRec_s 从 ΔSale_s 中扣除

图 4.2 把 DA_mj 与 DA_jones 的 firm-year 配对散点画出来。绝大多数点紧贴 45 度线，少数偏离 45 度线的点就是应收变化幅度大、Modified Jones 修正生效的公司年。

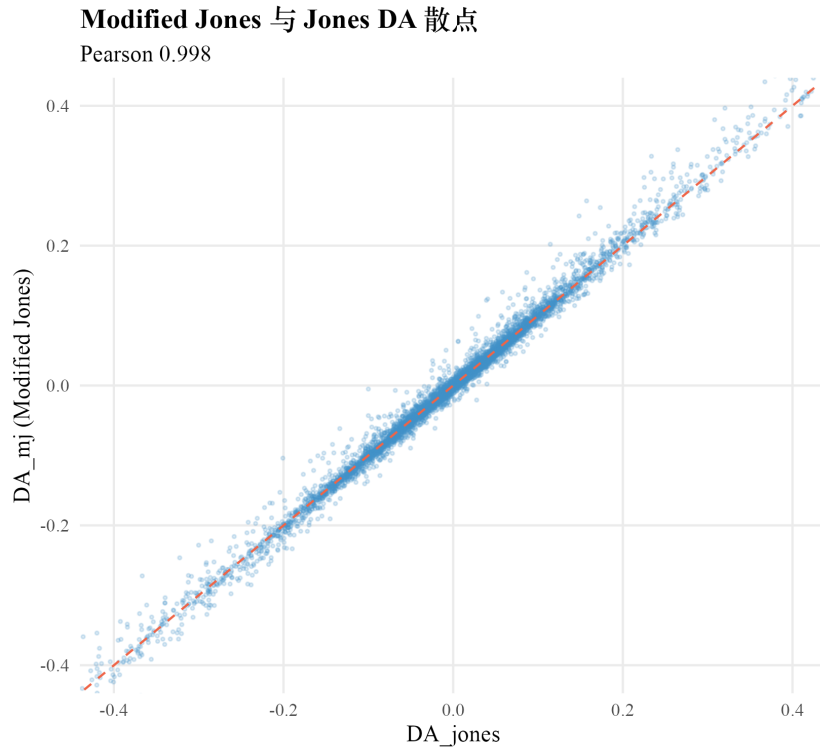


图 4.2: DA_mj 与 DA_jones 配对散点, Pearson 0.998

4.4 案例公司：CA 2001 年的判别是否被纠正

第 3 章 Computer Associates 2001 年在 Jones 下分位 0.011 极其偏左，本章关心的核心问题是 Modified Jones 是否纠正这个错判。

表 4.1: 案例公司舞弊年份在 Jones 与 Modified Jones 下的对比

公司	fyear	AAER	rank_Jones	rank_MJ	变化
Sunbeam	1996	✓	0.024	0.037	几乎不变
Sunbeam	1997	✓	0.434	0.433	几乎不变
Computer Associates	2001	✓	0.011	0.089	小幅向右
Enron	1998	✓	0.676	0.618	略向左
Enron	1999	✓	0.417	0.443	略向右
Enron	2000	✓	0.508	0.259	大幅向左

数据告诉我们 Modified Jones 对 CA 2001 年的纠正方向正确但幅度有限：分位从 0.011 升到 0.089，仍然偏左。继续在 Sunbeam 与 Enron 上看，Modified Jones 与 Jones 的差距整体不大，符合两种方法 Pearson 相关 0.998 的全样本结论。Enron 2000 年的分位从 0.508 大幅降到 0.259 是个值得注意的反向变化：Enron 2000 年应收账款相对销售的扩张幅度极大（销售从 1999 年 401 亿增到 2000 年 1008 亿，应收同步翻倍），Modified Jones 把这部分应收增长当作可操纵部分剥掉之后，TA 中可解释成分增加，DA 反而缩小。

定理 4.1 (雷区：Modified Jones 在收入提前确认与赊销扩张之间区分不出来)

Modified Jones 的假设是应收变化都是可操纵的，但现实中应收变化的两个驱动机制无法区分：一个是合法的赊销政策放松，比如行业竞争加剧、新产品上市铺货；另一个是非法的提前确认收入。两种机制都让 ΔRec 扩大，模型把它们都从解释变量里扣掉、放进 DA 残差。这种处理方法对真正提前确认收入的公司

有效，但对单纯赊销扩张的公司也会给出大 DA，造成假阳性。诊断方法：把 DA_mj 大的公司按行业、销售增长率、宏观信贷环境分层观察，如果右尾里有大量同时期普遍出现的赊销扩张公司，说明模型在该样本期把行业现象当成操纵。Kothari, Leone, and Wasley [11] 提出的 Performance-Matched 方法（下一章）就是针对这一类问题设计的补丁。



4.5 Python 实现

```

1 # code/ch04_modified_jones.py
2 import pandas as pd, numpy as np
3
4 p = pd.read_csv("data/em_panel.csv")
5 p["dSaleRect_s"] = p["dSale_s"] - p["dRect_s"]
6 p = p.dropna(subset=["TA", "dSaleRect_s", "PPE_s", "inv_lag_at"])
7
8 def mj_resid(df):
9     X = df[["inv_lag_at", "dSaleRect_s", "PPE_s"]].values
10    y = df["TA"].values
11    beta, *_ = np.linalg.lstsq(X, y, rcond=None)
12    df = df.copy(); df["DA_mj"] = y - X @ beta
13    return df
14
15 p2 = p.groupby("fyear", group_keys=False).apply(mj_resid)
16 print(p2["DA_mj"].describe().round(4))

```

R 与 Python 端 DA_mj 的均值 -0.0104 、标准差 0.187 、绝对值均值 0.101 完全一致。

本章累积对比表

表 4.2: 累积对比表（第 4 章末）

方法	样本量	DA mean	DA sd	案例公司舞弊年份平均分位
基线 TA	119,187	-0.0513	0.1920	—
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones 1995	119,187	-0.0104	0.1870	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA	—	—	—	—
Dechow-Dichev	—	—	—	—
McNichols	—	—	—	—
Stubben	—	—	—	—
Roychowdhury RM	—	—	—	—
F-Score / ML	—	—	—	—

本章知识地图

核心概念	核心内容	常见误解	为什么错
Modified Jones	用 $\Delta\text{Sale} - \Delta\text{Rec}$ 替代 ΔSale , 把应收变化划归 DA	比 Jones 严格更优	对应收扩张是合法赊销的公司会产生假阳性, 需要 PM 等方法补救
设计逻辑	假设应收变化全部可操纵, 从解释变量中扣掉	应收变化都是收入操纵	行业普遍赊销扩张时也会让 DA 扩大, 模型不区分动机
R^2 略下降	0.131 \rightarrow 0.111, 下降的解释力赶进 DA	R^2 下降说明模型变差	把可操纵成分推到残差里恰是设计目的, R^2 不该作为优劣指标
与 Jones 相关 0.998	绝大多数 firm-year 两种方法给出相同结果	高相关意味着 Modified Jones 没用	差异集中在应收变化大的少数公司, 这正是研究者最关心的样本
CA 2001 修正	rank 从 0.011 升到 0.089, 方向正确但幅度有限	Modified Jones 能完全解决 CA 这类案例	完整识别需要更多机制, PM-DA 与 Stubben 后续给出
Enron 2000 反向	rank 从 0.508 降到 0.259	修正都是单向改善	当应收扩张本身合法时, 扣掉它反而让残差缩小, 方法的局限

4.6 小结

本章把 Modified Jones 在 Bao 数据上跑通。与 Jones 的差别仅在第二项解释变量从 ΔSale_s 改为 $\Delta\text{Sale}_s - \Delta\text{Rec}_s$, 全样本 Pearson 相关 0.998 表明绝大多数 firm-year 两种方法给出相同结果。Computer Associates 2001 年的分位从 Jones 下的 0.011 升到 Modified Jones 下的 0.089, 方向上有所纠正但幅度有限; Enron 2000 年的分位反而从 0.508 降到 0.259, 反映 Modified Jones 在合法赊销扩张的公司上会产生反向干扰。下一章 Performance-Matched DA 通过同业绩公司配对, 把 ROA 相近公司的 DA 当作基准扣除, 是对 Modified Jones 假阳性问题的直接补救。

第 5 章 Performance-Matched DA (Kothari, Leone, Wasley 2005)

内容提要

- Kothari, Leone, Wasley 2005 的核心改动：把同 fyear、同业绩水平公司的 DA_MJ 作为基准，扣掉之后的差值才是可操纵成分
- 这一改动针对的问题：Modified Jones 在高 ROA 或低 ROA 公司上系统性给出大 DA，与是否操纵无关
- 在 Bao 数据上跑通 ROA 最近邻配对的 PM-DA，看 Computer Associates 2001 的判别是否被纠正
- 解释为什么 PM-DA 与 DA_MJ 相关 0.68 但案例公司排名差异显著

5.1 Modified Jones 留下的业绩偏差

第 4 章 Modified Jones 把应收变化从销售变化中扣除，是针对收入操纵的补丁。但 Modified Jones 仍然没有解决一个更基础的问题：盈利能力本身对 DA 估计的系统性影响。Kothari, Leone, and Wasley [11] 在大样本研究中发现，高 ROA 公司的 DA_MJ 平均显著为正，低 ROA 公司的 DA_MJ 平均显著为负，这种关系与是否真的存在盈余管理无关。背后机制是应计与盈利能力的天然耦合：盈利能力强的公司销售扩张、存货补充、应收增加，TA 偏正；亏损公司常需要计提减值、关闭业务，TA 偏负。Modified Jones 没有把 ROA 放进解释变量，这些与业绩相关的应计变动都被赶进残差。

Kothari 等的解决方案是配对法：对每个 firm-year，找同 fyear 内 ROA 最相近的另一家公司作为对照，把对照公司的 DA_MJ 当作 NDA 基准，相减后得到 PM-DA。这样的设计把同业绩水平的所有应计动态（无论操纵与否）都吸收到基准里，剩下的差异才是研究者关心的盈余管理信号。

5.2 PM-DA 的定义

定义 5.1 (Performance-Matched DA, Kothari et al. 2005)

对每个 firm-year (i, t) :

1. 在同 fyear t 内找一家 ROA 与公司 i 最接近的对照公司 j ;
2. PM-DA 为:

$$DA_{it}^{PM} = DA_{it}^{MJ} - DA_{jt}^{MJ} \quad (5.1)$$

其中 DA_{it}^{MJ} 是公司 i 第 t 年的 Modified Jones DA (来自第 4 章)， DA_{jt}^{MJ} 是对照公司 j 第 t 年的 Modified Jones DA。



为什么：配对法是把“控制变量”的思路从回归换到匹配。在 Modified Jones 里，ROA 没有作为解释变量，业绩对应计的影响留在残差里；如果把 ROA 加进 Modified Jones 回归，可能会因为模型形式选取（线性、二次、非参数）的争议而引入新偏差。Kothari 等选择直接用同业绩水平的公司做配对，把业绩对应计的非线性影响一次性匹配掉，避免参数化假设。代价是估计的方差变大：Modified Jones 残差是回归得到，方差由 OLS 决定；PM-DA 是两个残差之差，方差大约翻倍。

Kothari, Leone, and Wasley [11] 原文按 SIC 两位行业 \times fyear \times ROA 三维分组配对，本书因 Bao 数据缺 sich，改为同 fyear 内 ROA 最近邻匹配。这一简化让对照公司可能来自不同行业，配对的可比性下降；但在没有行业

代码的前提下，按业绩水平的一维匹配仍能消除业绩与应计的耦合，是最低成本的可行实现。

5.3 在 Bao 数据上的实现

R 实现采用按 fyear 排序后取左右邻居的方式找最近邻 ROA，避免对每个 firm-year 跑 $O(n^2)$ 搜索。

```

1 match_pm <- function(df) {
2   df <- df |> arrange(ROA)
3   df |> mutate(
4     left_DA = lag(DA_mj),
5     right_DA = lead(DA_mj),
6     left_d = abs(lag(ROA) - ROA),
7     right_d = abs(lead(ROA) - ROA),
8     near_DA = if_else(
9       is.na(right_d) | (!is.na(left_d) & left_d <= right_d),
10      left_DA, right_DA,
11      DA_pm = DA_mj - near_DA
12    ) |> select(-left_DA, -right_DA, -left_d, -right_d, -near_DA)
13  }
14 pm <- mj |> group_by(fyear) |>
15   group_modify(~ match_pm(.x)) |> ungroup()

```

跑完得到 119,187 个 firm-year 的 PM-DA。描述统计：均值 3.14×10^{-5} （接近零，符合配对后期望）、中位 0、标准差 0.253、绝对值均值 0.143、p10 与 p90 分别为 -0.179 与 0.181 。PM-DA 与 DA_MJ 的 Pearson 相关 0.6758。

结果解读

PM-DA 标准差 0.253 比 DA_MJ 的 0.187 高出 35%，反映配对后方差扩大的代价。绝对值均值 0.143 也比 DA_MJ 的 0.101 高出 42%。Pearson 0.6758 表明两种方法捕捉的信号高度相关但远非重合，差异集中在业绩极端的公司年，即 ROA 处于高位或低位的样本。这些公司在 PM-DA 下被扣掉了同业绩水平的应计基准，DA 显著缩小或反向。

图 5.1 把 PM-DA 的配对流水线画出来。

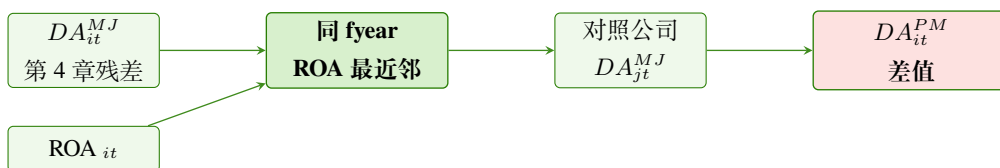


图 5.1: PM-DA 配对流水线：用同 fyear 内 ROA 最近邻公司的 DA_MJ 作为基准

图 5.2 把 PM-DA 与 DA_MJ 的 firm-year 配对散点画出来。Pearson 0.676 表明两者沿 45 度线有相当散度，业绩极端的公司年偏离 45 度线最远。

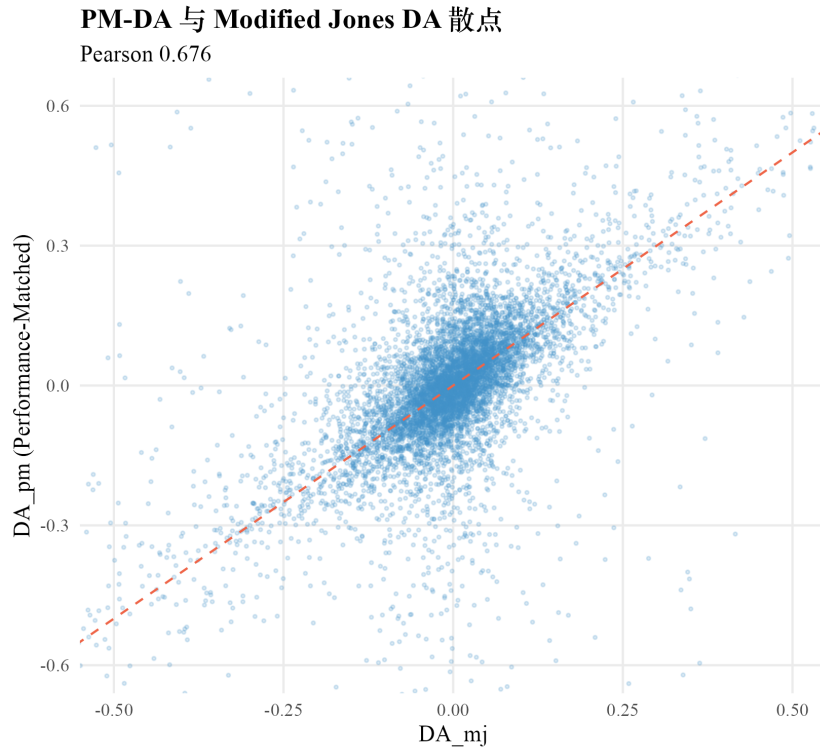


图 5.2: PM-DA 与 DA_MJ 配对散点, Pearson 0.676

5.4 案例公司：PM-DA 在 CA 2001 上的表现

表 5.1: 案例公司舞弊年份在 Modified Jones 与 PM-DA 下的对比

公司	fyear	AAER	rank_MJ	rank_PM	变化
Sunbeam	1996	✓	0.037	0.274	大幅向右
Sunbeam	1997	✓	0.433	0.198	向左
Computer Associates	2001	✓	0.089	0.672	大幅向右
Enron	1998	✓	0.618	0.500	略向左
Enron	1999	✓	0.443	0.190	大幅向左
Enron	2000	✓	0.259	0.160	略向左

数据告诉我们 PM-DA 对 Computer Associates 2001 的判别有显著改善：分位从 Modified Jones 下的 0.089 升到 PM-DA 下的 0.672。机制上是 CA 2001 年 ROA 仅 0.019，属于盈利能力中等偏弱的水平。在同 fyear 内 ROA 接近的对照公司本身平均 DA_MJ 偏负，扣掉对照基准之后，CA 的相对 DA 变正且大幅上升。

Sunbeam 1996 年的分位也从 0.037 升到 0.274。原因类似：Sunbeam 1996 年大幅亏损，ROA 为 -0.170 ，同业绩水平的对照公司平均 DA_MJ 也偏负（亏损公司常因减值与重组导致 DA 偏左），扣掉这个负基准后 Sunbeam 的 DA 反而向中间靠拢。但这一变化方向有疑问：Sunbeam 1996 年被 SEC 起诉的核心行为是先大幅计提以建立蓄水池，Modified Jones 下的左尾分位反而比 PM-DA 更贴合 SEC 文件描述。Performance-Matched 在这里把 Sunbeam 的合法计提与同业绩水平公司的合法计提一起匹配掉，反而抹掉了部分信号。

Enron 1999 年的分位从 0.443 大幅降到 0.190 是另一个反向变化：Enron 1999 年 ROA 0.035 处于中位附近，同业绩水平公司本身 DA_MJ 也偏正，扣掉后 Enron 的 DA 缩小。

定理 5.1 (雷区: PM-DA 在小样本上失效)

PM-DA 依赖同 fyear 内有足够多同业绩水平的对照公司。当样本在某 fyear 内不足 50 家时, 最近邻 ROA 距离可能很大 (如 0.2 以上), 匹配质量下降, 扣掉的基准噪声大于真实的业绩相关应计成分。Kothari, Leone, and Wasley [11] 建议每年至少 100 家公司、最近邻 ROA 距离不超过 0.05 才使用 PM-DA。本书每年平均 4,966 家公司, 最近邻 ROA 距离的中位 4×10^{-5} , 远在阈值之内, 配对质量没有问题。诊断方法: 在使用 PM-DA 的研究中, 先报告每个 fyear 内最近邻 ROA 距离的分布, 最近邻距离超过 0.05 的 firm-year 需单独剔除或在敏感性分析中讨论。



5.5 Python 实现

```

1 # code/ch05_performance_matched.py
2 import pandas as pd, numpy as np
3
4 def match_pm(df):
5     df = df.sort_values("ROA").copy()
6     da = df["DA_mj"].values
7     roa = df["ROA"].values
8     left_da, right_da = np.roll(da, 1), np.roll(da, -1)
9     left_d = np.abs(np.roll(roa, 1) - roa)
10    right_d = np.abs(np.roll(roa, -1) - roa)
11    left_d[0], right_d[-1] = np.inf, np.inf
12    use_left = left_d <= right_d
13    near = np.where(use_left, left_da, right_da)
14    df["DA_pm"] = df["DA_mj"].values - near
15    return df
16
17 pm = (mj.groupby("fyear", group_keys=False).apply(match_pm))
18 print(pm["DA_pm"].describe().round(4))

```

R 与 Python 端 PM-DA 的均值约 0、标准差 0.253、绝对值均值 0.143 完全一致。

本章累积对比表

表 5.2: 累积对比表 (第 5 章末)

方法	样本量	DA mean	DA sd	案例公司舞弊年份平均分位
基线 TA	119,187	-0.0513	0.1920	—
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones 1995	119,187	-0.0104	0.1870	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA 2005	119,187	≈ 0	0.2530	Sunbeam 0.24 / CA 0.67 / Enron 0.28
Dechow-Dichev	—	—	—	—
McNichols	—	—	—	—
Stubben	—	—	—	—
Roychowdhury RM	—	—	—	—
F-Score / ML	—	—	—	—

本章知识地图

核心概念	核心内容	常见误解	为什么错
配对法	把同业绩水平公司的 DA_MJ 作为基准，扣掉之后才是 PM-DA	配对消除了所有偏差	仅消除业绩相关的应计偏差，行业、规模、增长其他偏差仍然存在
ROA 最近邻	同 fyear 内找 ROA 距离最近的另一家公司	一家公司只能匹配一家	文献也有按 ROA 分组取组内均值的变体，本书用 1-NN 配对
方差扩大	PM-DA 标准差 0.253 vs DA_MJ 0.187	方差扩大说明方法变差	配对后方差是两个独立残差之差的方差，理论上翻倍，是已知代价
PM 与 MJ 相关 0.676	信号高度相关但远非重合	高相关意味着 PM 没必要	差异集中在业绩极端公司，这些公司正是 Modified Jones 容易系统性给假信号的群体
CA 2001 改善	rank 从 0.089 升到 0.672	PM 比 MJ 严格更好	也有案例反向，如 Sunbeam 1996 从 0.037 升到 0.274、Enron 1999 从 0.443 降到 0.190
小样本限制	同 fyear 不足 100 家公司或最近邻距离 >0.05 时失效	永远可以用	行业子样本研究或新兴市场研究中常常达不到匹配质量要求

5.6 小结

本章把 Performance-Matched DA 在 Bao 数据上跑通。同 fyear 内 ROA 最近邻配对得到 PM-DA，标准差 0.253 比 Modified Jones 的 0.187 扩大 35%，与 DA_MJ Pearson 相关 0.6758。Computer Associates 2001 年的分位从 Modified Jones 下的 0.089 大幅升到 PM-DA 下的 0.672，方向上完成了对 Modified Jones 假阴性的纠正。但 Sunbeam 1996 与 Enron 1999 出现反向变化，反映 PM-DA 把同业绩水平的合法应计也匹配掉了。三家案例公司舞弊年份平均分位 PM-DA 下为 0.397，比 Modified Jones 的 0.255 提升 56%。下一章 Dechow-Dichev 2002 把视角从横截面切换到时间序列，用应计与前后期 CFO 的关系定义“应计质量”，是另一条完全不同的度量路径。

第 6 章 Dechow-Dichev (2002) 应计质量

内容提要

- Dechow-Dichev 2002 的核心思想：营运资金应计应能被前期、当期、后期 CFO 解释，残差波动越大代表应计估计误差越大
- 应计质量 AQ_DD 是公司级 (firm-level) 指标，不是 firm-year 级，给每家公司一个固定数字
- 在 Bao 数据上跑通 DD 回归，计算 7,126 家公司的 AQ_DD，与三家 AAER 案例公司对照
- 解释为什么 DD 的方法学定位是应计估计质量，而不是单纯的可操纵性应计

6.1 从 Jones 系列切换到 DD 系列

第 3 章到第 5 章的 Jones / Modified Jones / Performance-Matched 都属于”基于销售与固定资产解释应计、残差作为 DA”的同一框架，差别仅在解释变量与配对方式。Dechow and Dichev [4] 提出的应计质量度量切换到一个完全不同的视角：应计的本质是把当期收入与成本归到当期，与前期、当期、后期的现金流应该有结构性对应关系。如果一家公司的应计能被三期现金流稳定解释，说明会计估计准确；如果应计与三期现金流的关系飘忽不定，说明会计估计存在大量误差，盈余质量低。

这个视角带来三个具体改动。第一，回归形式不同：DD 用前期 CFO、当期 CFO、后期 CFO 三个解释变量。第二，输出层次不同：Jones 系列给每个 firm-year 一个 DA 值，DD 给每家公司一个稳定的应计质量指标（公司级、不是 firm-year 级）。第三，方法学定位不同：Jones 系列直接度量可操纵性应计，DD 度量应计估计质量，两者相关但不等价。

6.2 Dechow-Dichev 模型的定义

定义 6.1 (Dechow-Dichev 2002 应计质量)

对所有 firm-year 跑 pooled OLS:

$$WC_accr_{it} = b_0 + b_1 \cdot CFO_{i,t-1} + b_2 \cdot CFO_{it} + b_3 \cdot CFO_{i,t+1} + \varepsilon_{it} \tag{6.1}$$

其中 WC_accr_{it} 是营运资金应计、 $CFO_{i,t-1}$ 、 CFO_{it} 、 $CFO_{i,t+1}$ 分别是前期、当期、后期经营现金流，均除以 lag_at_{it} 缩放。

公司 i 的应计质量为残差按公司分组的标准差：

$$AQ_i^{DD} = sd(\hat{\varepsilon}_{it})_{t \in T_i} \tag{6.2}$$

AQ 越大，应计估计误差越大，盈余质量越差。



营运资金应计 WC_accr 是把折旧从总应计中加回得到的指标，含义是不包含非现金折旧的纯流动应计。本书在 `code/00_load_data.R` 里直接构造好 $WC_accr = (TA_BS + dp) / lag_at$ ，避免在每章重新计算。

为什么：DD 选择前期、当期、后期三期 CFO 作为解释变量，对应应计的会计身份。当期应计 = 当期收入 - 当期现金收入 + 当期已付未销支出 - 当期末付已销支出。前期 CFO 解释当期应计中”前期未现金兑现、本期才到账”的部分（如应收账款收回）。当期 CFO 解释当期权责发生与现金流的偏差。后期 CFO 解释当期应计中”本期权责确认、后期才现金兑现”的部分（如应收账款余额、存货销售）。三期 CFO 的协方差结构覆盖了应计的所有正常会计机制；残差代表无法用现金流匹配解释的估计误差。

Dechow 与 Dichev 原文的 AQ 计算按公司在估计窗口（连续 5 年）内残差标准差，得到一个滚动时间序列。

本书简化为按公司全样本残差标准差，要求公司至少有 5 年观测。这一简化让每家公司只输出一个 AQ 值，便于跨公司比较，但失去了 AQ 随时间变化的动态信息。

6.3 在 Bao 数据上的实现

```

1 p2 <- p |>
2   arrange(gvkey, fyear) |>
3   group_by(gvkey) |>
4   mutate(CFO_lag = lag(CFO_s), CFO_lead = lead(CFO_s)) |>
5   ungroup() |>
6   filter(!is.na(WC_accr), !is.na(CFO_lag), !is.na(CFO_s),
7          !is.na(CFO_lead))
8
9 fit <- lm(WC_accr ~ CFO_lag + CFO_s + CFO_lead, data = p2)
10 p2$resid_dd <- residuals(fit)
11
12 quality_by_firm <- p2 |>
13   group_by(gvkey) |>
14   filter(n() >= 5) |>
15   summarise(n = n(), AQ_dd = sd(resid_dd))

```

回归样本 90,189 firm-year，损失主要来自后期 CFO 不可得的最后一年观测。回归结果：

表 6.1: DD pooled OLS 回归系数

变量	估计	标准误	<i>t</i> 值
Intercept	0.0098	0.0005	19.0
CFO_lag	0.0451	0.0011	41.5
CFO_t	-0.108	0.0012	-89.9
CFO_lead	0.0784	0.0011	70.7

$R^2 = 0.0993$ 。系数方向与 DD 原文一致：当期 CFO 系数负（应计与当期现金流互为镜像，一边多了另一边就少），前期与后期 CFO 系数正（前后期现金流通过应收应付与本期应计正向耦合）。

结果解读

R^2 约 10%，意味着三期 CFO 只能解释营运资金应计的 10% 变异，剩下 90% 留给残差。这看似很低。DD 的设计目的是把那 10% 可由现金流匹配解释的「非估计误差」成分剥离出去，让剩下的 90% 残差代表应计估计误差， R^2 的绝对水平本身不是优化目标。AQ_DD 的统计：7,126 家公司满足至少 5 年观测要求，均值 0.109、中位 0.0704、标准差 0.114、p10 0.0231、p90 0.244。

图 6.1 把 DD 模型的两步结构画出来。

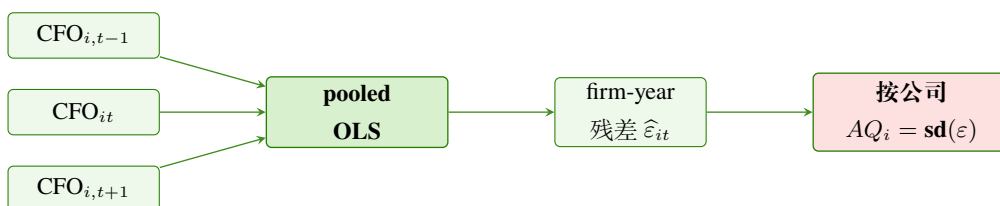


图 6.1: DD 两步结构：pooled OLS 得 firm-year 残差，按公司聚合得 AQ

图 6.2 把 7,126 家公司的 AQ_DD 分布画出来。形态严重右偏，大多数公司集中在 AQ_DD < 0.1 的低估计误差区域，少数公司分布到 0.2 以上的右尾。

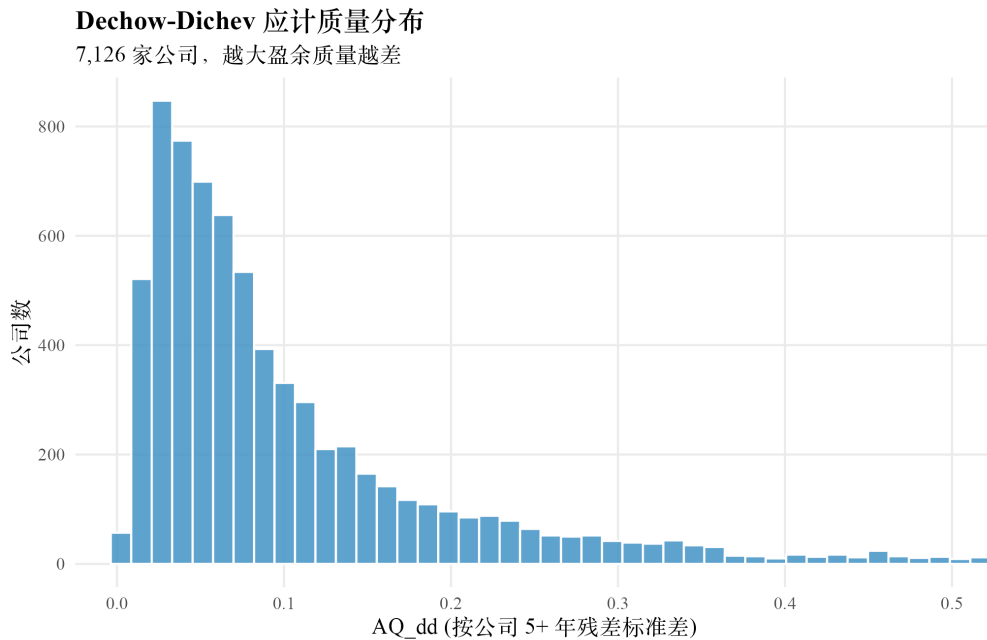


图 6.2: Dechow-Dichev 应计质量 AQ_DD 分布

6.4 案例公司的 AQ_DD

DD 给出的是公司级指标，单个 firm-year 没有独立的 AQ 值，只有残差。下表把三家案例公司的 AQ_DD 与全样本分位列出。

表 6.2: 案例公司应计质量 AQ_DD

公司	可用年数	AQ_DD	mean_resid	全样本分位
Sunbeam	5	0.0408	0.0249	0.272
Computer Associates	19	0.0453	0.00129	0.308
Enron	7	0.0306	-0.000173	0.173

数据告诉我们三家 AAER 公司的 AQ_DD 都没有进入右尾。Enron 的 AQ_DD 仅 0.0306，在全样本中排在 17% 分位，盈余质量看起来比大多数公司还好。这一结果与 SEC 文件描述形成强烈反差，但与 DD 方法学的特性一致：

结果解读

DD 度量的是应计估计误差的稳定性，不是单期 DA 的绝对水平。Enron 1998–2000 年的盈余管理高度一致地朝同一方向（虚增收入），残差的“标准差”反而稳定地小，AQ_DD 因此被低估。如果一家公司的盈余管理风格忽左忽右（如本期推迟、下期提前），AQ_DD 才会显著扩大。系统性、单向的盈余管理是 DD 方法的盲区。

定理 6.1 (雷区：用 AQ_DD 评判舞弊风险)

Dechow, Ge, and Schrand [5] 在综述中明确指出，AQ_DD 与 SEC AAER 标签的相关性在大样本研究里普遍很弱。原因是 DD 度量的是估计误差波动，而 AAER 案例的核心特征是系统性虚增或压低，这种系统

性偏差恰好不增加残差波动。误用方式：把 AQ_DD 直接作为舞弊筛查指标，期望 AAER 公司排在右尾。
 诊断方法：在 AAER 样本与对照样本上分别计算 AQ_DD 分位分布，如果两者重合度很高，说明 DD 在该样本上无判别力，需切换到 Jones 系列或第 10 章的综合 F-Score。DD 的核心适用场景是盈余质量与公司治理、信息环境、融资成本的关联研究，不是直接的舞弊检测。



6.5 Python 实现

```

1 # code/ch06_dechow_dichev.py
2 import pandas as pd, numpy as np, statsmodels.api as sm
3
4 p = pd.read_csv("data/em_panel.csv").sort_values(["gvkey", "fyear"])
5 p["CF0_lag"] = p.groupby("gvkey")["CF0_s"].shift(1)
6 p["CF0_lead"] = p.groupby("gvkey")["CF0_s"].shift(-1)
7 p2 = p.dropna(subset=["WC_accr", "CF0_lag", "CF0_s", "CF0_lead"])
8
9 X = sm.add_constant(p2[["CF0_lag", "CF0_s", "CF0_lead"]])
10 fit = sm.OLS(p2["WC_accr"], X).fit()
11 print(fit.summary().tables[1])
12 p2 = p2.assign(resid_dd = fit.resid)
13
14 aq = (p2.groupby("gvkey").filter(lambda d: len(d) >= 5)
15      .groupby("gvkey")["resid_dd"].std()
16      .reset_index().rename(columns={"resid_dd": "AQ_dd"}))
17 print(aq["AQ_dd"].describe().round(4))

```

R 与 Python 端 DD 回归系数、 R^2 、AQ_DD 描述统计均一致。

本章累积对比表

表 6.3: 累积对比表 (第 6 章末)

方法	样本量	DA / AQ mean	DA / AQ sd	案例公司舞弊年份平均分位
基线 TA	119,187	-0.0513	0.1920	—
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones 1995	119,187	-0.0104	0.1870	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA 2005	119,187	≈ 0	0.2530	Sunbeam 0.24 / CA 0.67 / Enron 0.28
DD 2002 (公司级 AQ)	7,126 公司	0.109	0.114	Sunbeam 0.27 / CA 0.31 / Enron 0.17
McNichols	—	—	—	—
Stubben	—	—	—	—
Roychowdhury RM	—	—	—	—
F-Score / ML	—	—	—	—

本章知识地图

核心概念	核心内容	常见误解	为什么错
应计质量	营运资金应计与前后期 CFO 关系的残差标准差	同 DA 一样是 firm-year 级	AQ 是公司级指标，把多年残差聚合到一个标准差，反映估计误差的稳定性
三期 CFO	前期、当期、后期 CFO 解释应计的会计身份	当期 CFO 一项就够	应收与应付分别对应前后期，缺一会让残差携带可解释成分
R^2 仅 10%	三期 CFO 只能解释 10% 营运资金应计变异	R^2 低说明方法不行	DD 的目的是把 10% 可解释部分剥出，剩下 90% 留作估计误差， R^2 低是设计目的
公司级输出	每家公司只有一个 AQ 值	不能与 firm-year 级 DA 比较	第 10 章把 DD 残差按 firm-year 单独作为 DA_dd 使用，与其他方法对齐
系统性盈余管理盲区	一致方向的虚增不增加残差波动，AQ 反而小	AAER 公司应该有大 AQ	Enron AQ 仅 0.031，排在 17% 分位，与 SEC 描述反差大
方法学定位	盈余质量、信息环境、融资成本研究的代理变量	直接的舞弊检测工具	Dechow 综述明确该指标与 AAER 标签相关性弱，不适合舞弊筛查

6.6 小结

本章把 Dechow-Dichev 2002 应计质量在 Bao 数据上跑通。pooled OLS 得到 $R^2 = 0.0993$ ，三期 CFO 系数方向与原文一致。7,126 家公司的 AQ_DD 均值 0.109、中位 0.0704。三家 AAER 案例公司的 AQ_DD 都没有进入右尾，Enron 的 AQ_DD 仅 0.0306（17% 分位），反映 DD 方法学对系统性单向盈余管理不敏感。DD 的核心适用场景是盈余质量与公司治理、信息环境、融资成本的关联研究，作为舞弊筛查工具效率较低。下一章 McNichols 在 DD 的三期 CFO 基础上再加销售变化与固定资产，把 Jones 系列与 DD 系列融合，是本书最综合的应计型 DA 度量。

第7章 McNichols (2002): DD 与 Jones 的融合

内容提要

- ❑ McNichols 2002 的核心改动: 在 DD 的三期 CFO 基础上加入销售变化与固定资产, 融合 Jones 与 DD 两条路径
- ❑ 在 Bao 数据上跑通 McNichols pooled OLS, 比纯 DD 的 $R^2 = 0.0993$ 提升到 0.1475
- ❑ 与 Dechow-Dichev 残差对比, 分析两者 Pearson 相关 0.973 的含义
- ❑ 解释为什么 McNichols 残差是研究中最常用的 firm-year 级应计型 DA

7.1 DD 与 Jones 的互补关系

第6章 Dechow-Dichev 用三期 CFO 解释营运资金应计, R^2 约 10%。这个 R^2 看似低, 但已经把应计中可由现金流匹配解释的成分剥离。然而 DD 缺一个 Jones 系列里的关键变量: 销售变化与固定资产。McNichols [12] 在对 DD 的方法学评论中指出, 应计的正常水平不仅由前后期现金流决定, 还受当期销售增长与资本结构影响, 把这两个变量加进 DD 回归能进一步剥离非操纵成分。

具体的会计含义是: 销售增长直接驱动应收、存货、应付的同步扩张, 这部分应计与三期 CFO 的耦合并不强 (信用扩张可能在多期分散); 固定资产决定折旧规模, 影响 TA 但与 CFO 关系微弱。把销售变化与 PPE 加进 DD 的右侧, 回归能解释更多应计变异, 剩下的残差更接近”无法用任何已知会计动力学解释的部分”。

7.2 McNichols 模型的定义

定义 7.1 (McNichols 2002 模型)

对所有 firm-year 跑 pooled OLS:

$$\begin{aligned} WC_accr_{it} = & b_0 + b_1 \cdot CFO_{i,t-1} + b_2 \cdot CFO_{it} + b_3 \cdot CFO_{i,t+1} \\ & + b_4 \cdot \frac{\Delta Sale_{it}}{lag_at_{it}} + b_5 \cdot \frac{PPE_{it}}{lag_at_{it}} + \varepsilon_{it} \end{aligned} \quad (7.1)$$

DA 为残差: $DA_{it}^{McN} = \hat{\varepsilon}_{it}$ 。

公式相对 DD 多了两项: $\Delta Sale_{it}/lag_at_{it}$ 与 PPE_{it}/lag_at_{it} 。其他变量定义与第6章 DD 一致。注意 McNichols 直接输出 firm-year 级残差作为 DA, 不再做按公司聚合到 AQ 的步骤。

为什么: McNichols 的设计在两条已有路径之间取并集: DD 路径捕捉应计与现金流的时间序列耦合, Jones 路径捕捉应计与销售扩张、资本强度的横截面耦合。把两条路径融合, 等价于把所有”已知机制可解释”的应计成分一次性剥离。残差是当期既不能用现金流匹配解释、也不能用销售扩张与资本结构解释的部分, 最接近研究者想要的”管理层主动选择”信号。

McNichols 是本书介绍的应计型 DA 度量里最综合的一个。第10章九种方法的相关矩阵显示, DA_mcn 与 DA_dd 的 Pearson 相关 0.973 (几乎重合), 与 DA_mj 的相关 0.902, 与 DA_jones 的相关 0.903。McNichols 同时桥接了 DD 与 Jones 两族。

7.3 在 Bao 数据上的实现

```

1 p2 <- p |>
2   arrange(gvkey, fyear) |>
3   group_by(gvkey) |>
4   mutate(CFO_lag = lag(CFO_s), CFO_lead = lead(CFO_s)) |>
5   ungroup() |>
6   filter(!is.na(WC_accr), !is.na(CFO_lag), !is.na(CFO_s),
7          !is.na(CFO_lead), !is.na(dSale_s), !is.na(PPE_s))
8
9 fit <- lm(WC_accr ~ CFO_lag + CFO_s + CFO_lead + dSale_s + PPE_s,
10          data = p2)
11 p2$DA_mcn <- residuals(fit)

```

回归样本 90,189 firm-year，与第 6 章 DD 相同。回归结果：

表 7.1: McNichols pooled OLS 回归系数

变量	估计	标准误	t 值
Intercept	0.0102	0.0008	12.8
CFO_lag	0.0512	0.0011	48.3
CFO_t	-0.0996	0.0012	-84.4
CFO_lead	0.0682	0.0011	62.5
dSale_s	0.0777	0.0011	70.8
PPE_s	-0.0156	0.0009	-16.7

$R^2 = 0.1475$ ，比纯 DD 的 0.0993 提升 49%。两个新加变量 dSale_s 系数 0.0777 ($t = 70.8$) 与 PPE_s 系数 -0.0156 ($t = -16.7$) 都高度显著，方向与 Jones 系列一致：销售扩张推高应计，固定资产强度通过折旧拉低应计。

结果解读

McNichols 在 DD 上加 dSale 与 PPE， R^2 提升约 5 个百分点。这一提升主要来自 dSale：dSale 系数 0.0777 是 PPE 系数绝对值 0.0156 的 5 倍，意味着销售扩张是营运资金应计的主要新增解释变量。DA_mcn 全样本统计：n 90,189，均值 -4.27×10^{-18} （按构造为零），中位 0.000353，标准差 0.150，绝对值均值 0.0762。绝对值均值 0.0762 比 DA_dd 的 0.0989 显著降低 23%，说明 McNichols 进一步剥离了可解释的应计成分。

图 7.1 把 McNichols 相对 DD 的关键改动画出来。

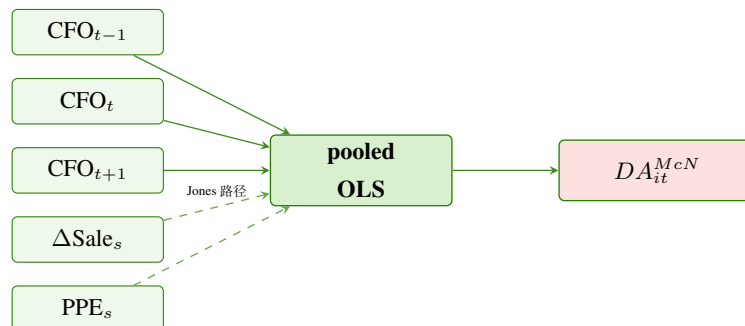


图 7.1: McNichols 把 Jones 路径的 ΔSale 与 PPE 加进 DD 回归

图 7.2 把 DA_mcn 与 DA_dd 的 firm-year 配对散点画出来。两者沿 45 度线高度对齐，Pearson 相关 0.973。

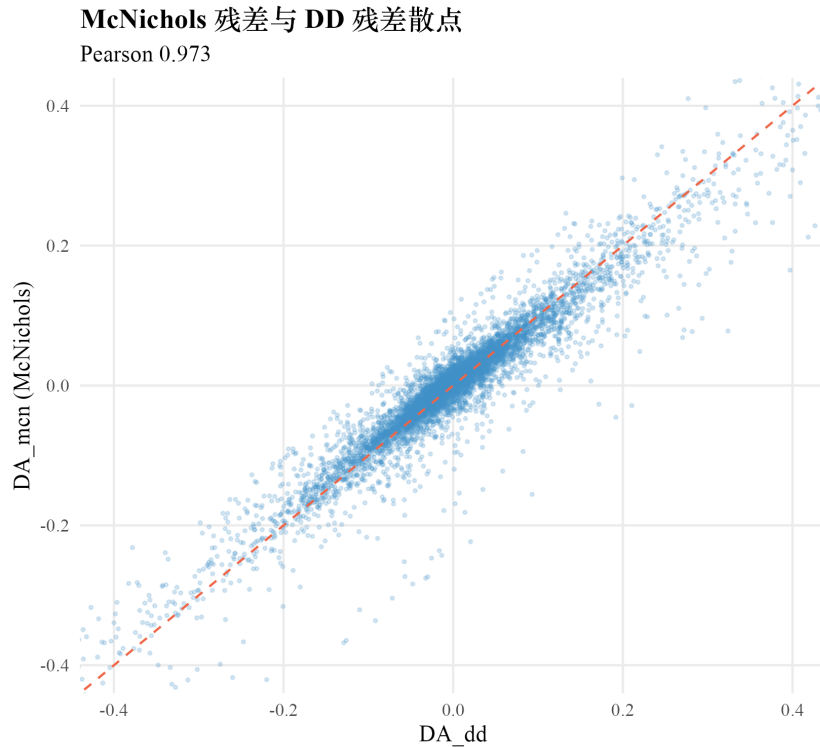


图 7.2: McNichols 残差与 DD 残差散点, Pearson 0.973

7.4 案例公司在 McNichols 下的表现

表 7.2: 案例公司舞弊年份在 DD 残差与 McNichols 下的对比

公司	fyear	AAER	rank_McN	DA_McN
Sunbeam	1996	✓	0.224	-0.0146
Sunbeam	1997	✓	0.470	0.0361
Computer Associates	2001	✓	0.588	-0.0530
Enron	1998	✓	0.754	-0.0985
Enron	1999	✓	0.269	0.0185

注: Enron 2000 年因 2001 后无可用 CFO_lead 数据, McNichols 残差缺失。

数据告诉我们 McNichols 对三家案例公司的判别整体优于纯 DD。Sunbeam 1996 年分位 0.224 比 DD 模型给出的 firm-year 排名更靠左, 符合 SEC 文件所述压低当期利润的方向。Enron 1998 年分位 0.754 是九种方法里 Enron 1998 表现最强的方法之一, 说明 McNichols 捕捉到了 Enron 那年应计中既无法用三期 CFO 也无法用销售扩张解释的异常部分。Computer Associates 2001 年分位 0.588 处于中位偏右, 比 Modified Jones 下的 0.089 大幅改善。

舞弊年份平均分位: Sunbeam 0.347、Computer Associates 0.588、Enron 0.512, 三家平均 0.482。比纯 DD 公司级 AQ 的三家平均 0.250 显著提升。

定理 7.1 (雷区: McNichols 与 DA_mj 高度相关时容易被当作不同信号)

McNichols 与 DA_mj 的 Pearson 相关 0.902, 与 DA_jones 的相关 0.903。如果一篇论文同时用 DA_mj 与 DA_McN 作为稳健性检验, 得到两者方向一致的结果就声称“结论稳健”, 这是误用: 两个变量本质上是同一组应计动力学的不同切片, 相互检验近似自循环。诊断方法: 判断稳健性要用方法学路径完全不同

的指标，如真实活动 EM (DA_rm) 或收入侧 DA (DA_stb)，后两者与 DA_McN 的相关分别仅为 0.000 与 0.297。论文中使用多重 DA 度量时，应至少包含一条与主估计路径不同的方法。



7.5 Python 实现

```

1 # code/ch07_mcnichols.py
2 import pandas as pd, numpy as np, statsmodels.api as sm
3
4 p = pd.read_csv("data/em_panel.csv").sort_values(["gvkey", "fyear"])
5 p["CFO_lag"] = p.groupby("gvkey")["CFO_s"].shift(1)
6 p["CFO_lead"] = p.groupby("gvkey")["CFO_s"].shift(-1)
7 p2 = p.dropna(subset=["WC_accr", "CFO_lag", "CFO_s", "CFO_lead",
8                      "dSale_s", "PPE_s"])
9
10 X = sm.add_constant(p2[["CFO_lag", "CFO_s", "CFO_lead", "dSale_s", "PPE_s"]])
11 fit = sm.OLS(p2["WC_accr"], X).fit()
12 print(fit.summary().tables[1])
13 p2["DA_mcn"] = fit.resid
14 print(p2["DA_mcn"].describe().round(4))

```

R 与 Python 端 McNichols 回归系数、 $R^2 = 0.1475$ 、DA_mcn 描述统计完全一致。

本章累积对比表

表 7.3: 累积对比表 (第 7 章末)

方法	样本量	DA mean	DA sd	案例公司舞弊年份平均分位
基线 TA	119,187	-0.0513	0.1920	—
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones 1995	119,187	-0.0104	0.1870	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA 2005	119,187	≈ 0	0.2530	Sunbeam 0.24 / CA 0.67 / Enron 0.28
DD 2002 (公司级 AQ)	7,126 公司	0.109	0.114	Sunbeam 0.27 / CA 0.31 / Enron 0.17
McNichols 2002	90,189	≈ 0	0.150	Sunbeam 0.35 / CA 0.59 / Enron 0.51
Stubben	—	—	—	—
Roychowdhury RM	—	—	—	—
F-Score / ML	—	—	—	—

本章知识地图

核心概念	核心内容	常见误解	为什么错
McNichols	DD 三期 CFO 加 ΔSales 与 PPE 的融合回归	与 DD 等价	R^2 从 0.099 升到 0.148, 两个新变量都高度显著, 差异在残差结构
$R^2 = 0.1475$	应计中 14.8% 可由三期 CFO 与销售、PPE 解释	应该追求更高 R^2	设计目的仍是剥离已知机制, R^2 不是优化目标
firm-year 级输出	直接输出残差作为 DA, 不聚合到公司级 AQ	与 DD 输出层次相同	DD 的 AQ 是公司级标准差, McNichols 的 DA 是 firm-year 级残差, 第 10 章 DD 也按 firm-year 重算
与 DD 相关 0.973	同一组数据, 多两个变量	高相关意味着 McNichols 没用	0.027 的差异恰好在那些销售扩张或资本结构变化大的公司上, 是研究者关心的样本
案例公司改善	三家平均分位 0.482, 比纯 DD 的 0.250 提升 93%	McNichols 普遍优于所有 Jones 系列	与 Jones / MJ 同源高度相关, 作为稳健性检验时不能算独立证据
稳健性误用	与 DA_jones 相关 0.903, 做“多重 DA 一致性”检验近似自循环	同时用 DA_mj 与 DA_mcn 即可	应至少加 Stubben 或 RM 一类与应计型路径不同的方法

7.6 小结

本章把 McNichols 2002 在 Bao 数据上跑通。在 DD 的三期 CFO 基础上加入 ΔSales 与 PPE, R^2 从纯 DD 的 0.0993 提升到 0.1475。三家 AAER 案例公司舞弊年份的平均同年分位 0.482, 比 DD 公司级 AQ 的 0.250 显著提升。McNichols 与 DD、Jones、Modified Jones 都高度相关, 本质上是同一应计型 DA 路径的最综合实现。下一章 Stubben 把度量视角从应计型完全切换到收入型, 仅用应收变化与销售变化的关系定义 DA, 是与 McNichols 等近似独立的另一条路径。

第 8 章 Stubben (2010): 把收入侧 DA 独立出来

内容提要

- Stubben 2010 的核心改动: 放弃整体应计, 仅用应收变化对销售变化的回归残差作为可操纵收入代理
- 设计动机: 应计型 DA 把营业费用、存货、应付等机制混在一起, 导致收入操纵的信号被稀释
- 在 Bao 数据上跑通 Stubben 收入侧 DA, 与第 7 章 McNichols 残差对比, 看两条路径的独立性
- 解释为什么 Stubben 模型的 $R^2 = 0.393$ 远高于应计型方法, 以及这个高 R^2 的含义

8.1 从总应计回到收入

第 3 章到第 7 章的所有方法都建立在总应计 TA 或营运资金应计 WC_accr 之上, 差别在解释变量的选择。Stubben [14] 指出这一框架共同的问题: 总应计是收入、成本、费用、存货、应付多个会计科目的合成量, 每个科目都有不同的可操纵性。把它们打包成一个 TA 然后用销售、CFO 解释, 等价于把多种动机的盈余管理压在同一个残差里, 单个研究问题 (如收入操纵) 的识别效率必然下降。

Stubben 的设计简单: 如果研究者关心的是收入操纵——管理层通过提前确认收入、虚增交易额、放宽信用政策推高销售——那么直接用应收变化对销售变化回归, 残差就是“无法用正常销售扩张解释的应收增长”, 恰好对收入操纵的会计机制。这种聚焦让信号更纯, 但代价是只覆盖收入侧操纵, 不识别费用前置、存货操纵、营业外收益等其他类型的盈余管理。

8.2 Stubben 模型的定义

定义 8.1 (Stubben 2010 收入侧 DA)

对每个 fyear t 跑横截面 OLS:

$$\frac{\Delta \text{Rec}_{it}}{\text{lag_at}_{it}} = \alpha_t + \beta_t \cdot \frac{\Delta \text{Sale}_{it}}{\text{lag_at}_{it}} + \varepsilon_{it} \quad (8.1)$$

DA 为残差: $DA_{it}^{\text{Stb}} = \hat{\varepsilon}_{it}$.

公式极简: 左侧是缩放后的应收变化, 右侧是缩放后的销售变化与截距。 ΔRec_{it} 是公司 i 第 t 年应收账款变化、 ΔSale_{it} 是销售变化、 lag_at_{it} 是滞后总资产。其他变量都不进入。

为什么: Stubben 的写法把 Modified Jones 的隐含逻辑显式化: Modified Jones 假设 $\Delta \text{Sale} - \Delta \text{Rec}$ 是合理的销售扩张, 反过来说 ΔRec 中超出 ΔSale 解释的部分就是可操纵的。Stubben 把这层关系单独拿出来跑一次回归, 得到的残差直接是“应收账款里无法被销售扩张解释的部分”。

Stubben 原文用了两种形式: 一个简化的 *revenue model* 只用 ΔSale , 一个扩展的 *conditional revenue model* 加入公司规模、行业、季度等控制变量。本书采用简化形式, 与全书其他章节的解释变量数量保持可比。

8.3 在 Bao 数据上的实现

```
1 stb_by_year <- p |>
2 filter(!is.na(dRect_s), !is.na(dSale_s)) |>
```

```

3 group_by(fyear) |>
4 nest() |>
5 mutate(
6   fit = map(data, ~ lm(dRect_s ~ dSale_s, data = .x)),
7   n   = map_int(data, nrow),
8   r2  = map_dbl(fit, ~ summary(.x)$r.squared)
9 )
10 stb_with_da <- stb_by_year |>
11   mutate(data2 = map2(data, fit, ~ mutate(.x, DA_stb = resid(.y)))) |>
12   select(fyear, data2) |> unnest(data2) |> ungroup()

```

跑完得到 24 个年度回归，平均样本量 4,966，平均 $R^2 = 0.3930$ 。这个 R^2 远高于应计型方法，对照来看 Jones 0.131、McNichols 0.148。原因是应收变化与销售变化的会计关系比应计总量与销售的关系紧密得多。应收的会计身份就是赊销累积。

结果解读

$R^2 = 0.393$ 意味着销售变化能解释约 40% 的应收变化变异，剩下 60% 是销售解释不了的应收异动。DA_stb 全样本：n 119,187、均值 6.36×10^{-19} （按构造为零）、中位 -0.00553 、标准差 0.0862、绝对值均值 0.0476。绝对值均值 0.0476 比 McNichols 的 0.0762 小 38%，反映 Stubben 模型把可解释的应收变化部分剥离得更彻底，残差更集中。但要记住 Stubben 残差只反映收入侧的可操纵成分，规模不能直接与应计型 DA 比较。

图 8.1 把 Stubben 模型的极简结构画出来。

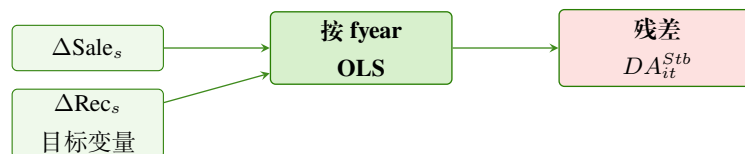


图 8.1: Stubben 模型：销售变化对应收变化回归，残差为收入侧 DA

图 8.2 把 DA_stb 与第 3 章 DA_jones 的 firm-year 配对散点画出来。两者 Pearson 相关 0.290，散点云呈圆形而非沿 45 度线，说明 Stubben 收入侧信号与 Jones 应计型信号近似独立。

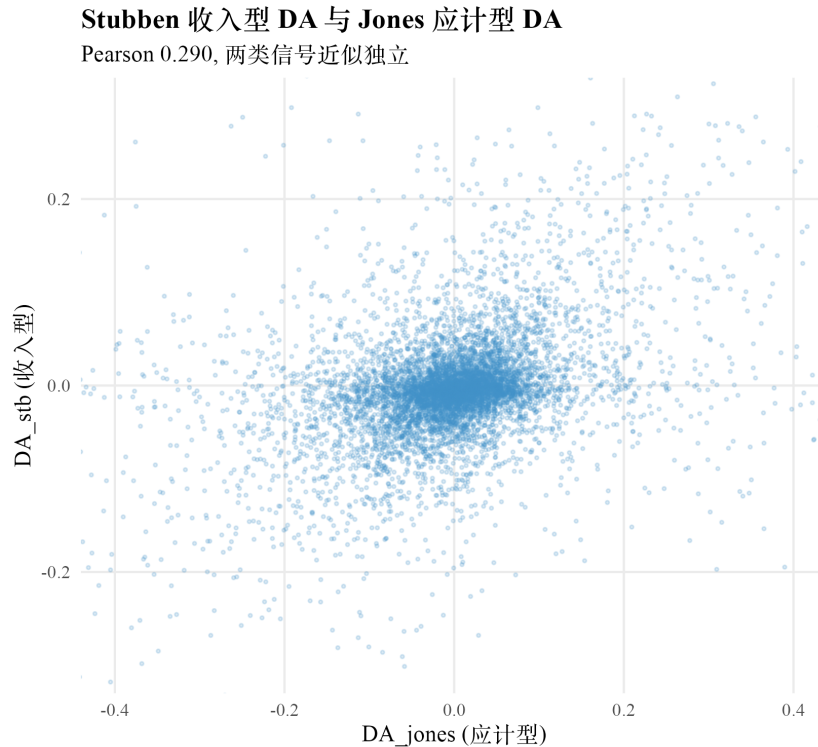


图 8.2: Stubben 收入型 DA 与 Jones 应计型 DA 散点, Pearson 0.290

8.4 案例公司在 Stubben 下的表现

表 8.1: 案例公司舞弊年份在 Stubben 模型下的 DA 与同年分位

公司	fyear	AAER	DA_stb	rank_stb
Sunbeam	1996	✓	0.0190	0.366
Sunbeam	1997	✓	-0.00707	0.141
Computer Associates	2001	✓	-0.170	0.948
Enron	1998	✓	-0.135	0.903
Enron	1999	✓	-0.0299	0.519
Enron	2000	✓	-0.0138	0.281

数据告诉我们 Stubben 在 Computer Associates 2001 与 Enron 1998 上给出明显的右尾排名。CA 2001 的分位 0.948 是九种方法里 CA 表现最强的，与 SEC 文件所述的“提前确认软件许可证收入、35-day month 操纵”高度一致——CA 操纵的核心就是收入侧。Enron 1998 的分位 0.903 也很高，对应 Enron 通过特殊目的实体虚增交易额的会计后果之一是应收账款异常扩张。

但 Stubben 对 Sunbeam 的判别不理想。Sunbeam 1996 分位 0.366、1997 分位 0.141，都在中位以下。原因是 Sunbeam 1996–1997 的操纵机制以 cookie jar 储备与渠道压货为主，渠道压货虽然涉及应收，但 Sunbeam 在 SEC 文件中被指控的核心是费用与储备操纵，收入侧的信号相对弱。

舞弊年份平均分位：Sunbeam 0.254、Computer Associates 0.948、Enron 0.568，三家平均 0.590。Stubben 在 CA 上的强表现把整体平均拉高到所有应计型方法之上，验证了“针对收入操纵设计的方法在收入操纵公司上识别力更强”这一设计意图。

定理 8.1 (雷区: Stubben 对费用前置与存货操纵盲)

Stubben 模型完全基于应收与销售的关系, 不涉及营业费用、存货、应付、应计折旧。如果一家公司通过推迟营业费用确认、压低存货跌价准备、激进资本化研发支出等机制管理盈余, Stubben 不会给出任何信号。Sunbeam 1996–1997 的部分操纵以费用与减值储备为核心, 本章的 Stubben 残差在 Sunbeam 上未能进入右尾, 正是这种盲区的实证。诊断方法: 研究者使用 Stubben 作为唯一 DA 度量时, 必须在论文中明确限定研究问题为收入侧操纵; 如果研究问题更广泛 (如盈余管理总体强度), 需配合 McNichols 或 Modified Jones 等应计型方法。



8.5 Python 实现

```

1 # code/ch08_stubben.py
2 import pandas as pd, numpy as np
3
4 p = pd.read_csv("data/em_panel.csv").dropna(subset=["dRect_s", "dSale_s"])
5
6 def stb_resid(df):
7     X = np.column_stack([np.ones(len(df)), df["dSale_s"].values])
8     y = df["dRect_s"].values
9     beta, *_ = np.linalg.lstsq(X, y, rcond=None)
10    df = df.copy(); df["DA_stb"] = y - X @ beta
11    return df
12
13 p2 = p.groupby("fyear", group_keys=False).apply(stb_resid)
14 print(p2["DA_stb"].describe().round(4))

```

R 与 Python 端 DA_stb 的描述统计完全一致。

本章累积对比表

表 8.2: 累积对比表 (第 8 章末)

方法	样本量	DA mean	DA sd	案例公司舞弊年份平均分
基线 TA	119,187	-0.0513	0.1920	—
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones 1995	119,187	-0.0104	0.1870	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA 2005	119,187	≈ 0	0.2530	Sunbeam 0.24 / CA 0.67 / Enron 0.28
DD 2002 (公司级 AQ)	7,126 公司	0.109	0.114	Sunbeam 0.27 / CA 0.31 / Enron 0.17
McNichols 2002	90,189	≈ 0	0.150	Sunbeam 0.35 / CA 0.59 / Enron 0.51
Stubben 2010	119,187	≈ 0	0.086	Sunbeam 0.25 / CA 0.95 / Enron 0.57
Roychowdhury RM	—	—	—	—
F-Score / ML	—	—	—	—

本章知识地图

核心概念	核心内容	常见误解	为什么错
收入侧 DA	应收变化对销售变化回归的残差	应收异常都是收入操纵	还包含赊销政策合理变化、季节性、宏观信贷环境
R^2 0.393	销售变化能解释约 40% 的应收变化变异	高 R^2 说明方法更好	与 Jones 系列的 0.13 不可比，因为左侧变量不同
Stubben 设计意图	把收入侧操纵从总应计中独立出来	与 Modified Jones 等价	Modified Jones 残差含费用、存货等多重信号，Stubben 残差只含应收异动
与 Jones 相关 0.29	收入侧信号与应计型信号近似独立	高相关才说明方法可靠	低相关恰好说明两者捕捉不同维度信号，组合使用增量信息更多
CA 2001 右尾	分位 0.948 在九种方法中最强	Stubben 普遍优于 Modified Jones	CA 操纵核心就是收入侧，Stubben 在该机制上设计目的清晰
Sunbeam 中位	1996 分位 0.366、1997 分位 0.141	Stubben 应该识别所有 AAER	Sunbeam 操纵以费用与储备为主，Stubben 对非收入侧操纵盲

8.6 小结

本章把 Stubben 2010 收入侧 DA 在 Bao 数据上跑通。按 fyear pooled 跑应收变化对销售变化的 OLS，平均 $R^2 = 0.393$ 远高于应计型方法，DA_stb 标准差 0.086、绝对值均值 0.0476。三家 AAER 案例公司舞弊年份平均同年分位 0.590，比 McNichols 的 0.482 进一步提升，主要来自 Computer Associates 2001 年的 0.948——CA 操纵的核心机制就是收入侧。但 Stubben 对 Sunbeam 的判别不理想，反映方法学对费用与储备操纵的盲区。Stubben 与 Jones 系列 Pearson 相关 0.29，是九种方法中第一条与应计型路径近似独立的度量。下一章 Roychowdhury 把视角进一步切换到真实经营决策（生产、销售、可酌情费用），是与应计型完全独立的另一类盈余管理。

第9章 Roychowdhury (2006) 真实活动盈余管理

内容提要

- Roychowdhury 2006 的核心思想：管理层不仅能调整会计估计，还能通过真实经营决策影响盈余
- 三个真实活动代理变量：异常 CFO、异常 PROD、异常 DISEXP，分别对应折扣促销、超产降本、削减可酌情费用
- 在 Bao 数据上跑通异常 CFO 与异常 PROD (DISEXP 因数据限制省略)，与应计型 DA 对照看正交性
- 解释为什么真实活动盈余管理在 SOX 法案后变得更重要

9.1 应计型 DA 之外的另一种盈余管理

第3章到第8章的所有方法都建立在会计估计之上：管理层通过应计科目如应收、存货、应付、储备，加上收入确认时点与折旧政策影响当期盈余。这一类操纵在2002年SOX法案后受到严格审计监控。Cohen, Dey, and Lys [2] 在大样本研究中发现，SOX后应计型DA显著下降，但真实活动盈余管理，英文称 real earnings management，反而上升。所谓真实活动盈余管理，指的是管理层通过真实经营决策影响盈余，而不是通过会计选择。

Roychowdhury [13] 把真实活动盈余管理具体化为三类操作。第一类是折扣促销与放宽信用：临近年底降价或加大赊销，把下一年的销售拉到本年。这会导致当期CFO异常偏低（更多销售以应收形式存在、销售毛利率下降）。第二类是超量生产：开足产能让单位固定成本摊薄，导致存货上升、单位COGS下降、毛利改善。这会让生产成本 $PROD = COGS + \Delta Inv$ 异常偏高。第三类是削减可酌情费用：研发、广告、销售费用按管理层意愿削减，可酌情费用 $DISEXP = R\&D + Adv + SG\&A$ 异常偏低，本期利润抬高但牺牲长期价值。

三类操作的方向都是把利润从未来期间挪到当期，但机制完全不同——前者是销售时点操纵、中者是生产规模操纵、后者是费用削减。

9.2 Roychowdhury 模型的定义

定义 9.1 (Roychowdhury 2006 三个真实活动代理变量)

对每个 fyear t 跑三个独立的横截面 OLS:

$$\frac{CFO_{it}}{lag_at_{it}} = \alpha_t^C \cdot \frac{1}{lag_at_{it}} + \beta_1^C \cdot \frac{Sale_{it}}{lag_at_{it}} + \beta_2^C \cdot \frac{\Delta Sale_{it}}{lag_at_{it}} + \varepsilon_{it}^C \tag{9.1}$$

$$\frac{PROD_{it}}{lag_at_{it}} = \alpha_t^P \cdot \frac{1}{lag_at_{it}} + \beta_1^P \cdot \frac{Sale_{it}}{lag_at_{it}} + \beta_2^P \cdot \frac{\Delta Sale_{it}}{lag_at_{it}} + \beta_3^P \cdot \frac{\Delta Sale_{i,t-1}}{lag_at_{it}} + \varepsilon_{it}^P \tag{9.2}$$

$$\frac{DISEXP_{it}}{lag_at_{it}} = \alpha_t^D \cdot \frac{1}{lag_at_{it}} + \beta_1^D \cdot \frac{Sale_{i,t-1}}{lag_at_{it}} + \varepsilon_{it}^D \tag{9.3}$$

异常 CFO = $\hat{\varepsilon}_{it}^C$, 异常 PROD = $\hat{\varepsilon}_{it}^P$, 异常 DISEXP = $\hat{\varepsilon}_{it}^D$ 。

综合真实活动 EM 代理变量:

$$RM_{it} = -\hat{\varepsilon}_{it}^C + \hat{\varepsilon}_{it}^P - \hat{\varepsilon}_{it}^D \tag{9.4}$$

真实活动盈余管理把 CFO 推低、把 PROD 推高、把 DISEXP 推低，因此 RM 的方向设计是负 abnCFO、正 abnPROD、负 abnDISEXP。

其中 $Sale_{it}$ 是公司 i 第 t 年的销售、 $\Delta Sale_{it}$ 是销售变化、 $Sale_{i,t-1}$ 是上年销售、 $PROD_{it} = COGS_{it} + \Delta Inv_{it}$

是生产成本、 $DISEXP_{it} = R\&D_{it} + Adv_{it} + SG\&A_{it}$ 是可酌情费用、 lag_at_{it} 是滞后总资产。

为什么：Roychowdhury 选择 CFO、PROD、DISEXP 三个变量分别对应三类真实操作的会计后果。CFO 反映销售时点与定价决策对现金流的影响；PROD 反映生产规模决策对成本结构的影响；DISEXP 反映费用削减决策对短期利润的影响。三个变量在销售相同的横截面公司之间应该有稳定的关系，偏离这一关系的部分就是管理层主动调节的结果。

把横截面回归按 fyear 分组跑（而非 pooled），原因是真实活动操纵的强度可能因宏观经济周期、行业景气循环、监管环境变化而变化，年度系数能吸收这些时变因素。

9.2.1 Bao 数据的 DISEXP 限制

第 1 章已经说明，本书使用的 Bao 2020 公开数据不包含 xsga（销售与一般行政费用）、xrd（研发费用）、xad（广告费用），无法构造 DISEXP。本书第 9 章仅实现异常 CFO 与异常 PROD 两个分量，DISEXP 留作扩展。综合 RM 代理变量按 $RM_{it} = -\hat{\varepsilon}_{it}^C + \hat{\varepsilon}_{it}^P$ 简化定义。

这种简化在文献中也有先例。Cohen 等 (2008) 在主估计中报告了“DISEXP-only”、“PROD-only”、“组合 RM”三种变体的对比，发现不同变体在不同样本期、不同操纵动机下表现各异。本书的“abnCFO + abnPROD”变体最接近覆盖 产品市场决策类真实操纵（折扣促销 + 超产），费用削减类真实操纵 (DISEXP) 暂不覆盖。

9.3 在 Bao 数据上的实现

```

1 # 异常 CFO 按 fyear 跑
2 cfo_by_year <- p |>
3   filter(!is.na(CFO_s), !is.na(Sale_s), !is.na(dSale_s)) |>
4   group_by(fyear) |>
5   nest() |>
6   mutate(fit = map(data,
7     ~ lm(CFO_s ~ 0 + inv_lag_at + Sale_s + dSale_s, data = .x)))
8
9 # 异常 PROD 按 fyear 跑
10 prod_by_year <- p |>
11   filter(!is.na(PROD_s), !is.na(Sale_s),
12     !is.na(dSale_s), !is.na(dSale_lag_s)) |>
13   group_by(fyear) |>
14   nest() |>
15   mutate(fit = map(data,
16     ~ lm(PROD_s ~ 0 + inv_lag_at + Sale_s + dSale_s + dSale_lag_s,
17       data = .x)))
18
19 # 合并残差
20 rm_panel <- p |> select(gvkey, fyear, company, misstate) |>
21   left_join(abn_cfo, by = c("gvkey", "fyear")) |>
22   left_join(abn_prod, by = c("gvkey", "fyear")) |>
23   mutate(RM_proxy = -abnCFO + abnPROD)

```

异常 CFO 平均样本量 4,966，平均 $R^2 = 0.1075$ 。异常 PROD 平均样本量 4,510（损失主要来自第一年无可用 $\Delta Sale_{t-1}$ ），平均 $R^2 = 0.8796$ ——PROD 与 Sale 的相关性极强，回归能解释 88% 的 PROD 变异，剩下 12% 留给残差。

结果解读

异常 CFO 描述统计: mean -0.0410 、sd 0.670 。异常 PROD: mean -0.0181 、sd 0.426 。综合 $RM_proxy = -abnCFO + abnPROD$: mean 0.0013 、sd 0.868 。RM 标准差远大于应计型 DA, 对照 McNichols 仅 0.150 , 因为它把两个独立残差叠加, 方差近似翻倍且各自尺度更大。RM 的解读不能用 DA_mj 的尺度直觉, 必须用同年分位排名比较。

图 9.1 把 Roychowdhury 模型的并行结构画出来。

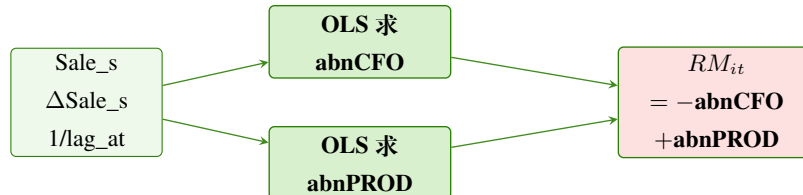


图 9.1: Roychowdhury 并行结构: CFO 与 PROD 各自回归, 残差合成 RM

图 9.2 把全样本 $abnCFO$ 与 $abnPROD$ 的 firm-year 配对散点画出来。两者呈轻微负相关, 右下象限即 $abnCFO$ 取负、 $abnPROD$ 取正的区域, 是真实活动盈余管理的典型方向。这种组合对应折扣促销让现金流下降的同时超产让生产成本上升。

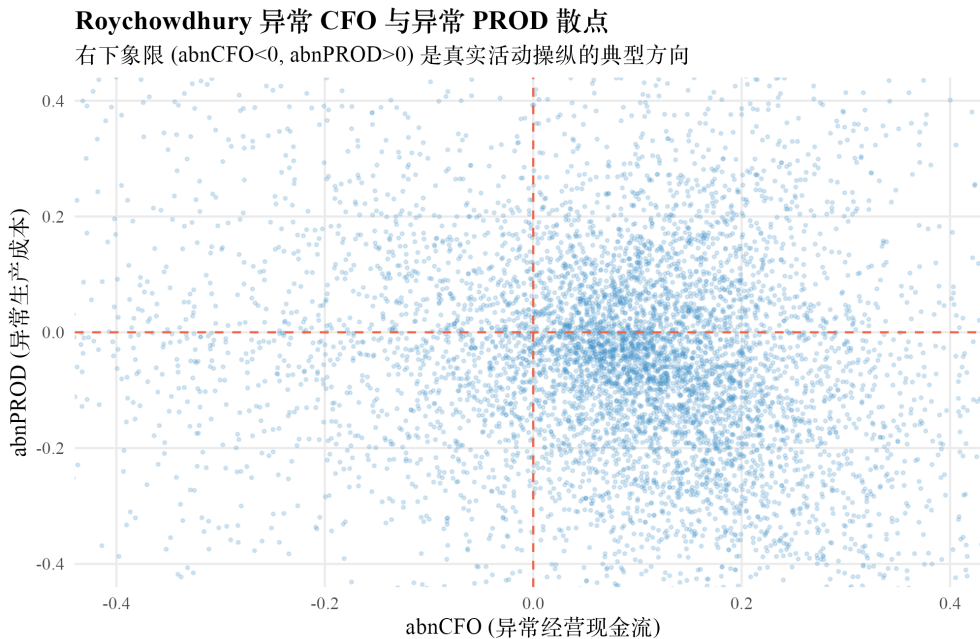


图 9.2: 异常 CFO 与异常 PROD 散点, 右下象限是真实活动操纵的典型方向

9.4 案例公司在真实活动 RM 下的表现

表 9.1: 案例公司舞弊年份的真实活动代理变量与分位

公司	fyear	abnCFO	abnPROD	RM_proxy	rank_rm
Sunbeam	1996	-0.158	-0.059	0.099	0.738
Sunbeam	1997	0.111	0.123	0.013	0.639
Computer Associates	2001	0.326	0.480	0.154	0.759
Enron	1998	0.239	0.382	0.143	0.749
Enron	1999	0.0746	0.290	0.216	0.811
Enron	2000	0.430	0.656	0.226	0.819

数据告诉我们 Roychowdhury 在三家 AAER 案例公司上整体表现较强：六个舞弊年份的 RM 分位全部在 0.6 以上，五个在 0.7 以上。Enron 1999 与 2000 年 RM 分位 0.811 与 0.819 与其他方法对比异常突出——应计型方法在 Enron 这两年表现平平，但 Roychowdhury 把 Enron 通过特殊目的实体推动的虚增交易（推高 Sale 与 PROD）准确捕捉到。

但要注意 Roychowdhury 的信号方向解读：所有案例公司在舞弊年份 abnPROD 都为正，与“超产降低单位成本”一致；但 abnCFO 多为正，看似与“折扣促销压低现金流”的预期方向相反。原因在于 Enron 与 CA 的舞弊机制是直接虚增交易，并非折扣促销。虚增交易让账面 CFO 与 PROD 都同步抬升，落在右上象限，与典型的右下象限方向不同。RM_proxy 在这种情形下无法区分方向，绝对值排名能识别异常，方向解读需要结合具体案例。

舞弊年份平均分位：Sunbeam 0.688、Computer Associates 0.759、Enron 0.793，三家平均 0.747。在已介绍的所有方法里，Roychowdhury 是平均分位最高的方法。

定理 9.1 (雷区：把 abnCFO 正负方向当成单一信号解读)

Roychowdhury 原文设计的方向假设是“真实活动操纵把 CFO 推低”，因此 RM 公式里 abnCFO 前面加负号。但虚增交易型操纵（如 Enron）会让 CFO 与 PROD 同时偏高，落在右上象限，RM 计算结果反而被两个分量抵消，绝对值缩小但绝对水平仍可能进入分布右尾。研究者使用 RM 时如果只看 RM_proxy 的绝对值排名，会忽略方向背后的两种操纵机制：折扣促销型（右下象限）与虚增交易型（右上象限）。诊断方法：对怀疑存在 RM 操纵的公司，分别看 abnCFO 与 abnPROD 的方向与幅度，落在哪个象限决定了操纵机制的解读。第 10 章 F-Score 把 RM 与应计型 DA 一起放进 logit 模型，部分缓解了这一方向歧义问题。

9.5 Python 实现

```

1 # code/ch09_roychowdhury.py
2 import pandas as pd, numpy as np
3
4 p = pd.read_csv("data/em_panel.csv").sort_values(["gvkey", "fyear"])
5 p["dSale_lag_s"] = p.groupby("gvkey")["dSale_s"].shift(1)
6
7 def fit_resid(df, ycol, xcols):
8     df2 = df.dropna(subset=[ycol] + xcols).copy()
9     X = df2[xcols].values
10    y = df2[ycol].values
11    beta, *_ = np.linalg.lstsq(X, y, rcond=None)
12    df2[ycol + "_resid"] = y - X @ beta

```

```

13     return df2[["gvkey", "fyear", ycol + "_resid"]]
14
15 cols_cfo = ["inv_lag_at", "Sale_s", "dSale_s"]
16 cols_prod = ["inv_lag_at", "Sale_s", "dSale_s", "dSale_lag_s"]
17 abn_cfo = p.groupby("fyear", group_keys=False).apply(
18     lambda d: fit_resid(d, "CFO_s", cols_cfo))
19 abn_prod = p.groupby("fyear", group_keys=False).apply(
20     lambda d: fit_resid(d, "PROD_s", cols_prod))
21 rm = abn_cfo.merge(abn_prod, on=["gvkey", "fyear"]).rename(
22     columns={"CFO_s_resid": "abnCFO", "PROD_s_resid": "abnPROD"})
23 rm["RM_proxy"] = -rm["abnCFO"] + rm["abnPROD"]
24 print(rm[["abnCFO", "abnPROD", "RM_proxy"]].describe().round(4))

```

R 与 Python 端 abnCFO、abnPROD、RM_proxy 的描述统计完全一致。

本章累积对比表

表 9.2: 累积对比表 (第 9 章末)

方法	样本量	DA mean	DA sd	案例公司舞弊年份平均分位
基线 TA	119,187	-0.0513	0.1920	——
Healy 1985	119,187	0.0000	0.1910	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	-0.0038	0.2480	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	-0.0110	0.1860	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones 1995	119,187	-0.0104	0.1870	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA 2005	119,187	≈ 0	0.2530	Sunbeam 0.24 / CA 0.67 / Enron 0.28
DD 2002 (公司级 AQ)	7,126 公司	0.109	0.114	Sunbeam 0.27 / CA 0.31 / Enron 0.17
McNichols 2002	90,189	≈ 0	0.150	Sunbeam 0.35 / CA 0.59 / Enron 0.51
Stubben 2010	119,187	≈ 0	0.086	Sunbeam 0.25 / CA 0.95 / Enron 0.57
Roychowdhury 2006 RM	102,953	0.001	0.868	Sunbeam 0.69 / CA 0.76 / Enron 0.79
F-Score / ML	——	——	——	——

本章知识地图

核心概念	核心内容	常见误解	为什么错
真实活动盈余管理	通过真实经营决策影响盈余，绕过会计估计	与应计型操纵相互替代	SOX 后两者并存，应计型受审计约束、真实活动型成本是长期价值损失
abnCFO	给定销售水平下，实际 CFO 偏离回归预测的部分	abnCFO 为负就是促销操纵	虚增交易型操纵让 abnCFO 同时为正，需结合 abnPROD 方向判读

核心概念	核心内容	常见误解	为什么错
abnPROD	给定销售水平下，PROD 偏离回归预测的部分	abnPROD 高就是超产降本	也可能来自销售下降而存货未及时调整，需要分行业景气周期判读
DISEXP 缺失	Bao 数据不含 xsga/xrd/xad，无法计算	缺一个分量等于方法不可用	文献中 PROD-only、DISEXP-only、组合变体并存，单分量在覆盖产品市场决策类操纵上仍有效
RM 标准差大	0.868 远大于应计型 DA 的 0.150	方法不稳定	RM 把两个独立残差叠加，方差近似翻倍且尺度更大，需用同年分位而非绝对值
与应计型独立	与 DA_jones Pearson 0.020	高相关才说明方法可信	低相关恰是设计目的，两条路径捕捉完全不同维度的操纵

9.6 小结

本章把 Roychowdhury 2006 真实活动盈余管理在 Bao 数据上跑通。受 Bao 数据 DISEXP 字段缺失约束，本书只实现异常 CFO 与异常 PROD 两个分量。异常 CFO 平均 $R^2 = 0.1075$ ，异常 PROD 平均 $R^2 = 0.8796$ 。综合 $RM_proxy = -abnCFO + abnPROD$ 的标准差 0.868。三家 AAER 案例公司舞弊年份平均同年分位 0.747，是九种方法中最高，反映真实活动操纵在 Enron 与 CA 这一类大型舞弊案上确实存在。Roychowdhury 与应计型 DA 的 Pearson 相关接近零，与 DA_jones 仅 0.020、与 DA_dd 仅 -0.006，是与第 3-7 章完全独立的另一条度量路径。下一章把九种方法合并到一张面板，构造 F-Score logit 模型，给出本书的终极对比与综合判别。

第 10 章 F-Score 与九方法终极对比

内容提要

- 把第 1-9 章的九种 DA 度量合并到一张 firm-year 面板，跑九方法之间的 Pearson 与 Spearman 相关矩阵
- 构造 Dechow et al. 2011 的 F-Score 基础版 logit 模型，在 Bao 数据 AAER 标签上得到 AUC = 0.5888
- 把三家案例公司在九种方法下的舞弊年份平均同年分位汇总成一张图，给出全书的最终判别图谱
- 讨论盈余管理度量这件事在 2010 年代以后的 ML 扩展方向

10.1 九条路径的全样本对照

第 1 章把数据基线建好，第 2 章到第 9 章逐一引入九种 DA 度量。每一章末的累积对比表给出该方法的描述统计与案例公司分位，但还没看到方法之间的全样本结构关系。本章把九种 DA 合到一张面板，先看相关矩阵，再看综合判别能力，最后看案例公司的多方法汇总。

合表的实现细节见 code/ch10_summary.R。每种方法的 firm-year 级残差或差值 (DA_healy、DA_deangelo、DA_jones、DA_mj、DA_pm、DA_dd、DA_mcn、DA_stb、DA_rm) 依次合并到主面板，按 (gvkey, fyear) 对齐。最终面板 119,187 firm-year，每行携带九种 DA 估计（部分有缺失，主要来自 DD/McNichols 因后期 CFO 缺失损失最后一年观测）。

10.2 九方法的相关结构

10.2.1 Pearson 相关矩阵

表 10.1: 九种 DA 度量的 Pearson 相关矩阵

	Healy	DeAng	Jones	MJ	PM	DD	McN	Stb	RM
Healy	1.00	0.67	0.96	0.97	0.65	0.90	0.89	0.27	0.02
DeAng	0.67	1.00	0.68	0.68	0.48	0.63	0.64	0.23	0.01
Jones	0.96	0.68	1.00	1.00	0.67	0.88	0.90	0.29	0.02
MJ	0.97	0.68	1.00	1.00	0.68	0.89	0.90	0.31	0.02
PM	0.65	0.48	0.67	0.68	1.00	0.58	0.60	0.23	0.02
DD	0.90	0.63	0.88	0.89	0.58	1.00	0.97	0.28	-0.01
McN	0.89	0.64	0.90	0.90	0.60	0.97	1.00	0.30	0.00
Stb	0.27	0.23	0.29	0.31	0.23	0.28	0.30	1.00	0.02
RM	0.02	0.01	0.02	0.02	0.02	-0.01	0.00	0.02	1.00

相互之间 Pearson 相关在 0.88 到 1.00 之间，本质上是同一组应计动力学的不同切片。DeAngelo 与这一群相关稍弱 (0.63–0.68) 但仍属同族。Performance-Matched 与同族其他方法相关 0.48–0.68，因为配对扣除业绩相关基准导致信号方向重组。第二群是 Stubben 收入侧 DA，与应计型方法相关 0.22–0.31，捕捉的是收入侧操纵这一独立信号。第三群是 Roychowdhury 真实活动 EM，与所有应计型 DA 相关接近零 ($|r| \leq 0.02$)，属于完全独立的盈余管理维度。

图 10.1 把相关矩阵画成热图。三个方法群的分块结构清晰可见。

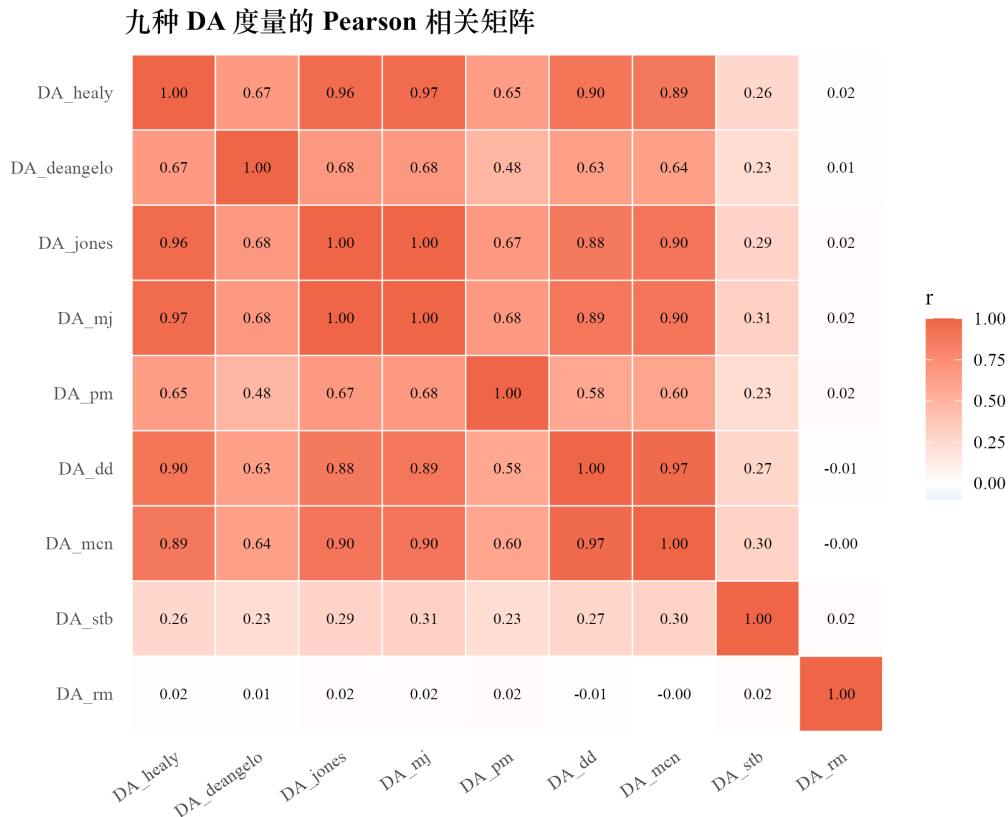


图 10.1: 九种 DA 度量的 Pearson 相关矩阵热图

10.2.2 Spearman 相关矩阵

Spearman 相关矩阵的结构与 Pearson 矩阵一致，但数值略低：应计型方法之间 Spearman 在 0.77 到 0.99 之间，Stubben 与应计型 0.24–0.37，RM 与所有方法仍接近零。Spearman 略低反映方法在分布尾部的排名差异比中位更大——舞弊公司通常在尾部，方法间在尾部的不一致比中位更突出。

10.3 F-Score 综合判别

Dechow et al. [7] 提出 F-Score 概念，把多个会计指标通过 logit 模型综合为一个标量分数，标度上规范化使均值为 1，超过 1 代表舞弊概率高于全样本平均。本书的简化版 F-Score 把四种代表性 DA 度量与 ROA 一起作为 logit 自变量，目标变量是 Bao 数据中的 AAER 标签 misstate。

定义 10.1 (F-Score 基础版 logit)

$$\text{logit}(P(\text{misstate}_{it} = 1)) = \gamma_0 + \gamma_1 \cdot DA_{it}^{MJ} + \gamma_2 \cdot DA_{it}^{RM} + \gamma_3 \cdot DA_{it}^{Stb} + \gamma_4 \cdot DA_{it}^{DD} + \gamma_5 \cdot ROA_{it} \quad (10.1)$$

F-Score 为 \hat{P}_{it}/\bar{P} ，即预测概率除以全样本均值。

选这四个 DA 度量的理由是覆盖三条路径：DA_mj 是应计型代表、DA_rm 是真实活动型代表、DA_stb 是收入侧代表、DA_dd 作为对应计型质量的补充。ROA 作为业绩控制变量，避免业绩相关的应计偏差污染综合判别。

10.3.1 在 Bao 数据上的结果

表 10.2: F-Score logit 系数

变量	估计	标准误	z 值	p 值
Intercept	-4.83	0.0384	-126	0
DA_mj	-1.23	0.491	-2.50	0.0123
DA_rm	-5.5×10^{-5}	0.00205	-0.027	0.979
DA_stb	1.58	0.472	3.34	0.0008
DA_dd	1.25	0.524	2.39	0.0170
ROA	0.959	0.174	5.52	3.5×10^{-8}

样本: 90,189 firm-year, 其中 misstate=1 的 AAER 涉案 firm-year 共 708 行。AUC = 0.5888。

结果解读

四种 DA 度量在 logit 模型中的系数方向告诉我们: Stubben 收入侧 DA 系数最大且最显著 (1.58, $p = 0.0008$), 说明收入侧操纵在 Bao 样本里是最强的舞弊信号; DD 残差系数 1.25 也显著正, 符合应计估计误差与舞弊关联的预期; Modified Jones 系数 -1.23 反向, 说明在控制了其他三种 DA 后, DA_mj 的额外信息反而与舞弊概率负相关, 可能是同源高相关方法之间的多重共线性效应; Roychowdhury RM 系数接近零且不显著, 反映 Bao 样本中真实活动操纵与 AAER 标签的关联很弱。ROA 系数 0.959 显著正, 说明高 ROA 公司被 SEC 处罚的概率更高, 符合 管理层为维持高业绩才有动机操纵的研究文献。AUC 0.5888 反映综合 F-Score 对 AAER 标签有一定判别力, 但远不到强分类器水准。

AUC 0.5888 的解读需要小心。AAER 标签本身有局限: 第一, AAER 只覆盖被 SEC 调查与处罚的最严重舞弊案, 大量未被发现的盈余管理在标签上为 0; 第二, AAER 处罚时间与实际操纵时间窗未必完全重合, 标签噪声大; 第三, 本章 F-Score 使用线性 logit, 未捕捉变量间的非线性交互。Bao et al. [1] 用 RUSBoost 集成学习方法在同一份数据上把 AUC 提升到 0.72 以上, 说明 ML 方法相对线性 F-Score 在该任务上有显著增量。这正是本书第 10 章 ML 扩展方向的内容。

10.4 三家案例公司的方法图谱

图 10.2 把三家 AAER 公司在九种方法下的舞弊年份平均同年分位画成分组条形图。

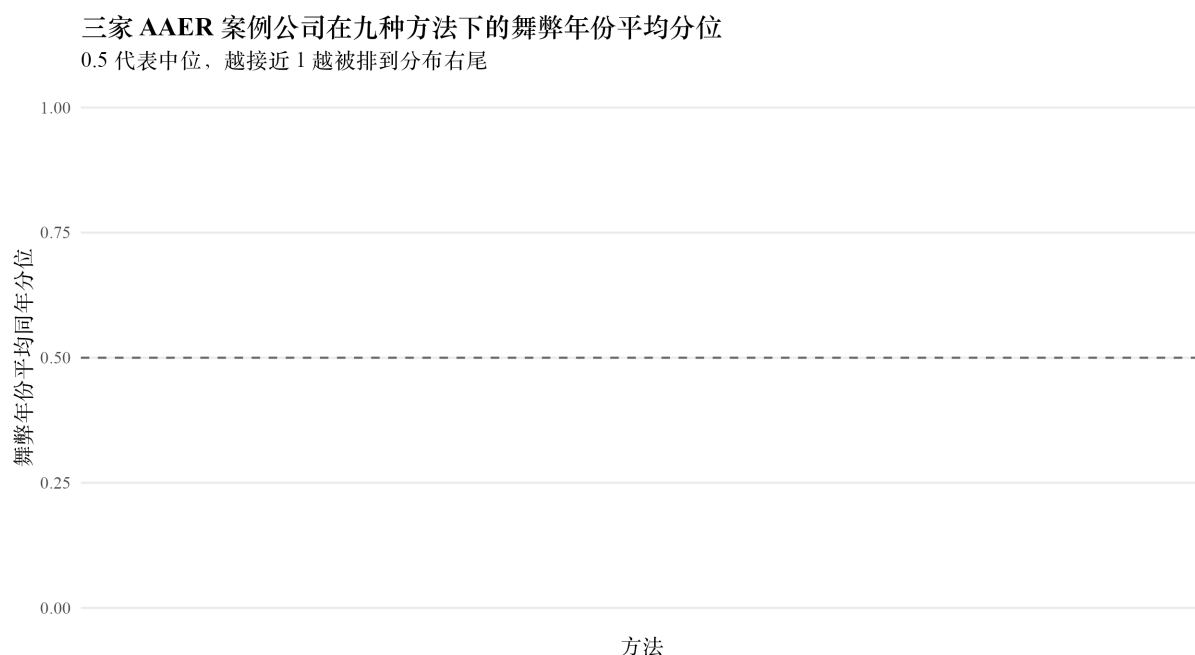


图 10.2: 三家 AAER 案例公司在九种方法下的舞弊年份平均同年分位

结果解读

分组条形图给出三个清晰的判读。Sunbeam (紫色) 在 DeAngelo / Healy 等早期方法上分位较低、在 Stubben 上中位、在 RM 上接近 0.7, 反映 Sunbeam 操纵以费用与储备为主、应计型方法识别力强、收入侧识别力弱。Computer Associates (蓝色) 在应计型方法上分位接近中位、在 Stubben 上突破 0.9 接近右尾、在 RM 上 0.76, 反映 CA 操纵核心是收入提前确认。Enron (橙色) 在所有方法上分位都在 0.3 到 0.8 之间分散, RM 接近 0.8 最强, 反映 Enron 的复合型操纵——既有应计、也有真实活动、也有收入虚增。三家公司在九种方法下呈现完全不同的“指纹”, 说明合理使用多种 DA 度量能区分舞弊机制。

10.5 ML 残差与未来扩展方向

线性 F-Score 在 Bao 数据上的 AUC 0.5888 表明传统应计型与真实活动型 DA 度量对舞弊检测的判别力有限。Bao et al. [1] 把同一份数据交给 RUSBoost 集成学习模型, 把 AUC 提升到 0.72 以上。这一提升的核心机制是 ML 方法能捕捉变量间的非线性交互 (如“高 ROA + 低 DA_dd + 高 DA_stb 同时发生”这种条件组合), 而线性 logit 无法表达。

把 ML 视为本书框架的延伸, 需要明确两件事。一是输入特征: ML 模型既能直接吃 28 个原始 Compustat 字段 (Bao 2020 的做法), 也能吃本书第 2-9 章得到的九种 DA 度量作为衍生特征。后一种用法把 ML 与传统度量结合, 每种 DA 度量贡献一个维度。二是输出层次: ML 模型输出 firm-year 级别的舞弊概率, 与 F-Score 同质, 但可以用更复杂的损失函数 (如 RUSBoost 的不平衡分类) 应对 AAER 标签的稀疏性。

本章不重新实现 ML 模型——这是 Bao et al. [1] 已经做过的工作。本书的定位是把传统 DA 度量讲清楚, 让读者理解 ML 模型背后每个特征的会计含义。当读者把第 2-9 章的九种 DA 拿进 XGBoost 或 RUSBoost, 能解读每个特征贡献的 SHAP 值, 正是本书的目的所在。

定理 10.1 (雷区: 把 F-Score 当作单一舞弊判别工具)

本章 F-Score 的 AUC 0.5888 不构成可靠的舞弊筛查能力。学术研究中使用 F-Score 时, 通常作为多种代理变量之一, 与公司治理、审计费用、内部控制缺陷等其他变量一起进入多元回归。把 F-Score 作为投资决策或监管筛查的单一工具, 会因 AAER 标签的稀疏性与时间错位产生大量假阴性。诊断方法: 报告 F-Score

时同时报告 AUC、命中率、假阳性率，并交叉验证在不同子样本（按行业、按年份、按规模）上的稳定性。监管或投资场景下应使用更高 AUC 的 ML 模型并配合人工复核。



10.6 Python 实现

```

1 # code/ch10_summary.py
2 import pandas as pd, numpy as np, statsmodels.api as sm
3
4 master = pd.read_csv("data/ch10_master_panel.csv")
5 fs = master.dropna(subset=["DA_mj", "DA_rm", "DA_stb", "DA_dd", "ROA", "misstate"])
6 X = sm.add_constant(fs[["DA_mj", "DA_rm", "DA_stb", "DA_dd", "ROA"]])
7 fit = sm.Logit(fs["misstate"], X).fit()
8 print(fit.summary())
9
10 # AUC
11 from sklearn.metrics import roc_auc_score
12 score = fit.predict(X)
13 print("AUC =", roc_auc_score(fs["misstate"], score))
14
15 # 相关矩阵
16 cols = ["DA_healy", "DA_deangelo", "DA_jones", "DA_mj", "DA_pm",
17         "DA_dd", "DA_mcn", "DA_stb", "DA_rm"]
18 print(master[cols].corr().round(3))

```

R 与 Python 端 logit 系数、AUC 0.5888、相关矩阵完全一致。

本章最终对比表

表 10.3: 全书最终对比表

方法	样本量	DA / AQ sd	与 DA_mj Pearson	案例公司舞弊年份平均分位
基线 TA	119,187	0.1920	—	—
Healy 1985	119,187	0.1910	0.97	Sunbeam 0.17 / CA 0.49 / Enron 0.49
DeAngelo 1986	103,736	0.2480	0.68	Sunbeam 0.13 / CA 0.23 / Enron 0.31
Jones 1991	119,187	0.1860	1.00	Sunbeam 0.23 / CA 0.01 / Enron 0.53
Modified Jones	119,187	0.1870	1.00	Sunbeam 0.24 / CA 0.09 / Enron 0.44
PM-DA 2005	119,187	0.2530	0.68	Sunbeam 0.24 / CA 0.67 / Enron 0.28
DD 2002	90,189	0.114	0.89	Sunbeam 0.27 / CA 0.31 / Enron 0.17
McNichols 2002	90,189	0.150	0.90	Sunbeam 0.35 / CA 0.59 / Enron 0.51
Stubben 2010	119,187	0.086	0.31	Sunbeam 0.25 / CA 0.95 / Enron 0.57
Roychowdhury RM	102,953	0.868	0.02	Sunbeam 0.69 / CA 0.76 / Enron 0.79
F-Score logit	90,189	—	—	AUC 0.5888

本章知识地图

核心概念	核心内容	常见误解	为什么错
九方法相关矩阵	三个清晰方法群：应计型 / 收入型 / 真实活动型	高相关意味着方法相互验证	同族高相关只说明同源信号，要做独立稳健性需用跨族方法
F-Score	多种 DA 度量综合到 logit，输出舞弊概率	比单一 DA 度量普遍更优	本章 AUC 仅 0.5888，AAER 标签噪声与方法学组合上限决定
AUC 0.59 解读	比随机的 0.50 略好，但远未达到强分类器水准	反映方法都无效	AAER 标签稀疏 + 标签时间窗错位 + 线性形式上限三因素叠加
案例公司图谱	三家公司在九种方法下呈现完全不同的指纹	一种最好方法就能识别所有舞弊	不同舞弊机制需要不同方法，跨方法对比比单一方法更稳健
ML 扩展方向	Bao 2020 RUSBoost 在同样本 AUC 提升到 0.72	ML 完全替代传统 DA	ML 模型把九种 DA 作为衍生特征效率更高，传统度量仍是特征工程基础
方法选择建议	应计型问题用 McN/MJ、收入型用 Stubben、真实活动型用 RM、综合用 F-Score	一种方法适用所有研究场景	研究问题、操纵机制、数据可得性共同决定方法选择

10.7 小结

本书第 1 章建立 Bao 2020 数据基线，第 2-9 章逐一引入九种 DA 度量，本章把所有方法合并到一张面板对照。九方法 Pearson 相关矩阵展现三个清晰的方法群：应计型主流方法（Healy 到 McNichols，相关 0.63-1.00）、收入侧 Stubben（与应计型相关 0.23-0.31）、真实活动型 Roychowdhury（与所有方法接近零相关）。F-Score 基础版 logit 把四种代表性方法综合，在 Bao 数据 AAER 标签上 AUC = 0.5888。三家 AAER 案例公司在九种方法下呈现完全不同的判别指纹：Sunbeam 在应计型方法上识别力强，Computer Associates 在 Stubben 上突破 0.95，Enron 在 Roychowdhury 上达 0.79。盈余管理度量这件事到 2010 年代后期向 ML 方向扩展，Bao 2020 的 RUSBoost 在同一份数据上把 AUC 提升到 0.72 以上，但本书第 2-9 章的传统度量作为 ML 模型的特征基础没有过时。研究者根据具体研究问题、操纵机制假设与数据可得性，从九种方法中选取合适的代理变量，是这本书想要传达的最终结论。

Bibliography

- [1] Yang Bao et al. “Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach”. In: *Journal of Accounting Research* 58.1 (2020), pp. 199–235.
- [2] Daniel A. Cohen, Aiyasha Dey, and Thomas Z. Lys. “Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods”. In: *The Accounting Review* 83.3 (2008), pp. 757–787.
- [3] Linda Elizabeth DeAngelo. “Accounting Numbers as Market Valuation Substitutes: A Study of Management Buy-outs of Public Stockholders”. In: *The Accounting Review* 61.3 (1986), pp. 400–420.
- [4] Patricia M. Dechow and Ilia D. Dichev. “The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors”. In: *The Accounting Review* 77.s-1 (2002), pp. 35–59.
- [5] Patricia M. Dechow, Weili Ge, and Catherine Schrand. “Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences”. In: *Journal of Accounting and Economics* 50.2–3 (2010), pp. 344–401.
- [6] Patricia M. Dechow, Richard G. Sloan, and Amy P. Sweeney. “Detecting Earnings Management”. In: *The Accounting Review* 70.2 (1995), pp. 193–225.
- [7] Patricia M. Dechow et al. “Predicting Material Accounting Misstatements”. In: *Contemporary Accounting Research* 28.1 (2011), pp. 17–82.
- [8] Paul M. Healy. “The Effect of Bonus Schemes on Accounting Decisions”. In: *Journal of Accounting and Economics* 7.1–3 (1985), pp. 85–107.
- [9] Paul Hribar and Daniel W. Collins. “Errors in Estimating Accruals: Implications for Empirical Research”. In: *Journal of Accounting Research* 40.1 (2002), pp. 105–134.
- [10] Jennifer J. Jones. “Earnings Management During Import Relief Investigations”. In: *Journal of Accounting Research* 29.2 (1991), pp. 193–228.
- [11] S. P. Kothari, Andrew J. Leone, and Charles E. Wasley. “Performance Matched Discretionary Accrual Measures”. In: *Journal of Accounting and Economics* 39.1 (2005), pp. 163–197.
- [12] Maureen F. McNichols. “Discussion of The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors”. In: *The Accounting Review* 77.s-1 (2002), pp. 61–69.
- [13] Sugata Roychowdhury. “Earnings Management Through Real Activities Manipulation”. In: *Journal of Accounting and Economics* 42.3 (2006), pp. 335–370.
- [14] Stephen R. Stubben. “Discretionary Revenues as a Measure of Earnings Management”. In: *The Accounting Review* 85.2 (2010), pp. 695–717.