

北方清洁取暖项目减排效果再评估

基于机器学习方法

邹恬华

Renmin University of China

Oct, 2024

- 评估北方清洁取暖项目对于空气污染的影响
- 政策背景：清洁取暖，5 批（2017-2022）88 个城市
- 边际贡献：更全面的评估、更丰富的政策细节、更前沿的实证方法
- 识别策略：使用机器学习的两步法
 - ▶ Step 1: 获得不能被气象和周期性因素解释的污染数据
 - ▶ Step 2: 异质性处理效应稳健的 / TWFE staggered DID
- 实证结果
 - ▶ 显著缓解了空气污染，取暖季缓解效应更大，非取暖季也有缓解效应
 - ▶ 政策强度、冬季气温越高的地区处理效应更高，煤改气最有效
 - ▶ 考虑政策强度（Continuous DID）、溢出效应、事前趋势，结果稳健
 - ▶ 成本收益比不及预期

1 政策背景

2 相关文献与边际贡献

3 数据

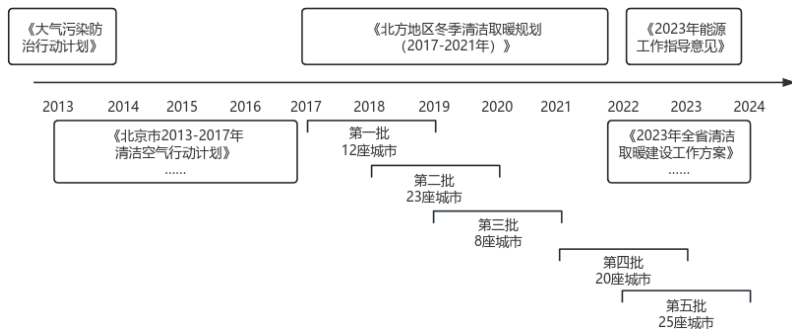
4 识别策略

5 实证结果

6 结论

政策背景：时间线

● 时间线（五批共 88 个城市）¹

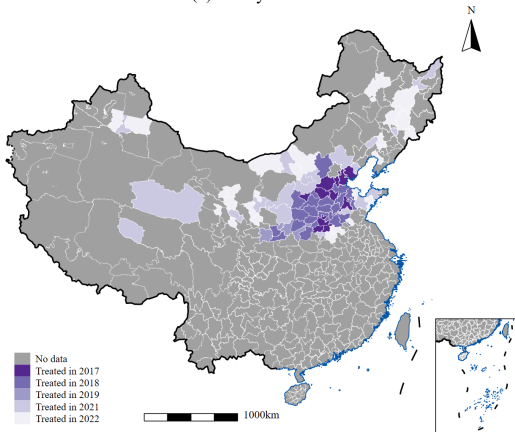


¹ 感谢赵晨同学对此图的贡献

政策背景：时间线

● 优先“2+26 重点城市”²

(c) Policy timeline

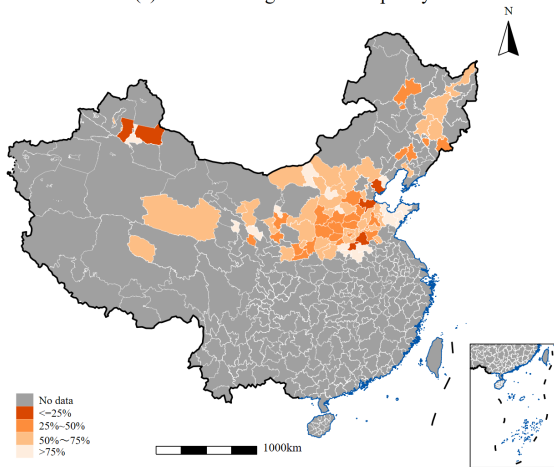


² 根据《京津冀及周边地区 2017 年大气污染防治工作方案》（以下简称《工作方案》），北京市，天津市，河北省石家庄、唐山、廊坊、保定、沧州、衡水、邢台、邯郸市，山西省太原、阳泉、长治、晋城市，山东省济南、淄博、济宁、德州、聊城、滨州、菏泽市，河南省郑州、开封、安阳、鹤壁、新乡、焦作、濮阳市（以下简称“2+26”城市）被确定为京津冀大气污染传输通道城市。

政策背景：政策实施前的清洁取暖水平

- “2+26 重点城市” 事前清洁取暖水平较低

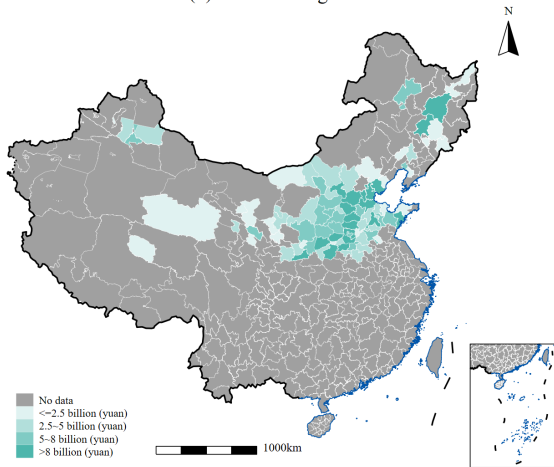
(a) Clean heating ratio before policy



政策背景：政策实施力度

● “2+26 重点城市” 资金投入大

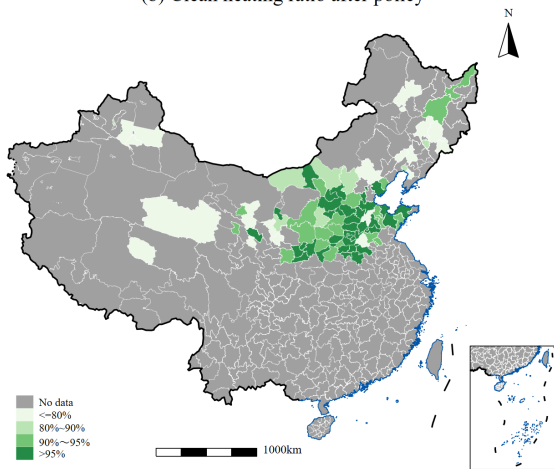
(d) Total funding value



政策背景：政策实施前的清洁取暖水平

- “2+26 重点城市” 事后清洁取暖水平较高

(b) Clean heating ratio after policy



政策背景：措施

● 北方地区冬季清洁取暖规划（2017-2021 年）

▶ 因地制宜选择供暖热源

- ★ 可再生能源供暖：地热/生物质能/太阳能
- ★ 天然气
- ★ 电供暖
- ★ 工业余热
- ★ 清洁燃煤集中供暖

▶ 全面提升热网系统效率

▶ 有效降低用户取暖能耗

● 实际实施过程中

▶ 也结合“洁净煤 + 节能环保炉具”³

▶ 由于资金不足，部分地区并没有实现 100% 清洁取暖改造

³ 2019 年 7 月，国家能源局综合司征求《关于解决“煤改气”“煤改电”等清洁供暖推进过程中有关问题的通知》意见的函：继续坚持民生为重，保障群众安全温暖过冬。……对于偏远山区等暂不能通过清洁供暖替代散烧煤供暖的，重点利用“洁净煤 + 节能环保炉具”等方式替代散烧煤。

政策措施如何减少空气污染

- 直接去除：工业余热
- 使用更清洁的能源：天然气、可再生能源供暖、洁净煤（散煤）
- 集中处理：电供暖、清洁燃煤集中供暖
- 减少能源使用量：提升热网系统效率、降低用户取暖能耗

- 1 政策背景
- 2 相关文献与边际贡献
- 3 数据
- 4 识别策略
- 5 实证结果
- 6 结论

为什么研究清洁取暖政策重要？

- 减少室内外环境污染，带来健康收益
 - ▶ 空气污染具有巨大的健康成本 (Zheng et al., 2019)
 - ▶ 外部性，政策干预有必要
- 室内空气污染，影响集中于农村居民和女性
 - ▶ 家用炉灶不完全燃烧，污染高一个数量级以上 (Wu et al., 2022)
 - ▶ 住宅部门单位煤炭消费导致的过早死亡人数是电力和工业部门的 40 倍 (Yun et al., 2021)
 - ▶ 对于农村人口而言，大约 90% 的空气污染来自于室内空气污染。同时，由于女性更多地在室内工作，女性会更多地得益于清洁取暖项目 (Meng et al., 2019)。

为什么研究清洁取暖政策重要？

- 打破收入不平等-能源贫困-健康不平等的恶性循环
 - ▶ 能源梯子 (energy ladder) 理论认为，随着收入的上升，家庭会逐渐转换使用更清洁的能源 (Baldwin, 1988)
 - ▶ 能源堆叠理论进一步细化，认为随着收入的上升，尽管家庭加大清洁能源的消费占比，但同时也会混合使用多种能源。
 - ▶ 收入较低的家庭倾向于使用污染较高的能源类型 (Wang et al., 2019)，造成收入不平等、能源贫困、健康不平等的恶性循环。

中国清洁取暖项目效果评估

- 显著减少空气污染 (Weng et al., 2022; Zhang et al., 2020)
- 健康收益覆盖前期成本
 - ▶ “2+26” 试点的成本与收益。从成本来看，清洁供暖总成本高达 431 亿元，政府和居民分别承担总成本的 44% 和 56%；从效益上看，清洁取暖项目带来健康经济效益约 1098.5 亿元 (Feng et al., 2021)
- 技术路线：热泵与生物质性价比较高
 - ▶ 一项针对中国北方地区的研究表明，生物质锅炉是生物质资源丰富的农村地区的最佳解决方案，而空气热泵则被认为适用于缺乏生物质或有灵活供暖需求的农村地区 (Deng et al., 2021)
 - ▶ 热泵首次投入的成本高，但通常具有最低的运营成本，因此在长期使用方面具有竞争力 (Zhou et al., 2022)

居民幸福感与家庭特征和政策实施细节相关

- 常见的抱怨：冷 + 贵

- ▶ 居民认为清洁取暖项目的温暖程度低于原来的散煤取暖 (Wu et al., 2020)。煤改气后，农民的取暖成本几乎翻了一番 (Yan et al., 2020)

- 幸福感与收入密切相关

- ▶ 对于高收入和中等收入地区的家庭，清洁取暖项目对室内温度、室内空气污染和生活满意度都有正向影响。但在低收入地区，该政策具有部分有效性：低收入家庭更有可能仍然使用散煤提供部分取暖，因此对室内空气污染、幸福感的影响较小 (Barrington-Leigh et al., 2019)

- 政策实施细节也会对幸福感产生影响

- ▶ 感知到的财政补贴越高，对清洁供暖的满意度就越高 (Gong et al., 2020)
- ▶ 与被动公告相比，面对面和重复传播计划信息对参与的影响更好 (Wang and Xie, 2023)
- ▶ 向家庭提供信息会导致对计划的更积极评价，而将计划作为一项任务来执行会降低家庭的满意度 (Xie and Zhou, 2021)

边际贡献

- 更全面的评估
 - ▶ 以往文献多局限于部分区域，或“2+26”城市，但清洁取暖是一个 package
- 更丰富的政策细节
 - ▶ 88 个城市的自评估报告 + 实地调研，丰富的实施细节，可以衡量政策强度、比较不同技术路线的优劣
- 更前沿的实证方法
 - ▶ 机器学习方法：空气污染数据受到气象（高频）和周期性（非线性）因素影响

- 1 政策背景
- 2 相关文献与边际贡献
- 3 数据
- 4 识别策略
- 5 实证结果
- 6 结论

数据

- 1998-2022 年 6 月区县级别日度数据
- 污染物数据
 - ▶ 来源：NASA 卫星反演数据
 - ▶ 变量：PM2.5、黑碳、SO2
 - ★ PM2.5：直径小于等于 2.5 微米的颗粒物。一些基于北方地区的研究发现，燃煤是北方地区 PM2.5 的最大来源。由于可入肺对人体危害大。
 - ★ 黑碳：来自化石燃料、生物燃料和生物质的不充分燃烧，是 PM2.5 中最主要的一种。是造成气候变暖的重要原因之一，对人体健康也有不利影响。
 - ★ SO2：工业排放的二氧化硫中占比最大的行业就是电力、热力生产和供应业 (27.6%) ⁴
 - ▶ 处理：由栅格数据汇总到区县层面

⁴ 第二次全国污染源普查公报

数据

- 气象数据

- ▶ 来源：GSOD（全球每日地表摘要）数据集
- ▶ 变量：气温、压强、降水量、风速、露点
- ▶ 处理：站点数据插值为栅格数据（气象数据通常很连续，所以进行插值是合理的），再汇总到区县层面

- 自评估报告（非公开数据）

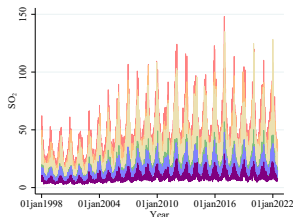
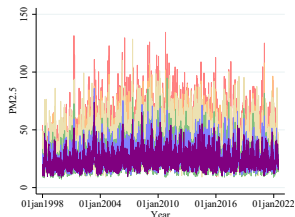
- ▶ 初始/事后清洁取暖水平，资金投入细则，各类措施改造的户数/面积

- 近似

- ▶ treatment group 包括三个兵团，以所在地表示。
- ▶ 去掉了北京市的观测，主要是因为北京早在 2013 年就进行了试点，而正式的政策从 2017 年才开始。如果直接将北京市受处理年份设置为 2013 年的话，政策的长期影响将只来自北京市的 variation。

描述性统计（原始数据）

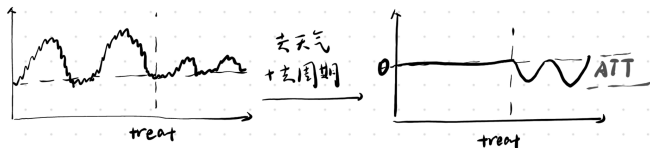
- 政策优先在污染更严重的地区实施
- 污染的周期性明显（受供暖季节影响）、波动大（受气象因素影响）
- 处理组有略微下降的趋势



- 1 政策背景
- 2 相关文献与边际贡献
- 3 数据
- 4 识别策略**
- 5 实证结果
- 6 结论

识别策略：两步法

- Step 1: 清除污染物数据中被气象和周期性因素影响的部分⁵



- ▶ 得到“干净”的污染物数据
- ▶ 直觉：污染中不能被气象和周期性因素解释的部分
- ▶ Step 1.1: 只使用 pre-treatment 数据训练，对于对照组，在时间段 20% 和 80% 之间随机选择处理时间
- ▶ Step 1.2: 对全样本预测，计算残差，作为被解释变量

⁵ 控制了气象因素（湿度、气压、温度、风速、露点）以及周期性因素（距离 1 月 1 日的天数、距离春节的天数、星期几，Factor variables, 类似固定效应）

识别策略：两步法

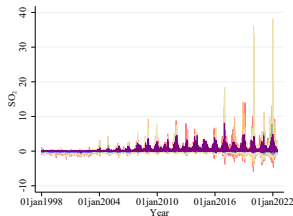
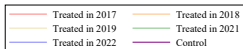
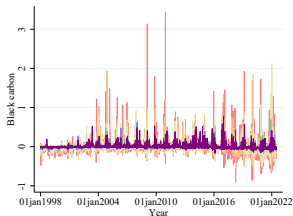
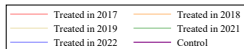
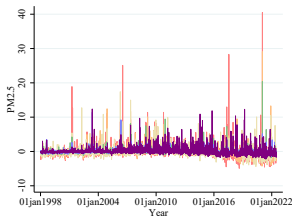
● Step 2: HTE robust DID (when possible)

- ▶ $Y_{it} - \hat{Y}_{it} = \beta D_{it} + X_{it} + \gamma_i + \lambda_t + \varepsilon_{it}$
- ▶ 非试点的北方地区区县作为对照组
- ▶ 数据聚合到年度层面（事件研究法直观/less computationally intensive）
- ▶ 当上述方式不可行时，使用传统的 TWFE staggered DID
 - ★ e.g., 需要考虑政策强度（Continuous DID），需要检验异质性
 - ★ 此时无需将数据聚合到年度层面，直接使用日度数据
 - ★ 此时可以控制 county-specific time trend，均控制⁶
- ▶ 如无特殊标注，标准误均聚类到城市层面

⁶ 在事件研究法的图中可以看到一个趋势，因此控制 county-specific time trend 可能更佳。有关事前趋势的进一步正式讨论见稳健性部分的讨论（honest DID）。

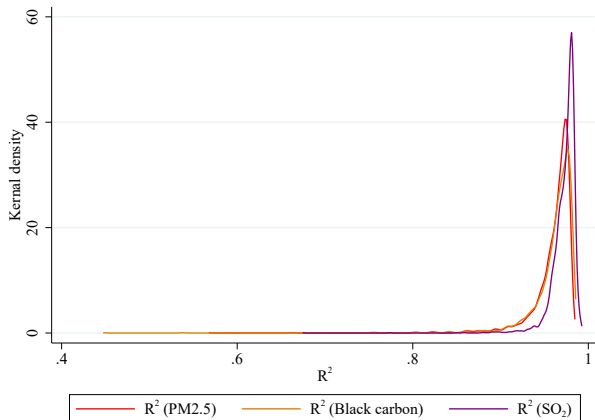
描述性统计（机器学习残差）

- 有政策效果（处理组在受处理后有向下的波谷）



模型表现

- OOB (out-of-bag) 拟合优度接近于 1



对比回归残差和机器学习残差

- 机器学习得到的残差变动更小，有效性优于回归
- 可以看出在政策实施时的下降趋势

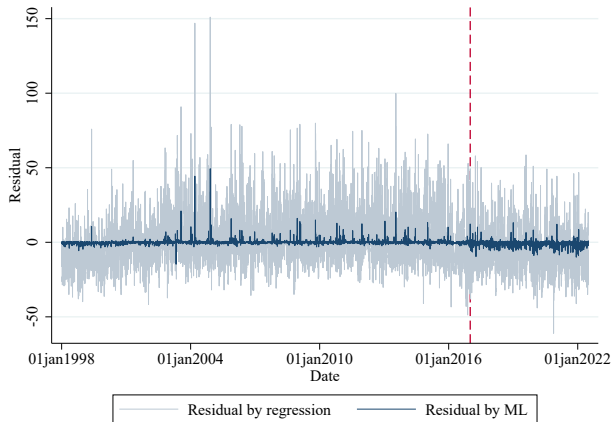


Figure 1: 山东省济南市历下区 PM2.5 的残差

- 1 政策背景
- 2 相关文献与边际贡献
- 3 数据
- 4 识别策略
- 5 实证结果**
- 6 结论

本文实证总框架

- Baseline

- ▶ 政策对空气污染的影响
- ▶ Continuous DID (考虑政策强度)

- 异质性⁷

- ▶ 政策强度
- ▶ 基础经济状况 (人均 GDP)
- ▶ 冬季气温
- ▶ 技术路线 (清洁取暖改造 vs 建筑节能改造? 煤改气 vs 煤改电?)

- 稳健性

- ▶ 溢出效应、事前趋势
- ▶ 考虑机器学习预测对于标准误的影响 (bootstrap)

- 成本收益分析

⁷ 疑问：以下因素是高度相关的，政策强度越高的地方通常采取更清洁的技术路线、有更好的基础经济状况、更高的冬季气温。

政策对空气污染的影响

- 政策缓解空气污染⁸
- 取暖季的处理效应更高⁹

	(1)	(2)	(3)
	Whole year	Heating season	Non-heating season
PM2.5	-0.304*** (0.048)	-0.394*** (0.075)	-0.260*** (0.057)
Black carbon	-0.039*** (0.005)	-0.075*** (0.012)	-0.018*** (0.003)
SO2	-0.019 (0.037)	-0.143* (0.078)	0.061*** (0.021)
County fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	34,300	34,300	34,300

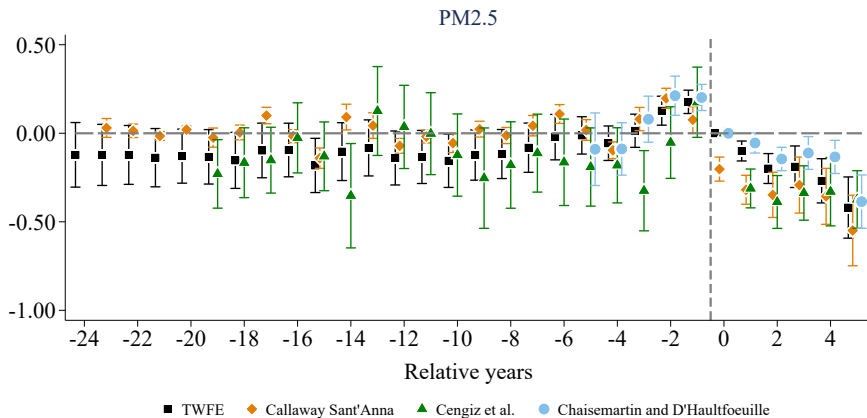
Notes: 每个系数来自单独的 csdid 估计。数据聚合到年度。供暖季取 11.15-3.15。

⁸ 这里系数的绝对值不是很大, PM2.5 的均值在 30 左右。

⁹ 非取暖季仍有效应的原因可能是：1. 政策进行了房屋保暖改造，改善了房屋的隔热，使得夏季空调使用更少，所以减少发电带来的空气污染；2. 北方农村同时使用散煤进行取暖和烹饪，政策可能通过促进清洁炉灶的使用，或通过促进清洁煤的使用，减少烹饪产生的空气污染。

事件研究法

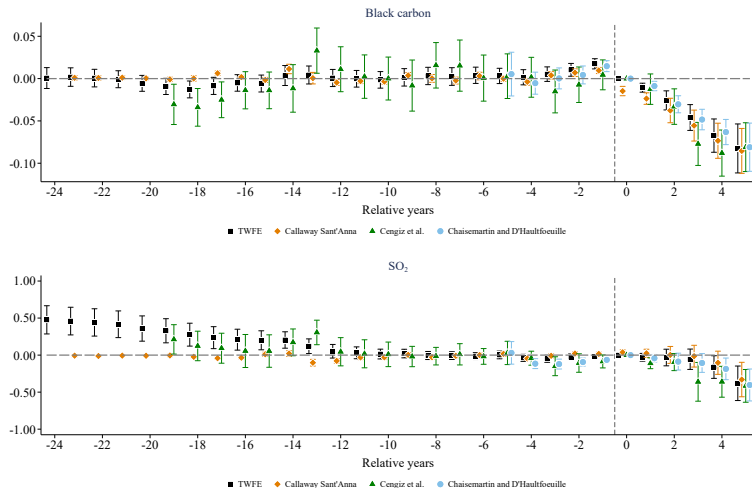
- 北方清洁取暖显著降低了 PM2.5 水平¹⁰



¹⁰ 为了保证可比性，此处的 TWFE 使用年度数据，没有控制 county-specific time trend.

事件研究法

- 北方清洁取暖显著降低了黑碳水平
- 北方清洁取暖降低了二氧化硫水平（不显著）



Continuous DID、政策强度/经济状况/冬季气温异质性

- Continuous DID：显著
- 政策强度、冬季气温越高的地区处理效应更高
- 处理效应与当地经济状况无关、城乡比例无关¹¹

	(1) TWFE	(2) Continuous	(3) Funding	(4) Winter_ave_temp	(5) lag_GDP_percap	(6) City_ratio
Policy	-0.1188** (0.052)	-0.0093*** (0.003)	-0.0243 (0.062)	-0.1484*** (0.055)	0.0550 (0.070)	-0.0707 (0.061)
Policy × Char			-0.0084** (0.004)	-0.0201*** (0.006)	-0.0014 (0.001)	-0.1239 (0.101)
Char				0.0199** (0.008)	0.0008 (0.001)	
Constant	-0.0363*** (0.004)	-0.0365*** (0.003)	-0.0356*** (0.004)	0.0552 (0.034)	-0.0514 (0.036)	-0.0370*** (0.004)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,074,972	12,074,972	12,074,972	12,074,972	8,544,371	11,995,763
R-squared	0.046	0.046	0.046	0.046	0.066	0.046

Notes: TWFE 估计。日度数据。

¹¹ 这可能是因为，尽管改造一户农村家庭的处理效应可能更大，但政策力度更多集中于城市（城区/县城/农村平均改造户数：
57/21/39）

技术路线异质性

- 清洁取暖改造和建筑节能改造没有显著差别 (Column 1)
- 在清洁取暖改造中, 煤改气最有效¹² (Column 2-6)

	(1) 建筑 vs 热源	(2) 热电联产	(3) 工业余热	(4) 煤改气	(5) 煤改电	(6) 其他
Policy	-0.0347 (0.116)	-0.2092*** (0.071)	-0.1180** (0.054)	0.0269 (0.058)	-0.1733** (0.067)	-0.1479** (0.060)
Policy × Char	-0.1090 (0.140)	0.2867* (0.150)	-0.2892 (0.805)	-0.4642*** (0.128)	0.1902 (0.143)	0.2658 (0.260)
Constant	-0.0364*** (0.004)	-0.0358*** (0.004)	-0.0360*** (0.004)	-0.0351*** (0.004)	-0.0359*** (0.004)	-0.0358*** (0.004)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,074,972	11,960,559	11,960,559	11,960,559	11,960,559	11,960,559
R-squared	0.046	0.046	0.046	0.046	0.046	0.046

Notes: TWFE 交互项模型。日度数据。特征变量被固定效应吸收。

¹² 天然气基本不排放烟尘和二氧化硫

改而不用？

- 自评估报告中对“改而不用”情况进行了统计：

- ▶ 第一档为“改而未用”，属于基本无人居住，采暖季家庭整体用电/气量极低（日均用电总量 $<1\text{kwh}$ ，日均用气总量 <0.05 立方米）；
- ▶ 第二档为“基本不用”，属于有人居住但基本不用电/气取暖，采暖季日均用电/气量与非采暖季没有差异，与非采暖季相比（可以9月份日均值计），采暖季日均用电/气量增量极低或为负值（日均用电增量 $<2\text{kwh}$ ，日均用气增量 <1 立方米）；
- ▶ 第三档为“低水平使用”，属于使用电/气取暖，但用电/气量低于正常使用水平，采暖季日均用电/气量增量介于第二档和第四档之间；
- ▶ 第四档为“正常使用”，属于正常使用电/气取暖、满足取暖需求，与非采暖季相比，采暖季日均用电/气量增量符合各地实际需求（各地根据实际情况合理确定清洁取暖正常使用时的日均用电、用气增量）。

溢出效应

- 溢出效应

- ▶ 政策实施地区较为集中，溢出效应可能较小
- ▶ 即使存在溢出效应，会导致处理效应被低估

- 回归检验：将样本进一步区分为

- ▶ 处理组 (T)
- ▶ 与处理组接壤的对照组 (Border C)¹³
- ▶ 不与处理组接壤的对照组 (Clean C)

¹³ 与多个处理组城市相邻的，以最早的为准。忽略了与北京市相邻的城市。

溢出效应

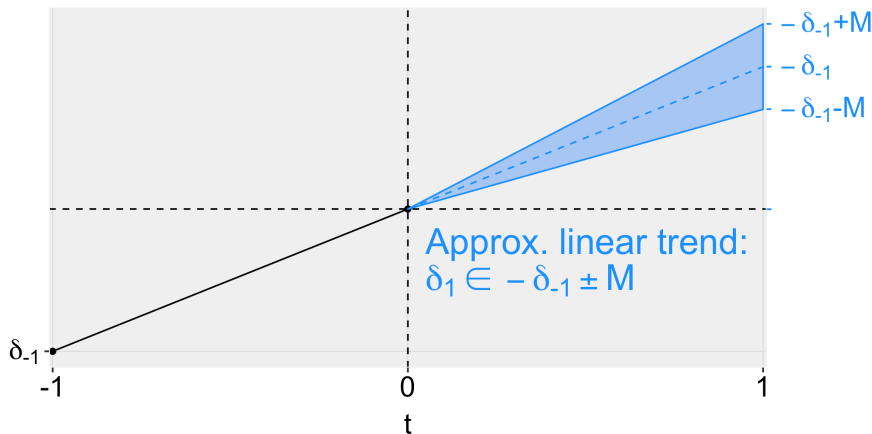
- 仅使用 Clean control 时，ATT 更大 (Column 2)
- 接壤市仅黑碳有处理效应 (Column 3)

	(1) T-C (Baseline)	(2) T-Clean C	(3) Border C-Clean C
PM2.5	-0.304*** (0.048)	-0.415*** (0.060)	-0.105 (0.084)
Black carbon	-0.039*** (0.005)	-0.047*** (0.005)	-0.017** (0.008)
SO2	-0.019 (0.037)	-0.026 (0.038)	-0.033 (0.076)
County fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	34300	27725	13925

Notes: 每个系数来自单独的 csdid 估计。数据聚合到年度。

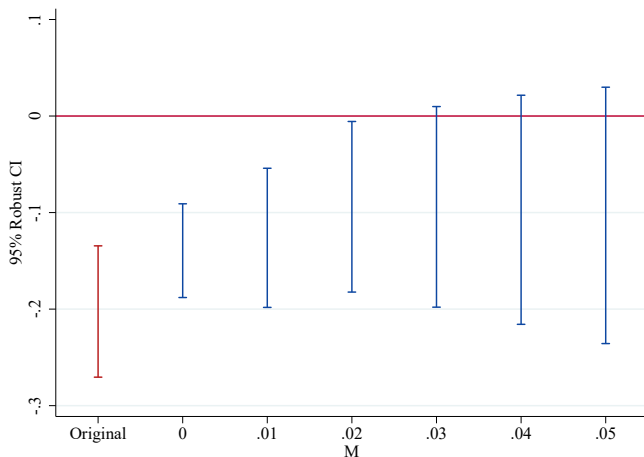
事前趋势：Honest DID (Rambachan and Roth, 2022)

- The post-treatment violations of parallel trends cannot deviate too much from a linear extrapolation of the pre-trend.
- In particular, we can impose that the slope of the pre-trend can change by no more than M across consecutive periods



Honest DID: csdid

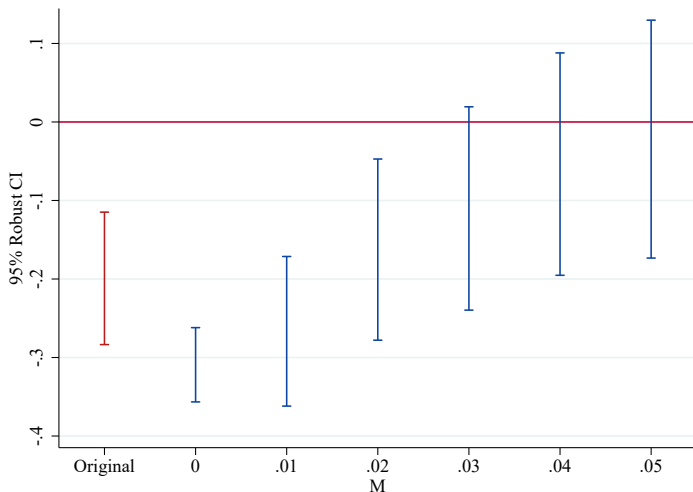
- 存在向下的事前趋势 ($M=0$ 的系数绝对值小于 original)
- 可以容忍 $M=0.02$ 的偏差¹⁴



¹⁴ post 1 期的系数，不是 ATT，下同

Honest DID: TWFE

- 存在向上的事前趋势 ($M=0$ 的系数绝对值大于 original)
- 可以容忍 $M=0.02$ 的偏差



标准误

- 在上述的分析中，标准误只考虑了第二步的误差，而没有考虑第一步（机器学习）
- 使用 bootstrap 方法考虑机器学习预测对于标准误的影响 (Burlig et al., 2020)
 - ▶ 对每个区县，进行有放回的抽样；使用上述样本进行预测，获得残差值。上述过程重复 20 次。
 - ▶ 对上述残差值再次进行 100 次抽样，每次从各区县的 20 个样本中抽一个再汇总，获得 100 个样本，进行基准回归，获得 100 个系数
 - ▶ 将上述系数的标准差视为标准误

标准误

● Bootstrap 标准误与聚类标准误类似

- ▶ 第二列中列出的是区县层面的聚类标准误，以保证可比性
- ▶ 理论上 Bootstrap 标准误要更大，今后可能需要更多次的抽样

	(1) Bootstrap	(2) Baseline
PM2.5	-0.365*** (0.011)	-0.304*** (0.018)
Black carbon	-0.040*** (0.001)	-0.039*** (0.002)
SO2	0.019** (0.009)	-0.019 (0.014)
County fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	34300	34300

Notes: 每个系数来自单独的 csdid 估计。数据聚合到年度。Column (2) 标准误聚类到区县层面。

附录：直接使用原始数据的回归结果（csdid）

- 标准误较大，系数符号比较混乱

	(1) Whole year	(2) Heating	(3) Non-heating
PM2.5	-1.590* (0.844)	-0.411 (0.873)	-1.537 (1.306)
Black carbon	-0.059*** (0.016)	-0.013 (0.048)	-0.113*** (0.029)
SO2	0.010 (0.186)	1.223*** (0.373)	-0.855*** (0.286)
County fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	35,450	35,450	35,450

Notes: 每个系数来自单独的 csdid 估计。数据聚合到年度。出于 csdid 的限制，仅控制了气象因素而没有控制周期性因素。

- 1 政策背景
- 2 相关文献与边际贡献
- 3 数据
- 4 识别策略
- 5 实证结果
- 6 结论**

成本收益分析

- 收益 = Δ 空气污染的经济成本¹⁵ = 3354.3 亿元¹⁶
 - ▶ Economic Cost = $\Delta \text{Mort} \times \text{VSL}$ ¹⁷
 - ▶ $\Delta \text{Mort} = y_0 \left(\frac{\text{RR}-1}{\text{RR}} \right) \text{Pop}$
 - ▶ $\text{RR}(z) = \exp \left(\frac{\theta \log \left(\frac{z}{\alpha} + 1 \right)}{1 + \exp \left(\frac{-(z-\mu)}{\nu} \right)} \right)$, where $z = \text{PM}_{2.5} - 2.4 \mu\text{g}/\text{m}^3$ ¹⁸
- 成本
 - ▶ 财政总投入：6206.694 亿元，其中：
 - ★ 中央投入：945.664 亿元
 - ★ 省市县投入：1826.019 亿元
 - ★ 社会投入：3080.689 亿元

¹⁵ 方法论遵循 Yin et al. (2021)。考虑了 PM2.5 对下呼吸道感染、缺血性心脏病、中风、慢性阻塞性肺病、肺癌导致的过早死亡的影响。

¹⁶ 可能低估的原因：1. 只考虑了一种污染物 (PM2.5) 的影响；2. 只考虑了减少过早死亡的收益，没有考虑健康状况改善的收益；3. 对室内、急性的空气污染缓解的程度可能大于区县层面的 PM2.5 指标

¹⁷ $\text{VSL} = 1.49 \times \text{人均 GDP (72.0885 千元)} \times \text{预期寿命 (78.6 岁)}$ (Cai et al., 2022)

¹⁸ 各参数的值来自 Burnett et al. (2018)。政策开始时处理组的平均 PM2.5 为 33.19371, 2022 年平均 PM2.5 为 29.65139