

# 欠发达地区政府数据开放政策对技术创新的效应分析： 基于理性——渐进主义视角

汇报人：张斌



湖南农业大学  
HUNAN AGRICULTURAL UNIVERSITY



# 目 录

- 01 | 问题的提出
- 02 | 文献简述
- 03 | 研究设计
- 04 | 实证结果
- 05 | 稳健性检验
- 06 | 一个政策分析的解释框架
- 07 | 结论



## 01 问题的提出

近年来，随着数字经济时代的到来，数据的价值与重要性开始逐渐显现出来。其中，**政府数据**由于蕴含着重要的价值而被社会广泛关注。2015年，国务院印发《**促进大数据发展行动纲要**》，成为我国数据开放领域的政策性、纲领性文件。之后，各个省、自治区、直辖市和地级市等政府纷纷响应，制定本行政区域内的开放政府数据法规条例。截止到2021年下半年，我国共有193个省级和城市的地方政府上线了开放政府数据平台。可以说，政府数据开放的政策支持和平台支撑已经完全搭建起来，**为相关政府数据的利用提供了保障。**



## 01问题的提出

在2018年至2021年的《中国地方政府数据开放报告》中，贵州省四年的累计分值为214.97分，排名全国第三，保持在全国的第一梯队，其中以贵阳市为代表的大数据发展更是在学界和业界被广泛关注。而且据《2021全球重要城市开放数据指数》报告显示，贵阳市排名全球第五，并获得数据“开放质量之星”的称号。

似乎贵阳市的大数据发展政策俨然获得了成功，已经走出了一条经济欠发达地区利用大数据发展政策带动自身技术发展创新的道路。但是真实情况是否依旧是这样呢？其数据开放政策和技术创新之间的关系究竟是怎样？信息公开对于技术创新的促进作用已被证实，但是与信息公开政策有着紧密联系的数据开放政策是否依然有类似的作用呢？



## 02文献简述

目前关于技术创新效应的相关研究虽然较多，但是信息化环境下的数据对其产生何种影响仍然没有一个直观的研究和解释框架，特别是政府数据对技术创新的影响作用机理究竟是怎么样的仍然没有回答。因此，本次研究从合成控制法的角度出发，探究数据开放政策和技术创新之间的关系，尝试为解决上述问题提供思路。



## 03 研究设计

### (1) 研究方法

采用合成控制法，其优点主要表现为：

第一，作为一种非参数的方法，是传统的双重差分法的扩展；

第二，合成的控制组权重由数据所决定，有效的降低了由于研究者主观选择或者判断而带来的误差；

第三，由于相关对照组是由多个实际的对象通过合成得来的，因此可以清晰地反应各个对象在“反事实框架”中的贡献。



## 03 研究设计

### (2) 模型设定

模型中假定有 $M+1$ 个观测区域，其中“1”指的是受政策影响的区域，即实验组（处理组），本次研究中指的是贵阳市；而“ $M$ ”指的是未受政策影响的区域，即对照组（控制组）。实验组和对照组在同一时间段内受到政策干预的情况均可被观测到，该时间段记为 $T$ ，那么采用 $T_0$ 来表示政策实施的年份。采用 $Y_{it}^N$ 表示 $i$ 地区在 $t$ 时刻未收到干预的情况，用 $Y_{it}^L$ 表示 $i$ 地区在 $t$ 时刻受到干预的情况，其中， $i=1, 2, 3, \dots, M+1$ ， $t=1, 2, 3, \dots, T$ 。那么， $Y_{it}^L - Y_{it}^N$ 就是 $i$ 地区受政策的影响，即实验效应，用 $\alpha_{it}$ 表示。



## 03研究设计

### (2) 模型设定

在政策未实施之前，所有 $M+1$ 个区域都不会受到干扰。因此，对于 $t < T_0$ 的时间段来看， $Y_{it}^L$ 和 $Y_{it}^N$ 两者之间是相等的；但是当 $T_0 < t < T$ 的时间段内，存在 $\alpha_{it} \neq 0$ 。此时引入虚拟变量 $D_{it}$ ，当地区 $i$ 在 $t$ 时刻实施政策的时候，该虚拟变量为1，否则为0。那么，在 $t$ 时刻观测到的地区 $i$ 的实施结果为 $Y_{it} = D_{it}Y_{it}^L + (1 - D_{it})Y_{it}^N$ ，即 $Y_{it} = Y_{it}^N + \alpha_{it}D_{it}$ 。





## 03研究设计

### (2) 模型设定

当 $t$ 处于大于 $T_0$ 时期的时候（即2017年），由于只有1个区域受到影响，因此 $\alpha_{1t} = Y_{1t}^L - Y_{1t}^N = Y_{1t} - Y_{1t}^N$ 。而 $Y_{1t}$ 是处理组的实际结果，是可以观测到的， $Y_{1t}^N$ 是处理组在未受到政策影响下的结果，是无法观测到的。因此，需要构建“反事实框架”来对该变量进行估计，即 $Y_{it}^N = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it}$ 。

其中， $\delta_t$ 表示时间趋势，是一个 $1 \times r$ 维的无法被直接观测到的共同因子； $\theta_t$ 是一个 $1 \times r$ 维的未知参数； $Z_i$ 是一个 $r \times 1$ 维的控制变量，不受政策干预的影响； $\lambda_t$ 是一个 $1 \times F$ 维的不可观测的共同因子； $\mu_i$ 是 $F \times 1$ 维的不可观测的地区固定效应； $\varepsilon_{it}$ 是一个标准误差，其均值为0。

$$\sum_{j=2}^{M+1} w_j Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{M+1} w_j Z_j + \lambda_t \sum_{j=2}^{M+1} w_j \mu_j + \sum_{j=2}^{M+1} w_j \varepsilon_{jt}$$



## 03研究设计

### (3) 变量定义

表 1 变量类型及变量名称

变量类型	变量名称	选择依据
被解释变量	专利授权量	文献 31-33
	财政总收入（取对数）	文献 35
	高等学校数量	文献 40
	人均 GDP 数量	文献 34-35
	一般公共预算支出	文献 35/38
预测变量	互联网宽带接入数	文献 38
	第二产业增加值	文献 38
	城镇化率	文献 36/38/41
	高等学校大专及以上学历学生的数量	文献 36/39/40



## 03 研究设计

### (3) 数据来源与样本选取

使用合成控制法，需要保证在政策实施之前已经存在相当规模的时期数，从而可以提升数据的解释力度，拟合实验组的相关特征。**2017年**，贵阳市人大颁布了《贵阳市政府数据共享开放条例》，将政府数据开放工作纳入了法制化管理的轨道。将此次条例的颁布作为一次“准自然实验”，即 $T_0=2017$ 。将研究样本的时间段选择为2012年至2021年，其中2012年至2016年为政策颁布时间前的窗口期；2017年至2021年为政策颁布之后的时期。

数据均来自于国家统计局的《中国城市统计年鉴》以及各市统计局公开的国民经济和社会发展统计公报，其中无法查到的数据，采用网络搜索进行补全，依然无法获取到的，以该市该变量在被调查年份时间段内的平均值进行补全。



## 03研究设计

### (3) 数据来源与样本选取

样本选取：

- 排除直辖市（与贵阳市行政级别不对等）；
- 将河南省、广东省、重庆市、江苏省、河北省以及浙江省的省会城市排除（全国首批政务数据开放共享国家标准试点地区名单，上述各个省份均包含在该名单中，难免会与贵阳市产生同质性的效果。）；
- 西藏与香港、澳门以及台湾地区的数据难以获得，所以也将拉萨市、香港、澳门与台北市予以排除



## 04实证结果

### (1) 虚拟贵阳

表 2 产生“虚拟贵阳”的城市及其权重

城市	权重	城市	权重
昆明	0.177	兰州	0
沈阳	0	太原	0
哈尔滨	0.108	呼和浩特	0
长沙	0	西安	0.099
合肥	0	长春	0.087
济南	0	福州	0
乌鲁木齐	0	西宁	0
南昌	0	成都	0
武汉	0	银川	0
桂林	0.176	海口	0.354

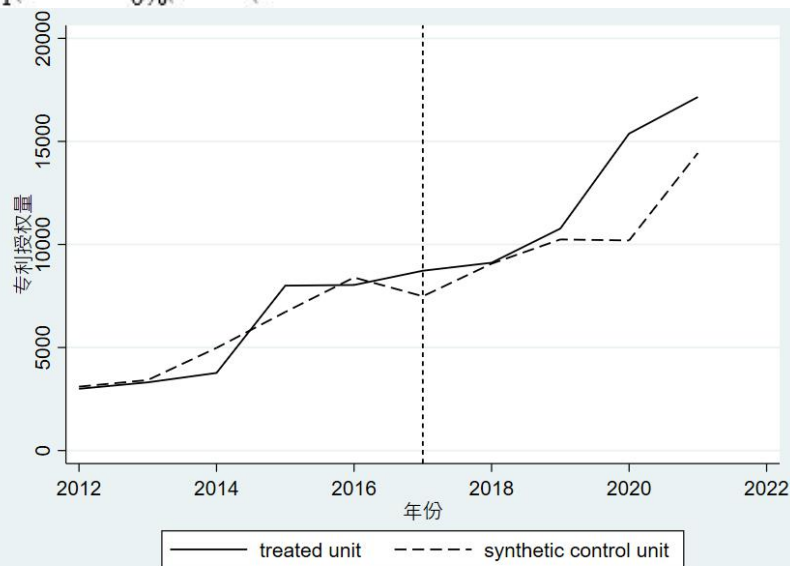


## 04 实证结果

### (2) 效应分析

表3 “真实贵阳”与“虚拟贵阳”数值及其相差率

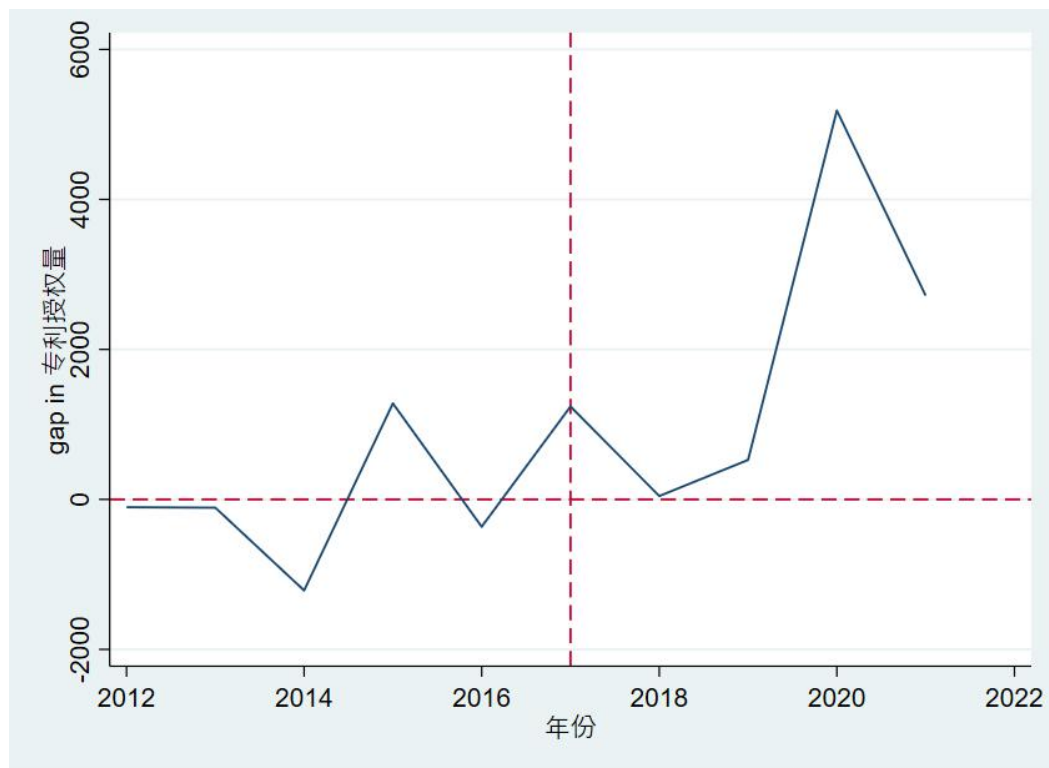
变量	真实贵阳	虚拟贵阳	相差率
财政总收入	2.810751	2.64861	-6%
高等学校数量	27.5	29.00583	5%
人均 GDP 数量	5.22	5.22851	0%
一般公共预算支出	467.1333	459.45	
互联网宽带接入数	172.7433	142.32	
第二产业增加值	1155.62	1161.3	
城镇化率	0.7222167	0.6021	
普通高等学校大专及以上学历学生的数量	351900	35237	





## 04 实证结果

### (2) 效应分析

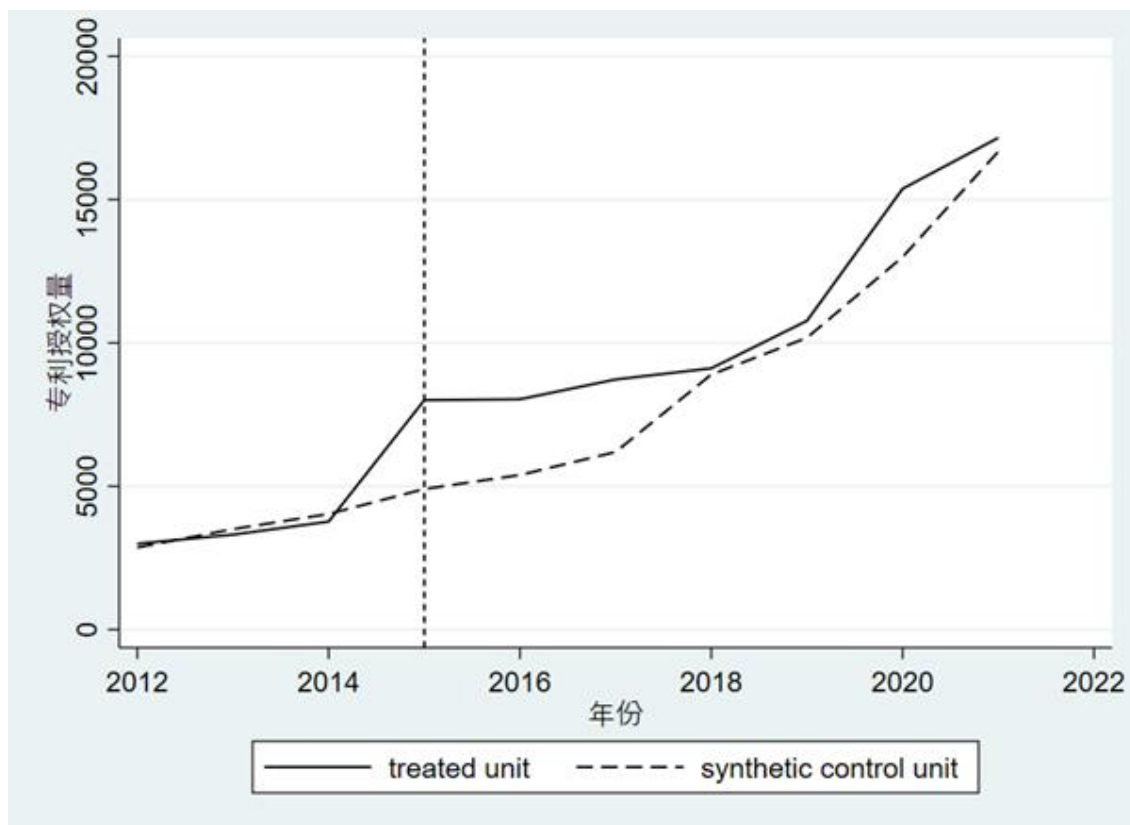


真实贵阳与虚拟贵阳在专利授权量上的差距



## 05稳健性检验

### (1) 安慰剂检验——改变时间

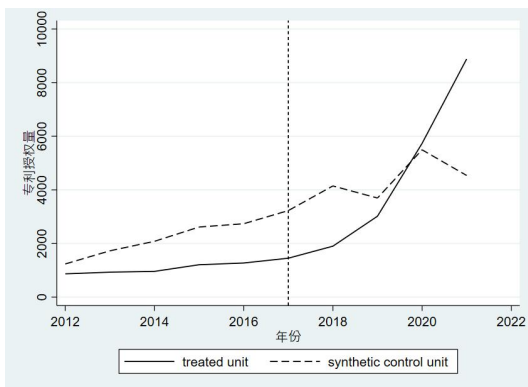




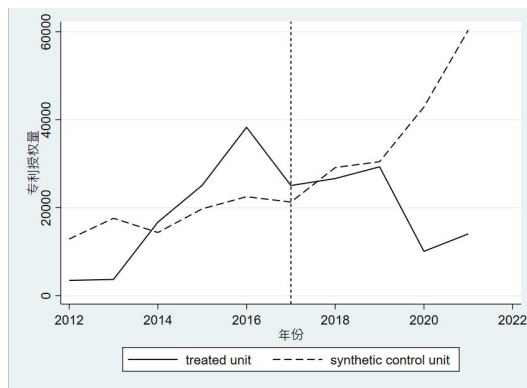


## 05稳健性检验

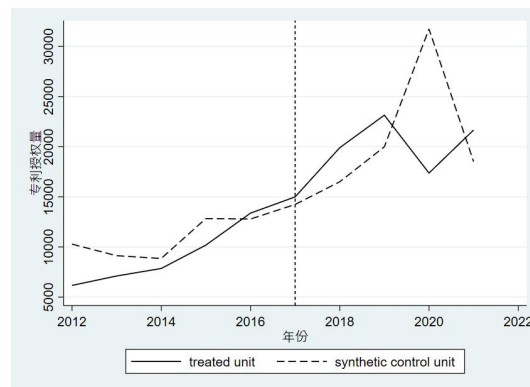
### (1) 安慰剂检验——改变对照组



海口



西安

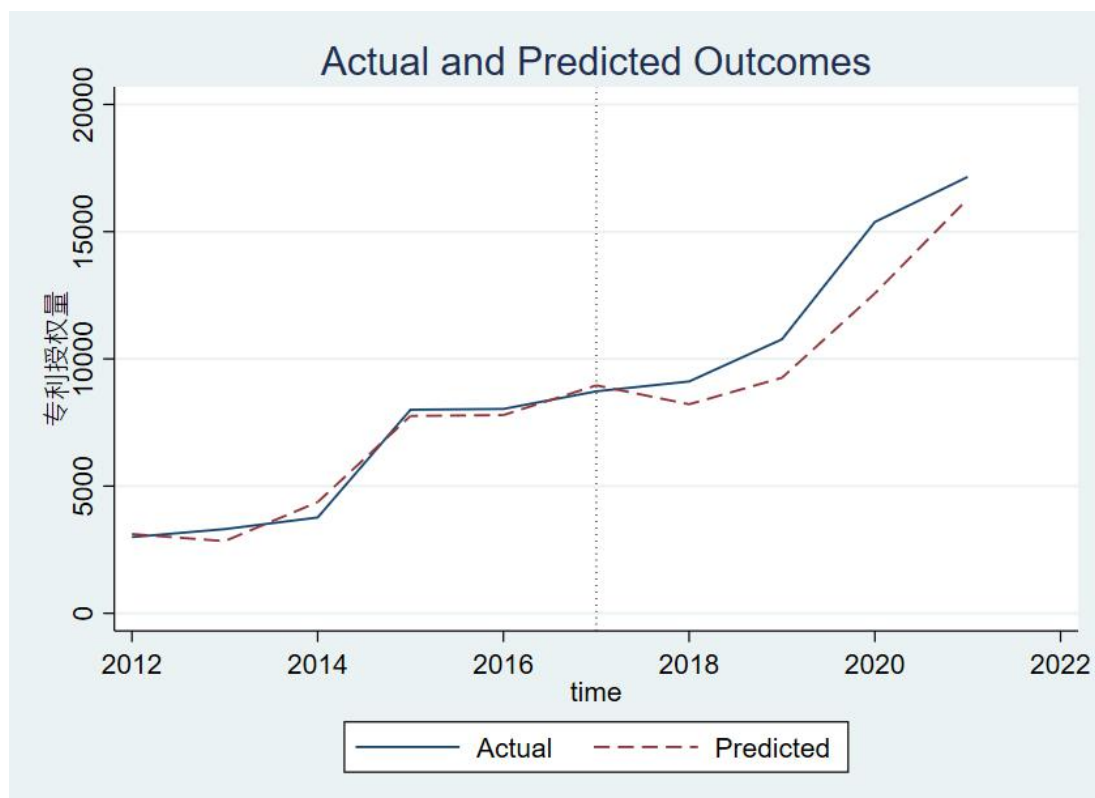


长春



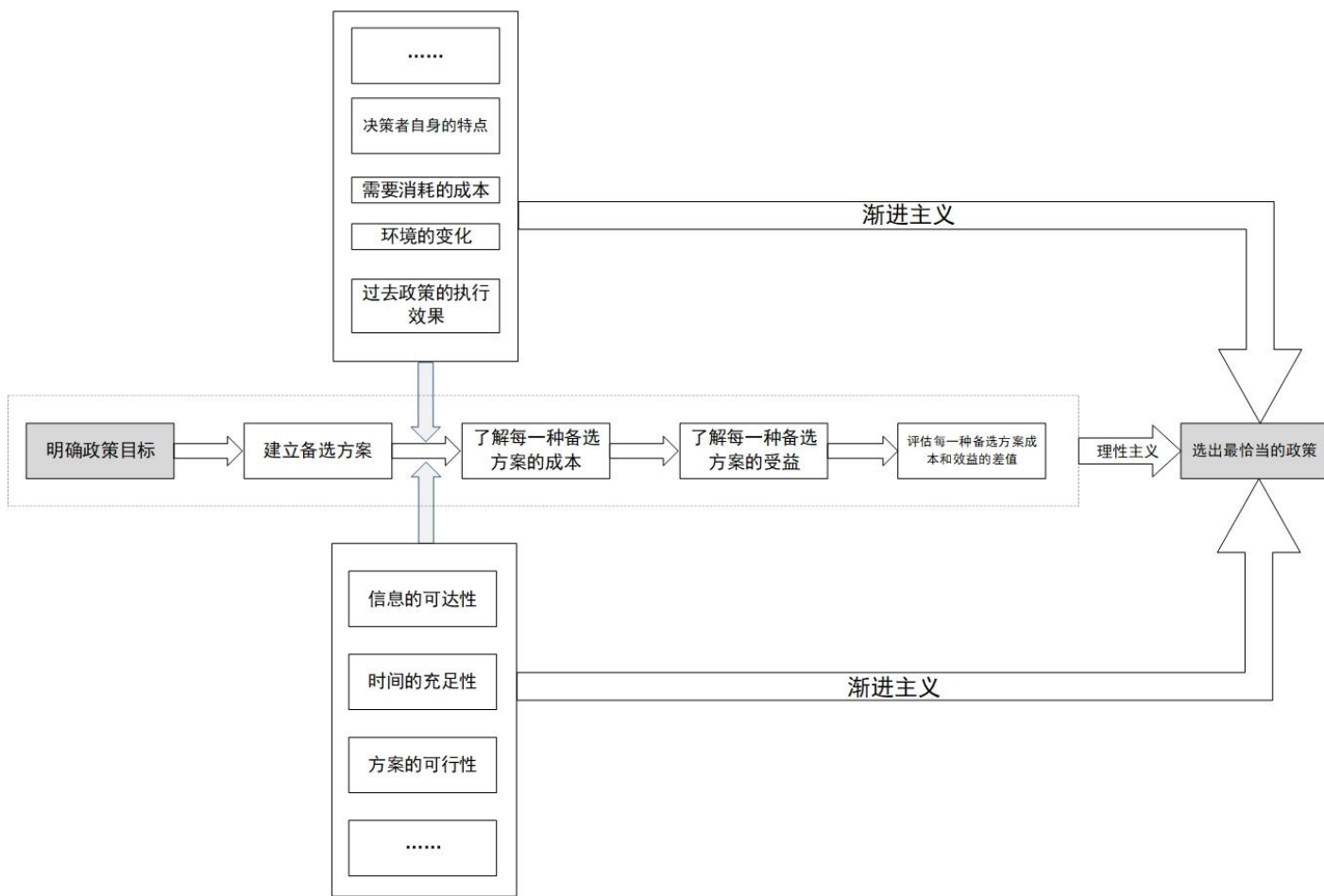
## 05 穩健性檢驗

### (2) 回歸控制檢驗





## 06 一个政策分析的解释性框架





## 07 结论

第一，从实际数据的分析来看，在经济欠发达地区深挖政府数据中蕴含的价值，确实能够在一定程度上助推当地的技术创新效应。

第二，对于政策的实施，政府需要兼顾“理性”与“渐进”主义两个方面。

第三，数据开放政策能够对技术创新起到一定的推动作用，不仅需要政府的支持，同时外部环境也是其发挥作用的重要推动力。



谢谢！  
请各位专家批评指正！