

分类号 C8/422
U D C _____

密级 公开
编号 10741



硕士学位论文
(专业学位)

论文题目 时变权重合成控制法及其应用研究

研究生姓名: 张颖

指导教师姓名、职称: 牛成英、教授

学科、专业名称: 统计学、应用统计

研究方向: 大数据分析

提交日期: 2024年6月3日

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 张颖 签字日期： 2024年6月3日

导师签名： 牛成英 签字日期： 2024年6月3日

导师(校外)签名： _____ 签字日期： _____

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定， 同意（选择“同意”/“不同意”）以下事项：

1.学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2.学校有权将本人的学位论文提交至清华大学“中国学术期刊（光盘版）电子杂志社”用于出版和编入CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分内容。

学位论文作者签名： 张颖 签字日期： 2024年6月3日

导师签名： 牛成英 签字日期： 2024年6月3日

导师(校外)签名： _____ 签字日期： _____

Time-varying weight synthesis control method and its application

Candidate : ZHANG Ying

Supervisor : NIU Chengying

摘 要

合成控制法通过构建一个合成控制模型来估计干预组在未干预状态下的潜在结果,比较干预组与控制组之间的干预效应差异,从而评估干预措施或政策变化的因果效应。该方法通过对控制组样本单元进行加权平均或适当地线性组合产生一个合成的控制组,解决了基于反事实结果因果效应分析中的大样本限制条件。但合成控制组的过程中,常见加权方法不同时期采用权重不变,而由于研究时期的拉长,各单元在不同时期的影响程度发生变化,单元的合成权重也应相应改变才更符合实践情况。

首先,本文将控制组样本单元合成权重看作随时间变化的函数,根据预干预期间干预单元潜在结果观测值与合成控制组潜在结果加权值均方误差最小确认预干预期的时变合成权重矩阵。该方法可以不受时间期数的限制,在研究期每一个时间节点上都能找到最接近干预单元的合成控制单元,极大地降低了干预单元潜在反事实结果的估计误差,提高了因果效应分析结果的准确性。

其次,考虑到合成控制法一般应用于面板数据,政策干预存在时间上的滞后效应,本文提出一种组合的灰色预测模型,组合全信息 $GM(1,1)$ 、新信息 $GM(1,1)$ 和新陈代谢 $GM(1,1)$ 模型,将预干预期时序数据按照 3:1 分为训练集和验证集,根据验证集模型误差最小自动选择最优方法来预测合成控制单元在干预期的时变权重,经验证本文所提出的方法更加稳健有效。

最后,应用该方法对我国碳排放权交易试点政策的效果评价进行实证分析。基于 2005 年至 2019 年的各省级行政区年度面板数据,根据时变合成组和试点地区碳排放强度差值,判断碳排放权交易是否有效抑制了该地区二氧化碳排放。

结果表明,本文提出的方法更适合处理面板数据的时间趋势,拥有更小的合成误差及更宽泛的应用场景,并且保持了防外推优势。应用该方法找到的合成控制组与干预单元在未被干预时潜在结果基本重合,模型预测结果通过安慰剂检验证明其更加精确。

关键词: 时变权重 合成控制法 灰色预测模型 低碳政策 碳排放权交易

Abstract

Synthetic control method approach evaluates the causal effect of the intervention or policy change by constructing a synthetic control model to estimate the potential outcome of the intervention group in the non-intervention state and comparing the difference in the intervention effect between the intervention group and the control group. This method solves the large sample constraints in the causal effect analysis based on counterfactual results by weighted average or appropriate linear combination of control group sample units to produce a composite control group. However, in the process of synthesizing the control group, the weight of the common weighting method remains unchanged in different periods. Due to the extension of the research period, the influence degree of each unit changes in different periods, and the synthetic weight of the unit should also change accordingly to be more in line with the practice.

First of all, the synthetic weight of sample units in the control group is regarded as a function of time change, and the time-varying synthetic weight matrix of the pre-intervention period is confirmed according to the minimum mean square error of the observed value of the potential outcome of the intervention unit during the pre-intervention period and the weighted value of the potential outcome of the synthetic control group. This method is not limited by the number of time periods, and can find the synthesis control unit closest to the intervention unit at every time node in the study period, which greatly reduces the estimation error of the potential counterfactual results of the intervention unit and improves the accuracy of the causal effect analysis results.

Secondly, considering that synthetic control method is generally applied to panel data and policy intervention has time lag effect, this paper proposes a combined gray prediction model, which combines full

information, new information and metabolic model, and divides the time series data of pre-intervention period into training set and validation set according to 3:1. According to the minimum model error of the verification set, the optimal method is automatically selected to predict the time-varying weight of the synthetic control unit in the intervention period, and the proposed method is more robust and effective.

Finally, applying this method to analyze and evaluate the effect of China's carbon emission trading pilot policy. Based on the annual panel data of each provincial administrative region from 2005 to 2019, the difference between the time-varying composite group and the carbon emission intensity of the pilot region is used to determine whether carbon emission trading has effectively suppressed the carbon dioxide emissions of the region.

The results show that the time-varying weight synthesis control method is more suitable for processing the time trend of panel data, has smaller synthesis error and wider application scenarios, and maintains the advantage of anti-extrapolation. The potential results of the synthetic control group found by this method were basically identical with those of the intervention unit without intervention, and the prediction results of the model were proved to be more accurate by placebo test.

Keywords: Time-varying weights; Synthetic control method; Grey prediction model; Low carbon policy; Carbon emission trading

目 录

1 绪 论	1
1.1 研究背景与研究意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	1
1.2 研究现状.....	2
1.2.1 国外研究现状.....	2
1.2.2 国内研究现状.....	4
1.2.3 文献述评.....	5
1.3 研究内容与研究框架.....	6
1.3.1 主要研究内容.....	6
1.3.2 技术路线图.....	8
1.4 创新点.....	9
2 相关理论基础	10
2.1 因果推断.....	10
2.1.1 基本概念.....	10
2.1.2 因果推断分析框架.....	10
2.2 反事实框架.....	11
2.3 随机对照试验.....	11
3 基于时变合成权重的合成控制因果推断分析	13
3.1 合成控制法的提出与发展.....	13
3.2 合成控制基本思想与模型设定.....	14
3.2.1 模型基础设定.....	14
3.2.2 权重估计.....	15
3.3 时变权重合成控制法.....	16
3.3.1 预干预期权重估计.....	16

3.3.2 时变权重预测方法.....	18
3.3.3 组合灰色预测模型.....	22
3.4 模型检验.....	24
3.4.1 误差分析.....	24
3.4.2 安慰剂检验.....	32
3.5 本章小结.....	33
4 “双碳”目标下中国碳排放权交易试点政策效果评价	35
4.1 中国碳减排政策发展规划.....	35
4.2 碳减排影响机制.....	36
4.2.1 碳减排有效措施.....	36
4.2.2 碳排放权交易政策减排路径.....	37
4.3 指标选取与数据来源.....	38
4.3.1 研究对象.....	38
4.3.2 结果变量.....	39
4.3.3 影响因素.....	42
4.4 结果分析.....	44
4.4.1 试点地区实证结果.....	44
4.4.2 影响因素与误差分析.....	50
4.4.3 政策效果与生效机制.....	51
4.5 本章小结.....	52
5 结论与展望.....	53
5.1 全文总结.....	53
5.2 研究展望.....	53
参考文献.....	55
攻读硕士学位期间承担的科研任务及主要成果.....	60
致 谢.....	61

1 绪 论

1.1 研究背景与研究意义

1.1.1 研究背景

在政府决策与管理过程中，制定政策促进决策实施，量化评价政策实际效果对后期政策调整有着重要意义。基于反事实框架的因果效应分析方法被广泛应用与自然科学和社会可选研究中，为此，传统定量分析方法可以概括分为控制实验、随机实验、自然实验或准实验三个大类（吴滨，2021），都是基于一定的假定条件建立模型后对政策实际效果进行模拟估计。随机化实验是反事实框架的因果效应分析的黄金标准，但大量社会科学研究基于观察数据，难以保证样本的随机性，未解决这一局限性，研究者提出了各种行之有效的“因果效应”评价方法，如，双重差分、倾向得分匹配、断点回归、合成控制、工具变量等。

合成控制（synthetic control, SC）以多个控制单元的凸组合，作为所建立的合成控制单元，由数据驱动选择权重来确定模型。基本思想是当观察单位是少量单元时，未受影响的单元的组合往往比任何单独的未受影响的单元提供了更合适的对照。合成控制法通过数据驱动将选择对照组的流程正式化，这也为反事实框架下小样本比较研究的定量推理模式打开了大门。Athey 和 Imbens 将该方法称作“过去 15 年中政策评估文献中最重要的创新”。但该方法仍旧要求不能有未被观察到的混杂因素以及较长的干预前时期，并且在单元合成时权重固定不变。为了解决“高维度，小样本”问题，以及传统面板数据随时间发生变化的现象，本文考虑个体合成权重在干预前时期内的时变性，再利用干预前的时变权重序列预测干预后的个体合成权重，并将该方法用于“双碳”政策的效果评价中以检验其解决实际问题的可行性。

1.1.2 研究意义

本文的研究目的是将时变系数和合成控制的思想结合，提出一种精度更高的

因果推断方法，并将其用在实际问题中检验其解决实际问题的能力。合成控制法由于其可解释性和透明性，被广泛应用于经济学和社会科学的实证研究中，研究者在应用过程中也针对实际数据条件对其进行改进。但查阅文献发现，目前的方法改进中，还没有从合成权重时变性这一角度来进行过探讨，本文的研究对此进行了补充，具有一定的理论意义。

政策评价方法在假定条件下进行模拟，通常具有较严苛的应用条件，而经济事件往往是非实验设计的结果，因此长期以来该类方法在经济领域的应用具有一定难度。并且政策评价使用的数据基本为面板数据，为了结果的准确性，时期跨度要求较高。使用随时间变化的合成权重来改进原始合成控制法，可以减小干预前合成单元与干预单元之间的实际偏差，提高干预后政策评价的精度，并通过对影响因素权重的分析在后续政策调整时提供更合理的建议，具有一定现实意义。

1.2 研究现状

1.2.1 国外研究现状

合成控制法最先由 Abadie (2010) 等提出来估计加州烟草控制计划的效果，他们把合成控制估计法引入计量经济模型，使得以实验为依据的研究所通常使用的固定效应模型得以进一步推广 (吴滨, 2021)。此后根据处理的实际问题被不同学者加以完善和改进。例如, Dube (2015) 和 Kreif (2016) 等在应用过程中将其扩展到了多个处理单元。Doudchenko (2016) 等放宽权重限制, 允许个体合成权重为负, 并且不限制权重之和必须等于 1, 还允许合成控制估计量的水平发生恒定变化。Xu (2017) 针对混杂因素进行改进, 提出可以同时处理多个干预单元和多个政策断点的广义合成控制法。

近些年来, 合成控制法不断地得到改进与延伸。Abadie (2019) 等再提出了一种针对分类数据的惩罚性合成控制估计量, 结合了对处理单元和对其合成控制有贡献的每个单元之间的成对匹配差异的惩罚, 同样可以处理多个干预单元。Chernozhukov (2019) 等在合成控制的推理过程中考虑了不同的惩罚项, 提出一种精确稳健的保形推理方法, 该方法易于实现, 对非平稳数据有效, 并且具有良

好的小样本性能。Ben-Michael (2021) 等根据 Rubin (1973) 和 Quade (1982) 等的偏差校正技术, 对合成控制估计量进行修改, 使用未经处理的单元构建合成控制反事实, 用回归调整来减弱合成控制估计量的偏差。Gunsilius (2020) 则提出了分布式合成控制, 将合成控制推广到概率测度。Arkhangelsky (2021) 等结合合成控制法和双重差分法的思想, 提出了合成双差分模型。该方法不仅对控制组中的单元进行加权, 还对干预前的时间段进行加权, 不再要求数据必须遵循平行趋势假设。

此外, 合成控制作为数据驱动的因果推断方法, 不断被应用于各类因果效应分析的案例中。但在实际使用过程中, 学者们也发现了其局限, 并针对这些局限性提出了实际应用过程中应注意到的一系列事项和优化方法。例如, Kaul (2015)、Abadie (2015)、Botosaru (2019) 和 Irene (2019) 等分别在不同时期不同案例中, 在假定预处理期无限的情况下, 渐进地推导并论证了合成控制法的许多理论性质。一方面只要预干预期足够长, 且在这期间潜在结果存在完美平衡, 那么就不再要求必须找到完美的协变量平衡才能使模型有较高精度。另一方面则是不应该在所有可用的预干预期结果中同时使用协变量, 否则某些条件下协变量的影响程度可能会被降低。Hahn (2017) 和 Firpo (2018) 等在合成控制的背景下分别研究了排列检验的性质, 并提出了排列检验的几个推广。从重复抽样的角度分析了排列检验, 讨论了几种可供选择的方法, 并在检验结果的基础上贡献了置信集。Kinn (2018) 强调合成控制应考虑背景和数据性质, 否则贸然地应用该方法做评估是充满风险的。他根据合成控制估计量的理论基础, 分析提供了恰当适用该方法来进行因果效应估计的具体条件, 并解释了在这些条件不成立时使用者应该如何调整。Samartsidis (2019) 等从少数干预单元, 观察面板数据评估二元干预的因果关系, 回顾了因果推断的现有方法, 并且详细介绍了每种方法背后的假设和不同方法之间的联系。Li (2020) 利用投影理论, 推导了合成控制估计量和改进的综合控制 ATE 估计量的渐近分布, 证明了适当设计的子抽样方法可以用来获得置信区间并进行推理, 并以检验电子零售商开设实体陈列室的效果作为实证, 证明了方法的实用性。最后, Abadie (2021) 再次整合讨论了合成控制框架作为一种比较案例研究方法的优点, 具体地描述了提供可靠估计量以及可能应用失败

的各类情形。

1.2.2 国内研究现状

合成控制法的方法改进和理论基础论证在国外研究文献或国内学者发表的英文文献中居多，国内中文文献多是应用该方法来处理经济、社会、生态、发展、卫生等方面的实际问题。例如，姚永玲（2023）等采用合成控制法处理 2002—2020 年的面板数据，从创新、协调、绿色、开放和共享 5 个维度来识别北京城市副中心建设对迁入地高质量发展的影响，并对高质量发展不同维度的变化特点，提供针对性措施建议。任以胜（2023）等以中国首个跨省流域生态补偿机制试点区域新安江流域为研究对象，运用合成控制法评估新安江流域生态补偿政策对 2012—2020 年黄山市水污染强度的影响，发现生态补偿政策有效降低了黄山市水污染强度，效果具有一定的持续性但逐渐减弱，并利用时间安慰剂、空间安慰剂等进行了稳健性检验。蔡芸（2023）等以上海、北京、重庆、天津、深圳 5 个超大城市的公共卫生系统韧性作为考察项，采用合成控制法来评价流动人口卫生服务均等化试点政策对超大城市公共卫生系统韧性的影响。结果表明流动人口均等化试点政策对城市公共卫生系统韧性的影响随着时间的推进存在地区差异性。人口压力较大的城市，均等化试点政策会在一定程度上降低城市公共卫生系统韧性。覃若兰（2023）等从大科学装置集群、创新资源要素集聚、知识溢出和技术扩散、协同创新系统构建和创新生态系统构建等方面，探究综合性国家科学中心对区域创新能力的影响机理。他们以 2009—2021 年 22 个省级行政区的面板数据为基础，利用合成控制法验证了综合性国家科学中心对创新能力有一定的提升效应，并提出加大综合性国家科学中心的建设投入力度、促进创新链和产业链深度融合等政策启示。

当然，合成控制法的一系列改进算法和组合算法也同样被广泛应用。其中广义合成控制法因其能同时处理多个干预单元和多个政策断点，应用尤为广泛。例如，白仲林（2019）等根据 2014 至 2017 年中国各城市的月度面板数据，利用广义合成控制法研究发现了限购政策能够有效地抑制商品房价格过快上涨，对支持实体经济发展存在较显著的“回馈效应”，并且其效果因建筑面积和城市规模

等特征存在差异性。林辉（2022）等基于 1998-2018 年世界各国的融资数据,运用广义合成控制法研究了“一带一路”合作协议的签署对沿线国家产生的政策效应,实证发现同中国签署“一带一路”合作协议确实能够切实地提升沿线国家的融资环境。黄秀路（2023）等基于 2007—2020 年中国城市面板数据,利用广义合成控制法评估了绿色金融改革创新试验区设立对能源强度的影响效应,并分析其影响机制,提出充分考虑不同地区在制度环境、资源禀赋等方面的差异是进一步完善绿色金融改革创新试验区建设的关键。邢子怡（2023）依据 2001-2021 年 97 个设立国家高新区的城市和 131 个无国家高新区设立的城市的面板数据,采用广义合成控制法估计了国家高新区设立对城市创新的平均政策处理效应,证明了国家高新区的设立能够有效地提升城市整体创新能力,并结合倾向得分匹配和基于面板数据的匹配方法等因果推断方法验证了上述结果的稳健性。

此外,合成双重差分法自 2021 年被提出后,也逐步受到青睐。例如,刘传明（2019）等基于 1995 年至 2016 年中国省级面板数据,采用合成双重差分方法对碳交易试点的减排效应进行稳健性检验,得出碳排放权交易政策显著减少二氧化碳排放的结论。吴志祥（2024）等基于 2005 年至 2020 年长三角地区 153 个县域单元的面板数据,将合成双重差分法与渐进双重差分法结合,实证检验了区域一体化能显著促进县域经济发展活力提升,分析其促进效果主要依赖于经济联系机制、产业结构升级机制以及市场活力机制的发挥程度,并因地区差异、行政约束、工业基础等因素而有所不同。何露雪（2023）等基于因果推断理论,以珠三角地区受到应用研究项目资助和基础研究项目资助的企业为样本,运用合成双重差分法和双重差分法探究了科技专项资助与企业科技成果产出的关系,发现应用研究项目和基础研究项目资助都显著提升了企业的科研成果产出数量,并通过平行趋势检验、安慰剂检验、更换被解释变量检验、异质性检验等一系列方法论证了结论的稳健性。

1.2.3 文献述评

合成控制法自被提出开始,一直是因果推断的主流方法。与传统的回归方法相比,合成控制由于数据主导和权重限制等方式拥有透明度和防外推这两个特

性,并且由于合成控制的选取并不要求通过处理后的结果进行选择,因此研究人员在研究过程中设计方案时可以避免出现决策受到研究结论的影响。合成控制法作为双重差分模型的拓展,允许混杂因素的影响可随时间变化,但可观测变量对合成组的贡献值及其合成权重仍旧是根据干预前时期计算出的固定值。而作为经常被用来处理时序面板数据的方法,合成控制法可观测变量的时变性实际上是很有必要进行考虑的。

以往的研究学者们致力于通过采用调整惩罚项、偏差校正、投影理论等技术来改进合成控制法,提高其适用性。但他们的研究重心在不可观测混杂因素上,可观测协变量的合成权重仍为固定权重。另外理论上预干预期足够长可以无限提高合成精度,但在应用该方法解决实际问题时,受客观因素限制面板数据期数往往达不到要求。因此,将时变系数和组合灰色预测模型引入合成控制中,可以有效解决上述高维度、小样本以及面板数据随时间变化的问题。

1.3 研究内容与研究框架

1.3.1 主要研究内容

本文基于面板数据,考虑控制组单元在长期的干预前时期内,由于区域发展水平不同,其单元的潜在结果应随着时间而变化。因此在合成对照单元时,合成权重也应随时间变化,而非在干预前后的长时间内都是固定值。针对上述问题,本文提出变权重的合成控制法,并在实际应用中检验其解决实际经济学问题或社会科学问题的效果。主要包括以下3个重点研究内容:

其一,针对面板数据,考虑合成控制法中个体合成权重的时变性。在预干预期改进潜在结果的因子模型,将固定个体合成权重改进为时变权重矩阵。通过机器学习和最优化算法,根据核心影响因素相近以及合成组与干预组均方根误差最小,对干预前权重矩阵进行求解和检验。

其二,根据每个合成单元在验证后的预干预期时变合成权重值的特征,预测干预后的个体合成权重矩阵。考虑到干预前均为年度数据,且期数较少,采用全信息 $GM(1,1)$ 、新信息 $GM(1,1)$ 和新陈代谢 $GM(1,1)$ 组合的灰色预测模型算法,

预测干预期合成权重值。将每一个合成分量预干预期数据前 75% 作为模型训练集，后 25% 作为模型验证集，根据验证后模型误差最小来选择最优模型，从而预测干预后的时变权重，将预测结果和合成组潜在结果加权，得到干预组的反事实潜在结果。

最后，将以上两步结合成的变权重合成控制法用于我国 6 个省级行政区的“双碳”目标试点政策效果评价中，检验改进后的合成控制法解决实际问题的能力。根据问题分析影响碳排放的因素和“双碳”政策的具体措施，选定经济水平、能源结构、产业结构、城市化水平和技术进步水平 5 个方面的解释变量，碳排放强度则作为被解释变量。根据人均 GDP、人均资本存量、能源消费在 GDP 中的占比、化石能源消费占比、第二产业增加值占比、产业结构高级化水平、人口密度、人均私人汽车拥有量和研究与试验发展（R&D）经费投入强度 9 个拟定指标收集数据，并对变量进行预处理，缩小不同量纲的影响，并且弱化数据序列可能存在的异方差。通过相关性和回归系数检验，确定显著影响碳排放的变量为最终的解释变量。以其他没有实施相同或相似措施的城市作为控制组，利用改进后的方法确定权重合成城市，以“合成城市”和实际城市的被解释变量差值来分析试点省级行政区试行排放放权交易机制对抑制碳排放的实际效果。同时根据实证结果也可以检查该方法在处理实际问题上的可行性。

1.3.2 技术路线图

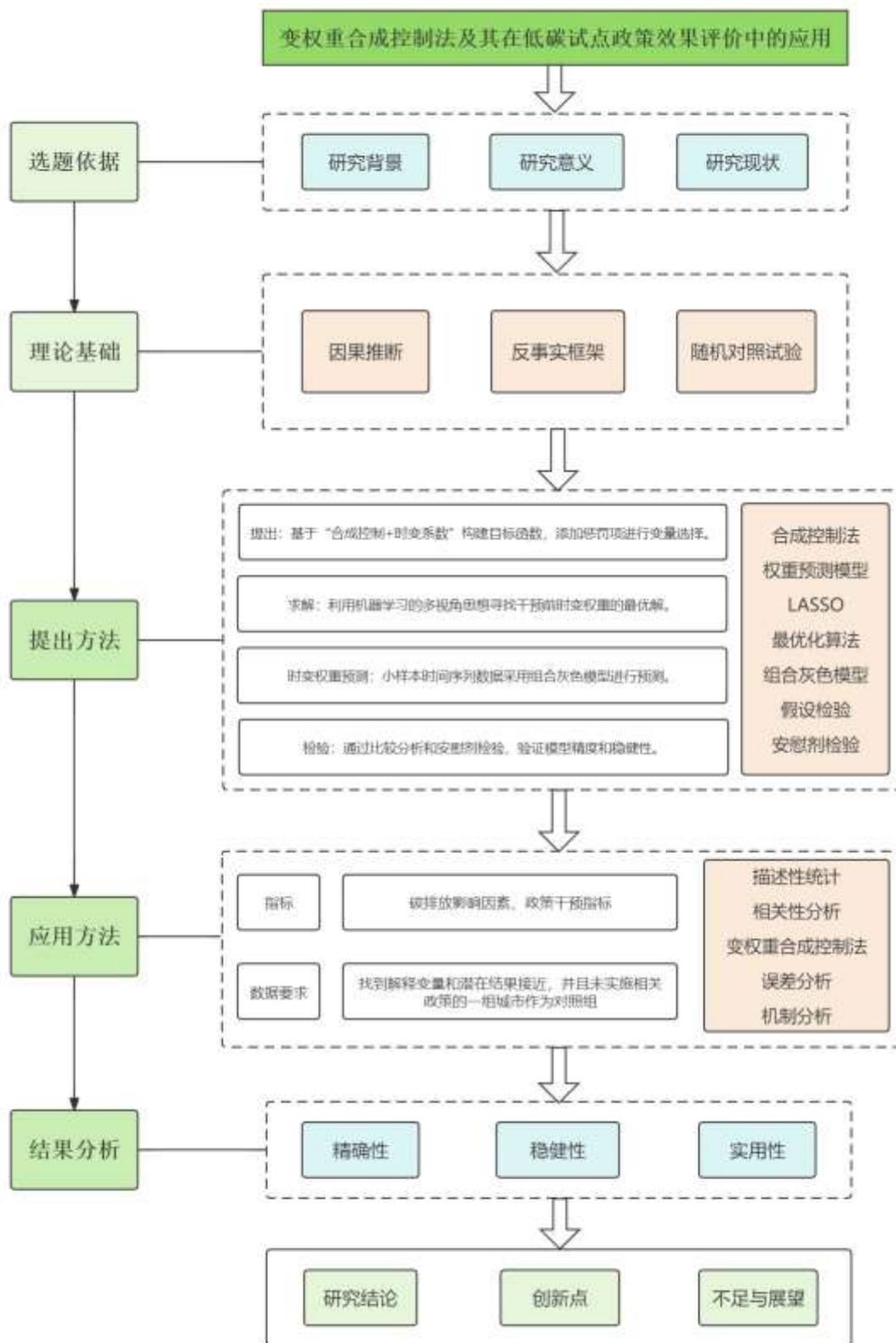


图 1.1 技术路线图

1.4 创新点

本文的目标是提出一种可靠的时变权重合成控制方法,通过与传统合成控制法对比,证明其结果拥有更高的精度,并在实际问题验证其可用。目前国内外学者们对合成控制法的研究改进,基本集中在如何处理未被观测的协变量,以及如何对不同因果推断方法进行综合应用上,较少对权重直接设限。本文的创新点在于根据显著影响因素计算个体合成权重时,不采用固定权重值,而是增加随时间变化这一特性,在方法上进行一定的改进。研究包括两部分重点内容,一是时变权重合成控制模型的构建和检验,二是将模型用于实际案例中分析其可用性。

时变权重合成控制法的构建分为两个步骤:一是预干预期时变合成权重的计算和检验,在构建目标函数时将固定权重改进为时变权重;二是干预期时变权重的预测,根据个体权重分量的时序特征,采用组合灰色模型进行权重预测。首先引入时变因素,将个体合成权重由固定向量 ω_j 改为时变矩阵 $\omega_j(t)$ 从而构建新的目标函数,同时添加惩罚项进行主要特征选取。利用机器学习找到预干预期内干预单元和控制组合成单元之间的均方根预测误差(MSPE)最小的权重作为预干预期个体合成权重的最优解。其次,根据求解出的时变合成权重和时序潜在结果的数据特征,采用改进的组合灰色预测模型从全数据GM(1,1)、新数据GM(1,1)以及新陈代谢GM(1,1)中自动选择最优方法来计算干预期权重分量,并通过权重非负且和为1的限制保证其防外推特性。最后,通过加州禁烟这一典型案例数据对时变权重合成控制法和传统合成控制法的应用结果进行对比检验方法可靠性,通过安慰剂检验进一步确认方法稳健性。

2 相关理论基础

2.1 因果推断

2.1.1 基本概念

首先给出因果推断中的三个基本概念：单元、干预和结果（王东明，2022）。单元是干预效果评价中的最小研究对象，在下文的加州禁烟案例中为各个州，在低碳试点案例中为我国的各个省级行政区。干预是对某一个或多个单元施加的扰动，在加州禁烟案例中为禁烟政策，在低碳试点案例中为6个试点省市的碳排放权交易政策。结果指潜在结果，是每个单元被干预或不被干预时，能实际观察到或不能实际观察到的结果变量的值。

根据干预措施，潜在结果可以分为观测结果和反事实结果两类。两类结果的差值即为施加干预措施后产生的效果，也称其为因果效应。

2.1.2 因果推断分析框架

自从相关系数的概念被提出，在因果的概念被明确界定之前，相关性和因果性常常是互换使用的，但是二者实际却有着本质的区别。在处理实际问题时，某些简单情况下，研究人员可以根据一般经验来判断或者确定干预效果和影响因素之间是否为因果关系和有多大的因果关系。然而当变量众多、问题复杂时，如何判断因果关系并使结果置信，就成了一件困难的事情。

因果性的一个重要特点是关系的单向方向性。因果推断则是根据某一效应发生的条件，对其因果关系作出结论的过程。因果推断的三种分析框架为反事实框架，潜在结果模型和结构因果模型。反事实框架可以看做是因果效应的开端，潜在结果模型和结构因果模型则是基于反事实的两个不同方向的因果推断分析框架（马忠贵，2022）。潜在结果模型在干预条件和结果变量已知的前提下，定量分析干预对潜在结果的因果效应，并对缺失的潜在结果进行补齐，使观察性研究的效果接近试验性研究。结构因果模型则是基于图论将干预和反事实潜在结果的因果关系转化成可以通过统计学来解决的问题。

2.2 反事实框架

反事实结果是干预的潜在结果。现实中,同一单元只能选择接受或者不接受干预,因此在随机对照实验中,无论是干预组还是对照组,都只能观测到接受干预的潜在结果和不接受干预的潜在结果其中之一,另一个未观察到的潜在结果即为反事实结果,即反事实结果实际为潜在结果中未被观测到的一项。

反事实潜在结果,指的是单元在我们设定的干预条件下,如果没有受到此项干预将会观测到的结果。其核心思想可以用一个例子来说明:

一位癌症患者每天使用一种新的抗癌药物,并记录他的癌细胞扩散程度 $Y_{D_i=1}$;如果他没有使用药物,自然地认为他应当有另一个癌细胞扩散程度值 $Y_{D_i=0}$ 。为了衡量新的抗癌药物与癌细胞扩散程度变化之间的因果效应,需要比较同一个人在使用和不使用新药物两种情况下的结果,即 $Y_{D_i=1}$ 和 $Y_{D_i=0}$ 。显然,患者同一天只能要么使用新药,要么不使用新药,即 D_i 只能有一个值。也就是说同时记录到 $Y_{D_i=1}$ 和 $Y_{D_i=0}$ 是不可能的,其中一个值必然会缺失。对于上述这名患者而言, $Y_{D_i=1}$ 是实际可观测到的结果, $Y_{D_i=0}$ 是需要估计的反事实结果,两者的差值则是患者使用新的抗癌药物所产生的有效治疗效果,也即因果效应。

与建立相关关系的统计模型不同,反事实框架下的因果推断不必在模型设定上增加太多假定条件,给应用者留下了更多诠释和补充空间(刘学,2020)。并且由于反事实框架中因果推断的各种具体方法之间存在联系和递进关系,研究者应用过程中可以根据实证需要充分考虑方法的选择、搭配和佐证。除此之外,反事实框架下的因果推断在效应评估之后,可以根据结果做机制分析,特别适用于认知行为相关的研究。

2.3 随机对照试验

推断因果关系最有效的方法是进行随机对照试验,也称自然实验,它可以说是因果推断模型中的标准原则。随机对照试验的基本方法是将研究对象随机分组,分配过程与实验参与者的特征无关,再对两个组实施不同的干预措施或者只对某一个组采取干预措施,从而对照效果的不同。随机对照试验能够最大程度地

避免试验在设计和实施中可能出现的各种偏差，平衡不可观测混杂因素，提高统计学检验的有效性，被公认为是评价干预措施的黄金标准。

但在随机对照试验过程中，如果完全忽略协变量的作用，仅使用比较分析来进行因果关系推断，分析结果可能会产生偏差。这些影响因果关系估计的协变量被统称为混杂因素。为了得到准确的因果推断，理想状态下将所有实验对象随机分配到干预组和对照组，再设计将正在研究的结果变量作为对照组和干预组在研究期内的唯一差异项，其余条件则保持一致，确保干预组和对照组的结果可比性。

然而实际情况中，研究者往往会寻找自然发生的实验或基于现有的观察数据来构造实验条件，要做到“完全随机”和“其余条件完全保持一致”几乎是不可能的。如何利用非随机的观测数据进行统计推断成了亟待解决的关键问题。因此基于观测数据进行一系列准实验的方法应运而生。研究者们转而致力于保证对潜在结果有显著影响的变量保持一致性，尽量弱化或消除其它不可观测混杂因素对潜在结果造成的影响差异。

3 基于时变合成权重的合成控制因果推断分析

3.1 合成控制法的提出与发展

尽管大多数数据科学都集中于使用统计学和机器学习的工具来进行预测,但许多问题根源都是因果关系。尤其在“相关系数”的概念被提出以后,相关性和因果性常常被互换使用,但不可忽略的是二者有本质的区别。相较于相关性而言,因果性最重要的一个特点是关系单向性,而关联性则是双向的。根据因果性的特点,因果推断即为根据某一效应发生的条件,对某一因果关系作出结论。因果推断打破了对数据相关关系的盲目迷信,强调在数据挖掘的基础上建立因果模型以量化因果效应。它在经济学、社会科学以及计算机科学等诸多领域都展现出了蓬勃的发展趋势,研究和挖掘因果关系的科学方法对各学科都具有一定的普适性。

合成控制法于 2010 年由 Abadie 等提出来估计加州烟草控制计划的效果,它以多个控制单元的凸组合作为所建立的合成控制单元,由数据驱动选择权重来确定模型,把合成控制估计法引入到计量经济模型中,使通常使用的以实验为依据的双重差分法得到进一步推广。自提出以来,合成控制法一直是因果推断的主流方法,同时学者们也针对应用过程中出现的问题不断对方法进行补充和改进。例如, Abadie、Dube 和 Kreif 等在应用过程中将其扩展到了多个处理单元。Doudchenko 等对权重放宽限制,允许权重为负且权重之和可以不等于 1,并且允许合成控制估计量的水平发生恒定变化。Chernozhukov 等在合成控制法的推理过程中针对未被观测因素考虑不同的惩罚项,提出一种精确稳健的保形推理方法,对非平稳数据有效且具有良好的小样本性能。Xu 针对混杂因素进行分析,将交互式固定效应模型与合成控制相结合提出广义合成控制法,使用结果的参数因子模型来预测反事实结果,可以同时处理多个干预单元和多个政策断点。Arkhangelsky 等结合合成控制法和双重差分的思想提出了合成双差分模型,不仅对控制组中的单元进行加权,还对干预前的时间段进行加权,不再要求遵循平行趋势假设。Gunsilius 将合成控制法推广到概率测度,提出了分布式合成控制方法。此外, Abadie 还和 Ben-Michael 等利用偏差校正技术,对合成控制估计量进行修改,使用回归调整来减弱合成控制估计量的偏差。

合成控制法通过数据驱动将选择对照组的流程正式化, 由于其可解释性和透明性, 该方法被广泛应用于经济学和社会科学的实证研究中, 针对的是长时期随时间发生变化的面板数据。但整理文献发现目前合成控制法的改进方向多是针对如何控制未被观测的混杂因素以及如何同时处理多个干预单元, 在权重上也倾向于考虑放弃“防外推”的特点, 放松权重要求。为了让合成控制法更好的处理时序面板数据, 本文考虑在合成控制中引入个体合成权重的时变性以提高合成精度减少评价误差。

3.2 合成控制基本思想与模型设定

合成控制法的基本思想是当观察单位是少量单元时, 未受影响的单元的组合往往比任何单独的单元提供了更合适的对照。使用该方法的前提是不能有未被观察到的混杂因素以及需要较长的干预前时期, 并且在干预前时期内进行单元合成时, 合成权重固定不变。

3.2.1 模型基础设定

设定观察到 $J+1$ 个地区, 不丧失普遍性的情况下, 只有第 1 个区域受到政策干预, 且该地区在 T_0 时期后不间断地被干预影响, 第 2 到 $J+1$ 个没有受到干预的地区组成控制组库。

任意单元 $i (i=1, 2, \dots, J+1)$ 在任一时间点 $t (t=1, 2, \dots, T_0, \dots, T)$ 的观测值表示为:

$$Y_{it} = Y_{it}^N + \alpha_{it} D_{it} \quad (3.1)$$

$$\text{其中 } D_{it} = \begin{cases} 1, & i=1 \text{ 且 } t > T_0, \\ 0, & \text{其它} \end{cases}$$

Y_{it}^N 表示区域 i 在 t 时刻没有受到政策干预时的潜在结果; D_{it} 为政策变量, 仅在单元 i 在 t 时刻受到政策干预时取值为 1, 其余情况取值为 0。因此政策评价的目标在于估计第 1 个地区在 T_0+1 至 T 时期的干预效果:

$$\alpha_{it} = Y_{it}^I - Y_{it}^N, \quad (3.2)$$

Y_{it}^I 表示区域 i 在 t 时刻受到政策干预时的潜在结果。一个地区在同一时间点只能处于被干预或未被干预其中一个状态，即 Y_{it}^I 和 Y_{it}^N 不可能同时被观测到，所以评价政策效果的关键在于如何得到被干预单元在干预期内假如没有受到干预的反事实结果 $(Y_{i70+1}^N, \dots, Y_{iT}^N)$ 。

假设 Y_{it}^N 由下列因子模型给出：

$$Y_{it}^N = \delta_t + \theta_t X_i + \mu_i \lambda_t + \varepsilon_{it}, \quad (3.3)$$

其中， δ_t 是未知公共因子，同一时间点对所有单元具有相同的影响。 X_i 是不受政策干预影响的 $(Q \times 1)$ 维可观测协变量向量 $(X_{i1}, \dots, X_{iQ})'$ ， θ_t 是其随时间变化的 $(1 \times Q)$ 维系数向量。 $\mu_i \lambda_t$ 是可观测协变量以外的混杂因素影响，控制不同单元的空间相关性，其中 λ_t 是不可观测时变因素， μ_i 为其系数向量。随机误差项 ε_{it} 是地区受到的暂时性冲击，不可观测，在地区层面满足 $E(\varepsilon_{it}) = 0$ 。

3.2.2 权重估计

为了对因果效应进行量化估计，对控制组库内的区域进行线性组合，合成一个虚拟的对照组，用来代替干预地区没有接受干预情况的反事实结果。

假设存在 $(1 \times J)$ 维的权重向量 $W = (\omega_2, \dots, \omega_{J+1})$ ，满足 $\omega_j \geq 0$ 且 $\sum_{j=2}^{J+1} \omega_j = 1$ 。

向量 W 的每组特定值代表一个潜在的合成控制，决定控制组的加权平均值，即和干预地区作对照的合成单元的潜在结果。干预地区与合成对照组的结果变量之间的差值，即为政策评价关心的因果效应值。因此合成权重 W 的选择对政策评价的结果起决定性作用。

为了得到最优的权重向量 $W^* = (\omega_2^*, \dots, \omega_{J+1}^*)$ ，要求：

$$\sum_{j=2}^{J+1} \omega_j^* Y_{jt} = Y_{it}, \quad (3.4)$$

即在预干预期 $(t = 1, 2, \dots, T_0)$ 内，干预地区的潜在结果观测值和合成单元的潜

在结果观测值加权平均值相同。只有当 0 至 T_0 时期合成单元的结果和干预区域未被干预时观察结果相等或差距在可接受范围内时，后面预测的干预效果才可信。

同时，为了将特征差异明显地区因插值而造成的偏差最小化，通过可观测协变量 X_i 将合成单元限制在与干预地区具有相似特征的地区内。要求

$$\sum_{j=2}^{J+1} \omega_j^* X_j = X_1, \quad (3.5)$$

即干预地区的协变量观测值和合成单元的协变量观测值加权平均值相同。

但是由于混杂因素和随机干扰的存在，合成单元和干预地区一般情况下无法达到完全一致，即通常没有一组权值能使得式 (3.4)、式 (3.5) 的等号在数据中完全成立。此时，一般是选择向量 W ，使得上述方程在可接受的误差之内近似成立，潜在结果欧式距离达到最小值时得到最优解 W^* ：

$$\begin{aligned} \min_{W, W^*} \sum_{t=1}^{T_0} \left\| Y_{1t} - \sum_{j=2}^J \omega_j Y_{jt} \right\|_2 \\ \text{s.t. } \omega_j \geq 0, \sum_{j=2}^{J+1} \omega_j = 1. \end{aligned} \quad (3.6)$$

以 $\sum_{j=2}^J \omega_j^* Y_{jt}^N$ 作为 $T_0 + 1$ 至 T 时期的反事实结果 Y_{1t}^N 。

由目标函数 (3.6) 求解出的权重向量 $W^* = (\omega_2^*, \dots, \omega_{J+1}^*)$ ，表示在 0 至 T 时期内，时间层面上个体合成权重为固定值，不随时间发生变化，这与时序数据的时变性质相悖。考虑合成控制法多被用来解决经济学和社会科学实证问题，处理的是随时间波动的面板数据，因此本文将目标函数中的权重向量改进为权重矩阵，来描述合成权重随时间变化这一性质。

3.3 时变权重合成控制法

3.3.1 预干预期权重估计

沿用 3.2.1 节的模型基础设定，引入时变合成权重 $\omega_j(t)$ 构建新的目标函数：

$$\begin{aligned} \min_{W, W^*} \sum_{t=1}^{T_0} \left\| Y_{1t} - \sum_{j=2}^J \omega_j(t) Y_{jt} \right\|_2 \\ \text{s.t. } \omega_j \geq 0, \sum_{j=2}^{J+1} \omega_j = 1. \end{aligned} \quad (3.7)$$

其中， $\omega_j(t)$ 值与时间点有关，求解后应为 $(T_0 \times J)$ 维矩阵

$$W_{T_0}^* = \begin{bmatrix} \omega_{1,2}^* & \cdots & \omega_{1,J+1}^* \\ \vdots & \ddots & \vdots \\ \omega_{T_0,2}^* & \cdots & \omega_{T_0,J+1}^* \end{bmatrix}。 \text{注意，} \omega_j(t) Y_{jt} \text{ 仅计算各控制组单元在同一时间点的潜}$$

在结果线性加权。同时引入 LASSO 惩罚因子，将不显著的合成单元权重系数压缩为 0，进行变量选择：

$$\begin{aligned} \min_{W, W^*} \sum_{t=1}^{T_0} \left\| Y_{1t} - \sum_{j=2}^J \omega_j(t) Y_{jt} \right\|_2 + \sum_{t=1}^{T_0} \lambda_t |\omega_j(t)| \\ \text{s.t. } \omega_j \geq 0, \sum_{j=2}^{J+1} \omega_j = 1. \end{aligned} \quad (3.8)$$

由此，第 1 个单元在 T_0 时期后假设没有受到政策干预的反事实结果为

$$Y_{1t}^N = \sum_{j=2}^J \omega_j^*(t) Y_{jt}^N。$$

由因子模型 (3.3) 可知，方法中设定 X_{iq} 和 Y_{it} 间是线性相关的。为了减少插值偏差，在求解 $W_{T_0}^*$ 之前通过式 (3.5) 将合成单元限制在 X_i 的值与干预地区相似的这些合成组地区内。引入矩阵 V 测量 X_i 之间的差异：

$$\begin{aligned} \left\| X_1 - \sum_{j=2}^{J+1} \omega_j^* X_j \right\| &= \sqrt{\left(X_1 - \sum_{j=2}^{J+1} \omega_j^* X_j \right)' V \left(X_1 - \sum_{j=2}^{J+1} \omega_j^* X_j \right)} \\ &= \sqrt{\sum_{q=1}^Q v_q \left(X_{1q} - \sum_{j=2}^{J+1} \omega_j^* X_{jq} \right)^2}. \end{aligned} \quad (3.9)$$

其中， V 限定为非负对角阵，表示各解释变量之间不相关， v_q 为对角线上的值。

X_j 为各解释变量在 T_0 时期前的平均值。

根据潜在结果距离最小将上述权重矩阵求解转为线性规划求参问题，计算固定合成权重 $W^*(V)$ 和最优解释变量权重 V ：

$$\begin{aligned} \min_{w, W^*(V)} & \left(X_1 - \sum_{j=2}^{J+1} \omega_j^* X_j \right)' V \left(X_1 - \sum_{j=2}^{J+1} \omega_j^* X_j \right) \\ \text{s.t.} & \quad \omega_j^* \geq 0, \sum_{j=2}^{J+1} \omega_j^* = 1. \end{aligned} \quad (3.10)$$

最终根据影响因素相近的原则,从控制组库中选择 K 个地区作为合成单元赋予时变权重。在 0 到 T_0 时期,根据潜在结果差异最小:

$$\begin{aligned} \min_{W, W_{T_0}^*} & \sum_{t=1}^{T_0} \left\| Y_{1t} - \sum_{k=2}^{K+1} \omega_k(t) Y_{kt} \right\|_2 \\ \text{s.t.} & \quad \omega_k(t) \geq 0, \sum_{k=2}^{K+1} \omega_k(t) = 1. \end{aligned} \quad (3.11)$$

确定最优时变合成权重 $W_{K \times T_0}^* = (\omega_1^*, \dots, \omega_{T_0}^*)'$ 。

3.3.2 时变权重预测方法

干预前的时变合成权重数据本质上是时序数据,反映了权重随时间不断变化的趋势。如果干预前期时间数较多或者包含季度数据,采用时间序列分析方法可以进行有效预测。

时间序列预测问题的核心是从数据中挖掘出权重随时间不断变化的规律,并利用其对干预后的合成权重数据做出估计。经典的时间序列模型包括移动平均模型(moving average model, MA)、自回归模型(auto-regressive model, AR)、自回归移动平均模型(auto-regressive moving average model, ARMA)(杨海民, 2018)。在已经引入差分过程、季节性、周期性以后,学者们又拓展出了差分自回归移动平均模型(auto-regressive integrated moving average, ARIMA)、季节性自回归移动平均模型(seasonal auto regressive integrated moving average, SARIMA)、季节性自回归集成移动平均模型(seasonal auto-regressive integrated moving average with exogenous variables, SARIMAX)和基于机器学习的神经网络模型(neural networks, NN)等。此外还有很多研究者针对特定问题,根据时序数据性质延伸出了一系列经典模型的组合和新的方法,这些方法已被运用到各类经济问题和社会问题中。例如,郑艳妮(2023)等基于 SARIMA、Holt-winters 与 Prophet 三种时间序列模型,预测了 2021 年武威市丙型肝炎月发病数,为武威市丙型肝炎防控政策的制定提供依据。赵凌霄等(2023)基于 EMD 和 Cat Boost

算法改进时间序列模型，提出了 EMD-Cat Boost-ARIMA 模型。该方法减少了 PM2.5 时间序列的非线性、高噪声、不平稳与波动性对预测的影响，能更精确的预测了 PM2.5 浓度，并以大连市 PM2.5 预测为例证明其相比较传统时间序列模型具有较高性能。

传统的时间序列预测方法主要是在确定时间序列参数模型的基础上求解出模型参数，并利用求解出的模型完成预测工作。例如 Box (1990) 等提出的“Box-Jenkins”方法一直非常流行，直到现在仍被不少学者参考引用。该方法对于给定的时间序列先确定适当参数值，然后通过有效估计算法计算具体的参数值，最后根据模拟结果检验参数准确性并进行改进。另外，最大似然估计法也经常用来进行时间序列模型参数的估计，根据观察到的样本值构建关于模型参数的概率密度，求解出使其最大的模型参数值 (Hamilton, 1994)，但该方法要求噪声符合一个具体分布。状态空间方程也可以用来实现时间序列的预测，把回归模型表示成动态系统方程形式，通过卡尔曼滤波来进行求解 (Durbin, 2012)。此外，随着计算机网络技术的广泛应用和普及，数据规模的急剧增长，深度学习也被用作实现时间序列预测的有效工具。例如，Anava (2013) 等提出了基于参数模型的在线时间序列预测算法 ARMA-ONS，通过在线算法对 AR 模型参数进行求解，并随新时间序列数据的到来更新模型参数。Binkowski (2017) 等受到标准的 AR 模型和 RNN 中对原始输出赋予不同的权重后再输出的门机制思路的启发，提出了一种 SOCNN 模型来解决多元非同步时间序列数据的预测问题。

如果干预前数据量不够多，那么对于小样本数据，如果需要预测的干预后时期较长，传统的时间序列预测方法可能会造成较大的偏差，因此学者们针对这类情况提出通过少量不完全信息建立数学模型做出预测的灰色预测系列方法。

灰色预测模型 (grey prediction model, GM) 是一种基于灰色系统理论的预测方法，它补充了小样本的时间序列预测领域，在过去几十年里得到了广泛的研究和应用。其预测对象要求数据具有准指数规律，并且数据非负，适用于年份数据预测，且期数较少的情况，广泛应用于经济、环境、工程、管理等领域。该理论是由中国科学家发展起来的一种描述不完备和不确定信息的数学方法。该理论认为，人类社会和自然界中的许多现象都是由不完备信息所导致的，因此需要一种特殊的数学方法来处理这些问题。

灰色预测模型的基本原理是通过对序列数据的累加、平均和相关分析，建立数学模型来进行预测。灰色预测模型主要包括 GM(1,1)模型、灰色关联度模型、灰色马尔可夫模型和灰色神经网络模型等。其中，GM(1,1)模型是最早也是最经典的灰色预测模型，它通过累加序列后建立一阶线性微分方程来进行预测。灰色关联度模型则是通过计算数据序列之间的关联度来进行预测，适用于确切知道数据之间存在线性关系的情况。灰色马尔可夫模型结合了马尔可夫链和灰色系统理论，可以处理具有时序特征的数据。灰色神经网络模型则将神经网络算法与灰色系统理论相结合，具有较强的非线性建模能力。灰色预测模型在经济预测、环境预测、工程管理等领域均具有广泛的应用。在经济领域，灰色预测模型可以用于经济增长预测、股市预测等。在环境领域，灰色预测模型可以用于空气质量预测、水资源管理等。在工程管理领域，灰色预测模型可以用于项目进度预测、质量控制等。

灰色预测模型是通过少量的、不完全的信息，对含有不确定因素的系统建立数学模型做出预测的一种小样本预测方法。灰色表明系统信息不完整，已知的仅有一部分信息，与之类似的概念还有白色和黑色。

灰色预测模型用等时距观测到的反映来预测对象特征的一系列数量值。该方法首先通过关联分析判断各因素之间时序发展趋势的相异程度，并对原始数据进行累加生成处理来寻找时序上的变动规律，生成有较强规律性的数据序列。然后建立相应的微分方程，预测序列未来的发展趋势。

灰色预测模型适用于年份数据预测，并且训练集期数较少的情况。如果期数较多或者包含季度数据，则直接采用时间序列分析更为合适。原始的灰色预测模型也称为全信息灰色预测模型，用 GM(m,n)表示，其中 m 代表阶数，n 代表预测对象的个数。一般来说最常用的是 GM(1,1)，即对单个预测变量进行一阶微分求解。本文也以 GM(1,1)为基础，构建全信息 GM(1,1)和新信息 GM(1,1)、新陈代谢 GM(1,1)两个改进算法的组合模型，对于预期时变个体合成权重进行预测。

将全部原始数据序列记为

$$x_k^{(0)} = (x_k^{(0)}(1), x_k^{(0)}(2), \dots, x_k^{(0)}(T_0))', \quad k = 1, 2, \dots, K, \quad (3.12)$$

对原数据序列累加得到一次累计序列

$$x_k^{(1)} = (x_k^{(1)}(2), \dots, x_k^{(1)}(T_0))', \quad (3.13)$$

通过 $x_k^{(1)}$ 得到紧邻均值生成序列

$$z_k^{(1)} = (z_k^{(1)}(2), \dots, z_k^{(1)}(T_0))'. \quad (3.14)$$

定义级比

$$\lambda(t) = \frac{x_k^{(1)}(t)}{x_k^{(1)}(t-1)}, \quad t = 2, \dots, T_0, \quad (3.15)$$

若级比序列均在 (1,1.5) 之间, 则一次累加序列具有准指数规律, 否则通过平移变换给原序列每个数据都加上一个常数 C , 再生成序列并做准指数检验, 直到数据满足准指数规律为止。

模型基本形式为

$$x_k^{(0)}(t) + az_k^{(1)}(t) = u, \quad (3.16)$$

通过最小二乘法, 使求解出的拟合函数值与原始数据平方差最小, 确定发展系数 $-a$ 和灰作用量 u 。原始 $GM(1,1)$ 在模型建立过程中融入了所有的已知数据, 在预测期较长时可能无法满足预测数据的精确性。因此以原始 $GM(1,1)$ 为基础, 陆续发展出了新信息 $GM(1,1)$ 和新陈代谢 $GM(1,1)$ 。

新信息 $GM(1,1)$ 是将前文全数据模型预测出的数据 $x_k^{(0)}(T_0+1)$ 不断加入原序列进行迭代, 再递推求得预测数据。新陈代谢 $GM(1,1)$ 则是将全数据模型预测出的数据 $x_k^{(0)}(T_0+1)$ 加入原序列, 这与新信息 $GM(1,1)$ 一致。但该方法会再进一步, 在原序列末尾增加数据的同时, 去掉第一个数据 $x_k^{(0)}(1)$, 依次循环, 递推求得预测数据, 整个过程中原序列长度保持不变。

三种方法都有其可用性, 实际应用中根据时间序列期数大小和干预期与预测期的分配来选择模型, 可以达到较好的效果。因此本文将这三种方法组合起来, 根据验证集的模型误差最小, 自动选择最优模型预测干预权重。

3.3.3 组合灰色预测模型

原合成控制法采用固定权重简单线性加权 $W^* Y_{jt}^N$ 预测合成单元在干预期 ($t = T_0 + 1, \dots, T$) 的潜在结果。引入时变性后, 单元层面上权重 ω_{kt}^* 随时间变化。考虑到本文中使用的实际案例收集的数据为年份数据, 且期数较少, 并且权重限制为 $\omega_{kt}^* \geq 0$ 且 $\sum_k \omega_{kt}^* = 1$, 因此拟定采用改进的组合灰色模型 $GM(m, n)$ 来拟合 $W_{K \times T_0}^* \cdot Y_{K \times T_0}$, 预测干预期的合成潜在结果 $\sum_K W_{K \times (T-T_0)} \cdot Y_{K \times (T-T_0)}$ 。具体步骤见下列算法 3.1。

算法 3.1 组合灰色模型预测算法

输入：时间期数 T ，预干预期 T_0 ，时变权重矩阵 $W_{K \times T_0}^*$ ，预干预期观测结果矩阵 $Y_{K \times T_0}$ ；

输出：干预期预测合成潜在结果 $\sum_K W_{K \times (T-T_0)} \cdot Y_{K \times (T-T_0)}$ 。

生成序列。将 $W_{K \times T_0}^* \cdot Y_{K \times T_0}$ 记为 K 个列向量 $x_k^{(0)} = (x_k^{(0)}(1), x_k^{(0)}(2), \dots, x_k^{(0)}(T_0))'$ ，

Step1: $k = 1, 2, \dots, K$ 。并对原序列 $x_k^{(0)}$ 做一次累加，记为 $x_k^{(1)} = (x_k^{(1)}(2), \dots, x_k^{(1)}(T_0))'$ ，得

到紧邻均值生成序列 $z_k^{(1)} = (z_k^{(1)}(2), \dots, z_k^{(1)}(T_0))'$ 。

准指数检验。定义级比 $\lambda(t) = \frac{x_k^{(1)}(t)}{x_k^{(1)}(t-1)}$ ， $t = 2, \dots, T_0$ 。若级比序列均在 $(1, 1.5)$ 之

Step2:

间，则一次累加序列具有准指数规律，否则通过平移变换给原序列每个数据都加上一个常数 C ，再生成序列并做准指数检验，直到数据满足准指数规律为止。

Step3: 全数据模型。模型基本形式为 $x_k^{(0)}(t) + az_k^{(1)}(t) = u$ ，通过最小二乘法使求解出的拟合函数与原始数据平方差最小，确定发展系数 $-a$ 和灰作用量 u 。

Step4: 新信息模型。将上述全数据模型预测出的数据 $x_k^{(0)}(T_0 + 1)$ 加入原序列，递推求得预测数据。

Step5: 新陈代谢模型。将全数据模型预测出的数据 $x_k^{(0)}(T_0 + 1)$ 加入原序列，去掉第一个数据 $x_k^{(0)}(1)$ ，依次循环，递推求得预测数据。

Step6: 择优应用。将原序列前 75% 作为训练集，后 25% 作为验证集。上述三个模型中，选择测试集均方误差 $SSE = [x_k^{(0)}(t) - \hat{x}_k^{(0)}(t)]^2$ 最小的模型进行后续 $T_0 + 1 \sim T$ 个数据，即求得 $\sum_K W_{K \times (T-T_0)} \cdot Y_{K \times (T-T_0)}$ 。

残差检验。定义相对误差 $\varepsilon_k(t) = \frac{|x_k^{(0)}(t) - \hat{x}_k^{(0)}(t)|}{x_k^{(0)}(t)} \times 100$ 和级比偏差

Step7: $\eta(t) = \left| 1 - \frac{1-0.5\hat{a}}{1+0.5\hat{a}} \frac{1}{\lambda(t)} \right|$ ，平均误差 $\bar{\varepsilon}_k(t) = \frac{1}{T_0-1} \sum_{t=2}^{T_0} |\varepsilon_k(t)|$ ，根据平均误差和级

比偏差大小检验模型的拟合效果。

3.4 模型检验

3.4.1 误差分析

采用 Abadie 评估“加州烟草控制计划”效果的数据实现上述权重矩阵的求解过程。1989 年，99 号提案将加州每包香烟的消费税提高了 25 美分，并将税收收入用于健康和反吸烟教育预算，资助反吸烟媒体宣传活动，并促使全州颁布了清洁室内空气条例。Abadie 等采用 1970 年至 2000 年期间的年度州级面板数据，以 1989 年至 2000 年期间没有采用过大规模禁烟策略的 38 个州组成控制组库 ($J = 38, T_0 = 19, T = 31$)，以年人均香烟消费量为结果变量 Y_{it} ，以人均累计收入、香烟零售价格、15~24 岁青年人口占比、人均啤酒消费量作为 4 个主要解释变量 $X_{iq} (q=1,2,3,4)$ ，以间断的三年吸烟消费量(1975 年，1980 年和 1988 年)作为补充解释变量 $X_{iq} (q=5,6,7)$ ，研究相较于未实施提案前，政策提出后 Y_{it} 的变化值。

根据上文方法，由 1980 年至 1988 年解释变量平均值差距最小作为合成地区挑选条件。引入的解释变量权重矩阵 V 对角线上值为

$$v_q = (1, 157.762, 58.859, 46.646, 114.077, 0.671, 998.962),$$

由式 (3.10) 求解出最优固定权重向量

$$W^*(V) = (0.169, 0.070, 0.198, 0.231, 0.332)。$$

确定控制组包含 5 个与加州影响因素相近的合成地区：Colorado、Connecticut、Montana、Nevada 以及 Utah，对应最优时变权重矩阵为 (5×19) 阶矩阵：

$$W_{t_0}^* = \begin{bmatrix} 0.200 & 0.199 & 0.195 & 0.194 & 0.198 & 0.199 & 0.198 & 0.198 & 0.199 & 0.192 & 0.191 & 0.183 & 0.182 & 0.184 & 0.182 & 0.186 & 0.186 & 0.191 & 0.193 \\ 0.200 & 0.201 & 0.207 & 0.207 & 0.205 & 0.204 & 0.204 & 0.204 & 0.203 & 0.204 & 0.200 & 0.203 & 0.201 & 0.196 & 0.190 & 0.188 & 0.188 & 0.180 & 0.178 \\ 0.199 & 0.202 & 0.201 & 0.202 & 0.202 & 0.201 & 0.201 & 0.200 & 0.200 & 0.201 & 0.200 & 0.197 & 0.192 & 0.197 & 0.193 & 0.199 & 0.203 & 0.213 & 0.206 \\ 0.207 & 0.180 & 0.169 & 0.170 & 0.177 & 0.183 & 0.180 & 0.177 & 0.181 & 0.159 & 0.161 & 0.151 & 0.151 & 0.153 & 0.164 & 0.149 & 0.151 & 0.133 & 0.138 \\ 0.194 & 0.218 & 0.228 & 0.226 & 0.218 & 0.213 & 0.216 & 0.222 & 0.217 & 0.243 & 0.247 & 0.266 & 0.274 & 0.270 & 0.271 & 0.277 & 0.273 & 0.283 & 0.286 \end{bmatrix}。$$

合成加州和真实加州的潜在结果对比详见表 3.2、图 3.1，单元层面上权重 ω_{kt}^* 随时间变化如图 3.2 所示。

表 3.2 加州与合成加州在预干预期的潜在结果

年份	真实加州	固定权重合成加州		时变权重合成加州	
		合成加州	误差	合成加州	误差(10e-5)
1970 年	123.0	117.1	5.937	123.0	0.001
1971 年	121.0	118.8	2.153	121.0	0.135
1972 年	123.5	124.2	-0.738	123.5	0.001
1973 年	124.4	125.4	-1.024	124.4	0.030
1974 年	126.7	126.9	-0.174	126.7	0.100
1975 年	127.1	127.0	0.146	127.1	0.127
1976 年	128.0	127.8	0.225	128.0	0.111
1977 年	126.4	125.7	0.728	126.4	0.095
1978 年	126.1	124.9	1.196	126.1	0.095
1979 年	121.9	122.9	-0.999	121.9	0.014
1980 年	120.2	120.5	-0.271	120.2	0.041
1981 年	118.6	120.2	-1.644	118.6	-0.239
1982 年	115.4	116.9	-1.524	115.4	0.102
1983 年	110.8	111.4	-0.588	110.8	0.067
1984 年	104.8	103.5	1.306	104.8	-0.007
1985 年	102.8	103.3	-0.467	102.8	0.138
1986 年	99.7	99.9	-0.177	99.7	0.024
1987 年	97.5	99.7	-2.207	97.5	0.029
1988 年	90.1	91.6	-1.527	90.1	0.019

计算潜在结果的平均绝对误差值，固定权重合成加州误差为 23.031，时变权重合成加州误差为 $1.37e-5$ 。另外对比表 3.2 中两列误差值也可以看出，时变个体合成权重加权之后的加州在预干预期内每一个时间节点上都具有更小的偏差，更接近真实加州的年人均吸烟量。以上数据足以说明使用本文提出的时变个体合成权重改进合成控制法，能让预干预期内合成精度得到有显著提升。

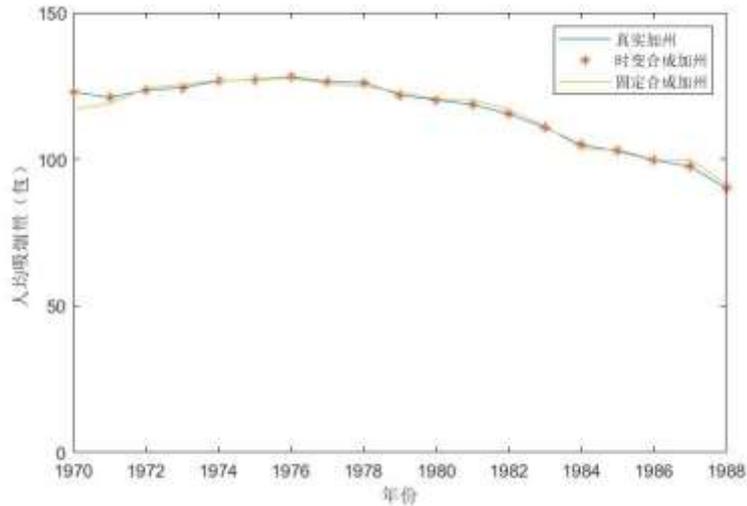


图 3.1 合成加州在预干预期的潜在结果对比图

图 3.1 直观展示了真实加州、固定权重合成加州与时变个体合成权重合成加州在预干预期的年人均吸烟量。理想状态下，合成加州应与真实加州的潜在结果完全一致即折线图完全重合或近似重合，才能保证合成加州的准确性为干预期结果对比提供可靠依据。实际结果可以看出，时变个体合成权重合成的加州与真实加州图形基本重合，固定权重合成的加州在 1970 年、1971 年、1981 年、1982 年以及 1987 年均有明显偏离，直观表明相比固定合成权重，时变合成权重具有更高的合成精度。预干预期合成单元时变权重变化如图 3.2 所示。

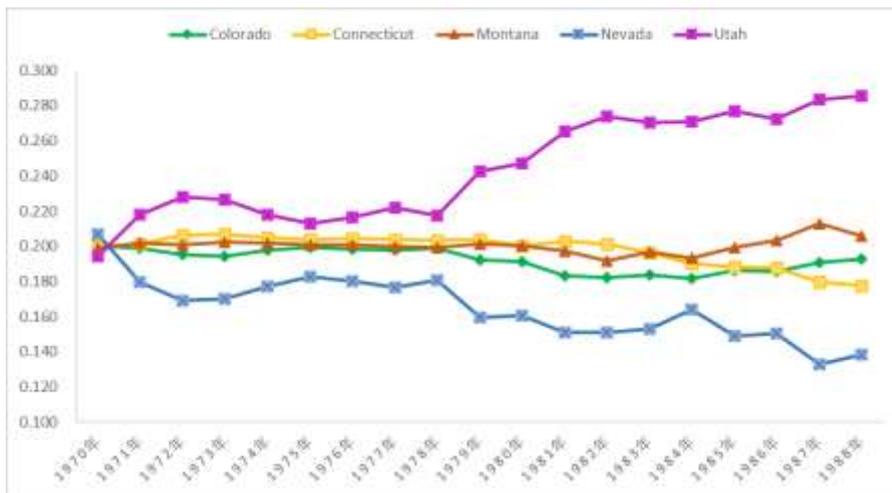


图 3.2 预干预期合成单元时变权重变化图

观察加州禁烟案例中计算出的预干预期合成组个体时变权重值，发现权重分

量在时间上变化明显。Utah 的权重最高，并随时间不断上升，表明该州有极大概率在干预期对合成加州的影响持续增加。Colorado、Connecticut 和 Montana 三个州权重在 0.2 上下轻微波动，干预期权重也可能维持在该水平。Nevada 与 Utah 的时间趋势相反，它在 5 个合成州中权重最小，并随时间进一步降低，该州在干预期对合成加州的潜在结果影响可能仍然较弱。

用时变权重序列对该地区潜在结果值加权，其结果如图 3.3 所示。控制组单元潜在结果加权值 $\omega_k^*(t)Y_{kt}$ 具有明显的时变性，对合成加州潜在结果的贡献度随时间变化趋于集中。

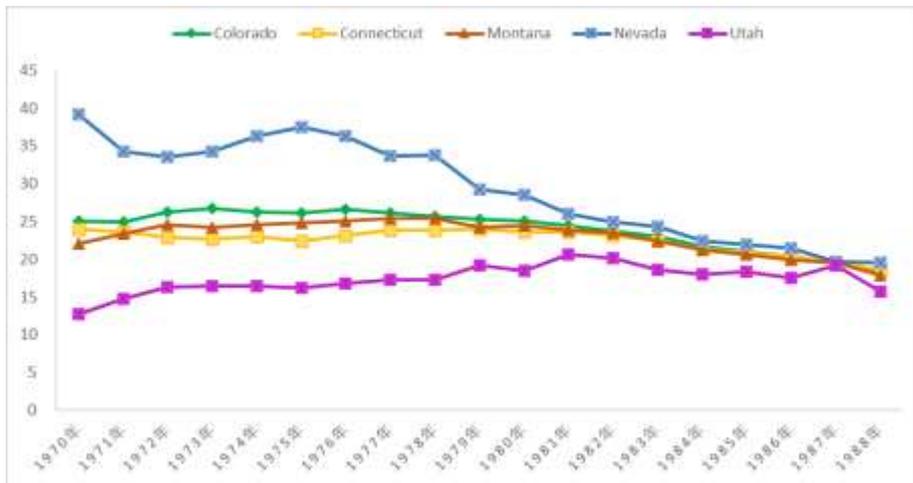


图 3.3 预干预期合成单元潜在结果加权

由于本案例中只需预测单个单元的合成权重值，因此基础灰色预测模型选定为 $GM(1,1)$ 。根据单元在预干预期的潜在结果加权 $\omega_k^*(t)Y_{kt}$ 的数据特征，由算法 3.1 比较其在三种灰色预测模型中的误差大小，选择验证误差最小的模型作为最终的权重预测方法并计算 5 个合成城市在干预期 1989 年至 2000 年的潜在结果合成分量。

注意到级比

$$\lambda(t) = \frac{x_k^{(1)}(t)}{x_k^{(1)}(t-1)} = \frac{x_k^{(0)}(t)}{x_k^{(1)}(t-1)} + 1, \quad (3.17)$$

定义原始序列的光滑比

$$\rho(t) = \frac{x_k^{(0)}(t)}{x_k^{(1)}(t-1)}, \quad (3.18)$$

显然，光滑比逐渐趋于 0。由于级比序列要求在(1,1.5)之间，一次累加序列才具有准指数规律，因此只要保证光滑比小于 0.5，即通过准指数规律检验。此案例中，五个合成城市光滑比如图 3.4 所示。可以看出，除去前两个时期外，光滑比小于 0.5 的数据占比为 100%，即通过检验可以进行下一步模型拟合和预测。

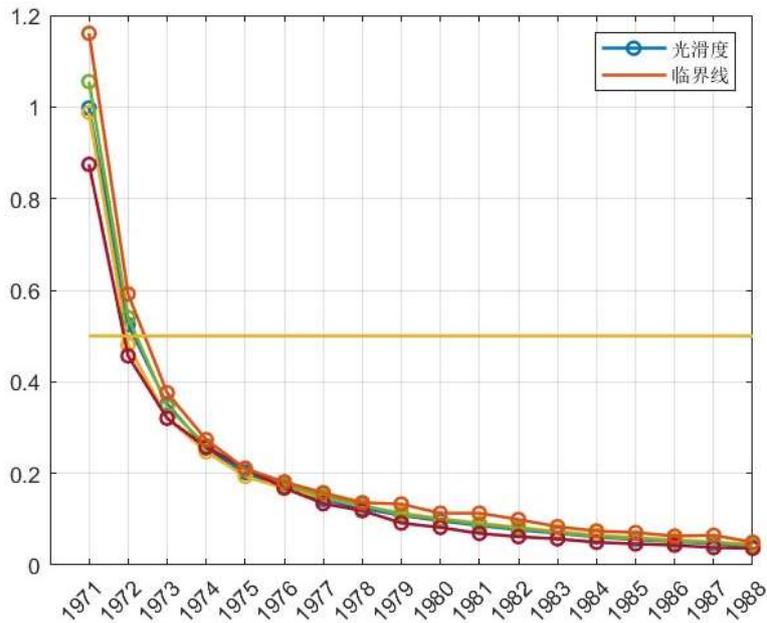


图 3.4 准指数规律检验

基于 1970 至 1988 年 Colorado、Connecticut、Montana、Nevada 和 Utah 这 5 个城市的潜在结果加权值 $\omega_k^*(t)Y_{kt}$ ，以前 75% 数据量（1970 年至 1983 年）14 期数据为模型训练集，后 25% 数据量（1984 年至 1988 年）5 期数据为模型验证集。用全数据 $GM(1,1)$ ，新信息 $GM(1,1)$ 和新陈代谢 $GM(1,1)$ 分别计算出模型拟合均方误差如表 3.3 所示。

表 3.3 均方误差结果

模型	Colorado	Connecticut	Montana	Nevada	Utah
全数据 $GM(1,1)$	62.427	56.510	76.439	55.912	81.123
新信息 $GM(1,1)$	62.423	56.510	76.439	55.793	81.092
新陈代谢 $GM(1,1)$	44.884	57.292	57.802	22.352	81.298

根据表 3.3 所示结果，Colorado 最小误差 44.884，Montana 最小误差 57.802，Nevada 最小误差 22.352，三个城市适用新陈代谢 $GM(1,1)$ 模型预测干预期数据；Connecticut 最小误差 56.51023，Utah 最小误差 81.092，两个城市适用新信息 $GM(1,1)$ 模型预测干预期数据；使得组合模型拟合误差最小，预测结果更准确。采用以上组合模型，计算出的合成组干预期潜在结果分量 $\hat{\omega}_k(t)Y_{kt}$ 如下表 3.4 所示，加和后得到时变个体合成权重合成加州的潜在结果。

表 3.4 组合灰色模型预测结果

年份	Colorado	Connecticut	Montana	Nevada	Utah
1989 年	20.122	20.495	20.099	19.957	19.056
1990 年	19.408	20.306	19.494	18.703	19.214
1991 年	18.785	20.119	18.966	17.344	19.372
1992 年	18.225	19.934	18.376	16.048	19.532
1993 年	17.634	19.750	17.798	15.041	19.694
1994 年	17.042	19.568	17.247	14.358	19.856
1995 年	16.550	19.388	16.741	13.802	20.020
1996 年	16.069	19.210	16.315	13.222	20.185
1997 年	15.616	19.033	15.963	12.918	20.352
1998 年	15.200	18.857	15.560	12.364	20.520
1999 年	14.848	18.684	15.258	11.921	20.689
2000 年	14.542	18.512	14.989	11.370	20.860

根据表 3.4 合成组 5 个城市在干预期 1989 年至 2000 年的潜在结果合成分量预测值，和该组城市在 1989 年至 2000 年未受到禁烟政策干预的潜在结果实际观测值，计算出未标准化干预期时变权重为

$$W_{T-T_0} = \begin{bmatrix} 0.227 & 0.222 & 0.208 & 0.206 & 0.199 & 0.191 & 0.194 & 0.193 & 0.192 & 0.187 & 0.187 & 0.199 \\ 0.204 & 0.222 & 0.232 & 0.239 & 0.250 & 0.255 & 0.244 & 0.253 & 0.251 & 0.250 & 0.255 & 0.259 \\ 0.233 & 0.230 & 0.229 & 0.212 & 0.207 & 0.196 & 0.185 & 0.187 & 0.180 & 0.175 & 0.185 & 0.199 \\ 0.145 & 0.136 & 0.150 & 0.146 & 0.139 & 0.136 & 0.137 & 0.134 & 0.135 & 0.121 & 0.115 & 0.122 \\ 0.334 & 0.360 & 0.362 & 0.355 & 0.350 & 0.356 & 0.385 & 0.374 & 0.357 & 0.485 & 0.471 & 0.513 \end{bmatrix}。$$

为了保证合成控制法的防外推特性，个体合成权重仍旧要求值非负且和为 1。在模型拟合出以上原始数据的变化趋势后，对于预期时变权重添加限制条件，求得标准化后时变权重为

$$W_{T-T_0}^* = \begin{bmatrix} 0.198 & 0.190 & 0.176 & 0.178 & 0.174 & 0.169 & 0.169 & 0.170 & 0.172 & 0.154 & 0.154 & 0.154 \\ 0.178 & 0.190 & 0.196 & 0.206 & 0.218 & 0.225 & 0.214 & 0.222 & 0.225 & 0.205 & 0.210 & 0.201 \\ 0.204 & 0.197 & 0.194 & 0.183 & 0.181 & 0.172 & 0.162 & 0.164 & 0.161 & 0.143 & 0.152 & 0.154 \\ 0.127 & 0.116 & 0.127 & 0.126 & 0.121 & 0.120 & 0.119 & 0.117 & 0.121 & 0.099 & 0.095 & 0.094 \\ 0.293 & 0.307 & 0.307 & 0.307 & 0.306 & 0.314 & 0.336 & 0.328 & 0.320 & 0.398 & 0.389 & 0.397 \end{bmatrix}。$$

线性加权后得到时变个体合成权重合成加州在 1989 年至 2000 年的合成潜在结果，其与真实加州和固定权重合成加州的潜在结果对比如表 3.5。

表 3.5 加州与合成加州在干预期的潜在结果

年份	真实加州	固定权重合成加州	时变权重合成加州
1989 年	82.400	89.913	87.288
1990 年	77.800	87.412	83.001
1991 年	68.700	82.189	80.066
1992 年	67.500	81.603	79.523
1993 年	63.400	81.188	78.514
1994 年	58.600	80.731	77.623
1995 年	56.400	78.492	75.543
1996 年	54.500	77.464	74.538
1997 年	53.800	77.678	75.259
1998 年	52.300	74.366	67.766
1999 年	47.200	73.540	67.172
2000 年	41.600	67.341	62.153

使用组合灰色模型预测五组数据后，与模型选择时类似，同样地以原始数据前 75% 数据量即 1989 年至 1997 年作为模型训练集，后 25% 数据量即 1998 年至 2000 年作为模型验证集，来检验模型拟合程度。相对误差如图 3.5 所示。

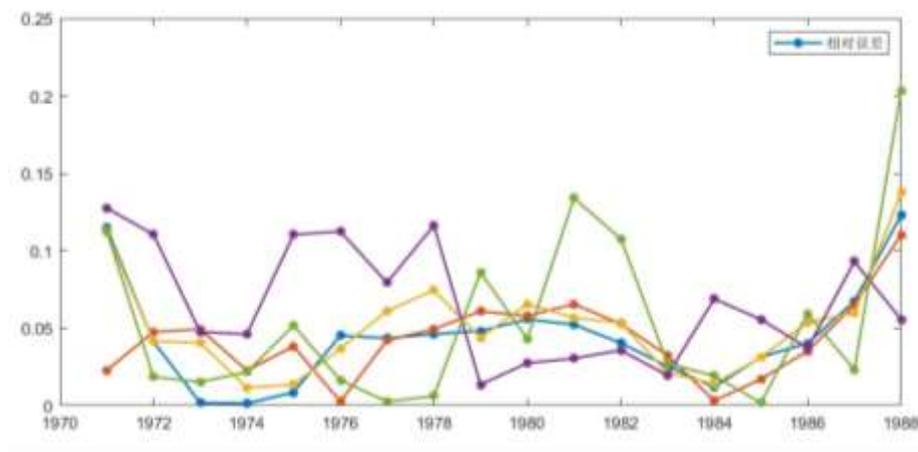


图 3.5 相对误差

时变个体合成权重合成加州在 1970-1988 年的原始数据，组合灰色模型拟合数据，以及 1989-2000 年的预测数据如图 3.6 所示。对比结果有充分的可信度。结合表 3.4 和图 3.5，Colorado、Connecticut、Montana、Nevada 和 Utah 五组干预期潜在结果合成分量预测数据的平均值为 17.685，预测数据验证误差值均在 0.2 以下，说明组合灰色模型具有良好的拟合效果，预测数据也具有较高的可信度。

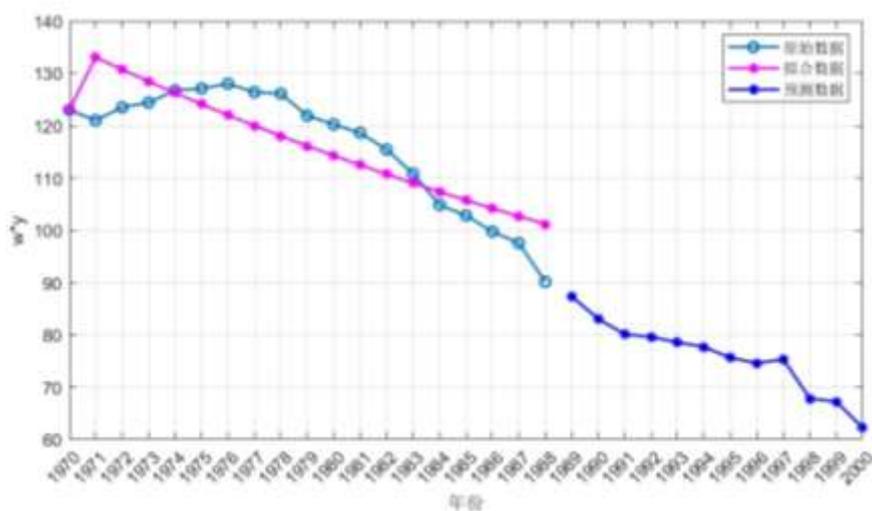


图 3.6 预测曲线

将 Colorado、Connecticut、Montana、Nevada 和 Utah 五个城市的干预期潜在结果合成分量预测数据加和，得到 1989 年至 2000 年时变个体合成权重合成加州

的潜在结果 $\sum_K W_{K \times (T-T_0)} \cdot Y_{K \times (T-T_0)}$ 。最终时变权重合成加州的禁烟政策干预效果和固定权重合成加州禁烟政策干预效果对比如图 3.7 所示。

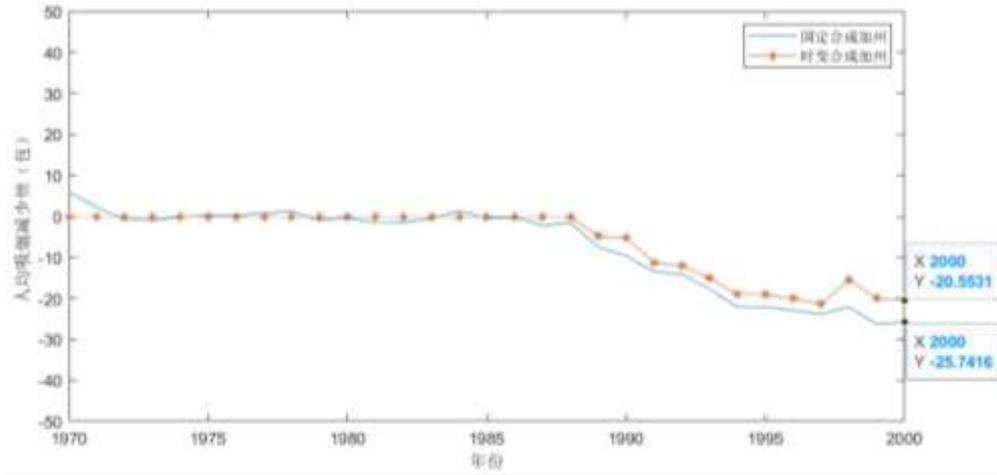


图 3.7 合成加州政策效果

1970 年至 1988 年禁烟政策还未实施，政策干预效果为零，即图 3.7 中该时期内年人均吸烟减少量应为 0。由上图可以看出，固定权重合成加州在此期间内人均吸烟减少量在零值上下浮动，时变权重合成加州基本稳定在零值，可以说明时变权重合成加州确实具有更高的准确性。

1989 年至 2000 年，时变权重合成加州和固定权重合成加州在转折点上具有一致性，但由时变权重测算出的“99 号提案”禁烟政策效果一直小于固定权重测算出的政策效果。比较 2000 年加州与两个合成加州的吸烟消费量，固定权重合成加州的年消费量差值为 26 包，时变权重合成加州的年消费量差值为 21 包。

3.4.2 安慰剂检验

为了进一步检验模型的准确性，本文参照 Abadie 等的做法，对控制组库中 38 个州反复应用时变权重合成控制法，进行一系列置换检验。38 次迭代中，依次将控制组中每一个未受干预的州假定为干预单元，再将加州转移至控制组库中。通过 2000 年的人均吸烟减少量结果对比，计算出每一组置换检验的政策干预估计效果，最终安慰剂检验结果如图 3.8 所示。

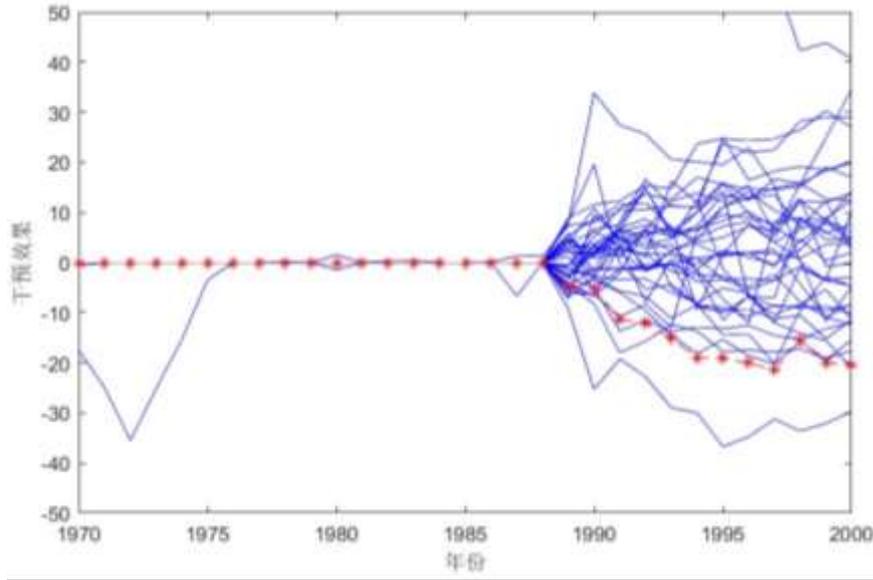


图 3.8 安慰剂检验结果

理想状态下，图 3.8 中加州作为干预组的干预效果（红色）应该最为明显，即图 3.8 中假设其他城市作为干预组的干预效果（蓝色）应在加州上方。分析图 3.8 的实际数据展示，可以看到 38 组安慰剂检验（蓝色）和一组实际案例结果（红色）共计 39 根线条中，仅有一根蓝色线条在红色线条下方，即仅有 1 组置换后的干预城市在 2000 年香烟销售减少量多于加州。

根据上述结果，假设检验 p 值为

$$p = \frac{1}{38} = 0.026, \quad (3.19)$$

通过置信水平为 95% 的假设检验。因此应用上述时变合成权重合成控制法得出的政策评价结果有良好的置信度。

3.5 本章小结

经济学和社会科学的相关研究，往往采用大量面板数据，合成控制法作为处理该类问题的常用方法，将时变性引入合成权重中可以提高合成精度，为后续量化评价干预效果提供更有力的理论基础。本文利用加州禁烟案例将时变合成控制法与原合成控制法进行对比，通过干预前数据拟合误差验证了本文提出的方法在合成单元时具有更小的误差。同时通过安慰剂检验验证了本文的方法在实际案例应用中也具有较高的置信度。

但是,在通过预干预期时变权重预测干预期权重时,本文暂时没有提出通用的方法,需要根据数据量以及合成单元权重的数字特征和变化规律寻找合适的预测模型。对于本文中出现的的小样本数据,通过组合灰色预测模型在全信息 $GM(1,1)$ 、新信息 $GM(1,1)$ 和新陈代谢 $GM(1,1)$ 三种方法中选择模型验证误差最小的方法来预测干预期合成权重。后续可以在该问题上继续研究。

4 “双碳”目标下中国碳排放权交易试点政策效果评价

4.1 中国碳减排政策发展规划

自 2005 年中共十六届五中全会首次提出了“建设资源节约、环境友好型社会”目标以来，构建人与自然和谐发展已成为中国经济社会发展的重要战略（陆贤伟，2017）。为此，我国政府也在积极探索节能减排的发展道路，国家发改委于 2012 年下发了《关于开展低碳省区和低碳城市试点工作的通知》（段娟，2019），以期通过低碳试点政策来累积低碳式发展的重要经验进而实现累积—推广—全面铺开的战略部署（杨帆，2021），从而推动落实中国政府承诺碳排放减排的目标。

中国作为负责任大国，一直积极推动全球碳减排工作。2014 年 9 月中国发布《国家应对气候变化规划（2014-2020 年）》，提出到 2020 年碳强度较 2005 年下降 40%~45% 的规划目标（杨秀汪，2021）。2015 年《巴黎协定》提出 2030 年碳强度将比 2005 年下降 60%~65%。2020 年 9 月 22 日国家主席习近平在第七十五届联合国大会上提出中国将力争在 2030 年前碳排放达到峰值，并争取 2060 年前实现碳中和，即为“双碳”目标。“双碳”目标让前期低碳试点政策实施效果备受关注。

2021 年 5 月 26 日，碳达峰碳中和工作领导小组在北京召开第一次全体会议。同年 10 月，中共中央、国务院两个重要文件《关于完整准确全面贯彻新发展理念做好碳达峰碳中和工作的意见》和《2030 年前碳达峰行动方案》相继出台，作为碳达峰碳中和“1+N”政策体系中的“1”，意见为碳达峰碳中和这项重大工作进行了系统谋划和总体部署，共同构建了中国碳达峰、碳中和“1+N”政策体系的顶层设计（周蒙蒙，2022），而重点领域和行业的配套政策也将围绕以上意见及方案陆续制定。2022 年 8 月，科技部、国家发展改革委、工业和信息化部等 9 部门印发《科技支撑碳达峰碳中和实施方案（2022—2030 年）》，统筹提出支撑 2030 年前实现碳达峰目标的科技创新行动和保障举措，并为 2060 年前实现碳中和目标做好技术研发储备。

4.2 碳减排影响机制

4.2.1 碳减排有效措施

减少二氧化碳排放的政策主要分为两类，一类是政府基于行政命令干预碳排放的环境规制手段，另一类则是基于市场行为的碳税和碳排放权交易措施。党的十八届三中全会提出“使市场在资源配置中起决定性作用”的重大理论创新，不断深化市场化体制改革，而碳排放权交易制度将这一理念贯彻于环境政策机制设计之中。

总结各国二氧化碳减排政策和研究者们对碳减排机制的评价，中国低碳政策抑制碳排放的主要有效路径可以总结为以下五点。

(1) 激励能源转型和节能减排措施

试点城市低碳政策通过提供激励措施，鼓励企业和居民采取节能减排措施，促进能源转型。例如，政府可以给予节能减排技术的研发和应用以财政支持，提供税收优惠政策，引导企业和居民减少碳排放。

(2) 加强能源管理和监管

试点城市低碳政策通过加强能源管理和监管，推动企业和居民实施碳排放的监测和报告。政府可以要求企业和居民定期报告碳排放情况，并对不符合要求的进行处罚，从而促使其采取减排措施。

(3) 推动清洁能源发展

试点城市低碳政策鼓励清洁能源的开发和利用，减少对高碳能源的依赖。政府可以提供补贴或优惠政策，鼓励企业和居民使用可再生能源，如太阳能、风能等，以替代传统的化石燃料，从而减少碳排放。

(4) 建立碳市场和碳交易机制

试点城市低碳政策可以建立碳市场和碳交易机制，通过碳排放权的交易来推动减排行为。政府可以设立碳排放配额，并将其分配给企业，企业可以在市场上进行碳排放权的买卖，从而实现碳排放的减少。

(5) 提高公众意识和参与度

试点城市低碳政策通过宣传和教育活动，提高公众对低碳生活方式的认识和意识。政府可以组织宣传活动，鼓励居民采取低碳出行、节约用水、减少垃圾等

行为，从而推动碳排放的减少。

这些机制相互作用，形成了一个综合的政策体系，为实现低碳经济和可持续发展提供了重要支持。

4.2.2 碳排放权交易政策减排路径

2017年12月，经国务院同意国家发展改革委印发了《全国碳排放权交易市场建设方案（电力行业）》，这标志着中国碳排放交易体系完成了总体设计并正式启动。2020年年底，生态环境部出台《碳排放权交易管理办法（试行）》，印发《2019-2020年全国碳排放权交易配额总量设定与分配实施方案（发电行业）》，正式启动全国碳市场第一个履约周期。2021年7月16日，全国碳排放权交易市场启动上线交易。发电行业成为首个纳入全国碳市场的行业，纳入重点排放单位超过2000家。2021年元旦起，全国碳市场发电行业第一个履约周期正式启动。2021年7月16日，全国碳排放权交易市场开市。

中国碳交易试点市场已运行多年，碳交易试点政策是否促进了试点省份二氧化碳排放减少，碳交易试点政策若具有减排效应，其碳减排效应如何实现，以及碳减排效应有多大，这是本文关注的重点。本文研究的首批低碳试点城市，在2013年底宣布采用碳排放权交易减排。碳排放权交易是一种市场机制，旨在通过碳排放权的市场机制控制温室气体排放、推动绿色低碳发展，是实现碳中和的重要政策工具（刘紫薇，2023），主要通过以下方式达到实际抑制二氧化碳排放的效果。

（1）经济激励减排行为

碳排放权交易创建了一个经济激励机制，通过设立碳排放配额并将其分配给企业，企业需要购买额外的排放权来弥补超出配额的排放量。这激励企业采取减排措施，以降低其碳排放量，从而减少二氧化碳的排放。

（2）促进清洁技术创新

碳排放权交易为企业提供了经济上的动力，鼓励它们采用更清洁、低碳的生产技术和设备。企业为了降低碳排放成本，会积极寻求和应用清洁技术创新，从而减少二氧化碳的排放。

（3）优化资源配置

碳排放权交易使得碳排放成本成为企业经营成本的一部分。企业通过购买和出售碳排放权，可以根据自身的情况和需求进行灵活调整，从而实现资源的优化配置。这有助于推动高碳排放行业向低碳排放行业的转型，减少二氧化碳的排放。

(4) 促进国际合作和技术转移

碳排放权交易可以促进国际间的碳市场合作，使得低碳技术和经验在不同国家之间进行转移。发达国家可以通过向发展中国家转让碳排放权，来帮助后者实现碳减排目标，推动全球范围内的减排任务，减少全球二氧化碳的排放总量。

(5) 提高公众意识和参与度

碳排放权交易机制的实施有助于提高公众对气候变化的关注和碳减排的意识。公众对碳排放权交易的了解和参与度的提高，可以促使企业更加积极地采取减排措施，减少二氧化碳的排放。

综上，碳排放权交易通过经济激励减排行为、促进清洁技术创新、优化资源配置、促进国际合作和技术转移以及提高公众意识和参与度等机制，可以有效地抑制二氧化碳排放。这种市场机制为减缓气候变化、实现可持续发展提供了重要的支持。

4.3 指标选取与数据来源

4.3.1 研究对象

由于约束碳排放的《京都议定书》正式生效时间为 2005 年，并且我国政府也将 2005 年作为 2020 年单位 GDP 碳排放强度减排目标的对比期，因此本文以 2005 年作为研究期的初始年份。2020 年 9 月，我国明确提出在 2030 年前碳排放达到峰值并且 2060 年前实现碳中和的“双碳”目标，随后各地区陆续开展碳减排工作。2021 年 7 月，全国碳排放权交易市场开市，试点城市在政策上也不再具有独特性。因此本文将 2019 年作为研究期的结束年份。

2010 年 8 月，国家发改委发布《关于开展低碳省区和低碳城市试点工作的通知》，将广东、陕西、辽宁、湖北、云南、天津、重庆以及深圳、厦门、南昌、贵阳、保定、杭州这五省八市作为首批低碳试点区域。2011 年，按照“十二五”规划纲要关于“逐步建立碳排放交易市场”的要求，中国在北京、天津、上海、

重庆、深圳、广东及湖北这五市两省启动了碳排放权交易地方试点工作。2013年6月18日,深圳作为试点率先启动交易,到2013年底2014年初,7个地方试点碳市场陆续开始上线交易。因此将2014年作为政策干预初始年份。

本文选取首批参与碳排放权交易试点的五市两省(北京、天津、上海、重庆、深圳、广东及湖北)作为干预组。由于本文在省级行政区层面进行研究,深圳市行政规划上隶属于广东省,因此干预组最终确定为北京、天津、上海、重庆、广东及湖北。2016年12月,福建省启动碳交易市场,作为国内第8个碳交易试点地区。因此以港澳台、福建及干预组所涉及到的省级行政区之外的其他未采取任何碳减排政策的省地区组成控制组库。鉴于实际数据可获得性,删除没有统计数据西藏自治区。控制组库最终包括河北、山西、黑龙江、吉林、辽宁、江苏、浙江、安徽、江西、山东、河南、湖南、海南、四川、贵州、云南、陕西、甘肃、青海、内蒙古、广西、宁夏、新疆这23个省级行政区。

4.3.2 结果变量

量化一国经济和碳排放量之间的关系采用“碳排放强度”,这也是双碳目标下人们关注的重要指标。由于我国幅员辽阔、人口众多,各地区经济发展水平差异较为明显,本文采用碳排放强度作为研究的结果变量,一定程度上削弱经济差异化对碳排放量产生的影响,使得试点地区可以找到合适的合成组。二氧化碳排放强度 CI 以单位GDP的二氧化碳排放量衡量,一般情况下随着技术进步和经济增长而下降。根据中国统计年鉴得到干预组和控制组各省级行政区在2005年至2019年省内GDP。

鉴于我国尚未公布实际二氧化碳排放数据,本文采用各类能源的碳排放量估算实际碳排放值。首先要计算各类能源折算成标准煤的数量。第一步,计算各种能源的折算系数,其由能源实际含热值和标准燃料热值的比值求得。第二步,根据该折算系数将具有一定实物量的能源折算成标准燃料的数量。参照2015年中国能源统计年鉴附录,各类能源折标准煤参考系数如表4.1。

表 4.1 各类能源折标准煤参考系数

能源品种	平均低位发热量 (kcal/kg)	折标系数
原煤	5000	0.714
焦炭	6800	0.971
原油	10000	1.429
汽油	10300	1.471
煤油	10300	1.471
柴油	10200	1.457
燃料油	10000	1.429
天然气	9300	1.330

资料来源：《中国能源统计年鉴（2015）》附录

其次，根据查阅国家发展和改革委员会能源研究所和《国家温室气体排放清单指南》，得到原煤、焦炭、原油、汽油、煤油、柴油、燃料油和天然气 8 种化石能源的碳排放系数。其值如下表 4.2 所示。

表 4.2 各类能源碳排放系数

能源品种	碳排放系数
原煤	0.756
焦炭	0.855
原油	0.586
汽油	0.554
煤油	0.571
柴油	0.592
燃料油	0.619
天然气	0.448

资料来源：国家发展和改革委员会能源研究所和《国家温室气体排放清单指南》

最终碳排放强度的计算公式为：

$$CI = \frac{\sum E_p C_p L_p}{GDP}, \quad (4.1)$$

其中 p 表示能源的种类，共包括原煤、焦炭、原油、汽油、煤油、柴油、燃料油以及天然气 8 类。 E_p 表示 p 类能源的消费量， C_p 表示能源对应的标准煤折算系数， L_p 表示能源对应的碳排放系数，参照《国家温室气体排放清单指南》如表

4.2, *GDP* 表示地区生产总值。根据以上计算过程, 得到 2005 至 2019 年各省级行政区碳排放强度如表 4.3, 这也是本章节最为关注的潜在结果变量。

表 4.3 各省级行政区年度 CO₂ 排放强度

地区	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
北京	1.722	1.519	1.332	1.2	1.126	0.983	0.796	0.735	0.61	0.587	0.528	0.448	0.403	0.382	0.326
天津	3.198	2.979	2.69	2.092	2.009	1.999	1.789	1.58	1.464	1.298	1.219	1.064	1.019	1.043	1.403
上海	2.501	2.179	1.894	1.762	1.64	1.573	1.447	1.357	1.332	1.123	1.057	0.939	0.88	0.812	0.718
重庆	2.615	2.517	2.292	2.29	2.185	1.988	1.793	1.553	1.2	1.153	0.918	0.838	0.789	0.757	0.659
广东	1.716	1.621	1.472	1.311	1.318	1.252	1.195	1.084	1.008	0.935	0.879	0.818	0.772	0.735	0.659
湖北	3.634	3.481	3.143	2.592	2.433	2.264	2.096	1.85	1.445	1.325	1.166	1.049	0.988	0.926	0.845
河北	5.734	5.402	4.969	4.412	4.372	3.975	3.74	3.498	3.278	3.014	3.111	2.896	2.71	2.626	2.704
山西	13.275	12.754	10.664	8.629	8.498	7.292	6.581	6.376	6.253	6.355	7.292	7.06	6.252	6.179	6.421
黑龙江	4.561	4.277	4.022	3.645	3.685	3.31	2.925	2.816	2.531	2.466	2.25	2.224	2.154	2.135	2.694
吉林	5.09	4.699	4.015	3.427	3.095	2.899	2.721	2.38	2.099	1.968	1.649	1.547	1.518	1.557	2.062
辽宁	6.19	5.755	5.191	4.347	4.044	3.652	3.239	3.003	2.643	2.516	2.426	3.166	3.1	3.049	3.375
江苏	2.545	2.383	2.147	1.86	1.748	1.624	1.579	1.464	1.362	1.242	1.194	1.126	1.005	0.924	0.877
浙江	2.235	2.152	2.016	1.796	1.743	1.552	1.41	1.275	1.204	1.119	1.063	0.958	0.915	0.827	0.761
安徽	3.681	3.467	3.206	3.044	2.947	2.543	2.232	2.047	1.99	1.897	1.8	1.621	1.515	1.417	1.153
江西	2.892	2.655	2.411	2.038	1.943	1.831	2.023	1.483	1.436	1.342	1.314	1.203	1.136	1.081	0.98
山东	3.815	3.594	3.395	3.055	2.903	2.777	2.427	2.412	2.126	2.117	2.198	2.139	2.056	1.933	2.135
河南	4.016	3.9	3.555	3.057	2.889	2.636	2.491	2.121	1.936	1.803	1.603	1.449	1.288	1.201	0.984
湖南	3.435	3.144	2.813	2.276	2.118	1.83	1.665	1.456	1.275	1.127	1.05	0.984	0.924	0.88	0.804
海南	1.781	2.324	3.516	3.108	3.003	2.627	2.534	2.333	1.948	1.957	2.025	1.796	1.584	1.548	1.448
四川	2.926	2.782	2.558	2.367	2.372	2.013	1.66	1.523	1.414	1.355	1.102	0.988	0.864	0.766	0.718
贵州	9.087	9.068	7.867	5.912	5.895	5.051	4.508	4.105	3.611	3.044	2.688	2.511	2.196	1.851	1.675
云南	5.11	4.881	4.261	3.676	3.681	3.319	2.784	2.494	2.15	1.785	1.521	1.385	1.329	1.353	1.089
陕西	4.486	4.519	4.11	3.645	3.546	3.389	3.032	3.013	2.855	2.756	2.678	2.534	2.314	2.029	2.087
甘肃	6.698	6.07	5.688	4.949	4.569	4.181	3.964	3.628	3.348	3.123	3.041	2.759	2.686	2.546	2.438
青海	4.314	4.54	4.096	3.993	3.831	3.061	2.925	3.076	3.023	2.599	2.286	2.502	2.35	2.105	2.011
内蒙古	7.984	7.406	6.581	5.94	5.641	5.188	5.263	4.945	4.535	4.424	4.373	4.352	5.157	5.506	6.135
广西	2.576	2.374	2.21	1.817	1.824	1.799	1.807	1.784	1.618	1.481	1.297	1.237	1.294	1.237	1.254
西藏	11.933	11.032	9.855	8.314	8.135	7.71	8.25	7.959	7.693	7.334	7.198	6.584	7.447	7.715	8.288
宁夏	5.914	5.763	5.423	5.057	5.773	5.086	4.949	5.036	5.126	5.188	5.331	5.375	5.07	4.708	4.492
新疆	1.722	1.519	1.332	1.2	1.126	0.983	0.796	0.735	0.61	0.587	0.528	0.448	0.403	0.382	0.326

根据表 4.3 中研究期 2005 年至 2019 年北京市、天津市、上海市、重庆市、广东省和湖北省 6 个试点行政区的碳排放强度绘制折线图(图 4.1)。可以看出,

除天津市在 2005 年至 2017 年碳排放强度持续下降,但在 2018 年至 2019 年突然升高以外,北京市、上海市、重庆市、广东省和湖北省在 2005 年至 2019 年的碳排放强度均保持持续下降趋势。

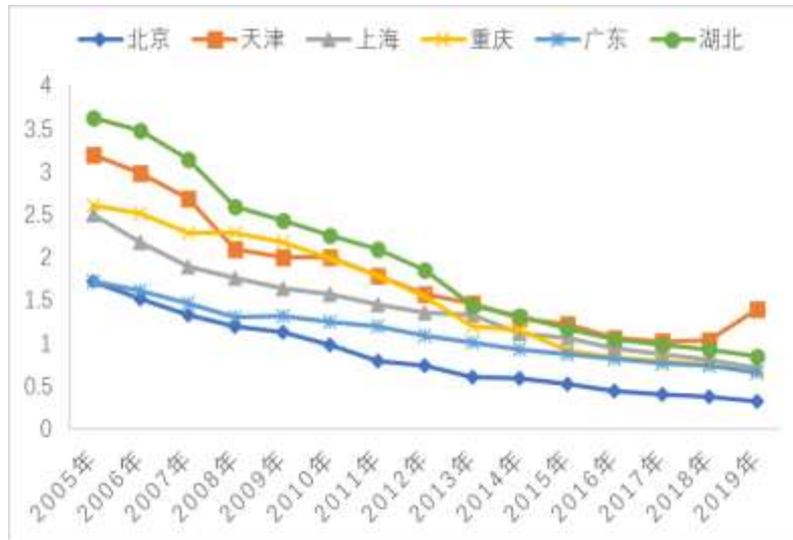


图 4.1 试点省级行政区研究期内二氧化碳排放强度

那么这些城市在 2014 年至 2019 年碳排放强度下降,到底是试行的碳排放权交易对二氧化碳排放量产生了实际抑制效果呢,还是 2005 年约束碳减排后,所有城市碳排放强度均降低的大趋势呢?本章通过前文提出的时变权重合成控制法来研究这一问题。

4.3.3 影响因素

一般情况下,无论是经济发达的省份还是经济发展欠佳的省份,CO₂强度指标是随着时间而下降的。总结杨秀汪(2021)、黄志平(2018)以及管志贵(2019)等学者的研究,碳排放强度主要取决于化石能源的碳排放系数、结构、强度以及其在能源消费总量中的比例,同时也受城市技术进步、经济增长、产业结构变化、农业工业化和城市化进程与规模等影响。

本文参考前人经验,将影响碳排放强度的解释变量定位在经济发展水平与资本存量,能源强度和结构,产业结构和深化程度,人口密度和交通因素以及技术进步五个方面,最终影响因素如表 4.4 所示。其中,经济发展水平由人均 GDP

表示，资本存量由人均资本存量表示。资本存量参考单豪杰老师（2008）提出的计算公式：

$$\text{本期资本存量} = \text{上期资本存量} * (1 - 10.96\%) + \text{本期固定资产形成总额} \quad (4.2)$$

能源强度用能源消费在 GDP 中的占比表示，能源结构由煤炭能源消费占比表示。产业结构由第二产业在 GDP 中的占比表示，产业深化程度由高级化水平，即第三产业增加值与第二产业增加值的比值表示。交通因素由人均民用汽车拥有量表示。技术进步由研究与试验发展（R&D）经费投入强度表示。

表 4.4 碳排放强度影响因素

序号	指标	单位
X1	人均 GDP	元
X2	人均资本存量	元
X3	能源消费在 GDP 中的占比	万吨/亿元
X4	化石能源消费占比	%
X5	第二产业增加值占比	%
X6	高级化水平	-
X7	人口密度	人/平方公里
X8	人均私人汽车拥有量	辆
X9	研究与试验发展（R&D）经费投入强度	-

通过查阅《中国统计年鉴》《中国能源统计年鉴》《中国城市统计年鉴》和万得数据库等来源，碳排放权交易试点地区和合成组库省级行政区的上述指标在预干预期 2005 年至 2013 年的平均值如下表 4.5 所示。

表 4.5 预干预期影响因素平均值

地区	X1	X2	X3	X4	X5	X6	X7	X8	X9
北京	69470.209	112003.336	0.384	23.678	21.528	3.510	1260.110	0.199	5.646
天津	65844.492	123999.888	0.532	36.596	47.799	1.010	2550.288	0.113	2.460
上海	73627.667	120727.370	0.483	21.681	40.847	1.394	3258.376	0.074	2.843
重庆	24792.889	43774.822	0.702	49.659	45.662	0.987	1630.417	0.036	1.216
广东	41460.446	52225.651	0.382	24.586	49.138	0.918	2505.784	0.071	1.651
湖北	25218.973	41041.676	0.733	56.606	46.699	0.891	1813.922	0.033	1.489
河北	26396.778	45092.817	1.021	70.036	47.374	0.819	2364.523	0.063	0.762
山西	23526.093	40654.884	1.295	68.407	59.128	0.595	2806.620	0.063	0.951
黑龙江	25175.996	40810.004	0.681	34.284	50.235	0.635	4087.376	0.045	1.053
吉林	28829.111	65866.360	0.786	53.531	37.497	1.245	1825.913	0.050	0.945
辽宁	38306.464	68497.101	0.815	47.583	50.022	0.795	1789.961	0.062	1.542
江苏	47833.222	72023.845	0.429	42.808	52.636	0.752	1983.347	0.064	1.982
浙江	47442.378	74404.154	0.420	27.240	51.509	0.822	1746.870	0.092	1.704
安徽	18778.633	28054.259	0.636	57.617	47.175	0.918	2138.236	0.032	1.274
江西	19235.556	30203.284	0.510	58.280	52.560	0.670	4405.407	0.028	0.882
山东	37529.771	61772.582	0.638	49.713	51.995	0.700	1394.031	0.066	1.557
河南	22186.499	40472.875	0.649	59.954	52.685	0.627	5301.503	0.039	0.828
湖南	22204.333	31920.890	0.662	58.082	44.706	0.966	2994.936	0.029	1.040
海南	21697.556	37399.941	0.565	12.306	26.569	1.706	2598.687	0.041	0.319
四川	19417.333	29853.311	0.660	46.523	46.221	0.854	2466.649	0.039	1.401
贵州	12142.218	22281.147	1.233	68.961	40.542	1.098	2992.464	0.029	0.602
云南	15111.667	27653.067	0.869	60.694	40.246	1.084	3099.659	0.046	0.602
陕西	24317.333	43054.396	0.642	49.889	51.900	0.767	5349.497	0.045	2.184
甘肃	14865.095	25162.916	1.018	47.109	45.918	0.913	4138.492	0.030	1.027
青海	21844.416	52375.278	1.249	34.578	39.459	1.257	2302.360	0.052	0.611
内蒙古	41257.269	81869.998	0.999	54.666	41.969	1.060	799.362	0.068	0.499
广西	18543.023	36077.987	0.627	52.768	38.825	1.066	1420.368	0.030	0.557
宁夏	23590.382	57793.339	1.529	49.206	46.847	0.920	1522.456	0.061	0.726
新疆	23490.172	46347.243	1.155	38.784	45.150	0.830	4287.928	0.056	0.418

4.4 结果分析

4.4.1 试点地区实证结果

根据前章提出的时变权重合成控制法和组合灰色模型预测算法,北京市、天津市、上海市、重庆市、广东省和湖北省 6 个碳排放权交易政策试点省级行政区在 2019 年的干预效果如表 4.6 所示。其中,北京市和广东省在 23 个未实行低碳

政策的省级行政区组成的合成组库中没有找到误差范围内的合成组, 仅和浙江省相对最为相似。天津市、上海市、重庆市和湖北省 4 个省级行政区在合成组库中均找到了误差范围内的合成对照组。

表 4.6 试点地区政策干预结果

地区	碳强度实际值	合成地区碳强度	碳强度减少量
北京	0.326	0.761	0.435
天津	1.403	0.904	-0.499
上海	0.718	0.853	0.135
重庆	0.659	0.895	0.236
广东	0.659	0.761	0.102
湖北	0.845	0.945	0.100

天津市的合成组为浙江省和内蒙古自治区 2 个地区, 在预干预期的影响因素权重为

$$V = (1.000, 105.189, 105.501, 120.222, 106.167, 109.686, 87.467, 77.444, 52.829),$$

时变合成权重为

$$W_{T_0}^* = \begin{bmatrix} 0.832 & 0.843 & 0.852 & 0.929 & 0.932 & 0.877 & 0.902 & 0.917 & 0.922 \\ 0.168 & 0.157 & 0.148 & 0.071 & 0.068 & 0.123 & 0.098 & 0.083 & 0.078 \end{bmatrix}.$$

根据 $W_{T_0}^*$ 使用改进的组合灰色模型分别预测两个合成城市的 $W_{T-T_0}^*$, 模型选择误差为

$$SSE = \begin{bmatrix} 0.006 & 0.006 & 0.005 \\ 0.026 & 0.025 & 0.017 \end{bmatrix},$$

浙江省和内蒙古自治区的最小拟合误差值分别为 0.005 和 0.017, 验证的最优模型均为新陈代谢 $GM(1,1)$ 。因此两个合成城市都适用新陈代谢 $GM(1,1)$ 来预测干预期权重, 其预测结果为

$$W_{T-T_0}^* = \begin{bmatrix} 0.943 & 0.939 & 0.931 & 0.949 & 0.968 & 0.973 \\ 0.057 & 0.061 & 0.069 & 0.051 & 0.032 & 0.027 \end{bmatrix}.$$

最终天津和时变权重合成天津在研究期的碳排放强度如图 4.2 所示。

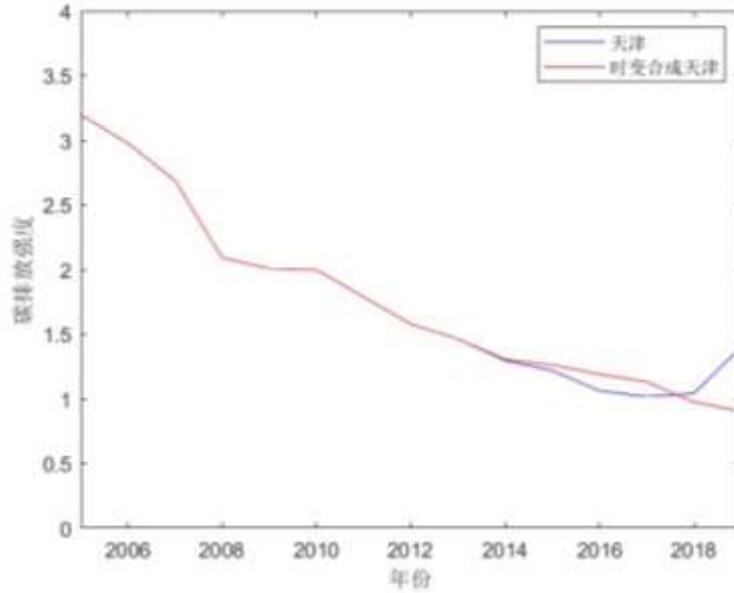


图 4.2 天津和合成天津历年碳排放强度

上海市的合成组为江苏省和浙江省 2 个地区，在预干预期的影响因素权重为

$$V = [1.000, 6.328, 3.919, 4.520, 1.829, 1.819, 3.346, 3.332, 5.827],$$

时变合成权重为

$$W_{T_0}^* = \begin{bmatrix} 0.858 & 0.117 & 0.010 & 0.010 & 0.010 & 0.292 & 0.219 & 0.434 & 0.810 \\ 0.142 & 0.883 & 0.990 & 0.990 & 0.990 & 0.708 & 0.781 & 0.566 & 0.190 \end{bmatrix}.$$

根据 $W_{T_0}^*$ 使用改进的灰色预测组合模型分别预测两个合成城市的 $W_{T-T_0}^*$ ，模型选择误差为

$$SSE = \begin{bmatrix} 0.392 & 0.431 & 0.152 \\ 0.624 & 0.623 & 0.496 \end{bmatrix},$$

江苏省和浙江省的最小拟合误差值分别为 0.152 和 0.496，验证的最优模型均为新陈代谢 $GM(1,1)$ ，预测结果为

$$W_{T-T_0}^* = \begin{bmatrix} 0.270 & 0.538 & 0.630 & 0.735 & 0.732 & 0.795 \\ 0.730 & 0.462 & 0.370 & 0.265 & 0.268 & 0.205 \end{bmatrix}.$$

最终上海和时变权重合成上海在研究期的碳排放强度如图 4.3 所示。

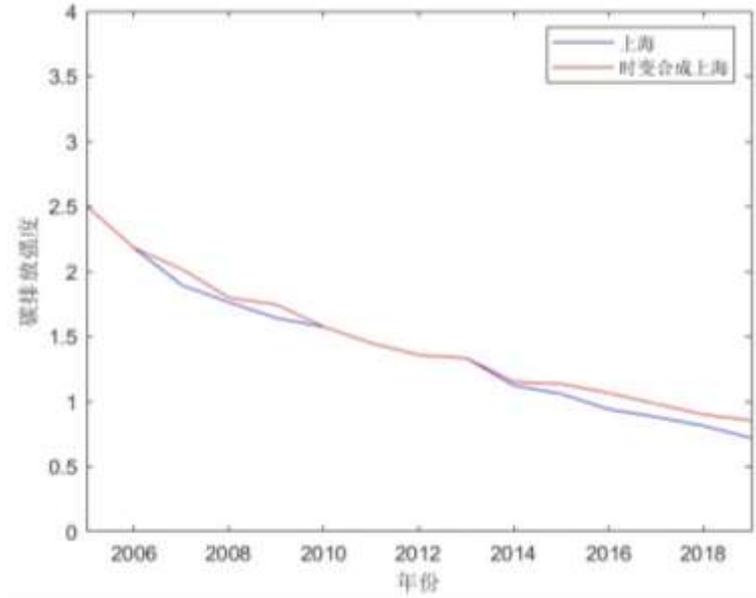


图 4.3 上海和合成上海历年碳排放强度

重庆市的合成组为江苏省、安徽省、海南省和广西壮族自治区 4 个地区，在预干预期的影响因素权重为

$$V = [1.000, 732.952, 0.194, 905.301, 0.268, 352.932, 275.595, 13.664, 28.760],$$

时变合成权重为

$$W_{T_0}^* = \begin{bmatrix} 0.250 & 0.279 & 0.493 & 0.318 & 0.334 & 0.368 & 0.545 & 0.824 & 0.970 \\ 0.234 & 0.142 & 0.055 & 0.185 & 0.170 & 0.166 & 0.115 & 0.050 & 0.010 \\ 0.267 & 0.298 & 0.045 & 0.180 & 0.166 & 0.156 & 0.084 & 0.036 & 0.010 \\ 0.249 & 0.281 & 0.408 & 0.317 & 0.330 & 0.309 & 0.256 & 0.091 & 0.010 \end{bmatrix},$$

根据 $W_{T_0}^*$ 使用改进的灰色预测组合模型分别预测两个合成城市的 $W_{T-T_0}^*$ ，模型选择误差为

$$SSE = \begin{bmatrix} 0.69586 & 0.69587 & 0.772 \\ 0.17421 & 0.17419 & 0.192 \\ 0.09264 & 0.09258 & 0.126 \\ 0.24466 & 0.24454 & 0.202 \end{bmatrix},$$

江苏省最小拟合误差值为 0.69586，验证的模型为全信息 $GM(1,1)$ ；安徽省和海南省两组数据的最小拟合误差值分别为 0.17419 和 0.09258，验证的模型为新信息 $GM(1,1)$ ；广西壮族自治区最小拟合误差值为 0.202，验证的模型为新陈代谢

GM(1,1)。组合灰色模型预测结果为

$$W_{T-T_0}^* = \begin{bmatrix} 0.778 & 0.839 & 0.871 & 0.917 & 0.943 & 0.953 \\ 0.066 & 0.056 & 0.048 & 0.038 & 0.030 & 0.028 \\ 0.049 & 0.036 & 0.030 & 0.023 & 0.017 & 0.013 \\ 0.107 & 0.069 & 0.051 & 0.022 & 0.010 & 0.007 \end{bmatrix}。$$

最终重庆和时变权重合成重庆在研究期的碳排放强度如图 4.4 所示。

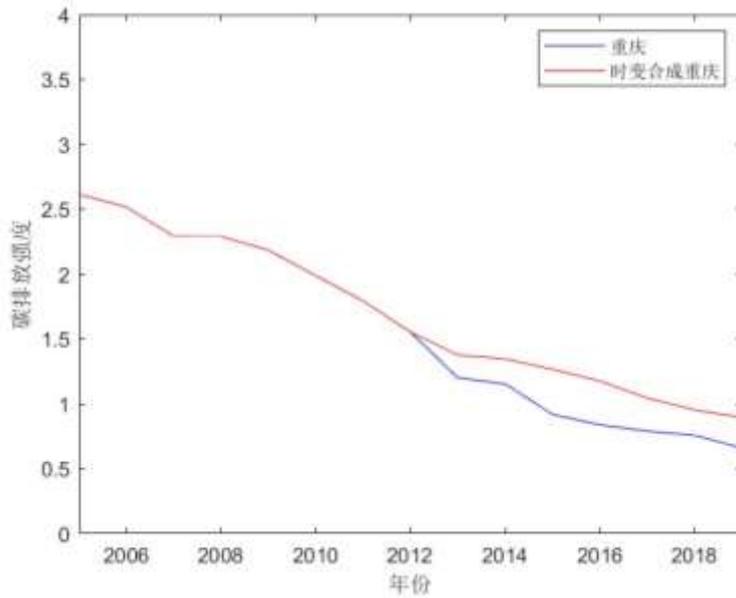


图 4.4 重庆和合成重庆历年碳排放强度

湖北省的合成组为吉林省、辽宁省、江苏省和四川省 4 个地区，在预干预期的影响因素权重为

$$V = [1.000, 0.003, 3.688, 0.032, 994.167, 324.463, 13.184, 0.005, 953.134],$$

时变合成权重为

$$W_{T_0}^* = \begin{bmatrix} 0.175 & 0.191 & 0.203 & 0.172 & 0.178 & 0.188 & 0.192 & 0.181 & 0.041 \\ 0.142 & 0.158 & 0.162 & 0.127 & 0.122 & 0.142 & 0.163 & 0.130 & 0.028 \\ 0.356 & 0.339 & 0.333 & 0.414 & 0.435 & 0.380 & 0.322 & 0.349 & 0.594 \\ 0.326 & 0.313 & 0.303 & 0.287 & 0.266 & 0.291 & 0.322 & 0.339 & 0.338 \end{bmatrix}。$$

根据 $W_{T_0}^*$ 使用改进的灰色预测组合模型分别预测两个合成城市的 $W_{T-T_0}^*$ ，模型选择误差为

$$SSE = \begin{bmatrix} 0.07636 & 0.07649 & 0.083 \\ 0.06454 & 0.06479 & 0.072 \\ 0.09369 & 0.09376 & 0.099 \\ 0.00415 & 0.00416 & 0.003 \end{bmatrix},$$

吉林省、辽宁省和江苏省 3 组数据的最小拟合误差分别为 0.07636、0.06454 和 0.09369，验证的最优模型为全信息 $GM(1,1)$ ；四川省最小拟合误差值为 0.003，验证的最优模型为新陈代谢 $GM(1,1)$ 。组合灰色模型预测结果为

$$W_{T-T_0}^* = \begin{bmatrix} 0.129 & 0.126 & 0.116 & 0.097 & 0.080 & 0.053 \\ 0.091 & 0.076 & 0.049 & 0.041 & 0.034 & 0.027 \\ 0.471 & 0.454 & 0.470 & 0.490 & 0.506 & 0.532 \\ 0.310 & 0.344 & 0.365 & 0.372 & 0.381 & 0.388 \end{bmatrix}。$$

最终湖北和时变权重合成湖北在研究期的碳排放强度如图 4.5 所示。

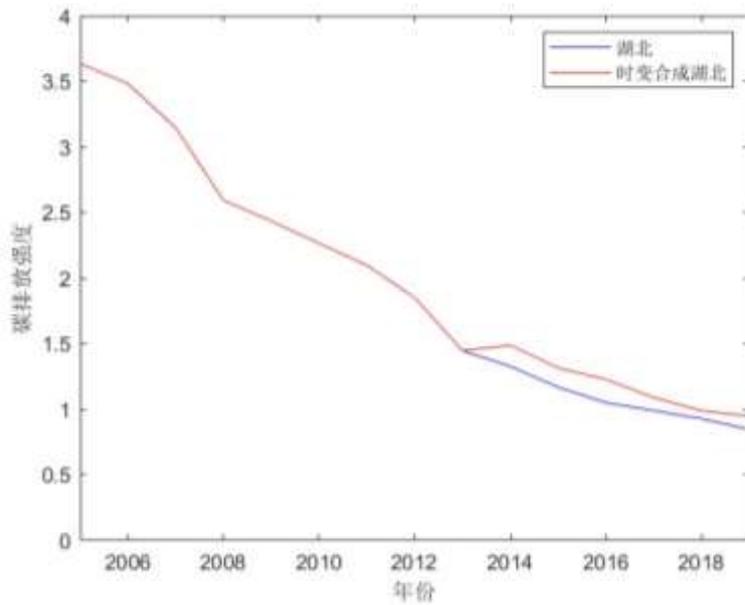


图 4.5 湖北和合成湖北历年碳排放强度

湖北省在以上 4 个试点城市中，合成效果与预测最好。预干预期时变权重合成湖北省的预测合成潜在结果与真实湖北省的观测结果基本重叠。政策干预后，时变权重合成湖北省的碳排放强度折线持续居于真实湖北上方，表明低碳政策减排效果显著。

4.4.2 影响因素与误差分析

6个干预行政区的合成误差值如下表4.7中列“E”所示。将人均GDP (X1) 作为基数, 人均资本存量 (X2)、能源消费在GDP中的占比 (X3)、化石能源消费占比 (X4)、第二产业增加值占比 (X5)、高级化水平 (X6)、人口密度 (X7)、人均私人汽车拥有量 (X8) 和研究与试验发展 (R&D) 经费投入强度 (X9) 8个影响因子在寻找最优对照组的过程中, 相对权重分别如下表4.7所示。

表 4.7 合成误差和影响因子权重

地区	E	X1	X2	X3	X4	X5	X6	X7	X8	X9
北京	3.213	1.000	3.703	2.744	2.712	1.585	1.172	2.767	4.350	4.470
天津	0.255	1.000	105.189	105.501	120.222	106.167	109.686	87.467	77.444	52.829
上海	0.074	1.000	6.328	3.919	4.520	1.829	1.819	3.346	3.332	5.827
重庆	0.332	1.000	732.952	0.194	905.301	0.268	352.932	275.595	13.664	28.760
广东	1.474	1.000	0.182	483.066	2.649	4.040	4.954	0.629	1.486	26.497
湖北	0.095	1.000	0.003	3.688	0.032	994.167	324.463	13.184	0.005	953.134

由于各影响因素单位不同, 取值差距较大, 只分析相对权重可能无法正确推断各地区的碳减排措施侧重。因此考虑影响因素在合成对照地区过程中的实际作用量, 将预干预期内影响因素观测值与合成过程中的相对权重相乘, 得到各因素对碳排放强度的实际影响分量, 计算结果具体数值见下表4.8。

表 4.8 各因素实际影响分量

地区	X1	X2	X3	X4	X5	X6	X7	X8	X9
北京	69470	414763	1	64	34	4	3486	1	25
天津	65844	13043435	56	4400	5075	111	223067	9	130
上海	73628	763926	2	98	75	3	10902	0	17
重庆	24793	32084858	0	44957	12	348	449335	0	35
广东	41460	9510	185	65	199	5	1576	0	44
湖北	25219	127	3	2	46426	289	23915	0	1419

根据表4.8可以看出, 整体来说经济水平在合成对照组时影响显著, 城市化水平次之。人均GDP (X1)、人均固定资本存量 (X2) 和人口密度 (X7) 影响分量数值巨大, 在6个地区中都几乎起到决定性作用。另外, 对于天津市、重庆

市和湖北省 3 个行政区，能源结构和产业结构同样作用明显，化石能源消费占比和第二产业增加值占比影响分量数值较大。值得注意的是，对比研究与试验发展（R&D）经费投入强度（X9）数值占比可以发现，湖北省是 6 个试点省市中唯一在技术进步程度上影响明显的地区。

溯源未成功找到合成对照组的北京市和广东省，从表 4.8 可以看出，主要是人均 GDP 和人均资本存量影响分量值过于突出。除此之外，这两个地区的人口密度影响分量值也远超其他影响因素，因此在要求的误差内找不到经济水平和城市化水平接近的合成组。

4.4.3 政策效果与生效机制

四个应用时变权重合成控制法找到对照组的试点城市，比较图 11 至图 14 中干预后的碳排放强度走势可以发现，试行碳排放权交易政策确实对该地区的二氧化碳排放产生了抑制作用。

天津市在 2018 和 2019 两年，实际碳排放强度陡然减小，因此从 2019 年单年度数据来看，或许政策减排效果并不突出，但对比开始实施政策的 2014 年至 2017 年时变合成天津数据来看，天津市二氧化碳排放强度在政策提出后是有明显改善的。天津市在碳减排过程中，经济水平、能源结构、产业结构和城市化水平平均有明显影响，人均 GDP、人均资本存量、化石能源消费占比、第二产业增加值占比和人口密度影响力比较平均。研究与试验发展（R&D）经费投入强度影响相对较弱，技术进度程度在天津市的减排政策中并没有显著体现。

上海市由于其经济发展居于国内领先地位，从各因素的影响程度来看，和北京非常类似。上海的经济水平和城市化水平尤为突出，人均 GDP、人均资本存量和人口密度几乎组成了所有的碳强度影响值。在 2006 年至 2010 年寻找合成组城市时，根据解释变量没有找到碳排放强度完全重合的合成组，但最终合成误差 0.074，仍在可接受范围内。在 2014 年实施低碳政策后，对比时变合成上海的碳排放强度来看，该城市的低碳政策对碳排放持续存在抑制效果。

重庆市在预干预期 2005 年至 2012 年找到了碳排放强度几乎一致的合成组。在碳减排过程中，人均资本存量影响最为突出，人口密度次之，化石能源消费占比和人均 GDP 也有明显影响。政策干预后，对比时变合成重庆的数据，重庆市

低碳政策减排效果十分明显。

湖北省与前几个地区有所不同，他的碳减排措施在第二产业增加值占比上表现明显，人均 GDP 和人口密度次之。同时，湖北省研究与试验发展（R&D）经费投入强度影响值 1419，远高于其他 5 个省级行政区，说明该省技术进步水平在碳减排措施中也被重视，并对减排成果起到了一定的促进作用。

4.5 本章小结

随着全球气候变化日益严重，低碳政策已成为各国政府应对气候变化的重要手段。低碳政策不仅有助于减少温室气体排放，降低气候变化风险，还有助于推动可持续发展和经济增长。2010 年我国正式发布关于开展低碳试点工作的通知，低碳政策成为我国社科领域内一个重要课题而备受关注。2013 年底，中国第一批碳排放权交易试点工作正式施行，到 2021 年，我国碳排放权交易市场已在全国正式开市。

从试点城市的实证结果来看，经济和城市化水平对碳排放强度的影响至关重要。北京市和广东省主要由于经济水平即人均 GDP 和人均资本存量与其他地区差异突出，没有成功找到合成对照组。天津市、上海市、重庆市和湖北省四个省级行政区，虽然减排措施侧重点不一样，但通过建立碳市场和碳交易机制，确实有效地抑制了碳排放。

5 结论与展望

5.1 全文总结

本文将个体权重时变性引入到合成控制的计算过程中。在干预预期，该处理方法可以让合成组和干预组几乎完全重合，在比较案例中极大地缩小了控制组和对照组的误差。此后，通过改进的组合灰色模型来预测干预后时变权重变化趋势，并将每个时间节点上不同个体合成权重值收缩到原始合成控制法的“非负且和为1”权重限制中。和原始的合成控制法比较，本文提出的方法拥有更高的精度和更灵活宽泛的适用场景，在减小误差的同时，并不损失合成控制保有的“透明度”和“防外推”两个优点。

时间序列预测方法虽然经历了很长时间的发展，但数据规模对预测方法的效率影响仍然极大。变权重合成控制法中的权重预测期为干预后时期，已知时序数据则为干预前合成权重序列。鉴于年度数据的可获得性，其数据跨度一般较小，因此用灰色模型来预测干预期数据更为合理。但同样的，灰色模型预测精度也受时间序列跨度和干预期和预测期的分配方式所影响，在运用该方法预测干预后权重时，本文根据数据量进行灰色预测模型的调整，尽可能在多组合成单元的小样本数据上得到最为精确的预测结果。

另外，实证部分通过剖析解释变量的影响值，可以明确知道各试点城市采取不同措施完成减排目标时的侧重。结合影响程度分析各城市低碳政策的效果，其结果更具准确性和可复制性。2020年“双碳”目标提出后，其余未参与碳排放权交易试点的地区，参考6个试点地区在主要指标上的影响分量分布，可以根据与本地区的措施相似程度，采取相同或类似的方案来完成本地区的碳减排工作。

5.2 研究展望

在今后的研究中，本文所研究的内容还可以从以下几个角度展开深入讨论。

其一，由于计算的复杂性，没有考虑解释变量权重的时变性。在干预预期，各影响指标并不完全采用一样的权重 V 来计算时变合成权重 W_{T_0} ，这样可以保证

预干预期和干预期的个体合成权重都更具有准确性。

其二，可以进一步引入新的小样本时间序列预测方法。本文根据加州禁烟案例和我国碳排放权交易试点案例实证数据预干预期短，干预期长的特性，采用组合的灰色预测模型来计算干预期时变合成权重。在模型选取时，只在全数据 $GM(1,1)$ 、新信息 $GM(1,1)$ 和新陈代谢 $GM(1,1)$ 这三种方法中选择最小误差模型。后续可以继续预测方法部分深入挖掘，提出兼容其他数据类型的更稳健、应用更广泛的预测方法。

参考文献

- [1] Abadie A. Using Synthetic Controls to Evaluate an International Strategic Positioning Program in Uruguay: Feasibility, Data Requirements and Methodological Aspects1[J]. Journal of Economic Literature,2021, 59(2):391-425.
- [2] Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program[J]. Journal of the American Statistical Association. 2010,105(490):493-505.
- [3] Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. Comparative Politics and the Synthetic Control Method[J]. American Journal of Political Science,2015, 59 (2):495-510.
- [4] Abadie, Alberto, and Jérémy L' Hour. A Penalized Synthetic Control Estimator for Disaggregated Data[J]. Journal of the American Statistical Association, 2019, 116(536):1817-1834.
- [5] ANAVA O, HAZAN E, MANNOR S, et al. Online Learning for Time Series Prediction[J]. Journal of Machine Learning Research,2013,30:172-184.
- [6] Arkhangelsky, Dmitry, Susan Athey, David A, Hirshberg, Guido W, Imbens, Stefan Wager. Synthetic Difference in Differences. Working Paper,2019.
- [7] Ben-Michael Eli, Avi Feller, Jesse Rothstein. The Augmented Synthetic Control Method[J]. Journal of the American Statistical Association,2021,1789-1803.
- [8] BINKOWSKI M, MARTI G, DONNAT P. Autoregressive Convolutional Neural Networks for Asynchronous Time Series[J]. Submission and Formatting Instructions for ICML,2017.
- [9] Botosaru, Irene, and Bruno Ferman. On the Role of Covariates in the Synthetic Control Method[J]. Econometrics Journal,2019,22 (2):117-30.
- [10]BOX G, JENKINS G. Time Series Analysis, ForecastingandControl[M]. Holden-Day, 1990.

- [11]Chernozhukov, Victor, Kaspar Wüthrich, and Yin chu Zhu. Practical and Robust t-Test Based Inference for Synthetic Control and Related Methods[J]. Research Papers in Economics,2019.
- [12]Chernozhukov, Victor, Kaspar Wüthrich, and Yinchu Zhu. An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls[J]. Journal of the American Statistical Association,2021,536(116):1849-1864.
- [13]Doudchenko, Nikolay, and Guido W. Imbens. Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis. NBER Working Paper,2016.
- [14]Dube, Arindrajit, and Ben Zipperer. Pooling Multiple Case Studies Using Synthetic Controls: An Application to Minimum Wage Policies[J]. Social Science Electronic Publishing,2015.
- [15]DURBIN J, KOOPMAN S J. Time Series Analysis by State Space Methods [M]. Oxford: Oxford University Press, 2012.
- [16]Firpo, Sergio, and Vitor Possebom. Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets[J]. Journal of Causal Inference, 2018,6(2).
- [17]Gunsilius, Florian. Distributional Synthetic Controls[J]. Econometrica,2023, 91(3).
- [18]Hahn, Jin Yong, and Ruoyao Shi. Synthetic Control and Inference. Econometrics,2017,5 (4).
- [19]HAMILTON J. Time Series Analysis [M]. Princeton: Princeton University Press, 1994.
- [20]Jiawei C, Man H E C, Seiler M J, et al. Household tenure choice and housing price volatility under a binding home-purchase limit policy constraint[J]. Journal of Housing Economics, 2018, 41(9):124-134.
- [21]Kaul A, Klößner, Stefan, Pfeifer G, et al. Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates[J]. Mpra Paper, 2015.
- [22]Kinn D. Synthetic Control Methods and Big Data[J]. Papers, 2018.
- [23]Kreif, Noémi, Richard Grieve, Dominik Hangartner, Alex James Turner,

- Silviya Nikolova, and Matt Sutton. Examination of the Synthetic Control Method for Evaluating Health Policies with Multiple Treated Units[J]. Health Economics, 2016,25 (12):14-28.
- [24]Li, Kathleen T. Statistical Inference for Average Treatment Effects Estimated by Synthetic Control Methods[J]. Journal of the American Statistical Association, 2020,115(532):68-83.
- [25]Osman, M. M., Syuid S., Bachok S., Khalid N. Factors influencing housing prices among local people: a study in Perak, Malaysia[J]. Advanced Science Letters, 2017, 23(1):165-168.
- [26]Quade, Dana. Nonparametric Analysis of Covariance by Matching. Biometrics, 1982,38(3):597-611.
- [27]Samartsidis, Pantelis, Shaun R. Seaman, Anne M. Presanis, Matthew Hickman, and Daniela De Angelis. Assessing the Causal Effect of Binary Interventions from Observational Panel Data with Few Treated Units[J]. Statistical Science, 2019,34(3):486-503.
- [28]Thomas Jansson. Housing choices and labor income risk[J]. Journal of Urban Economics,2017,(99):107-119.
- [29]Xu, Yiqing. Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models[J]. Political Analysis,2017,25(1):57-76.
- [30]2030年前碳达峰行动方案[N]. 人民日报, 2021-10-27(007).
- [31]白仲林, 孙艳华, 高泽铭. 商品房限购政策的实体经济发展效应研究[J]. 统计研究, 2019, 36(11):37-48.
- [32]蔡芸, 韩惠敏. 均等化试点政策与超大城市公共卫生系统韧性评价[J]. 中国卫生事业管理, 2023, 40(10):730-735+750.
- [33]单豪杰. 中国资本存量 K 的再估算:1952~2006 年[J]. 数量经济技术经济研究, 2008, 25(10):17-31.
- [34]段娟. 中国绿色低碳发展道路的实践探索及其启示[J]. 宁夏社会科学, 2019(06):27-34.
- [35]管志贵, 田学斌. 碳市场政策能否减少区域碳排放?——基于中国省级面板数

- 据的实证研究[J]. 内蒙古社会科学(汉文版), 2019, 40(01):126-132.
- [36]何露雪, 林力佳, 郭利. 基于DID和SDID的科技专项资助与企业科技成果关系研究——以珠三角为例[J]. 科技和产业, 2023, 23(11):89-97.
- [37]黄秀路, 武宵旭, 袁圆等. 绿色金融改革的节能效应与机制[J]. 中国人口·资源与环境, 2023, 33(08):27-36.
- [38]黄志平. 碳排放权交易有利于碳减排吗?——基于双重差分法的研究[J]. 干旱区资源与环境, 2018, 32(09):32-36.
- [39]科技支撑碳达峰碳中和实施方案(2022—2030年)[J]. 中国科技奖励, 2022(12):13-20.
- [40]林辉, 孙煦初. “一带一路”倡议对沿线国家融资环境的促进作用[J]. 经济与管理评论, 2022, 38(01):33-46.
- [41]刘传明, 孙喆, 张瑾. 中国碳排放权交易试点的碳减排政策效应研究[J]. 中国人口·资源与环境, 2019, 29(11):49-58.
- [42]刘学. 反事实框架下因果推论的原则和机制[J]. 东南大学学报(哲学社会科学版), 2020, 22(04):98-109+157.
- [43]刘紫薇, 袁怀宇. “双碳”目标下我国碳权交易政策评价研究[J]. 投资与合作, 2023, (05):163-165.
- [44]陆贤伟. 低碳试点政策实施效果研究——基于合成控制法的证据[J]. 软科学, 2017, 31(11):98-101+109.
- [45]马忠贵, 徐晓晗, 刘雪儿. 因果推断三种分析框架及其应用综述[J]. 工程科学学报, 2022, 44(07):1231-1243.
- [46]任以胜, 龙一鸣, 陆林. 流域生态补偿政策对受偿地区水污染强度的影响——以新安江流域为例[J]. 经济地理, 2023, 43(11):181-189.
- [47]王东明, 陈都鑫. 因果推断: 起源和发展[J]. 控制工程, 2022, 29(03):464-473.
- [48]吴滨. 政策评价方法综述[J]. 统计与管理, 2021, 36(06):15-22.
- [49]吴志祥, 计小青, 许泽庆. 区域一体化与县域经济发展活力——基于长三角扩容的准自然实验[J]. 华东经济管理, 2024, 38(01):14-25.
- [50]邢子怡. 高质量发展背景下国家高新区设立对城市创新的影响研究[J]. 统计

- 与管理, 2023, 38(07):118-128.
- [51]杨帆, 王永佳. 内蒙古低碳试点工作现状与对策建议[J]. 环境与发
展, 2021, 33(01):216-221.
- [52]杨海民, 潘志松, 白玮. 时间序列预测方法综述[J]. 计算机科
学, 2019, 46(01):21-28.
- [53]杨秀汪, 李江龙, 郭小叶. 中国碳交易试点政策的碳减排效应如何?——基于
合成控制法的实证研究[J]. 西安交通大学学报(社会科学版), 2021,
41(03):93-104+122.
- [54]姚永玲, 陈兴涛. 北京城市副中心建设促进当地高质量发展了吗? ——基于
合成控制法的检验[J]. 地理科学进展, 2023, 42(11):2071-2083.
- [55]赵凌霄, 李智扬, 屈磊磊. 基于 EMD 和 CatBoost 算法的改进时间序列模型——
以大连市 PM_{2.5} 预测为例 [J/OL]. 南京林业大学学报(自然科学
版):1-8[2023-12-28].
- [56]郑艳妮, 赵玉锐, 梁莉萍等. 基于 SARIMA、Holt-winters 与 Prophet 三种时间
序列模型的 2021 年武威市丙型病毒性肝炎发病预测[J]. 疾病预防控制通
报, 2023, 38(05):49-53+68.
- [57]中共中央国务院关于完整准确全面贯彻新发展理念做好碳达峰碳中和工作
的意见[N]. 人民日报, 2021-10-25(001).
- [58]周蒙蒙. 双碳目标下的碳交易管理[J]. 现代国企研究, 2022(08):76-79.

攻读硕士学位期间承担的科研任务及主要成果

科研项目:甘肃省中央引导地方科技发展资金项目《城市计算方法体系构建及甘肃智慧城市应用》(项目编号:YDZX20216200001876)。

已完成论文:牛成英,张颖,闫新宇.基于多源数据的城市医疗设施布局优化——以兰州市主城区为例[J/OL].西华师范大学学报(自然科学版),2024.

致 谢

感谢导师在研一帮助我确定题目和研究方向，让我有充足的时间查阅文献学习方法，并在之后持续关注我的论文进展进行学习指导，这是我完成这篇硕士论文的基石，学生朽木拖沓至今才勉强写完。

感谢师门和同学三年关照，让我安心度过了这几年异地求学生活。感谢学校和老师在疫情期间提供物资，随时关注身体情况。感谢家人朋友持续的关心，让我无数次收到鼓励再投入学习中。感谢自己有勇气在工作数年后重新选择脱产学习，沉淀后重新出发。

前路漫漫，希望大家都能在自己选择的道路上踏实前行，期待下次遇见。