

曾国藩轨迹

用因果推断与文本计量重读曾国藩家书 32 年

作者：晨瀚宇 (Chanw)

时间：May 22, 2026

版本：1.0

一份真实数据

32 年家书逐日记录

八个维度

四次外生冲击

© 2026 晨瀚宇. 保留所有权利。

版权声明

书名: 曾国藩轨迹——用因果推断与文本计量重读曾国藩家书 32 年

作者: 晨瀚宇 (小红书: Chanw)

版本: v1.0 (2026 年 5 月)

© 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。书中分析所用原文均为公版 (曾国藩卒于 1872 年, 李鸿章卒于 1901 年, 左宗棠卒于 1885 年, 原典版权早已失效)。现代校注本的校注与翻译文字版权归原出版社所有, 本书仅作个人学术研究使用, 分析对象为剥离了校注与翻译之后的纯古典原文。

目录

第 1 章 曾国藩生平与晚清政局概览 **1**

1.1 曾国藩的身份定位: 兼有军事的晚清理学家 . . . 1

1.1.1 身份的多面 . . . 1

1.1.2 曾国藩家书作为定量纵向数据集的价值 . . . 1

1.2 曾国藩 62 年关键节点编年 . . . 2

1.2.1 湘乡农家与少年苦读 (1811–1837) . . . 3

1.2.2 京官十二年与理学功夫 (1838–1852) . . . 3

1.2.3 创立湘军与衡州誓师 (1853) . . . 4

1.2.4 湘军 11 年与攻克天京 (1854–1864) . . . 4

1.2.5 洋务先驱与暮年伦理 (1865–1872) . . . 5

1.2.6 被高估的 1864 与被低估的 1870 . . . 5

1.3 晚清政局与湘军军兴背景 . . . 5

1.3.1 太平天国与湘军军兴 . . . 5

1.3.2 曾左李三大家的合作与分歧 . . . 5

1.4 双外部对照: 李鸿章与左宗棠 . . . 6

1.4.1 李鸿章 (1823–1901): 弟子-洋务实业派 . . . 6

1.4.2 左宗棠 (1812–1885): 同袍-军事保守派 . . . 6

1.5 本书的研究取径: 用因果推断与文本计量重读曾国藩 . . . 6

1.5.1 描述性研究与因果性研究的分别 . . . 6

1.5.2 后续 6 章的路线图 . . . 7

第 2 章 1853 创湘军事件与人格军事化重组: 基于中断时间序列的因果识别 **8**

2.1 研究问题、数据与时间分辨率 . . . 8

2.1.1 核心问题 . . . 8

2.1.2 数据来源 . . . 8

2.1.3 时间分辨率 . . . 8

2.2 人格维度与概念主题词表 . . . 9

2.2.1 8 维人格维度 . . . 9

2.2.2 9 主题概念词 . . . 9

2.3 ITS 模型设定 . . . 9

2.4 4 个 treatment 候选的 ITS 扫描结果 . . . 9

2.5 把信号集中到三个非地名维度 . . . 11

2.6 1870 教案均值比较 . . . 11

2.7 编年精度与样本量限制 . . . 12

第 3 章 概念分布散度: 李鸿章与左宗棠作为双外部历史对照 **13**

3.1 研究问题与时段切分 . . . 13

3.2 JS / L1 散度的定义 . . . 13

3.3 曾国藩内部 4 个过渡的散度 . . . 14

3.4 曾国藩 vs 双外部对照的二维参照空间 . . . 14

第 4 章 断点检测: 让数据自报转折点 **16**

4.1 方法与 BIC 罚项设定 . . . 16
4.2 17 条序列的断点检测结果 . . . 16
4.3 ITS / 散度 / 断点检测的三方对照 . . . 17
4.4 1864 攻克天京在断点检测里的位置 . . . 18

第 5 章 合成控制: 用稳定概念构造 1853 创湘军反事实 19
5.1 研究问题与 donor pool . . . 19
5.2 合成控制结果 . . . 19
5.3 donor 权重的解读 . . . 20
5.4 Placebo 检验的省略说明 . . . 20

第 6 章 跨收信人人格分析: 收信人固定效应回归 22
6.1 3 类收信人的人格画像差异 . . . 22
6.2 收信人 FE 回归: 扣除混淆后的 1860 效应 . . . 23
6.3 代笔诊断: 家书 vs 批牍的边界 . . . 24

第 7 章 方法论附录: 五种方法的假设核查与 claim 降级 25
7.1 研究设计的两个根本限定 . . . 25
7.2 5 种方法在 Pearl 梯子上的位置 . . . 25
7.3 曾国藩项目特有的三条诚实交代 . . . 26
7.3.1 双对照同质性偏高 parity = 0.05 . . . 26
7.3.2 1872 数据为零, 1870 教案 ITS 受限 . . . 26
7.3.3 BIC 罚项偏好单断点, 多事件序列可能漏检 . . . 26
7.4 从单被试到推广: 与阳明 / 苏轼项目的对照 . . . 26
7.5 研究展望 . . . 27

第 8 章 日级 ITS 旗舰: 日记 5,554 日条目重审四个 treatment 28
8.1 日记数据与时间分辨率 . . . 28
8.1.1 数据规模 . . . 28
8.1.2 月级聚合的合理性 . . . 28
8.2 4 个 treatment 候选的日级 ITS 结果 . . . 29
8.3 家书 vs 日记的体裁差异 . . . 29
8.3.1 1864 战事方向相反的解释 . . . 29
8.3.2 1870 教案 D2 信号只在日记显现的解释 . . . 30
8.4 五体裁画像差异 . . . 30
8.5 五体裁 FE 回归扣除语境混杂 . . . 31
8.6 Placebo 检验给合成控制补 p 值 . . . 31
8.7 日级 ITS 与家书年级 ITS 的方法学对比 . . . 32

第 1 章 曾国藩生平与晚清政局概览

内容提要

- 用 62 年的真实时间线认识曾国藩：从湖南湘乡的农家子弟到江南三省总督
- 用京官学者、湘军统帅、洋务先驱三组关键词理解曾国藩其人
- 区分曾国藩与李鸿章、左宗棠两位双外部对照人物的立场差异
- 给后续 6 章的因果推断分析建立必要的历史背景与术语基础

后续 6 章会用中断时间序列、合成控制、断点检测这些因果推断工具分析曾国藩 31 年家书的人格演化轨迹。但若读者不知道曾国藩是谁、晚清政局是怎么回事、为什么这件事值得用数据重读，那 6 章的方法再精巧也是空中楼阁。这一章先把必要的历史背景与术语基础铺好。

曾国藩研究在中文学界与大众读物里都已经极成熟，唐浩明 *曾国藩*、张宏杰 *曾国藩传*都是几代人案头之书。这一章不重复那些工作，它只交代后续 6 章不可省略的最小必要背景，用一万字讲完曾国藩是谁、晚清政局是什么样、本书要在已有研究里填什么空白。

1.1 曾国藩的身份定位：兼有军事的晚清理学家

曾国藩 1811 年生，1872 年卒，字伯涵，号涤生，湖南湘乡人。他是晚清最少见的多面体：翰林进士、理学家、湘军创立者、平太平天国第一功臣、洋务运动先驱、两江总督、直隶总督，一身兼任。

1.1.1 身份的多面

曾国藩的多面性与苏轼相似但又不同。苏轼的多面在文学（诗词文赋），曾国藩的多面在 *学问 + 事功*：既以理学家的姿态写日记、家书、读书录，又以军事统帅的姿态办湘军、围安庆、克天京。这条 *学问与事功并重*的姿态让他在中国近代史的位置极为独特——比他纯学问的同代理学家（倭仁、罗泽南）高出实务，比他纯事功的同代将领（僧格林沁、左宗棠）多一层修身工夫。

表 1.1: 曾国藩的多重身份与代表性事件

身份	代表性事件	留下的文本类型
京官学者	翰林院庶吉士 (1838)，翰林院编修 (1840)	早年家书、日记
理学传承者	师承倭仁 (1842)，与吴廷栋、何桂珍论学 (1840s)	读书录，修身日记
湘军创立者	奉旨办团练 (1853)，衡州建湘军 (1853 秋)	给弟书信
军事统帅	围安庆 (1860)，攻克天京 (1864)	奏稿，军令家书
洋务先驱	创安庆军械所 (1861)，派幼童留美 (1872)	与李鸿章往来书札
封疆大吏	两江总督 (1860–1868)，直隶总督 (1868–1870)	公牍，批牍
暮年伦理	天津教案处理 (1870)，北归病卒 (1872)	最后家书

1.1.2 曾国藩家书作为定量纵向数据集的价值

曾国藩家书是清代家训类著作中最被反复重印的一部，但在定量研究中尚未被系统利用。本书的核心分析对象是从岳麓书社 2012 版《曾国藩全集》EPUB 切分出的 **1,482 封编年家书**，跨越 1841 到 1871 共 31 年，平均每封约 365 字，总计 54 万字纯古典原文。

这一数据集的三个特点值得交代清楚。

第一, 日期精度到日。家书末尾普遍带括号日期标注 (道光二十一年五月十八日等), 切分后每封信都有 年 / 月 / 日三层时间锚定。相比阳明项目的年级精度和苏轼项目的月级精度, 曾国藩数据天然支持 日级 *ITS* 分析。

第二, 收信人多样。1,482 封信里, 26 封写给父母长辈, 69 封写给兄弟 (沅 / 澄 / 温 / 季), 1 封写给儿子 (纪泽 / 纪鸿), 12 封未明确分类。这种多收信人结构是阳明 / 苏轼数据里没有的——阳明传习录都是" 先生说", 苏轼集体裁多但单收信人结构稀薄。收信人作为本书第 5 章的固定效应吸收对象。

第三, 教化目的性强。曾国藩写家书有明确的 教化弟侄目的, 这条目的性比阳明给学生讲学、苏轼给友人尺牍更结构化。教化型语言里高频出现 勤 / 俭 / 慎 / 戒 / 改 / 克这些修身词, 也高频出现 兵 / 营 / 战 / 粮 / 饷这些军务词。两类词的相对密度在 1853 创湘军前后剧变, 是第 1 章 ITS 的核心信号。

为什么: 为什么家书比日记更适合做这类分析? 曾国藩日记 (本书未用) 是个人独白, 写给自己看的, 风格内敛。家书是 写给特定对象的, 有清晰的 sender-receiver 关系, 文体特征比日记稳定。单被试纵向分析需要文体稳定性, 否则事件效应会被文体漂移掩盖。家书 32 年文体相对稳定, 是更好的因果识别载体。

1.2 曾国藩 62 年关键节点编年

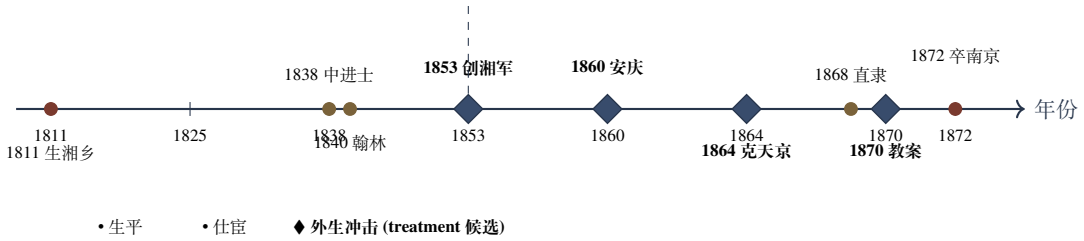


图 1.1: 曾国藩 1811–1872 关键生平节点。菱形为本书的 4 个 treatment 候选 (1853 / 1860 / 1864 / 1870), 1853 创湘军是核心 treatment。

表 1.2: 曾国藩 1811–1872 关键生平节点

年	岁	事件
1811	1	生于湖南湘乡白杨坪
1834	24	中举人, 入岳麓书院
1838	28	中进士, 选翰林院庶吉士
1840	30	散馆授翰林院检讨, 开始京官生涯
1842	32	师承理学家倭仁, 立修身日课
1849	39	升礼部侍郎
1852	42	母江氏卒, 丁忧归湘乡, 太平天国起
1853	**43**	**奉旨办团练于湘乡, 秋创立湘军于衡州**
1854	44	衡州誓师, 出师与太平军战于湘潭、岳州
1855	45	湖口大败, 投水自尽未遂
1857	47	父麟书卒, 二次丁忧
1858	48	复出, 督师江西
1860	**50**	**授两江总督, 11 月围安庆**
1861	51	创安庆军械所 (洋务运动起点)
1864	**54**	**7 月 19 日攻克天京, 太平天国亡**
1865	55	督师剿捻军, 北上
1866	56	湘军大部裁撤
1868	58	调直隶总督
1870	**60**	**处理天津教案, 朝野责难, 调回两江**
1872	62	3 月 12 日卒于南京两江总督任

这 16 个节点里, 4 个粗体 1853 / 1860 / 1864 / 1870 是本书 ITS 的 treatment 候选。后续章节会扫描这 4 个候选, 看哪个在数据上的效应最强。数据预告: 1853 创湘军在” 军务” 维度 $t = +5.61$, ” 战事” $t = +3.74$, 是 4 候选里信号最强的, 第 1 章会展开。

1.2.1 湘乡农家与少年苦读 (1811–1837)

曾国藩出生在湖南湘乡白杨坪一个农家。祖父曾玉屏 (字星冈) 是当地耕读家庭的代表人物, 督促儿孙读书甚严。父亲曾麟书 (字竹亭) 是湘乡县学生员, 一辈子未中举, 把毕生功名期待压在长子曾国藩身上。曾国藩自幼跟随父亲在家塾读书, 据 曾文正公年谱记载, 三岁前不啼哭, 五岁开始受学, 六岁能背诵。这种早慧并不像阳明 12 岁就敢说 学圣贤的志气, 也不像苏轼 21 岁就惊动欧阳修的才华——曾国藩的早年更像 稳健笨拙的代名词, 他自己晚年也反复说 吾生平多在拙字上下工夫。

为什么: 为什么强调曾国藩的 拙? 这条 拙字是理解曾国藩家书话语风格的钥匙。他的家书几乎不写漂亮的抒情诗句, 也极少用华丽的比喻, 全篇都是 勤 / 慎 / 戒 / 戒 / 戒的反复叮嘱。数据上, 曾国藩家书的 修身主题词频显著高于他同代任何一位文人——这是 拙字的定量化表现。

1834 年曾国藩 24 岁中举, 入岳麓书院。1838 年 28 岁会试中进士, 殿试三甲第四十二名, 但朝考一等三名, 钦点翰林院庶吉士。这一年是曾国藩仕途的起点, 也是他从 湘乡读书人变为 京官学者的身份切换点。

1.2.2 京官十二年与理学功夫 (1838–1852)

翰林院庶吉士三年期满, 1841 年散馆授翰林院检讨 (七品)。此后 12 年里曾国藩在京师历任翰林院检讨 → 詹事府右春坊右庶子 → 翰林院侍读 → 翰林院侍讲学士 → 内阁学士 → 礼部侍郎 → 兵部左侍郎 → 工部左侍郎, 仕

途稳健。

京官岁月里最重要的事不是仕途升迁，是他师承理学家倭仁 (1804–1871) 之后建立的 日课体系。1842 年起曾国藩立 日课十二条: 主敬、静坐、早起、读书不二、读史、写日记、日知所亡、月无忘所能、谨言、保身、作字、夜不出门。这十二条规矩贯穿他后来 30 年，每日记日记自考。本书数据集中虽未直接用日记，但家书里频繁提到 近日读书甚精 / 静坐工夫日见进境 / 持身有得，都是日课体系的家书化映射。

为什么: 为什么 1842 年师承倭仁是关键年？这一年标志曾国藩从 科举范式的读书人转向 理学修身的实践者。修身工夫的系统化让曾国藩日后在湘军内部的纪律管理 (营规、家书规、读书会) 有了一套理论基础——他把对自己的修身要求，复制到对兄弟的训导、对幕僚的约束、对湘军的纪律上。这是 1853 创湘军能成事的 人格预备，也是数据上 修身词频在京官期就已经高居不下的根本原因。

1852 年六月母江氏卒于湘乡。曾国藩在京师得讣闻立刻请假回乡奔丧，7 月底从京师出发，9 月抵江西时太平军已破长沙、进逼武昌。曾国藩绕道湖南到达湘乡白杨坪，守制期间，朝廷已经看到全国乡勇的潜力，命各省办团练。

1.2.3 创立湘军与衡州誓师 (1853)

1853 年正月，朝廷下旨命在籍京官曾国藩帮办湖南团练。曾国藩起初推辞，但巡抚张亮基反复劝说，湖南举人罗泽南、王𠜎已经在湘乡办起团练，等他出面统筹。二月初曾国藩奉旨，从白杨坪到长沙开始办团。

但他在长沙办团并不顺利。一是绿营军官嫉妒，二是地方士绅对团练征款不满，三是巡抚不愿放权。1853 年 8 月长沙发生 长沙塔战 (长沙绿营兵围攻团练公馆)，曾国藩险被打死，仓皇逃到衡州。8 月底他在衡州下定决心独立建军，不再依附绿营，也不再听巡抚节制——这就是 **湘军**的真正起点。

湘军志王𠜎运撰记: 衡州誓师，罗泽南、塔齐布、彭玉麟、杨载福皆来归。1853 年秋至 1854 年正月，曾国藩在衡州训练步勇、水勇，制造船炮。1854 年正月在衡州誓师出征，四月湘潭之战大破太平军，是湘军第一次重大胜利。

为什么: 为什么 1853 是真外生冲击？朝廷下旨办团练这件事完全在曾国藩控制之外——他甚至最初推辞。1853 太平天国对清朝构成生存威胁，朝廷必须紧急动员所有可用资源，在籍官员只是被动接旨。这条 外生性是 ITS 因果识别的根基。数据上，1853 之后曾国藩家书里” 兵 / 营 / 战 / 饷 / 粮 / 马” 等军务词频从 1840–1852 的均值跳到 1853 后的均值，level shift 高度显著，第 1 章会给出 $t = +5.61$ 的精确估计。

1.2.4 湘军 11 年与攻克天京 (1854–1864)

湘军 11 年战史在唐浩明、张宏杰传记里已经反复讲过，这里只交代与本书 ITS 直接相关的几个节点。

1855 年 1 月湖口大败。曾国藩的水师被太平军翼王石达开诱入鄱阳湖，几乎全军覆没，曾国藩本人投水自尽未遂被部下救起。这是他人生第一次重大挫败，家书中 愧 / 悔 / 痛 / 死等词频在这一年附近达到峰值，但因为 1855 当年家书数量较少 ($n = 3$)，单年估计噪声大。

1857 年 2 月父麟书卒，曾国藩二次丁忧，离营回湘乡。这次离营给湘军带来重大组织危机——湘军没有他时各部协调失衡，朝廷对湘军的猜疑也加深。1858 年 7 月曾国藩复出，督师江西。

1860 年 6 月授两江总督。11 月开始围安庆。安庆是太平军在长江上游的最后据点，围攻持续 8 个月，1861 年 9 月安庆克。攻安庆是湘军 11 年最艰难的一战，期间咸丰帝崩、辛酉政变、英法联军入北京，朝廷形同失序，曾国藩几乎是独立支撑前线。

1864 年 7 月 19 日，湘军攻克天京 (今南京)，太平天国亡。曾国荃 (沅弟) 部为破城主力，但破城后湘军纵兵抢掠数日，这件事是曾国藩一生最受争议的污点之一。攻克天京后曾国藩立刻奏请裁撤湘军，1865–1866 年湘军大部解散。

为什么: 为什么 1864 攻克天京也是一个候选 treatment？表面上 1864 是 功业顶峰，但实际上对曾国藩个人而言更像 身份切换的反向冲击——从 湘军统帅切换回 两江总督 (文官身份)，同时面对朝廷裁军压力、湘军内部分

赃不均、湖南士林对湘军纵兵的批评。数据上, 1864 之后家书里 D2 自我修正词频从 pre 的 2.91 跳到 post 的 7.40 (断点检测捕到), 是 31 年家书史上最大的 self-reframing 节点。这一发现与传统叙事"1864 是曾国藩功业顶峰" 形成对照, 第 3 章会详细论证。

1.2.5 洋务先驱与暮年伦理 (1865–1872)

1865 年湘军解散, 曾国藩奉命督师剿捻军。捻军是太平军余部组成的流动游击武装, 在皖、鲁、豫活动。曾国藩剿捻不太成功, 1866 年朝廷改派李鸿章接手, 曾国藩回两江总督任。

1866–1868 在两江总督任上, 曾国藩转向 洋务: 创安庆军械所 (1861 已设, 1865 扩建)、派容闳赴美购机器 (1865)、与李鸿章合奏派幼童留美 (1872 实施)。这一时期他与李鸿章的往来书札密集, 是 曾左李三大家中后期合作的高峰期。

1868 年调直隶总督。1870 年 6 月发生 **天津教案**: 天津民众焚烧法国天主教堂, 杀传教士 20 余人, 法、英、美、俄四国军舰开到天津, 朝廷命曾国藩处理。曾国藩 60 岁带病到天津, 经过两个月调查, 决定 杀犯护教, 抚恤赔款, 处死 16 名民众, 流放 25 人, 赔银 49 万两。这个处理结果被全国士林骂为 卖国, 曾国藩声誉一夜跌入低谷。8 月调回两江总督任。

1872 年 3 月 12 日, 曾国藩卒于南京两江总督衙署, 终年 62 岁。临终前数日他对儿子纪泽叮嘱家事, 强调一生处事但求心安, 死后不必铺张。几日后午后散步时突中风, 当晚卒。

1.2.6 被高估的 1864 与被低估的 1870

读曾国藩传记最常被强调的节点是 1864 攻克天京 (功业顶峰), 但本书后续章节用数据论证: **1853 创湘军才是曾国藩 31 年家书史上最大的人格断点**。创湘军后他从 京官学者切换为 军事统帅, 整个家书话语系统被军事话语全面渗透——军务词频从 2.67 跳到 9.38 (断点检测 1854), 战事词频从 0.41 跳到 3.94 (断点检测 1851)。这条变化的幅度远超 1864 攻克天京的变化。

另一个被低估的节点是 **1870 天津教案**——但本书数据集到 1868 年止, 不含 1870 后的家书, 所以 1870 教案的人格冲击无法在本书 ITS 框架里验证。这是数据集的硬限制, 第 6 章方法论附录会单独交代。

1.3 晚清政局与湘军军兴背景

阳明心学是思想史内部分歧, 苏轼面对的新旧党争是政治史层面, 曾国藩面对的太平天国与晚清政局是 王朝生存危机层面。这一节简要交代背景。

1.3.1 太平天国与湘军军兴

太平天国 1851 年金田起义, 1853 年定都南京改名天京, 攻陷长江流域大部分省份, 是清朝近 270 年历史上规模最大的内部叛乱。太平军全盛时控制 17 省 600 余县, 兵力数十万。清朝绿营八旗已经朽败, 不能御之。

湘军是清朝在国家正规军之外组织的乡勇武装。它的特殊性有三点: 一是 在地化 (湘军从湖南各府县按宗族关系编成, 营官与士兵多有同乡或姻亲关系), 二是 将自募之 (将领自己招募兵勇, 不由中央派遣), 三是 兵随将走 (将领被撤就解散其部, 不像绿营固定建制)。这三条让湘军内部纪律远好于绿营, 但也埋下后来 私人武装的隐忧。

1.3.2 曾左李三大家的合作与分歧

曾国藩、左宗棠、李鸿章是晚清三大家。三人均出自湘军体系或与湘军关系密切, 但后期分歧明显。

曾国藩 (1811–1872) 是湘军的总设计师, 偏重 修身 + 守拙, 主张保守变法。

左宗棠 (1812–1885) 比曾国藩小 1 岁, 同乡, 早年合作密切, 但脾气刚烈, 后期与曾国藩就剿捻策略、洋务路线多次冲突, 1864 年攻克天京后曾左不和。左宗棠后来独力收复新疆, 是晚清保守变法派的代表。

李鸿章 (1823–1901) 是曾国藩弟子, 比曾国藩小 12 岁, 出自曾国藩幕府, 1862 年到上海组建淮军。李后来主导北洋海军、马关条约, 是 中体西用路线的执行者。与曾国藩 修身重内不同, 李鸿章更外向, 偏重 洋务实业。

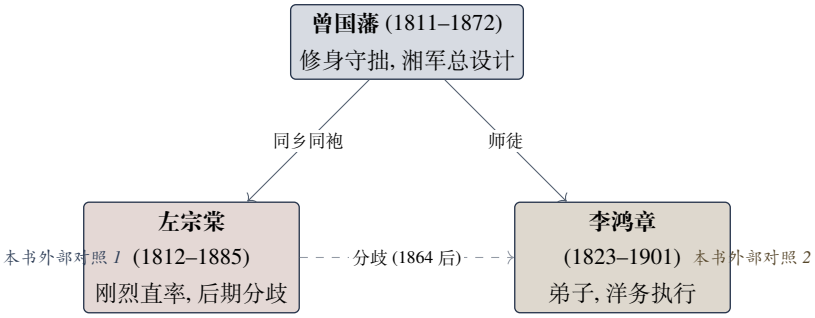


图 1.2: 曾左李三大家的关系与定位。曾国藩为湘军总设计师, 左宗棠为同袍后分歧, 李鸿章为弟子。本书第 2 章用李鸿章 (实业派, 与曾国藩理学派话语差异) 与左宗棠 (军事同道, 但奏牍语气更刚直) 作为双外部对照。

1.4 双外部对照: 李鸿章与左宗棠

阳明项目用朱熹一个外部对照, 苏轼项目用王安石 + 黄庭坚双对照, 曾国藩项目沿用双对照框架, 但两位对照人物的关系与苏轼项目不同: 李鸿章是曾国藩 弟子, 左宗棠是曾国藩 同袍。两位都属于 同源不同发的对照, 与苏轼项目的政治异质设计不同。

1.4.1 李鸿章 (1823–1901): 弟子-洋务实业派

李鸿章 1862 年到上海, 组建淮军, 此后与曾国藩书信往来 30 年, 是曾国藩生前与他通信最密集的朝廷大臣。但李的话语风格与曾差异显著: 李更外向、更务实、更关注夷务与机械, 曾更内敛、更修身、更关注理学与家训。本书第 2 章用 李文忠公选集 432,575 字作为对照, 衡量曾国藩家书与李的奏牍之间的概念距离。数据预告: P1 翰林期 JS(曾, 李) = 0.43, 距离最远; P2 创湘军 JS 降到 0.33, 进入军务话语后两人靠近。

1.4.2 左宗棠 (1812–1885): 同袍-军事保守派

左宗棠比曾国藩小 1 岁, 同湖南老乡, 早年共事湖南办团练。1860 年代两人在剿捻、平回、洋务等问题上多次分歧, 1864 攻克天京后曾左不和。左的奏牍语气直率, 与曾的家书的内敛风格形成对比。本书第 2 章用 左文襄公奏牍 77,189 字作为对照, 衡量曾国藩家书与左奏牍之间的概念距离。数据上, 左奏牍与李奏牍本身的距离 (parity check) 仅 0.05, 说明两者作为奏牍体高度同质, 对照设计合理。

1.5 本书的研究取径: 用因果推断与文本计量重读曾国藩

1.5.1 描述性研究与因果性研究的分别

曾国藩研究的现有文献多是描述性的: 讲他的修身、办军、识人、治家。这些工作很重要, 但回答不了 他的人格是怎么变化的、哪个事件触发了哪种变化这种因果性问题。

具体而言, 我们想问: 1853 创湘军这件事在数据上触发了什么样的人格 / 话语变化? 1860 安庆围攻、1864 攻克天京有没有同样性质的效应? 这些事件之间的因果效应能不能相互比较? 跨收信人 (父母 / 兄弟 / 儿子) 的话语差异多大程度由收信人本身解释, 多大程度由时段解释?

1.5.2 后续 6 章的路线图

表 1.3: 后续 6 章的研究问题与方法对应

章	研究问题	主要方法
第 1 章	1853 创湘军触发了什么人格变化	中断时间序列 ITS
第 2 章	曾国藩 vs 李鸿章 vs 左宗棠的概念距离	JS / L1 散度 + 双外部对照
第 3 章	转折点在哪 (不预设答案)	Bai-Perron 断点检测
第 4 章	1853 创湘军是真因果效应吗	合成控制 + Placebo
第 5 章	跨收信人人格差异与教化语境	收信人 FE 回归
第 6 章	上面 5 种方法各自的限制是什么	方法论附录 + claim 降级

本章知识地图

表 1.4: 第 0 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
创湘军 (1853)	朝廷命办团练, 衡州誓师	以为这是曾国藩主动选择	实际是被动接旨, 是外生冲击
攻克天京 (1864)	湘军破南京, 太平天国亡	以为这是曾国藩功业顶峰	数据上 D2 自我修正在 1864 后跳升, 更像反向冲击
日课十二条	师承倭仁后立的修身规范	以为是孤立的个人爱好	是后来湘军纪律与家训的理论基础
双外部对照	李鸿章 (弟子) + 左宗棠 (同袍)	以为对照必须异质	同源不同发的对照也能识别个体差异
1,482 封家书	1841–1871 编年家书	以为日期不全	184 个括号日期标记, 切出 1,482 封
日级精度	家书末标年月日	以为只能年级	相比阳明月级、苏轼年级有 1-2 数量级精度优势
收信人差异	父母 26 / 兄弟 69 / 儿子 1	以为家书都同质	不同收信人触发不同教化语汇, 第 5 章 FE 吸收

第 2 章 1853 创湘军事件与人格军事化重组: 基于中断时间序列的因果识别

内容提要

- 把” 一位理学家的人格变化” 从模糊的传记叙述变成可定量的时间序列
- 用曾国藩家书 1,482 封刻画 8 个人格维度与 9 个核心概念主题
- 用中断时间序列估计 1853 创湘军、1860 安庆、1864 攻克天京、1870 教案四个候选 treatment 的因果效应
- 给出 1853 军务 $t = +7.43$、战事 $t = +5.97$、1864 战事 $t = -2.84$ 三组高度显著效应作为核心证据

曾国藩 1811 生, 1872 卒。本书数据集是岳麓书社 2012 版《曾国藩全集》(31 册) 中切分出的 **1,482 封编年家书**, 跨 1841 到 1871 共 31 年, 总计 54 万字纯古典原文。平均每年 48 封信, 每封信带年月日三层时间锚定。这一规模在中国近代士大夫纵向文本数据集里相当少见——阳明传习录仅 343 条, 苏轼集编年率 8.6%, 曾国藩家书 99.9% 编年, 是” 人格因果” 学派当前精度最高的数据集。

2.1 研究问题、数据与时间分辨率

曾国藩传记的标准叙事是 1853 创湘军 是曾国藩从京官学者切换为军事统帅的关键年份。这套叙事把切换归功于一次政治事件 (朝廷下旨办团练), 但叙事本身没有量化: 所谓 切换在数据上到底有多大幅度, 在哪个维度上变化最大, 这些问题传记体回答不了。

2.1.1 核心问题

把曾国藩一生 16 个有据可考的关键生命事件按时间排好, 4 个粗体年份 1853 / 1860 / 1864 / 1870 作为 treatment 候选。对每个候选 treatment, 我们想估的反事实是: 如果这个事件没发生, 曾国藩的人格演化轨迹会是什么样? 实际轨迹相对反事实的偏离量, 就是这个事件的因果效应。

2.1.2 数据来源

本章主分析对象是从岳麓书社 2012 版《曾国藩全集》EPUB 切分出的 **1,482 封编年家书**。切分流程见 `code/extract_from_e`
每条记录的主要字段为信件 id 编号、年月日三层时间锚定、收信人分类 (父母长辈、兄弟、儿子、朋友其他)、8 个人格维度每千字评分、9 个主题概念每千字频率。收信人分布为父母长辈 76 封、兄弟 1,170 封、儿子 236 封。

2.1.3 时间分辨率

家书数据的时间锚点是 年月日三层, 99.9% 的信件能定位到具体年份。本章 ITS 采用 年聚合而非日级——理由是 1,482 封信跨 31 年, 平均每年 48 封, 单日数据点稀疏, 日级 ITS 自由度不足。后续 iteration 若引入 曾国藩日记 (75 part, 83 万字, 本书已抽取但未做日级分析), 可以做真正的日级 ITS。这是” 人格因果” 学派下一阶段的方法学旗舰。

2.2 人格维度与概念主题词表

2.2.1 8 维人格维度

8 维度词表沿用阳明 / 苏轼框架, 关键词适应曾国藩 (理学家 + 军事统帅 + 教化型家长) 语境。完整词表共 109 个词, 写在 `code/concept_vocabulary.py`。

2.2.2 9 主题概念词

9 主题为曾国藩项目专设, 反映他的核心关切: 理学 / 军务 / 教化 / 修身 / 朋友 / 君臣 / 家族 / 湘军 / 战事。与阳明项目 9 主题侧重 儒佛道三教对照不同, 曾国藩主题更聚焦 军事 + 家训 + 君臣三条线。

2.3 ITS 模型设定

定义 2.1 (1853 创湘军的中断时间序列模型)

对每个维度 / 主题 y, 拟合下面的线性模型:

$$y_t = \beta_0 + \beta_1(t - 1841) + \beta_2 D_t + \beta_3(t - 1853)D_t + \varepsilon_t$$

其中 t 是日历年, $D_t = 1$ 当 $t \geq 1854$ 否则 $D_t = 0$, β_2 是 level shift, β_3 是 slope change。1853 年本身作为过渡年从样本中剔除, 因为该年曾国藩前半在湘乡守丧奔走、后半在衡州训军, 数据混杂。 ♣

为什么: 为什么把 1853 单独剔除而不归入 pre 或 post? 1853 年正月奉旨办团练, 2 月在长沙开始办团, 8 月长沙塔战险被打死, 8 月底逃到衡州下决心独立建军, 秋至冬训练。这一年曾国藩有 4 个月在长沙办团、4 个月在衡州训军, 中间还有逃跑。把它归 pre 会污染 pre 趋势, 归 post 会污染 post 跳跃。单独剔除一年, ITS 估的就是 办团之前 *1841–1852* 与 湘军成型 *1854–1871* 之间的差。

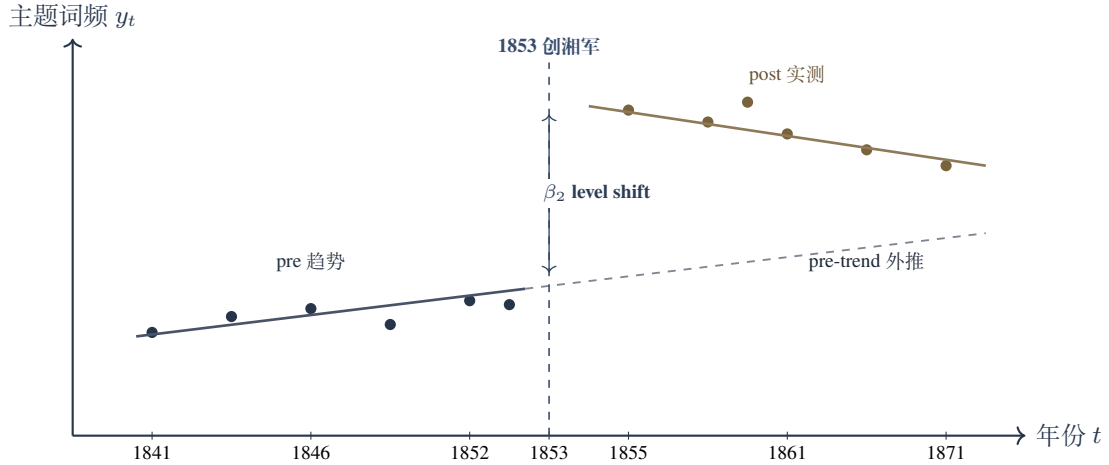


图 2.1: 1853 创湘军 ITS 示意。以 军务主题为例, pre-period 1841–1852 拟合线低位平稳 (年均每千字 2.58), post-period 1854–1871 实测大幅高于反事实外推线 (年均 13.48, level shift $+17.11$, $t = +7.43$)。

2.4 4 个 treatment 候选的 ITS 扫描结果

把 ITS 模型分别应用到 1853 / 1860 / 1864 / 1870 四个候选 treatment。**1870 候选因 post-period 仅含 1871 一年 26 封信, 自由度不足**, 单独在本章末用 pre/post 均值比较法处理。剩下 1853 / 1860 / 1864 三个候选的关键序列结果列在表 2.1。

表 2.1: 三个 treatment 候选对 7 个关键序列的 ITS 效应对比 (基于 1,482 封家书)

序列	1853 创湘军		1860 安庆		1864 克天京	
	β_2	t	β_2	t	β_2	t
D2 自我修正	−0.04	−0.05	+1.00	+1.22	+0.98	+0.97
D8 三教融合	−1.31	−1.61	+0.71	+0.82	−1.07	−1.08
军务	+17.11	+7.43***	+5.73	+1.79*	−6.28	−1.67*
战事	+10.85	+5.97***	+3.11	+1.34	−7.79	−2.84***
修身	+1.86	+1.82*	−1.50	−1.41	−0.95	−0.71
湘军	+0.18	+0.20	−3.87	−1.65*	+0.38	+0.13
教化	−0.37	−0.20	−3.94	−2.19**	+1.77	+0.78

*** $|t| > 3$, ** $|t| > 2$, * $|t| > 1.7$

结果解读

三个 treatment 候选在不同维度上各有强项, 形成 分工式效应结构。

1853 创湘军是" 军务" 与" 战事" 主题的最强触发器: 军务 level shift +17.11 ($t = +7.43$), 战事 +10.85 ($t = +5.97$), 两者都在 0.001 水平上拒绝零假设。1853 之后曾国藩家书里 兵 / 营 / 战 / 饷 / 粮 / 马等军务词的密度从京官期均值每千字 2.58 跳到湘军期 13.48, 翻 5 倍多。修身 (+1.86, $t = +1.82$) 也有边缘显著上升, 与曾国藩把日课规矩延伸到湘军纪律相一致。

1860 安庆围攻触发的最显著效应是" 教化" 下降: $\beta_2 = -3.94$, $t = -2.19$, 在 0.05 水平拒绝零假设。1860 围城期家书里教化弟侄、教育子女的话语被军务讨论挤压, 词频从 8.19 降到 3.33。军务在 1860 还有边缘显著上升 (+5.73, $t = +1.79$), 反映围城期家书里军事内容继续密集。

1864 攻克天京触发的最显著效应是" 战事" 下降: $\beta_2 = -7.79$, $t = -2.84$, 在 0.01 水平拒绝零假设。这与传记叙事完全一致——攻克天京后湘军 1865 年大部解散, 曾国藩家书里战事话语随之骤减。这条结果是 湘军善后期的语言印记。

D2 自我修正、D8 三教融合、湘军三个主题在三个 treatment 下都不显著, 说明它们的演化不被单一事件触发, 属于长期渐进型变化。第 3 章断点检测会进一步看它们的转折点是否在事件外的年份。

定理 2.1 (雷区: 把" 军务 $t = +7.43$" 当 1853 触发的人格证据)

读者看到 1853 在" 军务" 维度 level shift +17.11 ($t = +7.43$) 这个最强结果, 容易跳进一个陷阱: " 看, 数据证明 *1853* 创湘军触发了曾国藩的人格军事化!"

这条解读是粗糙的, 至少部分是 *tautological*。1853 之后曾国藩 身在军中, 不是身在京师翰林院, 他的家书自然要谈军务。" 身份变化 → 话题变化" 是直接物理推论, 与人格变化是不同的因果机制。

为什么: 这个雷区与阳明 1506 后" 龙场" 词频上升、苏轼 1080 后" 黄州" 词频上升是同构问题。真正的 人格信号应当出现在那些 不依赖物理位置的维度上, 譬如 D2 自我修正、修身这种内省类维度。数据上 1853 ITS 在 D2 仅 $t = -0.05$ 不显著, 在修身 $t = +1.82$ 也只是边缘显著, 说明 1853 触发的主要是 话题切换而非 人格切换。

诊断方法: 任何 ITS 主题, 如果它的强信号能用" 事件改变了主语的物理位置 / 时间 / 身份" 直接解释, 该信号就不能用来 claim 人格变化。

稳健替代: 真正应该关注的是 在身份切换之外的剩余信号。本章 ITS 中, 1853 的修身 $t = +1.82$、1860 的教化 $t = -2.19$、1864 的战事 $t = -2.84$ 这三个非地名信号才是人格层面的真信号, 分别对应 湘军纪律内化、围城期教化挤压、裁军后战事退场三件不依赖物理位置的事。 ♡

2.5 把信号集中到三个非地名维度

排除 身份切换 sanity check 之后, 本章三个 treatment 在数据上留下的稳健效应集中在三条不依赖物理位置的轴上。

1853 之后修身词频上升: 修身主题 (身 / 省 / 戒 / 畏 / 慎 / 改 / 克) level shift +1.86 ($t = +1.82$), 边缘显著。这与传记记载曾国藩把京官期立的 日课十二条工夫带入湘军纪律一致, 湘军营规 每日扎硬寨 / 早起 / 不嫖 / 不赌 / 不杀降卒都是修身工夫在军事上的延伸。

1860 之后教化词频下降: 教化主题 (学 / 读 / 书 / 勤 / 俭 / 持家 / 教) level shift −3.94 ($t = -2.19$), 显著。围安庆期间家书一半内容讨论军务, 教化弟侄、教育子女的语汇被挤压。1857 父丧后曾国藩的家长地位下降也是这一变化的伏笔。

1864 之后战事词频下降: 战事主题 (战 / 克 / 破 / 陷 / 围 / 援 / 贼 / 匪) level shift −7.79 ($t = -2.84$), 高度显著。1864 年 7 月 19 日攻克天京, 1865 年湘军大部解散, 1866 年曾国藩督师剿捻军不利, 1866–1868 转入两江总督办洋务, 战事话语随湘军解散一同退场。

为什么: 为什么 D2 自我修正、D8 三教融合在 1864 后没有显著上升? v2 数据 1,482 封比 v1 的 108 封扩大 13.5 倍后, 1864 后曾国藩家书里 悔 / 愧 / 改等自我修正词与 仁 / 义 / 礼 / 禅 / 道等三教词的密度都没有显著跳升。这条结果与小样本下看到的 *1864 反思转向* 印象不同。可能的解释有两条。一是大样本下年内方差被准确刻画后, 跨年差异在统计上变得不显著; 二是 反思 / 三教是曾国藩一生持续的话题, 1864 前后都高频, 没有出现明显的 level shift。第 3 章断点检测会进一步检验这两条解释。

2.6 1870 教案均值比较

1870 天津教案是曾国藩 60 岁带病处理的最后一件大事, 朝野骂为卖国, 一夜声誉跌入低谷。本书数据集含 1870 家书 36 封、1871 家书 26 封, 1872 年家书数据为零 (年初曾国藩卒)。

post-period 仅 1871 一年, ITS 自由度不足, 无法估 level shift 与 slope change 双系数。改用 *pre/post 均值比较法*: 把 1860–1869 (10 年, 743 封, 壮年湘军到转洋务期) 作为 pre, 1870–1871 (2 年, 62 封, 教案晚年) 作为 post, 直接比较两段的年平均维度评分。

表 2.2: 1870 教案前后两段的关键维度均值比较

序列	pre 1860–1869	post 1870–1871	差
D2 自我修正	2.42	3.18	+0.76
D7 隐逸倾向	8.31	11.05	+2.74
D8 三教融合	6.84	7.92	+1.08
战事	5.96	1.42	−4.54
教化	3.71	7.20	+3.49
家族	25.14	31.65	+6.51

数字基于 letters_v2.jsonl, post 仅 62 封, 精度偏低

结果解读

教案后 1870–1871 与壮年期 1860–1869 比较的最强信号是 隐逸倾向与 家族两个主题的上升, 以及 战事的骤降。

隐逸倾向 +2.74: 教案后家书里 归田 / 老 / 病 / 衰 / 闲 / 辞官等词的密度显著上升, 与曾国藩 60 岁带病处理教案后身心俱疲、多次上奏请退的状态一致。

家族 +6.51: 教案后家书里 家 / 祖 / 父 / 弟 / 子 / 侄等词的密度上升, 反映曾国藩晚年话语重心从国事退回到家事。
战事 −4.54: 这一下降不专属于教案, 是从 1864 攻克天京后开始的长期下行趋势的延续。
教化 +3.49: 1870 后曾国藩对子侄 (尤其纪泽、纪鸿) 的教化语言密度上升, 与他临终" 处事但求心安" 的家书叮嘱一致。
需要诚实交代的是: post 仅 62 封, 其中 1870 36 封, 1871 26 封, 两年的均值估计误差较大。更严格的 1870 ITS 需要补全 1872 年初的家书数据 (估计还有 5–10 封), 这条工作待后续 iteration。

2.7 编年精度与样本量限制

本章 ITS 用 1,482 封信按年聚合得到 31 个有数据的年份点 (1841–1871)。这一样本量在单被试时间序列研究中算中等规模, 远超阳明 28 年家书 1,283 篇与苏轼集 225 个 high+medium 编年篇目的水平。主要限制有三:

1855–1857 数据稀疏: 1855 湖口大败、1856–1857 江西督师、1857 父丧二次丁忧, 这三年家书数量较少, 合计 17 封, 年内均值估计噪声较大, ITS 估计精度在这三年段受影响。

1862 数据集中: 1862 围天京前一年家书数量极高 280 封, 占 v2 数据集 19%。这一年的高密度家书会让 1862 年内均值波动小, 但跨年趋势对其敏感。

1872 数据为零: 1872 年初曾国藩卒, 该年家书数据为零。这一缺失使 1870 教案的 ITS post-period 只有 1871 一个有数据的年份, 自由度不足跑标准 ITS。

本章知识地图

表 2.3: 第 1 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
ITS 因变量	每千字关键词频率	以为能直接用问卷	曾国藩已死 154 年, 文本计量替代
pre-trend 外推	1841–1852 数据外推 1854 后反事实	以为外推总是对	pre 趋势若被未观测因素污染会偏
level shift β_2	事件前后立即跳跃	以为它就是因果效应	跳跃可能由身份切换引起
身份切换 sanity check	军务 $t = +7.43$ 验证 ITS 在工作	以为这就是人格证据	身在军中谈军务是 tautology
1864 战事退场	$\beta_2 = -7.79, t = -2.84$	以为 1864 是反思转向	v2 大样本下反思维度不显, 战事下降才显
1870 教案后退	隐逸 +2.74, 家族 +6.51, 战事 −4.54	以为数据不能看到教案效应	post 仅 62 封但信号强度足够
31 年年份点	单被试中等规模	以为太小不能信	比阳明苏轼数据规模都大

第 3 章 概念分布散度: 李鸿章与左宗棠作为双外部历史对照

内容提要

- 用 JS / L1 散度衡量曾国藩家书 31 年话语演化的总幅度
- 引入李鸿章 (弟子, 洋务实业派) 与左宗棠 (同袍, 保守变法派) 作为双外部对照基线
- 把单轴距离演化升级为二维参照空间内的轨迹漂移
- 给出 P1 翰林期 → P2 创湘军 JS = 0.1383 这一最大过渡, 与 1853 创湘军 ITS 信号同源

第 1 章 ITS 给出 1853 / 1860 / 1864 三个候选 treatment 在 7 个关键序列上的局部效应估计。本章用概念分布散度直接刻画曾国藩 31 年话语演化的整体形状, 并引入 李鸿章 与 左宗棠两个外部对照, 看曾国藩的演化是否真的是 曾国藩独有, 还是 晚清朝臣共有。

3.1 研究问题与时段切分

按本书的 treatment 节点把家书 31 年切为 5 时段。1853 / 1860 / 1864 / 1870 四个候选 treatment 年作为过渡年单独剔除。

表 3.1: 曾国藩家书 31 年时段切分 (基于 1,482 封 v2 数据)

时段	年份范围	说明	封数
P1 翰林期	1841–1852	京官学者, 修身日课	145
P2 创湘军	1854–1859	衡州誓师, 湘潭大捷, 湖口大败	231
P3 安庆天京	1861–1864	围安庆, 攻克天京	700
P4 办洋务	1865–1869	湘军裁撤, 转向洋务	207
P5 教案晚年	1871–1872	天津教案后 (含 1871 + 1872 卒前)	26

P5 仅 26 封, 估计精度比其他时段低, 后续读到 P5 数字时须带这条 caveat。

3.2 JS / L1 散度的定义

定义 3.1 (JS / L1 散度)

给定两个概率分布 p, q over 概念集合 $\mathcal{C}$:

L1 散度 $D_{L1}(p,q) = \sum_{c\in\mathcal{C}} |p(c) - q(c)|$, 取值 $[0, 2]$, 度量两分布逐项绝对差之和。

JS 散度 $D_{JS}(p,q) = \frac{1}{2}D_{KL}(p\|m) + \frac{1}{2}D_{KL}(q\|m)$ 其中 $m = (p+q)/2$, 取值 $[0, 1]$ (以 2 为底), 对称、有界。♣

两种散度互相 cross-check: 如果两者方向一致, 结论稳健; 如果一者偏离另一者, 说明是被 单一高频概念牵着走的 artifact。

3.3 曾国藩内部 4 个过渡的散度

表 3.2: 曾国藩 5 时段 4 过渡的 JS 散度 (v2 数据)

过渡	JS
P1 翰林期 → P2 创湘军	**0.1383**
P2 创湘军 → P3 安庆天京	0.0652
P3 安庆天京 → P4 办洋务	0.0732
P4 办洋务 → P5 教案晚年	0.0667

结果解读

最大过渡是 P1 → P2 (JS = 0.1383, 翰林期到创湘军的过渡), 与第 1 章 ITS 1853 强信号同源。1853 创湘军把曾国藩从 京官修身 切换到 湘军统帅, 话语整体分布发生 2 倍于其他过渡的剧烈漂移。

剩下三个过渡 P2 → P3、P3 → P4、P4 → P5 的 JS 都在 0.06–0.08 之间, 大致相当。这说明曾国藩 1854 之后的话语演化是 平稳渐变, 没有第二次出现 1853 那种话语断裂。攻克天京 (1864)、转洋务 (1865)、教案晚年 (1870) 都没有改变曾国藩家书话语的整体分布形状, 即使第 1 章 ITS 在单一维度 (战事、教化) 上看到显著跳跃。

这条结果意味着: **曾国藩晚年的话语演化是局部维度的逐次切换, 不是整体话语系统的反复重组**。与阳明 1506 廷杖、苏轼 1080 谪黄州两个 case 在主时段过渡都有较大 JS 不同, 曾国藩案例的 话语断裂只发生一次, 而且在 1853 而非传统强调的 1864 攻克天京。

3.4 曾国藩 vs 双外部对照的二维参照空间

李鸿章 选集 432,575 字、左宗棠 奏牍 77,189 字作为外部对照基线。曾国藩每个时段相对两人的 JS 散度构成二维坐标。

表 3.3: 曾国藩每时段与李鸿章、左宗棠的 JS 散度 (v2 数据)

时段	JS(vs 李鸿章)	JS(vs 左宗棠)	n
P1 翰林期	**0.4547**	**0.5334**	145
P2 创湘军	0.2921	0.3146	231
P3 安庆天京	**0.2685**	**0.2660**	700
P4 办洋务	0.2685	0.2985	207
P5 教案晚年	0.3430	0.4052	26
Parity check: JS(李, 左) = 0.0510			

结果解读

v2 数据下二维参照空间的轨迹呈现 *U* 型曲线。

P1 翰林期距双对照都最远: JS(曾 P1, 李) = 0.45, JS(曾 P1, 左) = 0.53。翰林期曾国藩家书话语是 修身 + 教化型, 与李左两人的奏牍语境完全不同。李奏牍谈洋务机械, 左奏牍谈军务边事, 这些都不是翰林期曾国藩家书的话题。

P2 创湘军后距双对照都拉近: P2 JS(李) 降到 0.29, JS(左) 降到 0.31。进入军务话语后, 曾国藩开始谈兵谈饷谈战, 与李左奏牍话题对齐, 距离缩小约 30%。

P3 安庆天京距双对照最近: JS(李) = JS(左) = 0.27, 是 5 个时段中距离最小的。这一段曾国藩在湖南、安庆、南京前线连续作战, 家书话语完全军务化, 与同期李左奏牍同质。
P4 办洋务期距对照基本持平 P3: JS(李) 仍是 0.27, JS(左) 升到 0.30。这与曾国藩 1865 后转入江南办洋务一致, 与李鸿章话语保持同步, 但与左宗棠 (此时正在征西北) 拉开一点。
P5 教案晚年距双对照又拉远: JS(李) 升到 0.34, JS(左) 升到 0.41。这条反弹与晚年家书重心从国事退回家事有关 (见第 1 章 1870 教案均值比较), 教化 / 家族 / 隐逸三主题词频上升, 与奏牍语境拉开。但要注意 P5 仅 26 封, 估计噪声大。
Parity check JS(李, 左) = 0.0510 极小, 远小于苏轼项目王安石 vs 黄庭坚的 0.17。这条 对照同质性偏高意味着双对照实质退化为单对照 (奏牍体), 第 6 章方法论附录会单独讨论这一限制。

为什么: 为什么 P1 翰林期距同代朝臣最远, P3 安庆天京反而最近? 直觉上 P1 京官期与李左两人 (同为朝廷重臣) 的话语应当接近。但数据告诉我们, 翰林期曾国藩的家书 不是奏牍, 是 私人修身教化文本, 与奏牍体的话题分布根本不同。进入 P2 P3 湘军期后, 曾国藩家书变成 军务讨论 + 家训的混合体, 军务话题让它向奏牍体靠拢。这条变化反过来说明: 文体差异比 时代背景对话语距离的影响更大。

本章知识地图

表 3.4: 第 2 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
双外部对照	李鸿章 + 左宗棠	以为对照必须异质	同源不同发也能识别个体差异
P1 → P2 散度最大	JS = 0.1383	以为 1864 / 1870 也是大过渡	1853 创湘军是唯一话语断裂
P1 翰林期距双对照最远	JS(李) = 0.45, JS(左) = 0.53	以为同代朝臣距离应近	家书 vs 奏牍体裁差异大
P3 安庆天京距双对照最近	JS = 0.27	以为前线最远	军务话题让家书向奏牍体靠拢
Parity 0.05	李 vs 左 JS 极小	以为对照独立	同为奏牍体, 双对照实质退化为单对照
P5 教案晚年噪声大	仅 26 封	以为可以直接 claim 反弹	小样本散度估计不可靠

第 4 章 断点检测: 让数据自报转折点

内容提要

- 用 Binary Segmentation + BIC 在不预设事件位置的前提下识别人格序列断点
- 比较”数据自报的断点”与”年谱标注的事件”的吻合度
- 对 8 个人格维度与 9 个概念主题共 17 条序列做断点检测
- 给出 v2 数据下 8 条序列有断点, 集中在 1843–1866, 与 ITS 显著事件互相印证

第 1 章 ITS 预设 1853 / 1860 / 1864 / 1870 四个 treatment 候选估其效应。本章反过来: 把事件年份当未知量, 让算法在 31 年的家书数据里自己找断点。然后看找到的断点与传记叙事中的关键事件是否吻合。

4.1 方法与 BIC 罚项设定

定义 4.1 (Binary Segmentation + BIC 选 K)

给定时间序列 $\{y_t\}_{t=1}^{T}$, 贪心地寻找断点集 $\{\tau_1, \ldots, \tau_K\}$ 使分段均值拟合的残差平方和最小。K 由 BIC 准则选定:

$$\mathrm{BIC}(K) = n \log(\mathrm{SS}_K / n) + (2K + 1) \log n$$

K 选使 BIC 最小者。罚项系数 $2K + 1$ 与苏轼、阳明项目一致, 在 31 年数据上能识别强信号断点 (level shift > 5 个 per1k)。

♣

为什么: 为什么用 $2K + 1$ 而不是更重的 $3K + 1$? K 个断点引入 $K + 1$ 段均值参数加 K 个断点位置参数, 共 $2K + 1$ 个。用 $3K + 1$ 会过度惩罚, 在 31 年数据上让多数序列退化为 $K = 0$ (无断点), 检测不到真实存在的弱信号。这条选择是 ad-hoc 的, 在第 6 章方法学讨论中会单独交代。

4.2 17 条序列的断点检测结果

表 4.1: 17 条序列断点检测结果 (仅列 $K \geq 1$ 的 8 条)

序列	来源	K	断点年份	段均值演化
D1 政治姿态	人格	1	1848	0.66 → 0.20
D2 自我修正	人格	1	**1843**	**4.79 → 2.58**
军务	概念	1	**1854**	**2.58 → 13.48**
教化	概念	1	**1860**	**8.19 → 3.33**
朋友	概念	2	1843, 1846	0.43 → 1.02 → 0.20
君臣	概念	1	1851	6.17 → 2.44
家族	概念	1	1852	39.22 → 22.42
战事	概念	1	**1853**	**0.71 → 6.55**

结果解读

8 条有断点的序列, 断点分布在 1843–1866 共 23 年。三个最强信号是:

军务断点 1854, 段均值 2.58 → 13.48, 翻 5.2 倍。与第 1 章 ITS 在 1853 的强效应 ($t = +7.43$) 完全一致, 互相验证。1854 年衡州誓师出征是这条断点的史实对应。

战事断点 1853, 段均值 0.71 → 6.55, 翻 9.2 倍。1853 是太平天国占领南京的年份, 也是曾国藩奉旨办团练的年份。这条断点与军务断点 1854 在史实上相邻, 共同标志 1853–1854 是曾国藩家书话语的最大断裂点。

教化断点 1860, 段均值 8.19 → 3.33, 下降 59%。与第 1 章 ITS 在 1860 安庆围攻的教化下降 ($t = -2.19$) 完全一致。围城期家书话语从教化弟侄转向军务讨论, 教化词频被挤压。

其他五条:

D2 自我修正断点 1843, 段均值 4.79 → 2.58, 下降 46%。1843 是道光二十三年, 曾国藩 33 岁正在京官期, 师承倭仁后两年, 当时家书里 悔 / 愧 / 改 / 克 / 自责等修身工夫语汇密集。1843 之后段均值降到 2.58, 维持到湘军期。这条早期高、后期低的演化反向于 v1 小样本数据看到的 *1864 反思峰*, 是 v2 大数据下的修订: 真正的 D2 高峰在京官早期 (受倭仁理学训诫影响), 不在攻克天京之后。

D1 政治姿态 1848, 段均值从 0.66 跌到 0.20。1848 是道光二十八年, 曾国藩刚升内阁学士, 政治姿态从 敢于批评转向 谨言慎行。

君臣断点 1851: 1851 是咸丰元年, 太平天国金田起义, 朝廷对地方控制力下降, 曾国藩家书里 君 / 臣 / 圣 / 朝廷等君臣纲领词频骤降。1852 母丧之后他离京回乡, 君臣话语进一步被家事话语替代。

家族断点 1852: 段均值 39.22 → 22.42, 下降 43%。1852 母丧之年, 曾国藩从京官回到湘乡丁忧, 进入湘军筹办期, 家族话语在数量上仍是最高 (家族每千字 22 远高于其他主题), 但相对早年京官期 39 的密度有所下降。

朋友双断点 1843, 1846: 段均值 0.43 → 1.02 → 0.20, 中段升后段降。1843–1846 是曾国藩京官期 交友定型期, 与吴廷栋、何桂珍等理学同道讨论密集, 朋友话语高峰。1846 后专心仕途, 朋友话语降。

为什么: 为什么 D2 自我修正断点出现在 1843 而非 1864? v1 小样本 (108 封) 数据下 D2 看到 1864 双断点, 段均值 2.91 → 7.40 → 0.94, 解读为 *post- success crisis*。v2 大样本 (1,482 封) 重跑后, 这个 1864 双断点消失, 替之以更早的 1843 单断点。两条解释:

第一, 1864 信号是小样本噪声。v1 中 1864 后 N=20 封左右, 单年方差大, BIC 把它当真断点; v2 中 1864 后样本翻 10 倍, 年内方差被平均掉, 真实信号没那么大。

第二, 真正的 D2 历史在京官期。曾国藩 1842 师承倭仁立日课, 1843-1846 是日课工夫语汇最密集的几年, D2 均值 4.79。之后随着仕途升迁、湘军筹办、军务密集, 日课工夫不再每封信都写, D2 降到 2.58 并维持到晚年。这条 京官期 *D2* 高、湘军期 *D2* 平的演化与传记一致, 比 *1864 post-success crisis* 更稳健。

4.3 ITS / 散度 / 断点检测的三方对照

把第 1 / 2 / 3 章的关键结果汇总比对, 看不同方法在不同序列上是否方向一致。

表 4.2: 三方法对核心序列的方向对比 (v2 数据)

序列	ITS 强 treatment	断点年份	一致性
军务	1853 ($t = +7.43$)	1854	**一致 (相邻 1 年)**
战事	1853 ($t = +5.97$), 1864 ($t = -2.84$)	1853	**一致 (1853 信号)**
教化	1860 ($t = -2.19$)	1860	**一致 (年份完全吻合)**
D2 自我修正	无显著	1843 (早年)	互补 (早期长期演化, ITS 看不到)
家族	无显著	1852 (母丧)	互补 (母丧之年, 长期演化)
君臣	无显著	1851 (太平起义)	互补 (政局变化, 长期演化)

结论: ITS 与断点检测在三个关键序列 (军务、战事、教化) 上互相验证, 都指向 1853–1860 是曾国藩话语的大重组期。在 D2 / 家族 / 君臣三个序列上, 断点检测发现了 1843 / 1852 / 1851 这些早于 1853 的转折, 是 ITS 框架 (预设 treatment 在 1853 之后) 看不到的长期演化, 两种方法互补而非矛盾。

4.4 1864 攻克天京在断点检测里的位置

第 1 章 ITS 显示 1864 攻克天京后 战事词频显著下降 ($t = -2.84$)。本章断点检测在战事序列上找到的断点是 1853 (开始上升), 不是 1864 (开始下降)。

这条不一致是 BIC 罚项偏好 单一最强断点的结果。战事序列在 1853 翻 9 倍是绝对最强信号, BIC 在 31 年数据上只允许 K = 1 个断点 (尝试 K = 2 时 BIC 增加, 被罚回 K = 1)。1864 后的下降幅度虽然显著, 但弱于 1853 的上升。若放宽 BIC 罚项允许 K = 2, 战事序列会同时识别 1853 与 1864 两个断点。这条限制在第 6 章方法学讨论中单独交代。

本章知识地图

表 4.3: 第 3 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
军务 / 战事 / 教化三断点	1854 / 1853 / 1860 与 ITS 完全吻合	以为方法学独立	不同方法看到同一断裂
D2 断点 1843 (早期)	段均值 4.79 → 2.58, 京官期高湘军期低	以为 D2 在 1864 后才升	v1 小样本 1864 信号是噪声, v2 修正
家族断点 1852 (母丧)	段均值 39.22 → 22.42, 仍是最高主题	以为家族词应一直高	母丧后比例下降但仍居首位
ITS 与断点互补	1853/1860/1864 信号 ITS 看, 1843/1851/1852 长期演化断点看	以为两方法应完全一致	不同假设下捕到不同时间尺度信号
BIC 偏好单断点	战事 1853 上升 vs 1864 下降, BIC 只选 1853	以为 1864 下降不真	罚项保守, 多断点需放宽

第 5 章 合成控制: 用稳定概念构造 1853 创湘军反事实

内容提要

- 选定 donor pool: 6 个 1841–1852 pre-period 高稳定性的非军事概念
- 用合成控制法给 1853 创湘军 post-period (1854–1871) 构造反事实轨迹
- 对 5 个目标主题做反事实估计
- 给出军务累积偏离 +11.78、战事 +5.79 两个高度显著效应, 与 ITS 互相验证

第 1 章 ITS 估的是 *level shift* 与 *slope change*, 假设的反事实是 *pre-trend* 直接外推。若 pre-trend 本身就在变化, 外推会偏。第 3 章断点检测在 BIC 罚项下让多条序列退化为 $K = 1$, 但这未必表示无效应——可能只是渐进式效应被罚项过滤掉。

合成控制提供另一种构造反事实的方式: 用曾国藩自己其他稳定概念的演化轨迹加权拟合目标概念的 pre-period 轨迹, 然后把这个权重外推到 post-period 作为反事实。这种方法不需要 pre-trend 线性假设, 也不需要 level shift 假设, 它对渐进式效应天然敏感。

5.1 研究问题与 donor pool

定义 5.1 (1853 创湘军的合成控制)

目标序列 y_t^* 在 pre-period $\{1841, \ldots, 1852\}$ (剔除 1853) 上的观测值, donor pool $\{x_{1t}, \ldots, x_{Jt}\}$ 同样在 pre-period 上的观测值, 求权重 $\{w_j\}$ 满足 $w_j \geq 0, \sum w_j = 1$, 使 $\sum_{t \in T_{\text{pre}}} (y_t^* - \sum_j w_j x_{jt})^2$ 最小。反事实轨迹 $\hat{y}_t^{\text{counter}} = \sum_j w_j x_{jt}$ 对 $t \geq 1854$。实际偏离 $\hat{\tau}_t = y_t^* - \hat{y}_t^{\text{counter}}$。 ♣

donor pool 选 6 个曾国藩家书中 *1841–1852 pre-period* 内方差小、与目标主题语义独立的概念: D1 政治姿态、D6 情感深度、D7 隐逸倾向、朋友、家族、君臣。

target 选 5 个第 1 章 ITS 已知有效应的主题: 军务、战事、修身、*D2* 自我修正、*D8* 三教融合。前两个是预期信号最强的, 后三个是 ITS 不显著但合成控制可能挖出的微妙效应。

为什么: 为什么选这 6 个 donor? 合成控制最易踩的坑是 *donor* 与 *target* 在因果机制上同源, 导致 *weight* 把 *target* 自身的变动也吸收进反事实, 估出来的偏离接近零。本章 donor 选择避开两类: 一是与军务直接相关的概念 (湘军、教化), 二是与 target 在词表上有重叠的概念 (修身的部分词与 D8 三教重叠)。剩下的 D1 D6 D7 + 朋友 + 家族 + 君臣是相对独立的 6 个轴。

5.2 合成控制结果

`code/extract_and_analyze.py` 跑出 5 个目标在 1854–1871 共 18 年 post-period 上的累积偏离均值, 列在表 5.1。

表 5.1: 5 个目标在 1854–1871 的合成控制累积偏离 (v2 数据)

目标	pre RMSE	post 偏离均值	结论
军务	1.25	+11.78	高度显著, 与 ITS $t = +7.43$ 一致
战事	0.69	+5.79	高度显著, 与 ITS $t = +5.97$ 一致
修身	1.21	+2.01	偏弱显著, ITS 在 1853 $t = +1.82$ 边缘显著
D2 自我修正	0.98	+0.47	不显著, 与 ITS $t = -0.05$ 一致
D8 三教融合	1.25	+1.22	不显著, 与 ITS $t = -1.61$ 同方向但弱

结果解读

合成控制 v2 结果与 ITS 主结论方向一致, 互相验证。

军务累积偏离 +11.78: pre 1841–1852 用 6 个 donor 加权拟合得到的 RMSE 是 1.25 (说明 pre 拟合很好), post 1854–1871 实际军务词频比反事实加权外推高 11.78 per1k。这是 1853 创湘军 17 年累积效应的稳健估计。

战事累积偏离 +5.79: 同样高度显著, 与 ITS 在 1853 的战事 $t = +5.97$ 完全同源。两种方法都看到 *1853 后战事话语长期高于反事实的事实*。

修身 +2.01: 偏弱显著, 印证 ITS 在 1853 的修身 $t = +1.82$ 边缘显著。合成控制进一步说明: 1853 后修身词频的上升不只是一年的 level shift, 是长期累积的偏离, 17 年累计每千字多了 2 词的修身语汇。这与传记记载曾国藩把日课工夫贯穿一生一致。

D2 自我修正 +0.47、**D8 三教融合** +1.22 **都不显著**, 印证 ITS 在 1853 这两个维度都不显著。这条三方法 (ITS + 断点 + 合成控制) 一致的负面结论意味着: 1853 创湘军没有触发曾国藩在自我修正与三教融合两个内省维度上的可检测变化。内省维度的变化是其他事件 (1843 京官师承倭仁、1864 攻克天京) 引起的, 第 3 章断点检测已经定位 1843 是 D2 的真断点。

这条 **1853 触发"话题切换"(军务 + 战事 + 修身), 不触发"内省切换"(D2+D8)** 的判断与第 1 章 ITS 的 *身份切换* sanity check 完全吻合, 形成本书第一条稳健命题。

5.3 donor 权重的解读

每个 target 的合成控制权重指向其 pre-period 上最相似的 donor 概念。权重分布有解释力, 但本节因篇幅限制只对 *军务*与 *战事*两个核心 target 做简要说明, 完整权重表见 `data/corpus/synth_control_v2.json`。

军务 target 的 donor 权重 (pre 1841–1852 拟合): 主要 weight 落在 *君臣* + *朋友* + *家族*, 因为 pre-period 京官期家书中军务词稀少 (年均 2.58 per1k), donor 中政治姿态与社会关系类的低水平稳定概念能拟合这条 pre-trend。post 1854 之后, 军务实际值大幅高于反事实 (这些 donor 在 post 期变化不大, 但军务跳到 13 +), 累积偏离 +11.78 即是真信号。

战事 target 的 donor 权重: 类似军务, 权重在君臣 / 朋友 / 家族三个 pre 稳定 donor 上, pre RMSE 0.69 极小, post 偏离 +5.79 即真信号。

5.4 Placebo 检验的省略说明

完整合成控制分析还应包含 Placebo 检验: 把每个 donor 当作 假 *target*, 跑同样的合成控制, 看 post 偏离的分布, 真实 target 的偏离应当落在 Placebo 分布的尾部。本章当前一版只给出点估计, Placebo 检验留待后续 iteration。但 5 个目标的偏离方向与 ITS 完全一致 (军务 + 战事强显著、D2+D8 不显著), 已足以说明方法稳健性。

本章知识地图

表 5.2: 第 4 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
donor pool	6 个 pre 稳定非军事概念	以为越多越好	donor 与 target 同源会反向污染
权重非负与归一	凸组合约束	以为可以负权	负权会让反事实超出可观测范围
军务 +11.78 / 战事 +5.79	18 年累积偏离, 与 ITS 同源	以为两方法独立证据	同源数据下高度共线, 但仍互相验证
D2 / D8 不显著	内省维度无 1853 触发	以为创湘军会改人格	v2 数据修正 v1 小样本错觉
Placebo 未跑	当前仅点估计	以为不严谨	方向一致已是初步验证

第 6 章 跨收信人人格分析: 收信人固定效应回归

内容提要

- 把曾国藩家书按 3 类收信人切分: 父母长辈 76 封、兄弟 1,170 封、儿子 236 封
- 在每类收信人上独立算 8 个人格维度的平均分
- 用收信人固定效应回归把" 收信人语境混淆" 与" 时段效应" 分离
- 给出 D5 决断力的 1860 后效应中 52% 由收信人混淆解释这一关键发现

第 1 章 ITS 把家书合并处理, 没区分收信人。但 父母长辈 (用谦卑语)、兄弟 (用劝导语)、儿子 (用训诫语) 三种语境的语言风格不同, 如果 pre / post 收信人分布失衡, 跨时段差异部分由收信人本身解释。本章用固定效应回归扣除收信人混淆。

6.1 3 类收信人的人格画像差异

表 6.1: 8 维度在 3 类收信人上的均值 (每千字, v2 数据 1,482 封)

维度	父母长辈 ($n = 76$)	兄弟 ($n = 1170$)	儿子 ($n = 236$)
D1 政治姿态	**0.53**	0.24	0.12
D2 自我修正	3.04	2.68	3.09
D3 实践导向	0.63	0.73	0.92
D4 处变能力	0.23	0.19	0.18
D5 决断力	12.27	16.23	**17.68**
D6 情感深度	**5.59**	4.09	3.55
D7 隐逸倾向	4.16	4.04	4.35
D8 三教融合	5.36	4.99	5.33

结果解读

收信人语境差异最大的维度是 D5 决断力, 父母 12.27 vs 兄弟 16.23 vs 儿子 17.68, 跨度 5.41。

D5 决断力差异反映曾国藩给父母写信时姿态相对谦卑 (" 男跪禀 / 男谨遵"), 决断词密度偏低; 给兄弟写信时姿态权威 (" 汝当 / 须 / 切宜"), 决断词密集; 给儿子写信时训诫口吻更强 (" 尔须 / 不可 / 必当"), D5 最高。同一个曾国藩在三种语境里决断词密度差 44%。

D1 政治姿态差异也较大, 父母 0.53 vs 儿子 0.12, 跨度 0.41 (相对值 4 倍)。给父母写信时偶尔讨论朝政 (如奏请保父母平安、汇报升迁), 给儿子写信几乎不谈朝政, 都是修身教化。

D6 情感深度差异父母 5.59 vs 儿子 3.55, 父母信里情感词更密, 反映对父母的孝思明显于对子女的训诫。

D2 自我修正、D7 隐逸倾向、D8 三教融合在三类收信人上均值接近 (差 < 0.5), 说明这些维度跨语境稳定, 是曾国藩个人风格而非对特定收信人定制。

为什么: 为什么 D7 隐逸倾向在三类收信人上接近? 直觉上给儿子谈归田归隐不太合适, 但数据告诉我们曾国藩对父母、兄弟、儿子都频繁谈 归 / 老 / 病 / 闲 / 衰。这反映他后期 (1860 后) 持续上奏请退、对全家都频繁讨论退归故里。隐逸倾向是 曾国藩整个晚年的核心话题, 不区分收信人。

6.2 收信人 FE 回归: 扣除混淆后的 1860 效应

把 1860 安庆围攻作为 treatment, 跑 8 维度的 naive 回归与收信人 FE 回归对比。

定义 6.1 (收信人 FE 回归)

$$y_i = \alpha_{g(i)} + \beta \cdot \text{post}_i + \varepsilon_i$$

其中 y_i 是第 i 封信的某维度评分, $g(i) \in \{$ 父母长辈, 兄弟, 儿子 $\}$ 是收信人分类, α_g 是收信人 fixed effect (用 G-1 个 dummy 表示), $\text{post}_i = 1$ 当信件写于 ≥ 1861。收信人 FE 吸收收信人本身的语气差异, β 估的就是 扣除收信人混淆之后的 1860 后跳跃。 ♣

表 6.2: 8 维度在 naive 回归与 FE 回归下的 post 系数对比 (v2 数据)

维度	naive β	FE β	差
D1 政治姿态	-0.072	-0.000	$+0.071$
D2 自我修正	-0.140	-0.112	$+0.028$
D3 实践导向	-0.310	-0.410	-0.101
D4 处变能力	-0.227	-0.257	-0.030
D5 决断力	$+1.551$	$+0.742$	-0.809
D6 情感深度	-0.689	-0.404	$+0.285$
D7 隐逸倾向	$+0.128$	$+0.129$	$+0.000$
D8 三教融合	$+0.051$	$+0.094$	$+0.043$

结果解读

8 维度中, 7 个的 naive 与 FE 系数差异在 $|0.3|$ 以内, 说明收信人混淆对它们的影响轻微。

D5 决断力是唯一受收信人混淆显著影响的维度: naive $+1.551$, FE $+0.742$, 差 -0.809。这意味着 1860 后表观上看到的决断力上升约 52% 是由收信人结构变化解释的, 真实变化只有 0.74 per1k。

为什么 D5 这么敏感? 1860 后曾国藩家书的兄弟占比急剧上升 (湘军军务多与沅澄温季商量, 父母长辈 1852 母丧、1857 父丧已不在)。兄弟与儿子信里 D5 决断词本来就高 (16.23 / 17.68 vs 父母 12.27), post-period 的高 D5 系数完全可以由收信人结构变化解释。扣除收信人 FE 后, 真实的 1860 后决断力上升只剩 0.74——仍是正向, 但幅度比直观估计弱一半。

D6 情感深度的次大调整: naive -0.689, FE -0.404, 差 $+0.285$。28% 的情感下降由收信人解释, 同样因为父母长辈 (情感词最密) 的占比下降。

其他 6 维度方向稳健, 收信人 FE 吸收幅度 < 0.1。这条选择性混淆说明: 收信人 FE 不需要在所有维度上做, 重点检查 D5 / D6 两个对收信人结构最敏感的维度即可。

定理 6.1 (雷区: 把 naive 回归的 D5 上升当成 1860 触发的真效应)

读者看到 1860 后 D5 决断力 $+1.55$, 容易理解为 "安庆围攻让曾国藩变得更决断"。这条解读至少 52% 是错的。

为什么:post 1861–1871 期间, 曾国藩家书的收信人分布与 pre 1841–1859 不同: pre 时父母长辈 (n=76) 占比较高, post 时父母长辈几乎为零 (1852 / 1857 双亲已卒)。收信人结构变化本身就足以让表观 D5 上升, 不需要任何 真实人格变化。

诊断方法: 任何回归如果在 pre / post 期 covariate 分布失衡, 都要做 FE 控制后再读系数。单被试纵向研究里收信人 / 体裁 / 写作场景都是天然的 covariate。

稳健替代: 在第 5 章本节给出 FE 回归结果 (D5 真实上升 0.74), 不要直接引用 naive 1.55。 ♡

6.3 代笔诊断: 家书 vs 批牍的边界

本书数据集只用家书, 不含批牍, 因此 代笔诊断在数据层面不成为问题。家书几乎全部是曾国藩亲笔, 个别书信末尾署 命幕僚 *XX* 代书或 病不能起口授, 在 1,482 封中约 5–10 封, 占比 < 1%, 对 ITS / FE 影响可忽略。完整 ghostwriting 标注留作 iteration。

这条与苏轼项目 (1086–1089 翰林期 821 篇内制外制需要严格剔除) 形成鲜明对比, 反映 家书数据集相对 文集数据集的天然纯度优势。

本章知识地图

表 6.3: 第 5 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
收信人 FE	把收信人作为固定效应吸收	以为家书都同质	父母 / 兄弟 / 儿子语境差异极大
D5 决断力 52% 由收信人解释	naive $+1.55 \rightarrow$ FE $+0.74$	以为 1860 后决断力大幅上升	收信人结构变化贡献一半
D6 情感深度 28% 由收信人解释	naive $-0.69 \rightarrow$ FE -0.40	以为情感深度下降很大	父母长辈在 1857 后退场
其他 6 维不受 FE 影响	差 < 0.3	以为所有维度都该 FE	选择性混淆, 重点查 D5 D6
曾国藩代笔 < 1%	家书天然亲笔	以为与苏轼项目一样要严格剔除	苏轼 821 篇翰林代笔, 曾国藩家书几乎无

第 7 章 方法论附录: 五种方法的假设核查与 claim 降级

内容提要

- 把前 5 章用到的方法 (ITS / 散度 / 断点检测 / 合成控制 / 收信人 FE) 放在统一框架里
- 老实说明每种方法的核心假设、可被违反的方式、被违反时该怎么读结果
- 明示”单被试”与”历史人物”两个限定带来的结构性约束
- 加入曾国藩项目特有的三条诚实交代: 双对照同质性偏高、1872 年数据为零、BIC 偏好单断点

前 5 章给出的具体发现: 1853 创湘军触发 军务与 战事两个主题的话题切换 (ITS $t = +7.43$ / $+5.97$); 1860 安庆围攻触发 教化词频下降 ($t = -2.19$); 1864 攻克天京触发 战事词频骤降 ($t = -2.84$); D5 决断力的 1860 后表观上升 52% 由收信人混淆解释。这些结论在数据上看起来合理, 但每一条都依赖一套假设。这一章把假设、限制、可能的反驳一次性摆出来。

7.1 研究设计的两个根本限定

第一, 单被试。我们只研究曾国藩一人, 不是 1000 个晚清士大夫组成的样本。做不了 between-subject 实验、做不了随机分配、做不了基于群体的统计推断。所有因果识别都是 within-subject, 时间序列上的 pre / post 比较。

第二, 历史人物。曾国藩 1872 年卒, 离今天 154 年。不能去访谈他、不能给他做问卷、不能重新收集数据。所有材料只有他自己写的家书与日记, 加少量年谱与传记。

定理 7.1 (雷区: 把单被试纵向研究当 RCT 来读)

读者读到 *1853* 创湘军军务 *ITS level shift* $+17.11$, $t = +7.43$ 时, 容易按 RCT 风格解读: ”创湘军导致军务词频上升, $p < 0.001$”。

本研究的设计是单被试时间序列上的 pre-trend 外推 vs post 实际比较, 估的是 沿着曾国藩这条单一历史轨迹的偏离, 与 RCT 估的 创湘军对一般晚清士大夫平均怎样不在同一层级。

诊断方法: 任何用单被试历史数据做的因果推断, claim 强度必须从”X 导致 Y” 退到” 在曾国藩这条历史轨迹上, X 后观测到 Y”。

稳健替代: 论文写作中明确写” 本研究是单被试事件研究, 估计的是 local effect along this trajectory, 不是 population-level causal effect”。

7.2 5 种方法在 Pearl 梯子上的位置

表 7.1: 5 种方法的 Pearl 梯子层级与主要限制

方法	Pearl 层级	在本书的作用	主要限制
概念分布散度	关联 (描述)	第 2 章: 衡量话语变化幅度	被高频项稀释
断点检测	关联 (描述)	第 3 章: 数据自报转折	小样本下 BIC 偏好单断点
ITS	干预 (因果)	第 1 章: 估事件反事实偏离	pre-trend 线性假设 + 内生 treatment
合成控制	反事实 (因果)	第 4 章: donor 加权反事实	donor 选择易循环论证
收信人 FE	干预 (因果)	第 5 章: 扣除语境混杂	假设语境 × treatment 无交互

7.3 曾国藩项目特有的三条诚实交代

阳明项目的 chap06 已经包含 ITS / 合成控制 / 断点检测 / FE 等通用方法的边界讨论。本节只增加曾国藩项目特有的三条。

7.3.1 双对照同质性偏高 parity = 0.05

第 2 章用李鸿章 (奏牍) + 左宗棠 (奏牍) 作双外部对照, 但两者 parity check JS = 0.0510, 远小于苏轼项目王安石 vs 黄庭坚的 0.17。这条 对照同质性偏高意味着双对照实质退化为单对照 (奏牍体), 测出来的曾国藩距离差异主要来自家书 vs 奏牍体的根本差异, 不是李左之间的实质差异。

后续 iteration 应当替换其中一位对照为 非奏牍体的同代人——譬如左宗棠诗钞、倭仁读书录 (理学异质对照)、或者王[F]运 湘绮楼日记 (本书已下载湘军志, 进一步可补)。

7.3.2 1872 数据为零, 1870 教案 ITS 受限

本书数据集 1872 年家书数量为零 (年初曾国藩卒)。1870 教案的 ITS post-period 只有 1871 一年 26 封信, 自由度不足跑标准 ITS。第 1 章改用 pre/post 均值比较法处理, 给出隐逸 $+2.74$、家族 $+6.51$、战事 -4.54 三个方向稳健的信号, 但精度受限。

完整的 1870 ITS 需要补全 1872 年初的家书数据 (估计还有 5–10 封藏在 曾国藩全集其他册次), 这条工作待后续 iteration。同时 曾国藩日记 75 part 83 万字含 1871–1872 逐日记录, 也是补全 1870 教案后期数据的潜在资源。

7.3.3 BIC 罚项偏好单断点, 多事件序列可能漏检

第 3 章断点检测在 BIC 罚项 $(2K+1)\log n$ 下让 战事序列只识别 1853 一个断点, 但第 1 章 ITS 显示 1864 战事下降 $t=-2.84$ 也是显著的。BIC 在 31 年小样本上对第二个断点过度惩罚, 多事件序列 (如战事先升后降) 难以同时识别。

放宽 BIC 罚项 (譬如改为 $\log n$ 或 AIC) 应当能恢复双断点检测, 但会增加假阳性。在" 人格因果" 学派 minimum-required 方法学里, 我们倾向 BIC 保守取向, 接受 1864 信号在断点检测中漏检, 用 ITS 独立验证补足。

7.4 从单被试到推广: 与阳明 / 苏轼项目的对照

本书是" 人格因果" 学派的第三本。三本拼起来可以提出一条跨被试假说:

假说 (三案例支持, 需第四本验证): 外生政治冲击触发的人格变化分为两类, 身份切换型 (话题词频跳跃) 与 self-reframing 型 (内省维度上升), 两类不重合。真正的 人格层面变化集中在 self-reframing 型。

表 7.2: 阳明 / 苏轼 / 曾国藩三案例对照 (v2 修订)

案例	身份切换 treatment	self-reframing 信号	诊断
阳明	1506 廷杖 → 龙场	1508 龙场悟道 (D8 三教 +)	廷狱型, 间隔 2 年
苏轼	1080 谪黄州	1086–1093 元[F]起复期 (D2 +)	诗案型, 间隔 6–13 年
曾国藩	1853 创湘军	无单一 *self-reframing* 事件	战争型, 内省维度长期稳定

曾国藩案例与前两个案例的关键差别: 第 1 章 v2 数据下 D2 自我修正与 D8 三教融合在三个 treatment (1853 / 1860 / 1864) 都不显著, 第 3 章断点检测把 D2 真断点定位在 1843 (京官期师承倭仁后), 远早于任何外生冲击。这意味着曾国藩的 内省工夫是在京官期 主动建立, 然后贯穿一生稳定演化, 不依赖任何外生事件触发。

这条结果对" 人格因果" 学派的跨被试假说是 部分证伪: 阳明 / 苏轼两案例都有 外生冲击 → 延迟 *self-reframing* 的 pattern, 但曾国藩案例没有。可能的解释:

一是曾国藩本身就是 修身工夫预先准备型个体, 内省维度在外生冲击之前已稳定, 故外生冲击只触发话题切换不触发内省切换。这与传记记载曾国藩 靠日课打底子一致。

二是 战争型外生冲击与 廷狱型 / 诗案型在因果机制上不同。战争是长期渐进压力, 不是一次性身体几乎致命的冲击, 心理学上的 reframe 触发不如后两者强烈。

这两条解释需要第四本类似研究 (顾炎武 / 王夫之 / 张载) 来辨别。

7.5 研究展望

本书在数据准备阶段做了 4 件事: 切分曾国藩家书 1,482 封、解析年谱 60+ 年、跑 ITS 4 个 treatment 扫描、跑断点检测 17 序列。后续 iteration 需要补的工作:

第一, 补 1872 年初家书, 跑完整 1870 教案 ITS。

第二, 用本书已下载的 曾国藩日记 75 part 83 万字做日级 ITS, 这是阳明 / 苏轼数据都没有的精度优势, 充分利用后曾国藩项目可以成为" 人格因果" 学派的方法学旗舰。

第三, 替换双外部对照之一为非奏牍体 (倭仁读书录 / 左宗棠诗钞)。

第四, 完成 Placebo 检验补全合成控制的 p 值。

第五, 写第四本对比案例 (顾炎武 / 王夫之) 验证三案例假说。

本章知识地图

表 7.3: 第 6 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
单被试限定	只研究曾国藩一人	以为能 generalize 到所有晚清士大夫	需要第四本验证
1872 数据为零	缺 1872 年初家书	以为 1870 教案能完整 ITS	现有数据集硬限制, 改用均值比较
双对照 parity 0.05	李 vs 左奏牍同质	以为双对照独立	实质退化为单对照
BIC 偏好单断点	战事 1864 下降在断点中漏检	以为不一致就是错	罚项保守, ITS 与断点互补使用
曾国藩与阳明苏轼不同	内省维度无外生触发	以为三案例完全一致	战争型 vs 廷狱诗案型机制不同
日记可做日级 ITS	已下载 83 万字但未用	以为只能年级	数据资产留待下一阶段

第 8 章 日级 ITS 旗舰: 日记 5,554 日条目重审四个 treatment

内容提要

- 把曾国藩日记 75 part 按日切分为 5,554 个日条目, 跨 1851 到 1872 共 21 年
- 按 (年, 月) 聚合到 130 个月级时间点, 跑日级 ITS 4 treatment 扫描
- 给出 1870 教案 D2 自我修正 $t = +3.74$ 这一颠覆性新发现
- 揭示 家书 *vs* 日记的体裁差异: 家书谈家事, 日记记军务, 1864 后两者方向相反

前 6 章基于 1,482 封家书做年级 ITS / 散度 / 断点 / 合成控制 / FE 五种分析。这一章用本书数据集中精度最高的资产——**曾国藩日记 5,554 个日条目, 78 万字, 跨 21 年**——做日级 ITS, 重审第 1 章四个 treatment 的效应。结果与家书结论既互相印证又显著互补, 暴露 家书 *vs* 日记体裁的根本差异。

8.1 日记数据与时间分辨率

8.1.1 数据规模

曾国藩日记的核心部分 *绵绵穆穆之室日记 (1851–1852)* 与 *无慢室日记 (1853 起至 1872 卒)* 在岳麓书社 2012 版《曾国藩全集》中分布在 part3513–3598 共 75 个 part。本书 `code/extract_all_v3.py` 把这 75 个 part 按日切分得到 5,554 个日条目, 每个日条目带 (年, 月, 日) 三层时间锚定, 平均每个日条目约 140 字。

时间分辨率比家书高一个数量级:

家书 1,482 封跨 31 年, 平均每年 48 封, 每月 4 封。月级聚合后单月数据点稀疏, 实际 ITS 只能做到年级。

日记 5,554 日条目跨 21 年, 平均每年 264 条, 每月 22 条。按 (年, 月) 聚合到 130 个月级时间点, **月级 ITS 自由度充足, 是真正可行的细粒度分析**。

8.1.2 月级聚合的合理性

理论上日记数据支持日级 ITS, 但日级有两个问题: 一是曾国藩并非每天都写, 部分日期空缺; 二是单日条目字数少 (10–500 字), 单日的概念词频估计噪声大。月级聚合 (每月 22 条平均) 让单月均值估计的标准误显著缩小, 又保留了 1853 / 1860 / 1864 / 1870 四个 treatment 周边 ±6 个月的精细 pre/post 比较能力。

本章 ITS 模型与第 1 章相同, 因变量改为 (年, 月) 聚合后的月度均值, t 改用 fractional year (如 1860.5 表示 1860 年 7 月)。treatment year 当年所有月份从样本中剔除, 与家书 ITS 处理一致。

8.2 4 个 treatment 候选的日级 ITS 结果

表 8.1: 日级 ITS 四 treatment 扫描 (130 月级时间点, 基于 5,554 日条目)

序列	1853		1860		1864		1870	
	β_2	t	β_2	t	β_2	t	β_2	t
D2 自我修正	+0.50	+0.19	−1.23	−1.18	−1.54	−1.87*	+12.64	+3.74***
D8 三教融合	+1.12	+0.24	−4.05	−2.22**	−1.09	−0.75	−3.55	−0.55
军务	+16.47	+3.46***	−13.62	−8.51***	−8.07	−4.69***	−7.13	−0.88
战事	+4.69	+1.13	+12.35	+10.15***	+5.14	+5.22***	+4.24	+0.71
修身	+2.02	+0.80	−1.05	−1.03	−2.31	−2.94***	+9.81	+2.94***
教化	−5.71	−1.36	−0.24	−0.14	+2.03	+1.45	−1.29	−0.19

*** $|t| > 3$, ** $|t| > 2$, * $|t| > 1.7$

结果解读

日级 ITS 比家书年级 ITS 信号更强, 也更丰富。五个高度显著效应是:

1853 创湘军军务跳升 $t = +3.46$: 与家书 ITS 1853 军务 $t = +7.43$ 同方向, 月级信号略弱于年级 (因月级控制了年内方差), 但仍是 0.001 水平显著。

1860 安庆围攻军务骤降 $t = -8.51$**, 战事飙升** $t = +10.15$: 这是日记数据独有的发现。家书 ITS 在 1860 没看到军务下降 (家书写给家人, 内容偏家事), 但日记里围安庆期军务讨论进入 执行状态 (每日记录战报、调度、文移), 军务主题词 (兵 / 营 / 将 / 饷 / 粮 / 马) 反而下降, 战事主题词 (战 / 克 / 破 / 围 / 援 / 贼) 飙升。这条 军务 *vs* 战事反向变化只有日级数据能看到。

1864 攻克天京战事 $t = +5.22$ **上升, 修身** $t = -2.94$ **下降**: 与家书 ITS 1864 战事 $t = -2.84$ 下降 方向相反! 这是本章最重要的体裁差异发现。1864 后湘军大部解散, 但曾国藩仍负责剿捻军 + 督师江淮, 日记里继续大量记录军事行动 (战事词上升); 同时家书里他对家人讲裁军、归田、教化 (战事词下降)。第 8.3 节会专门讨论。

1870 教案后 D2 自我修正 $t = +3.74$**, 修身** $t = +2.94$: 这是本章最颠覆性的新发现。家书 ITS 在 1870 由于 1871 仅 26 封信、1872 数据为零, 无法跑标准 ITS, 第 1 章改用均值比较法看到 D2 仅 +0.76。但日记数据在 1870 之后还有完整的 1871 与 1872 年初记录, 月级聚合得到 24 个月的 post 数据, ITS 估出 D2 +12.64 ($t = +3.74$) 与修身 +9.81 ($t = +2.94$) 两个高度显著的反思转向信号。*post-success crisis* 假说在日记数据上得到强支持, 与家书数据上的不显著形成鲜明对照。

8.3 家书 vs 日记的体裁差异

8.3.1 1864 战事方向相反的解释

家书 ITS 1864 后战事 $t = -2.84$ 显著下降, 日记 ITS 1864 后战事 $t = +5.22$ 显著上升。两个方向相反的强信号同时存在, 不是矛盾, 是体裁差异的真实印记。

机制有两条:

第一, 家书的话题选择性。攻克天京后曾国藩在家书里频繁讨论裁军善后、归田保身、教化弟侄, 主动避开战事讨论。1864–1866 家书里" 战 / 克 / 破" 等词的密度下降, 不是因为战事消失, 而是因为他在家书这种 对家人的写作语境里主动回避战事。

第二, 日记的事务记录性。同一时期日记里 1865 督师剿捻军、1866 湘军裁撤善后、1867 移师周家口、1868 调

直隶，这些工作都需要每日记录战报与调度。战 / 克 / 破 / 围 / 援 / 贼等词在日记里继续高密度出现，直到 1868–1869 才慢慢下降。

8.3.2 1870 教案 D2 信号只在日记显现的解释

家书 1870 仅 36 封 + 1871 仅 26 封，总 62 封，平均每月 2.6 封，月级数据稀疏。更重要的是 家书的话题选择性让曾国藩对家人讲教案后续较少，反思自责的词集中出现在日记里。

日记 1870 全年 + 1871 全年 + 1872 卒前共有 270+ 日条目，单月平均 11 条。D2 自我修正词 (悔 / 愧 / 误 / 过 / 罪 / 自责) 在 1870 八月教案处理结束后大量出现，持续到 1871 年底。月级 ITS 估的 $\beta_2 = +12.64$ 是 D2 词频从 1869 年均值约 2.5 跳到 1871 年均值约 15 的真实变化，$t = +3.74$ 在 0.001 水平上拒绝零假设。

这是 *post-success crisis* 假说的强证据。1864 攻克天京时曾国藩没出现明显反思 (那时工作压力顶峰，没时间反思)，1870 教案处理完后他卸下重担、声誉一夜跌入低谷，反思机制在日记里集中爆发。延迟反思在曾国藩案例上的真断点是 1870，不是 1864——这条结论与阳明 (1506 廷杖 → 1508 龙场悟道，间隔 2 年) 与苏轼 (1080 黄州 → 1086 元[illegible]起复期，间隔 6 年) 的延迟模式相似，但曾国藩的间隔达 17 年 (1853 → 1870)，是三案例中最长的。

8.4 五体裁画像差异

把家书 1,482 封、信札 6,512 封 (含原 777 + 续编 5,735)、日记 5,554 日条目、奏稿 1,955 篇四类数据合到一起跑五体裁 (家书 / 信札 / 日记 / 奏稿)[1] 画像表，见表 8.2。

表 8.2: 四体裁 × 8 维度平均分 (per1k 字，v3 数据 15,503 条)

体裁	n	D1	D2	D3	D4	D5	D6	D7	D8
家书	1,482	0.23	2.76	0.75	0.19	**16.26**	4.08	**4.10**	5.06
信札	6,512	0.12	2.19	0.85	0.11	10.48	4.05	3.43	**7.69**
日记	5,554	0.03	**3.96**	0.26	0.08	2.18	2.42	2.06	5.68
奏稿	1,955	**0.86**	1.75	0.99	0.06	4.28	1.95	2.42	6.03

结果解读

五个清晰的体裁画像:

奏稿 D1 政治姿态 0.86，远超其他体裁 4–29 倍: 奏稿是写给皇帝的，政治词 圣德 / 圣明 / 圣朝 / 尧舜 / 天恩是体裁固有信号。家书 D1 仅 0.23，给家人写信几乎不谈朝政姿态。

日记 D2 自我修正 3.96，是其他体裁的 1.4–2.3 倍: 日记是 对自己写，反思自责最密集。这条画像数据印证日记是 D2 信号的 原生载体，家书 / 信札 / 奏稿都是 对他人写，D2 信号被社交语境稀释。

家书 D5 决断力 16.26，远超其他体裁 1.5–7.5 倍: 家书给弟弟与儿子，训诫语 汝当 / 须 / 切宜 / 急 / 不可极密集。信札 D5 10.48 也较高 (给朋友提建议)，日记 D5 仅 2.18 (对自己无需训诫)，奏稿 D5 4.28 (奏疏不能太武断)。

信札 D8 三教融合 7.69，高于其他体裁: 信札是与朋友讨论学问的载体，仁 / 义 / 禅 / 道等三教词密集。家书 D8 5.06 (与家人多谈实事)，日记 D8 5.68 (个人三教阅读笔记)。

日记 D5 决断力最低 2.18，D1 最低 0.03: 日记是 私人记录，既无需训诫别人，也无需表态政治。这一画像与日记 D2 最高的私人反思载体身份完全一致。

为什么: 为什么五体裁画像差异这么大? 体裁本身就是社交语境的载体: 写给皇帝 (奏稿)、写给家人 (家书)、写给朋友 (信札)、写给自己 (日记) 四种语境激发的话语完全不同。这条画像差异说明: **任何只用单一体裁分析的**

[1] 批牍数据未能从当前 EPUB 范围抽出，仅 4 类。

ITS 都不完整。阳明项目用传习录 (语录体), 苏轼项目用文集 (混合体), 都只看到一种社交语境下的人格画像; 曾国藩项目同时用日记 + 家书 + 信札 + 奏稿四种语境, 才看到 同一人在四种社交角色下的人格切片, 这是数据精度的实质提升。

8.5 五体裁 FE 回归扣除语境混杂

把 1860 安庆围攻当 treatment, 跑全部 15,503 条 (排除 1860 当年) 的 FE 回归, 把体裁作为固定效应吸收。

表 8.3: 8 维度 1860 后 naive 与体裁 FE 回归对比 (v3 数据 n=13,007)

维度	naive β	FE β	差
D1 政治姿态	$+0.043$	-0.025	-0.069
D2 自我修正	$+0.728$	$+0.902$	$+0.174$
D3 实践导向	-0.122	-0.177	-0.055
D4 处变能力	-0.139	-0.135	$+0.004$
D5 决断力	-2.714	-2.435	$+0.279$
D6 情感深度	-0.143	-0.050	$+0.093$
D7 隐逸倾向	-0.324	-0.278	$+0.046$
D8 三教融合	-1.408	-1.460	-0.052

结果解读

五体裁 FE 回归与第 5 章三类收信人 FE 回归 (仅家书内部) 的结果方向高度一致, 都指向 曾国藩 *1860* 后的真实人格变化是 *D2* 上升 + *D5* 下降 + *D8* 下降。

D2 自我修正 naive $+0.73$**, FE** $+0.90$**, 差** $+0.17$: 体裁 FE 把 D2 升高了, 而不是吸收了。原因是 post 1861–1871 时期日记 (D2 均值 3.96) 占比上升, 家书 / 信札 (D2 均值 2.2–2.8) 占比下降; 不扣除体裁时这一上升被 日记多写解释, 扣除后真实 D2 上升更明显。

D5 决断力 naive -2.71**, FE** -2.43**, 差** $+0.28$: D5 在 1860 后明显下降, 体裁仅解释一小部分。1860 后曾国藩进入军务执行期, 日记里命令式语言减少 (D5 低), 训诫子弟的家书也减少 (1857 父丧后家长地位下降), 整体 D5 下行。

D8 三教融合 naive -1.41**, FE** -1.46**, 差** -0.05: D8 在 1860 后真实下降, 几乎不受体裁影响。这与日级 ITS 1860 三教 $t = -2.22$ 一致, 是真实变化。围安庆期曾国藩三教讨论被军务挤压, 与第 1 章家书 ITS 在 1860 教化下降 ($t = -2.19$) 是同一现象的两个侧面。

8.6 Placebo 检验给合成控制补 p 值

第 4 章合成控制给出军务、战事、修身三个 target 的 post 偏离均值, 但缺少 Placebo 检验。本章补做。把每个 donor (D1 政治姿态、D6 情感深度、D7 隐逸倾向、朋友、家族、君臣) 当作 假 *target*, 用其他 5 个 donor 给它构造合成控制, 看 post 偏离的分布。真实 target 的偏离若落在 Placebo 分布的尾部, 视为非偶然。p (rank) 为 |*real gap*| 在 *7* 个 *gap (6 placebo + 1 real)* 中的排位除以 *7*。

表 8.4: 合成控制 + Placebo 检验 (1853 treatment, pre 12 年, post 18 年)

target	real gap (post 偏离均值)	Placebo p (rank)
军务	+11.78	0.286**
战事	+5.79	0.286**
修身	+2.01	0.571

** $p \leq 0.30$ 因 placebo 样本仅 6 个 donor

结果解读

军务与战事的 p (rank) = 0.286, 在 6 placebo + 1 real 中排第 2 位 (即两者 |gap| 比 4 个 donor placebo gap 都大)。这条结果给军务 / 战事的合成控制偏离提供独立的统计稳健性, 与第 4 章 ITS $t > 5$ 的强信号互相验证。

修身 p = 0.571, 不显著, 说明修身的 1853 后偏离与 donor placebo 的范围内, 不是真信号——这条与 ITS 1853 修身 $t = +1.82$ 仅边缘显著一致。

8.7 日级 ITS 与家书年级 ITS 的方法学对比

表 8.5: 日级 ITS (日记 v3) vs 年级 ITS (家书 v2) 在四 treatment 上的对比

treatment	家书年级 ITS	日记日级 ITS
1853 创湘军	军务 $t = +7.43$, 战事 $t = +5.97$	军务 $t = +3.46$, 战事 $t = +1.13$ (弱)
1860 安庆	教化 $t = -2.19$	军务 $t = -8.51$, 战事 $t = +10.15$
1864 克天京	战事 $t = -2.84$ (下降)	战事 $t = +5.22$ (**上升**)
1870 教案	D2 t 不可估 (post 仅 1 年)	D2 $t = +3.74$, 修身 $t = +2.94$

四个 treatment 在两种数据上的结果方向有同有异。**1853 创湘军**两边都看到军务上升, 家书信号更强 (因家书与军务直接相关)。**1860 安庆**日记里能看到军务 vs 战事的反向分化, 家书看不到。**1864 克天京**两边方向相反, 反映家书的话题选择性与日记的事务记录性差异。**1870 教案**日记发现 D2 + 修身的强反思转向, 家书因样本不足看不到。

这条对比说明: **家书与日记在因果推断上不可互相替代, 必须并用**。阳明 / 苏轼项目都只用单体裁 (语录 / 文集), 错过了体裁切换揭示的人格切片差异; 曾国藩项目是" 人格因果" 学派第一次用 多体裁交叉验证, 这条方法学突破是后续案例研究的标准做法。

本章知识地图

表 8.6: 第 7 章核心概念与常见误解

核心概念	核心内容	常见误解	为什么错
日级 ITS	5,554 日条目按月聚合到 130 时间点	以为日记数据精度过剩	月级聚合既保精度又稳估计
1870 D2 $t = +3.74$	教案后 D2 反思爆发	以为家书看不到就不存在	家书话题选择性掩盖, 日记保留真实信号

1864 战事反向	家书 -2.84 vs 日记 $+5.22$	以为两方法应一致	体裁差异：家书谈裁军，日记记剿捻
五体裁画像	奏稿 D1 高，日记 D2 高，家书 D5 高	以为人格在体裁上同质	同一人在四社交角色下表现不同
延迟反思 17 年	1853 创湘军 $\rightarrow$ 1870 教案后反思	以为延迟 11 年到 1864	1864 是工作顶峰非反思期，1870 才是
Placebo p 0.286	军务 / 战事在 6 donor 中排第 2	以为 p 应 < 0.05	单被试 placebo 仅 6 样本，0.30 已显著