

Propensity Score
Matching Methods:
Fundamentals and
Applications

倾向值匹配法的
概述与应用:

从统计关联到因果推论

苏毓淞 著



重庆大学出版社

Propensity Score
Matching Methods:
Fundamentals and
Applications

倾向值匹配法的 概述与应用:

从统计关联到因果推论

苏毓淞 著

图书在版编目(CIP)数据

倾向值匹配法的概述与应用:从统计关联到因果推论/苏毓淞著. --重庆:重庆大学出版社,2017.7
(万卷方法)

ISBN 978-7-5624-9632-8

I. ①倾… II. ①苏… III. ①统计分析—研究
IV. ①C813

中国版本图书馆 CIP 数据核字(2016)第 034121 号

倾向值匹配法的概述与应用:

从统计关联到因果推论

苏毓淞 著

策划编辑:林佳木

责任编辑:陈 力 版式设计:林佳木

责任校对:张红梅 责任印制:赵 晟

*

重庆大学出版社出版发行

出版人:易树平

社址:重庆市沙坪坝区大学城西路 21 号

邮编:401331

电话:(023) 88617190 88617185(中小学)

传真:(023) 88617186 88617166

网址:<http://www.cqup.com.cn>

邮箱:fxk@cqup.com.cn (营销中心)

全国新华书店经销

重庆市国丰印务有限责任公司印刷

*

开本:940mm×1360mm 1/32 印张:6.25 字数:140 千

2017 年 8 月第 1 版 2017 年 8 月第 1 次印刷

印数:1—3 000

ISBN 978-7-5624-9632-8 定价:32.00 元

本书如有印刷、装订等质量问题,本社负责调换

版权所有,请勿擅自翻印和用本书

制作各类出版物及配套用书,违者必究

万卷方法/社会科学研究方法·前沿与应用丛书

编委会

学术委员会(按汉语拼音为序)

风笑天 连大祥 刘涛雄 卢春龙 沈明明

汤维强 王天夫 张小劲 朱云汉

编委会(按汉语拼音为序)

陈 硕 洪 伟 呼和那日松 吕 杰

孟天广 庞 珣 任莉颖 苏毓淞

孙 龙 严 洁 郑 路 周 强

清华大学定量分析方法研究所 组织编写

总 序

中国当代社会科学各学科的研究方法多为西方舶来,这一事实不可回避。由此亦衍生出不尽如人意的另一个事实,即这些滥觞于西方的科学研究方法在国内学界的传播与应用,并未能像各个学科的一般理论和知识那样,为国内学者或学生广泛习得并恰当地使用。归根究底,造成这种现状的原因无外乎如下两点:一是中国社会科学诸学科长期以来重描述的研究范式;二是青年研究者在初入社科研究领域时(多为文科生)就已产生了视数学为畏途的心态。以此为背景的研究或秉承旧有的研究范式,对新的研究方法视而不见,或对西方学界主流的研究方法进行囫圇吞枣、生吞活剥式的使用。这都不利于正在迅速发展的中国社会科学与国际学术界交流。这个世纪开始,随着大批海外学子陆续回国任教,他们给社会科学学科发展带来很大冲击,也从方法上改变了学科研究的范式。

这里说的方法变革所造成的范式转变,并非简单的定量与定性方法之争。此处所谓新的范式是指通过严谨的研究设计,应用科学的方法(定量或定性方法皆属之),采集数据去从事研究。过去数十年国内学界所累积的学科知识,在这个新的研究范式下,已得到验证,甚至有突破性的理论创新。不过,任何转变都不可能一

蹴而就,固有研究范式的惯性及其导致的对于科学方法的疏离感,仍然如紧箍咒一般,时刻影响着科学研究方法在国内社会科学学科的发展。

“社会科学研究方法·前沿与应用”丛书的主要目的就是希望拉近学生与科学方法的距离,从而激发他们学习并应用这些方法。2012年暑期,重庆大学出版社雷少波编辑做客清华大学政治学系,言谈间,我提出组编一套介绍社会科学研究方法丛书的想法,有别于坊间同类介绍方法的丛书,这套丛书的最大特色在于将理论背后的“理论”呈现给读者,让初学者知其然,更能够知其所以然;同时使用实际的研究成果,提供最直观的应用案例,让读者了解到这些研究方法如何应用到以中国为主题的研究,通过主题的熟悉感,真正拉近读者与方法之间的距离,提升读者学习的热情,为有志于从事相关领域研究的人士提供扎实、恳切的引导及参考。

这套丛书编写的体例大致分为理论介绍、案例点评和软件实操展示三个部分。第一部分的理论介绍,主要是编写者结合自身学习该方法的历程,用最浅显易懂的方式(尽可能使用最少或最基本的数学),来介绍方法背后的理论与原理,以及应用该方法的场景,并重点介绍该方法的标准操作流程,以及提示该方法在操作上可能遇到的限制,为读者学习理论提供最大的便利。第二部分的案例点评是本书的亮点,编写者选取应用该方法的范文(优先选择以中文撰写和使用中国数据的论文)进行点评。这样选择的目的是

如前所述,是希望读者通过既有的研究,切实学习到该方法是如何在中国场景下落地的。编写者的点评首先发挥导读的作用,同时亦不会避重就轻,而是无褒贬之忌讳,既点出该方法在范文中使用的准确、巧妙之处,也会点出研究过程中存在的短板,不论是方法操作上的问题,还是研究设计本身的缺陷,亦或是数据上造成的罅隙,这也是编写者希望读者能够在实操过程中尽量避免的部分,毕竟方法的习得容易,但是要用得对、用得好,还是需要熟悉方法、经验丰富的资深使用者指导。此部分的点评就是希望用文字为读者建构一个面对面的学习场景,起到教师与学生一对一现场教学的效果。第三部分的软件实操展示,使用的软件主要还是坊间常用的 SPSS, STATA, R 三个软件。编写者使用数据范本,实际演示该方法通过不同软件的操作步骤,尽可能做到精准、简洁,同时兼顾可读性、可视性。原则上,这部分会至少涵盖两个软件操作展示,未来如读者有需求,我们将在丛书的网页以附录的方式,编写该方法在其他软件的操作步骤。

丛书总序的最后,必须感谢清华大学张小劲老师、王天夫老师、洪伟老师、孟天广老师,北京大学严洁老师、任莉颖老师,中国政法大学卢春龙老师,中国人民大学孙龙老师和远赴英国的呼和那日松老师,诸位同仁鼎力相助,为这套丛书的出炉提供了智力和时间上的大力支持。丛书编著过程中,经历了许多波折与挑战,若没有他们的投入,这套丛书的顺利问世几乎是不可想象的。特别感谢重庆大学出版社雷少波编辑的鼎力支持,我当初言谈间大胆

的设想方能实现;也特别感谢林佳木编辑、邹荣编辑不厌其烦地催稿,如果说雷编辑是伯乐的话,林、邹两位编辑则是催促着我们前进的鞭子。丛书的完成,这些文字背后的贵人缺一不可。

编著这套丛书,是为试图改进中国社会科学研究对于国际主流方法的使用状况,这也是上述诸位共谋此事背后的雄心壮志,但不积跬步无以至千里,任何实质性的转变都要依靠些微的积累和推进方能实现,落在实处,就是这一本本小书的质量。若能对有需要的读者有所启发,或能由此催生后续研究的跟进,就是编写者最大的心愿。由于编写者学识所限,这套丛书一定还存在诸多不足和有待完善之处,希望读者和相关方法的使用者及时反馈批评意见,以便于我们再版时加以订正和改进。

苏毓淞

于清华园

2016.10.5

序

2005 年春天,我在纽约大学 Adam Przeworski 教授的课上,第一次接触到倾向值匹配法,从此便与它结下了不解之缘。十年之后,“社会科学研究方法·前沿与应用”丛书召开第一次编写会,与会青年学者须各选择一个方法作为自己的主题,我当下就决定介绍这个方法。十年之前,倾向值匹配法在美国社会科学学界正处于上升期,许多学者认为因果推断有了新的解决方法,他们如获至宝,纷纷投入此一方法的使用,并为这个方法编写各类软件,以致坊间各式各样的“倾向值匹配方法”相关软件,一度如雨后春笋般层出不穷。不过,就如同 Przeworski 教授在那堂课后发表的工作论文中所指出的,这一方法在解决因果推断问题上存在局限性。学者们对这个方法的憧憬,其实是个美丽的误会;之后的数年间,也即我攻读博士期间(至 2009 年),我几乎见证了“倾向值匹配法”在学术界发展变迁的全过程。一方面,其所依据的理论不断产生新的突破和改进,学者提高了应用这个方法的标准和范式;然而,另一方面,他们的努力却为“倾向值匹配法”敲响了丧钟,它的绚烂,仿若一颗陨落的新星,暗淡了下来。正如学术研究中的许多方法一样,它们从最初喧嚣夺目、令人趋之若鹜的热闹,逐渐趋于平静和理智,这一过程让身处其中的我不禁扼腕叹息,同时又赞叹于学界专家学者们永无休止的探索精神,这或许正是“方法”和一

切学术研究的魅力所在吧！在这本书的前两章，我尽可能以由浅入深的方式，向你们介绍这个定量方法中一闪即逝的流星。

我想我可以下个定论，即倾向值匹配法在社会科学学界的盛行，与其隶属的因果推断理论有高度关系。20世纪以来，社会科学学者越来越希望能像自然科学学者一样，说因道果，确立研究的“科学性”与典范性，尽管他们的研究往往发现的是关联性，而非因果关系，不过他们依然为此付出了巨大的努力，不断挑战与审视自己的做法和观点。例如，早在1959年，美国社会学大师 Seymour Martin Lipset 博士即挑战了现代化理论，他认为关联不等于因果，现代化理论所主张的经济发展与民主化的关系，并不能构成因果关系。正是这样坦诚直面问题核心的精神，催促着社会科学学者们不断探究解决因果关系新的研究方法，同时也注定了“倾向值匹配法”这个宣称可以解决因果推断基本难题的研究方法，可以于21世纪初在社会科学学界横空出世！

如果读到这里，你打算合上这本书，从此远离倾向值匹配法，那你正好错过了这个在当代定量方法中最值得学习的方法。如果说倾向值匹配法是以最保守的方式进行变量间的因果推断，那么它的不足之处，同样也是其他方法推断因果受限的地方。所以，如果说学者对于这个方法的审慎规范，摧毁了这个方法的普适性，那么它的陨落同样也给其他研究方法带来本质上的冲击。简言之，我们再不能天真地认为回归分析结果就能给我们因果关系，甚至据此获得的关联性都有可能是错的，是谬误的（请详见本书关于共变量重合的讨论）！

是的,我的确是在进行负面营销!但我正是希望以这种方式提醒学习该方法的读者,倾向值匹配法并非是因果推断的灵药,多数学习该方法的入门使用者,认为使用倾向值匹配法获得的估计值就是因果效用,他们忽略了学者们后来在深挖该方法不足之处时,所开发的一连串检验理论,有些检验模块虽然已经有软件可以实现,有些检验却是待开发的,如果忽略掉后续的这些检验,行文时出现妄议因果的字眼,无疑是自开空门,给审稿人痛击文章短板的机会。所以请读者详读本书第2章的后半部分,以及第3章论文的点评,希望你们能从中认识到推论因果的难处,以及应用倾向值匹配法必须特别注意的地方。

什么是倾向值匹配法?就是将各个受测单元多维度的信息,使用统计方法简化成一维的数值,是为倾向值,然后据之进行匹配,匹配的目的就是要找寻实验组和对照组样本中拥有相同(或者相似)倾向值的样本,它们之间的差异,就是因果效用。这个原理非常简单,但聪明如你,肯定对于将多维度信息简化成一维度数值产生疑虑。什么是多维度的信息?如何简化它?这都是倾向值匹配法最为关键也是最容易出现问题的地方,本书的第1章将简要介绍为什么要对多维信息进行降维处理,第2章将更进一步讲解这背后的理论背景。书中行文时,不可避免要使用到一些数学公式来辅助理论论述,在这些推演过程的后续,我尽可能以最浅显易懂的文字对这些推演过程的原理进行说明,略过这些公式并不会妨碍对于理论的了解。

希望读者能够以更严谨的态度去进行因果推断的研究,这亦

是与读者分享本人多年来使用这一方法过程中得到的经验与教训。社会科学学界引入倾向值匹配法虽然无法彻底解决过去研究中无法推断因果的局限,但是关于这一方法的探究所带来的新的思维模式,却在很大程度上改变,甚至提高了学者们处理因果推断问题时的标准。正如我在本书第2章结语时总结的:归根究底,严谨科学的研究设计是进行因果推断的前设,倾向值匹配法是协助我们解决因果推断的一种研究方法,它不能帮助我们发现因果关系!换言之,任何看似完美的方法都无法为我们提供一劳永逸的解决方案,而围绕“倾向值匹配法”数年间的聚讼探索,却让我们见证了人文社会科学在近百年间愈加科学、严谨的进程,学者们试图以更有说服力的论证,来参与这个世界的方方面面。而兼具关怀与理论,感性与理性,正是人文社会科学得以生生不息,薪火相传的谜之所在。

最后感谢我的家人一直以来对我的支持,尤其是我的内人,在身体不适的情况下,仍然坚持为我校稿,没有她,书稿将充满生硬、佶屈聱牙的语句。如果书中存在任何错漏疏失之处,也请读者们反馈给出版社,我们将在本书再版之际进行更正。

苏毓淞

于西山脚下

2016.10.6

目 录

第1章 因果推论理论概述	/1
1.1 潜在结果模型	/4
1.1.1 因果态与实验分组	/6
1.1.2 因果推论的基本难题	/8
1.1.3 平均处理效用	/9
1.2 社会科学研究中的因果推论	/10
1.2.1 处理变量的可操作性	/11
1.2.2 单元同质性假定	/12
1.2.3 可忽略的处理分配假定	/13
1.2.4 控制混淆共变量	/16
1.2.5 稳定单元处理值假定	/18
1.2.6 共变量分布平衡与重合的要求	/20
1.2.7 勿控制处理分配后的变量	/25
1.3 小结	/27
第2章 使用倾向值匹配法估计因果效用	/29
2.1 倾向值	/32
2.1.1 倾向值的性质	/32
2.1.2 估计倾向值	/34
2.2 匹配法	/37
2.2.1 精确匹配法	/39
2.2.2 最近邻匹配法	/40
2.2.3 区间匹配法	/42
2.2.4 核匹配法	/43

2.2.5 马氏距离匹配法	/46
2.2.6 贪婪匹配法与最佳匹配法	/47
2.3 匹配后的检验	/48
2.3.1 共变量分布不平衡分析	/49
2.3.2 共变量分布不重合分析	/54
2.3.3 选择性偏差与敏感性分析	/57
2.4 匹配后估计平均处理效用	/67
2.4.1 使用回归模型估计平均处理效用	/68
2.4.2 平均处理效用的标准误差	/69
2.4.3 非二元处理下的平均处理效用	/70
2.5 小结	/70
第3章 社会科学案例选读	/75
3.1 案例评析重点	/76
3.2 教育学案例:计算机的使用对中学生数学成绩的影响	/76
3.3 心理学案例:独生和非独生子女情绪适应的差异	/89
3.4 政治学案例:候选人议题立场与选民投票抉择	/99
3.5 经济学案例:企业贸易形态与工资水平	/117
第4章 应用 R 软件和 STATA 软件实现倾向值匹配法	/135
4.1 R 程序包	/136
4.2 STATA 程序插件	/138
4.3 使用 R 进行倾向值匹配分析 LaLonde 数据	/139
4.4 使用 STATA 进行倾向值匹配分析 LaLonde 数据	/168
参考文献	/177

表目录

表 1.1	因果推论的基本难题	/8
表 1.2	二元处理下的虚拟实验数据	/9
表 2.1	儿童血铅水平的敏感性分析结果	/65
表 4.1	LaLonde 数据中的变量	/140

图目录

图 1.1	虚拟参加课后加强班前后数学成绩的差异	/21
图 1.2	虚拟参加课后加强班前后数学成绩的差异 (将实验组与对照组的比较限制在分布重合的 区块)	/23
图 1.3	比较实验组和对照组群体在教育程度和年龄 上平衡和重合的情况	/24
图 2.1	使用倾向值匹配法估计因果效用的标准作业 流程图	/31
图 2.2	图示最邻近匹配法对照组样本可替换和不可 替换的情况	/40
图 2.3	Cochran(1968)建议的五分位数和六分位数的 最佳分位点	/42
图 2.4	核匹配法常用的核及其对应的核函数和核密 度图形	/45
图 2.5	共变量分布在实验组和对照组平衡的情况	/52
图 2.6	共变量分布在实验组和对照组间缺乏平衡的情况	/53
图 2.7	比较民主国家与独裁国家人均 GDP 与 GDP 成长 率的关系	/55
图 4.1	使用 qqplot() 程序检验匹配前后共变量的平衡 情况	/148
图 4.2	使用 qqplot() 程序检验 GenMatch() 程序匹配前后	

	共变量的平衡情况	/155
图 4.3	使用散点图、分位数对比分位图、条状图检验 matchit() 程序匹配前后共变量的平衡情况	/161
图 4.4	使用折线图呈现 matchit() 程序匹配前 (All Data) 和匹配后 (Matched Data) 各个共变量 组间差值绝对值的变化	/164
图 4.5	使用条状图检验 psmatch2 程序匹配后实验组 和对照组倾向值重合的情况	/173