

职工医保门诊统筹政策对患者就医行为的影响*

——基于因果森林模型

袁理娜¹ 张欣燕¹ 宋小宛¹ 刘梓琦¹ 庞梓怡¹ 张涛^{1,2,Δ}

【摘要】目的 分析职工医保门诊统筹制度及其不同保障模式对参保职工就医行为选择的影响。方法 采用 2018 年中国健康与养老追踪调查数据,通过机器学习因果森林算法,测算职工医保门诊统筹制度对于门诊就医行为的影响。结果 门诊统筹政策使参保职工门诊就诊概率增加 3.3%。在无封顶线模式下,参保职工门诊就诊的概率增加 21.32%,选择医院就诊的概率增加 19.91%,门诊就诊的次数增加 0.2741 次。有起付线和封顶线模式下,患者选择医院门诊的概率下降 1.96%。结论 职工医保门诊统筹政策能够显著提高参保职工门诊利用水平,但需要关注高保障水平模式下引致的“道德风险”问题。

【关键词】 职工医保 门诊统筹 就医行为 机器学习 因果森林

【中图分类号】 R195.1

【文献标识码】 A

DOI 10.11783/j.issn.1002-3674.2024.06.027

自我国职工基本医疗保险制度(以下简称“职工医保”)建立以来,一直采用“统账结合”的管理模式。但个人账户抗风险能力薄弱和统筹账户报销缺位致使门诊患者承担沉重的疾病经济负担,甚至会诱导患者“小病拖、大病扛”,或为获得医保补偿发生“门诊转住院”的情况。数据显示,2021 年我国职工医保参保人员住院率为 17.7%,其中在职职工为 9.9%,退休人员为 39.5%,远高于经济合作与发展组织(organisation economic co-operation and development, OECD)国家平均住院率,但人均就诊人次数为 5.8 次,低于 OECD 国家平均水平^[1-2]。这反映了门诊保障水平低,门诊服务对住院服务产生了“挤入效应”,难以发挥门诊服务预防效应,导致不合理入院问题突出。

为此,2021 年 4 月国务院办公厅发布《关于建立健全职工基本医疗保险门诊共济保障机制的指导意见》,重点强调建立完善职工医保普通门诊费用统筹保障机制,逐步扩大由统筹基金支付的门诊慢特病种范围^[3]。目前,我国部分城市相继开展了职工医保门诊统筹探索实践,但是,门诊保障水平的提高是否会导致患者过度利用医疗服务,强化“趋高就医”动机,引致需方“道德风险”问题,继而带来医保基金的不合理使用和浪费,鲜有研究对此报告。鉴于此,本研究基于中国健康与养老追踪调查数据,采用机器学习中因果推断方法,探究职工医保门诊统筹制度及其不同保障模式对参保职工就医行为选择的影响,以期对门诊共济保障制度建立和完善提供重要参考。

资料与方法

1. 数据来源

本研究采用 2018 年中国健康与养老追踪调查(China health and retirement longitudinal study, CHARLS)数据,CHARLS 是由北京大学国家发展研究院组织开展的旨在收集一套代表中国 45 岁及以上中老年人家庭和个人的高质量微观数据,用以分析我国人口老龄化问题,推动老龄化问题的跨学科研究。根据研究目的,本文选择了 2018 年 CHARLS 数据中参加城镇职工医保的观测值,同时剔除在研究变量上存在缺失值的样本,最终纳入本研究的样本共计 2848 份。

2. 变量选择

(1) 被解释变量

本文聚焦于门诊统筹制度对于参保职工就医行为选择的影响,根据 CHARLS 数据中的现有变量,分别选择以下三个变量衡量就医选择行为^[4-6]:①是否门诊;②门诊就诊次数;③是否选择医院就诊。变量解释和赋值情况见表 1。

(2) 解释变量

根据样本所涉及的城市,利用政府网站确定样本城市在 2018 年是否实施了职工医保门诊统筹制度。在此基础上,根据是否设置了起付线和封顶线,将已实施门诊统筹制度的城市划分为三类:①无起付线;②无封顶线;③有起付线和封顶线。根据查阅结果,本研究共涉及的 105 个样本城市,共有 31 个城市实施了门诊统筹制度,其中 8 个城市未设置起付线,10 个城市未设置封顶线,13 个城市设置了起付线和封顶线^[6]。因此,本研究不仅比较是否实施门诊统筹之间就医选择差异,同时进一步比较不同门诊统筹模式与未实施门诊统筹之间的

* 基金项目:国家自然科学基金项目(72304082);2024 年国家级大学生创新创业项目(202410346050)

1. 杭州师范大学公共卫生学院(311121)

2. 浙江省“卫生健康科技创新与发展战略”软科学研究基地

Δ 通信作者:张涛, E-mail: lucky1230405@163.com

差异。

(3)控制变量

为了控制混杂变量对结果的影响,借鉴 Anderson

医疗服务利用模型框架^[7],控制变量为受访者的性别、年龄、婚姻状况、教育程度、职业、收入、居住地、自评健康状况和是否有慢性病,变量解释和赋值见表1。

表1 变量赋值和解释

类型	变量名称	赋值与解释
被解释变量	是否门诊	过去一个月是否看过门诊,0=没有,1=有
	是否选择医院就诊	若调查对象在过去一个月看过门诊,则进一步询问“过去一个月是否选择去综合医院、专科医院或中医院看门诊”,0=没有,1=有
	门诊就诊次数	过去一个月去不同医疗机构看门诊次数合计
解释变量	门诊统筹	0=未实施,1=实施
	门诊统筹模式	0=未实施,1=无起付线,2=无封顶线,3=有起付线和封顶线
控制变量	性别	1=男,2=女
	年龄	受访者年龄
	婚姻状况	1=已婚,2=其他
	教育程度	1=文盲,2=小学,3=中学,4=大学及以上
	职业	0=未工作,1=工作
	收入	受访者工作收入和各项转移支付收入合计
	居住地	1=城镇,2=农村
	自评健康	1=非常好,2=好,3=一般,4=差,5=非常差
	慢性病	0=没有,1=有

3.统计分析方法

传统方法对政策干预效果评估通常采用双重差分模型结合倾向性得分匹配,但是倾向值匹配方法处理控制变量的数量有限,难以解决政策干预组和对照组样本分配不随机以及控制变量分配不均等问题。此外,如果评估模型中控制变量以及未观测到的其他变量随时间发生变化,基于双重差分模型测算的政策效应稳健性也将受到质疑^[8]。

本研究借助机器学习因果森林模型,改进门诊统筹政策干预对照试验处理组样本选择偏差及其异质性问题,继而确定政策干预的因果效应。借鉴 Athey 提出因果森林算法测算职工医保门诊统筹政策干预的平均处理效应(average treatment effect, ATE)和个体平均处理效应(conditional average treatment effect, CATE),具体步骤如下^[8-9]:

(1)采用无放回抽样方式从原始样本 N 中随机抽取样本量为 s 的子集 b (默认 s 为原始样本量 50%),从而将原始样本划分为 T 和 E 两个子样本,分别作为训练样本集,即构建因果树结构和估计 ATE 的样本集。

(2)按照递归分区的方式对训练样本 T 构建根节点,并按照自上向下方式将根节点划分成左右两个不同子节点,按照上述规则对子节点继续分割,直到无法分割为止,继而形式因果树结构。在分割的过程中按照控制变量特征空间分成不同的子空间,以最大限度增加不同子空间 ATE 的异质性。如果分割节点过程中,无法增加拟合效果,则停止分割,并形成因果树的一个“叶子”。

(3)计算每个节点“叶子”的 ATE,即:

$$\hat{\tau}_i(x) = \frac{1}{|\{i: Z_i = 1, X_i \in L\}|} \sum_{i: Z_i = 1, X_i \in L} Y_i - \frac{1}{|\{i: Z_i = 0, X_i \in L\}|} \sum_{i: Z_i = 0, X_i \in L} Y_i \quad (1)$$

式(1)前半部分和后半部分分别表示处理变量 Z 为 1 和 0 时,该节点的被解释变量 Y 的平均效应的均值。

(4)按照上述(1)~(3)步骤重复 B 次,可形成由 B 棵树连接的因果森林。那么政策干预的效应可由下式计算可得:

$$\hat{\tau} = \frac{\sum_{i=1}^n \alpha_i(x) [Y_i - \hat{m}^{(-i)}(X_i)] [Z_i - \hat{e}^{(-i)}(X_i)]}{\sum_{i=1}^n \alpha_i(x) [Z_i - \hat{e}^{(-i)}(X_i)]^2} \quad (2)$$

其中 $\alpha_i(x)$ 由机器学习驱动形成的核(kernel), Z_i 为处理变量,用 $e(x) = P[Z_i | X_i = x]$ 代表倾向性得分值, $m(x) = E[Y_i | X_i = x]$ 表示政策干下的预测值。符号 $-i$ 表示“袋外”预测, $\hat{m}^{(-i)}(X_i)$ 和 $\hat{e}^{(-i)}(X_i)$ 表示不使用 Y_i 计算倾向性得分值和政策干预的预测值。

结果

1.基本情况

由表2可发现,实施门诊统筹地区的受访者为 31.6%,并且 18.0%受访者所在地区门诊统筹政策设置了起付线和封顶线。从医疗服务利用情况来看,17.1%受访者过去一个月在门诊就诊,13.9%的患者选择了医院就诊。从一般人口学特征和健康状况来看,受访者年龄的中位数为 62 岁,49.9%受访者为中学文化程度,79.2%受访者来自于城镇,患有慢性病的受访者达到 47.6%。

2.门诊统筹对就医选择行为影响

表 3 基于因果森林测算得出的门诊统筹对就医行为的 ATE 及其置信区间。由结果可知,实施门诊统筹政策比未实施门诊统筹政策地区,受访者门诊就诊的概率增加了 3.3%,选择医院门诊的概率增加 1.09%,但通过置信区间发现,部分个体的效应值 < 0,表明政

策干预对于就诊机构选择的影响存在较大的异质性。同理,在门诊次数方面,实施门诊统筹政策地区的受访者比未实施门诊统筹政策的地区少 0.0078 次,但是置信区间包含 0,表明干预政策对就诊次数的影响在不同个体之间差异较大。

表 2 调查对象的基本情况

变量	n/M	%/P95	变量	n/M	%/P95
是否门诊(n,%)			教育程度(n,%)		
是	486	17.1	文盲	120	4.2
否	2362	82.9	小学	660	23.2
门诊次数(次)(M,P95)	0	2	中学	1420	49.9
医院门诊(n,%)			大学及以上	648	22.8
是	395	13.9	职业(n,%)		
否	2453	86.1	工作	1267	44.5
门诊统筹(n,%)			未工作	1581	55.5
是	901	31.6	收入(元)(M,P95*)	33650	90000
否	1947	68.4	居住地(n,%)		
门诊统筹模式(n,%)			城镇	2257	79.2
未门诊统筹	1947	68.4	农村	591	20.8
无起付线	173	6.1	自评健康(n,%)		
无封顶线	215	7.5	非常好	392	13.8
有起付线和封顶线(n,%)	513	18.0	好	603	21.2
性别(n,%)			一般	1399	49.1
男	1618	56.8	差	344	12.1
女	1230	43.2	非常差	110	3.9
年龄(岁)(M,P95)	62	81	慢性病(n,%)		
婚姻状况(n,%)			有	1357	47.6
已婚	2305	87.9	无	1491	52.4
其他	345	12.1			

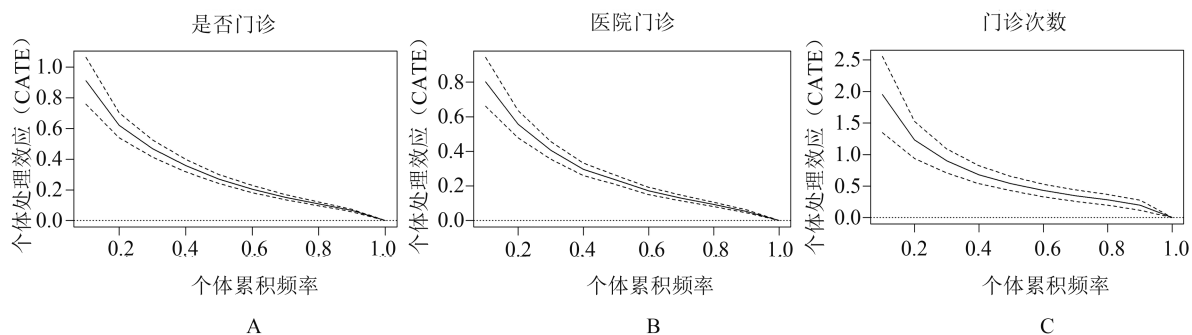
*:部分变量不满足正态分布,故采用 95%分位数(P95)描述

表 3 门诊统筹对就医行为的 ATE 和置信区间

	ATE	95% CI
是否门诊	0.0330	(0.0169,0.0491)
医院门诊	0.0109	(-0.0033,0.0142)
门诊次数	-0.0078	(-0.0667,0.0511)

图 1 描绘了所有 2848 个观测样本处理效应的预

测值分布情况,其中横轴代表个体累积频率,纵轴代表已排序的个体处理效应预测值。图 1A~C 分别代表门诊统筹政策干预下所有受访者是否门诊、医院门诊和门诊次数的预测值分布情况。由图可知,上述三个变量的个体政策干预效应的预测值均 > 0,但对于不同个体,政策效应的预测值存在较大差异。



注:虚线为 95%置信区间

图 1 门诊统筹对个体就医行为影响的预测值分布

3.不同门诊统筹模式政策效果的异质性

进一步观测不同门诊统筹模式对于个体就医行为

选择的影响(表 4),通过比较被解释变量的 ATE 和 95% CI 可以发现,相较于无统筹地区,在无封顶线模

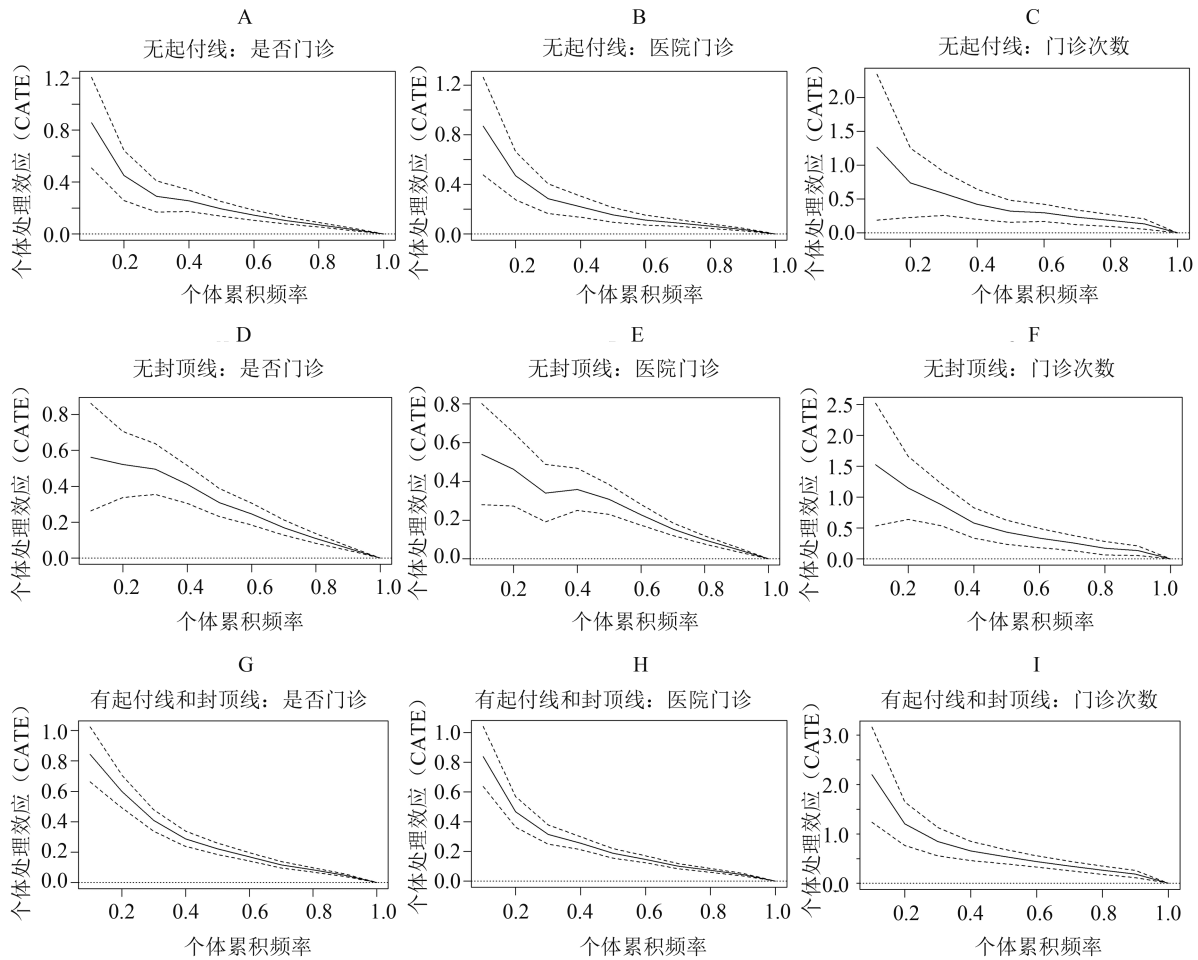
表 4 不同门诊统筹模式对就医行为的 ATE 和置信区间

门诊统筹模式	被解释变量	ATE	95% CI
无起付线	是否门诊	0.0004	(-0.0308, 0.0316)
	医院门诊	-0.0014	(-0.0316, 0.0288)
	门诊次数	-0.0457	(-0.1884, 0.0970)
无封顶线	是否门诊	0.2132	(0.1713, 0.2551)
	医院门诊	0.1991	(0.1567, 0.2415)
	门诊次数	0.2741	(0.1824, 0.3658)
有起付线和封顶线	是否门诊	0.0025	(-0.0162, 0.0212)
	医院门诊	-0.0196	(-0.0361, -0.0031)
	门诊次数	-0.0261	(-0.0963, 0.0441)

式下,受访者门诊就诊的概率增加 21.32%,选择医院就诊的概率增加 19.91%,门诊就诊的次数增加 0.2741。在有起付线和封顶线模式下,患者选择医院门诊的概率下降 1.96%,但是对于是否门诊和门诊次

数的影响方向存在异质性,即部分患者门诊就诊概率和次数会增加,也有部分患者则相反。在无起付线模式下,是否门诊、选择医院门诊和门诊次数的 ATE 置信区间包含 0,表明该统筹模式对于上述三个变量的影响具有不确定性,部分患者门诊就诊的概率,选择医院门诊的概率以及门诊就诊次数可能变大,也有部分患者会出现相反的变化。

图 2 描绘了不同门诊统筹模式下三项被解释变量预测值的分布情况,图 2A~C 表示无起付线模式,图 2D~F 表示无封顶线模式,图 2G~I 表示有起付线和封顶线模式。与图 1 结果相似,在不同门诊统筹模式下,所有个体的在是否门诊、选择医院门诊和门诊次数上的预测值均为正值,但对于不同个体,门诊统筹模式的政策效应具有显著的异质性特点。



注:虚线为 95% 置信区间

图 2 不同门诊统筹模式对个体就医行为影响的预测值分布

讨论与建议

本研究首次采用因果森林算法探究职工医保门诊统筹政策对于参保人门诊就医选择行为的影响,有效规避了传统的因果推断方法在样本分配不随机和协变量选择存在偏误等问题,为观测门诊统筹政策效应提供了有力的证据。研究的主要发现包括以下两个方

面:

第一,门诊统筹政策提高了参保职工的门诊服务利用水平。根据门诊统筹政策平均处理效应测算结果发现,参保职工利用门诊服务的概率显著提高,这一结果与既往研究一致^[4, 6, 10],这表明提高门诊保障水平之后,患者所需承担的医疗服务成本下降,激励患者选择门诊就医。对于是否选择医院门诊和门诊次数的影

响,虽然个体处理效应的预测值均为正值,但是平均处理效应的置信区间包含 0,表明门诊统筹政策对不同个体影响的方向存在差异。一方面可能是本研究测量的是过去一个月门诊次数,数值相对较小,且变异程度较大,因此导致个体处理效应的异质性较大。另一方面,门诊统筹政策的实施对于不同个体影响的异质性提示需要关注政策目标群体的受益公平性,以强化医保门诊统筹政策互助共济的普惠性功能。

第二,不同门诊统筹模式下参保职工门诊就医行为存在差异。进一步比较三种不同统筹模式对于参保职工就医选择行为的影响,结果显示,无封顶线报销政策对于门诊就医行为选择影响较大,不仅提高了门诊就诊的概率,同时还增加了选择医院就诊的概率以及门诊次数。取消报销封顶线能够很大程度上降低患者治疗负担,但是同时减少了对患者就医的约束。既往研究显示,成本分担变动对患者卫生服务利用产生显著影响^[11-13]。因此在无封顶线模式下,门诊高额费用的服务项目得到高比例的报销,可能会诱发患者过度医疗和“趋高就医”等道德风险。相较于有起付线和封顶线模式,参保职工选择医院门诊的概率下降,表明在门诊统筹政策下,通过对报销范围和比例设置相应限制,可有效对患者就医选择行为规范与约束。

综合上述发现,职工医保门诊统筹政策能够显著提高参保职工门诊利用水平,过去研究显示门诊服务和住院服务存在“替代关系”,提高门诊保障水平能够减少患者不合理入院率,减少医保资金的不合理使用。但是,在门诊统筹模式下,若缺少对报销比例和范围限制可能会导致道德风险,建议设置合理起付线、封顶线和分级分类的报销比例,同时改革门诊医保支付方式,并强化对医保基金的监管,规范诊疗服务行为。

本研究尚且存在以下两个方面的不足。第一,受限于 CHARLS 调查问卷,研究仅采用是否门诊,是否选择医院门诊以及门诊就诊次数三个变量衡量患者就医选择行为,后期可进一步对慢病患者就医选择行为进行细化测量和分析。第二,研究仅比较了无起付线,无封顶线和有起付线和封顶线三种门诊统筹模式下患

者就医选择行为的差异,后续可按照起付线和封顶线的金额将门诊统筹模式划分成多个类别,比较分析就医选择行为的差异。

参 考 文 献

- [1] 国家医疗保障局. 2021 年全国医疗保障事业发展统计公报[EB/OL].(2022-06-08). http://www.nhsa.gov.cn/art/2022/6/8/art_7_8276.html.
- [2] 国家卫生健康委员会. 中国卫生健康统计年鉴[M]. 北京: 中国协和医科大学出版社,2022.
- [3] 国务院办公厅. 关于建立健全职工基本医疗保险门诊共济保障机制的指导意见[EB/OL].(2021-04-22). http://www.gov.cn/xinwen/2021-04/22/content_5601352.htm.
- [4] 朱凤梅, 张小娟, 郝春鹏. 门诊保障制度改革:“以门诊换住院”的政策效应分析——基于中国职工医保抽样数据的实证检验[J]. 保险研究,2021(1): 73-90.
- [5] 朱凤梅. 职工医保门诊统筹对道德风险发生的影响:兼论患者成本分担变动的政策效果[J]. 北京社会科学,2021(9): 98-110.
- [6] 曹清华, 宋海伦. 城镇职工医保门诊统筹的政策效应分析:基于 CHARLS 数据的实证检验[J]. 社会保障研究,2022(4): 23-32.
- [7] 卢珊, 李月娥. Anderson 医疗卫生服务利用行为模型:指标体系的解读与操作化[J]. 中国卫生经济,2018,37(9): 6.
- [8] 胡尊国, 熊云晖, 邓理婕, 等. 区域协调发展政策效果的再评估:来自因果森林算法的异质性处理效应分析[J]. 经济学报,2022,9(2): 201-235.
- [9] Athey S, Tibshirani J, Wager S. Generalized Random Forests[J]. Annals of Statistics,2019, 47(2): 1148-1178.
- [10] He W. Effects of establishing a financing scheme for outpatient care on inpatient services: empirical evidence from a quasi-experiment in China[J]. European Journal of Health Economics,2022;23(1): 7-22.
- [11] Tirgil A, Gurol-urganci I, Atun R. Early experience of universal health coverage in Turkey on access to health services for the poor: regression kink design analysis[J]. Journal of Global Health,2018, 8(2): 020412.
- [12] Nishi T, Maeda T, Katsuki S, et al. Impact of the 2014 coinsurance rate revision for the elderly on healthcare resource utilization in Japan[J]. Health Economics Review,2021,11(1): 24-35.
- [13] Yoo KB, Ahn HU, Park EC, et al. Impact of co-payment for outpatient utilization among Medical Aid beneficiaries in Korea: A 5-year time series study[J]. Health Policy,2016,120(8): 960-966.

(责任编辑:郭海强)