

# 面板数据模型的秘密：从 DID、SC 到统一框架的最新进展

许文立 郭聆沂

(安徽大学经济学院; 深圳市自然资源和不动产评估发展研究中心)

[xuweny87@hotmail.com](mailto:xuweny87@hotmail.com); [lingyi0918@163.com](mailto:lingyi0918@163.com)

2023-12-12

[最新版下载: <https://wenddymacro.github.io/Wenddy-XU/>]

**摘要:** 面板数据是因果效应识别和推断中最常用的数据类型之一。近年来, 面板数据的因果推断方法有许多进展。本文结合最新的理论计量研究成果, 试图对面板数据的因果推断方法进行系统性归纳梳理, 以厘清面板数据模型的最新理论进展和实践中需要注意的一些问题。本文从双向固定效应模型、双重差分模型和合成控制法的估计量、识别假设入手, 简要回顾交叠处理下的异质性处理效应的偏误和稳健估计量, 梳理连续型 DID、多处理变量 DID、因子模型和共形推断, 以及双重差分和合成控制法的结合框架。本文以现实数据为例, 详细展示双重差分和合成控制法结合框架的应用, 以及在实践中可能面临的一些常见问题, 例如适用环境, 与双重差分法的配合问题等等。最后, 本文针对面板数据因果推断的应用给出了一些实践建议, 并讨论了一些有待进一步研究的开放性问题。

**关键词:** 面板数据; 双重差分; 合成控制法; 统一框架; 异质性处理效应

**分类号:** F224

**基金资助:** 教育部人文社会科学研究一般项目青年基金 (23YJC790168) 的资助

## 一、引言

近年来, 面板数据的因果推断已经出现了许多的新方法和应用 (Arkhangelsky and Imbens, 2023)。这些新方法的理论结果都极大地促进了研究者对社会科学因果推断的理解。一方面, 大部分文献主要关注于传统的双重差分法(DID)的最新进展, 例如, de Chaisemartin and D’Haultfoeuille (2022)、Roth et al. (2023)、刘冲等 (2022)、张子尧和黄炜 (2023)、许文立 (2023) 等对这一领域的最新进展做了很好的综述。虽然 DID 是应用研究最常用的因果识别方法之一, 但已有综述性文献最主要关注二值型处理变量的因果效应估计和推断, 忽略了连续或多值处理变量 DID 的最新理论进展。于此同时, 已有方法论综述也忽略了多处理变量的因果推断。

另一方面，合成控制法（SC）也是利用面板数据来识别因果效应的重要方法。它也是过去十几年来因果推断领域最重要的方法论进展（Athey and Imbens, 2017）。Abadie（2019）综述了 SC 的理论进展，包括偏误纠正、多处理个体等，也给出了一些实践应用建议。但相较于 DID 的最新方法论综述，SC 的方法论综述仍然较少，而且国内的学术界还没有关于 SC 最新理论进展的文献，尤其是关于因子模型在 SC 方法论中的作用，以及推断框架的进展。

最后，也是面板数据因果推断最前沿的发展方向之一——DID 和 SC 的融合与统一。通常来说，DID 更适合于有大量处理个体的政策评估环境，且它严重依赖于平行趋势假设，但平行趋势又不可检验，因此，从本质上来说，我们难以评判控制组选择是否合适。而合成控制法更适用于单一或少量处理个体的政策评估，且并不需要依赖于处理组和控制组间满足平行趋势，而是利用数据来合成一个控制组，从而匹配处理组。最近几年，理论计量经济学家开始融合 DID 和 SC 各自的优点，提出了一些新的面板数据因果推断方法，例如 Xu（2017）提出的交互固定效应反事实估计量，Athey et al.（2021）提出的矩阵完成估计量，以及 Arkhangelsky et al.（2021）提出的合成双重差分估计量。

综上所述，本文从以下几个方面补充了现有面板数据因果推断的最新方法论综述文献：第一，在简要回顾异质性处理效应稳健 DID 估计量后，综述连续型 DID 的异质性处理效应稳健估计量，多处理变量的传染性偏误及稳健估计量；第二，在简要回顾 SC 的理论进展后，重点综述因子模型在 SC 方法论中的作用，并概述共形推断（comformal inference）在 SC 推断中的应用；第三，重点综述 DID 和 SC 的融合，以及几种重要的统一方法论框架；第四，本文使用刘友金和曾小明（2018，中国工业经济，11:98-116）的数据来重新评估 2011 年在重庆和上海实施的房产税试点对产业转移的效应，并给出一些 SDID 的实践应用建议。

## 二、双向固定效应模型和双重差分模型

在面板数据模型中，DID 模型是经济学经验研究中最流行的方法之一（Currie et al., 2020）。它利用处理组和控制组之间的变动、处理前时期和处理后时期的变动，并比较两类变动之间的差异来得到平均处理效应。这种方法将控制组和处理前时期趋势作为处理组和处理后的反事实结果。其核心假设是处理组和控制组在政策干预前的平行趋势（Angrist and Pischke, 2008），并可以消除不可观测的时间和个体固定效应，因此，研究者常用双向固定效应模型来估计得到 DID 估计量。国内外有大量的经验研究使用这种方法，例如，Ashenfelter and Card（1985）、Bertran et al.（2004）、Angrist and Pischke（2015）、彭飞等（2020）、田淑英等（2022）、许文立和孙磊（2023）。<sup>1</sup>

---

<sup>1</sup> 关于 DID 最新进展的文献回顾，可以参见 Jonathan Roth, Pedro H.C. Sant'Anna, Alyssa Bilinski, John Poe（2022）. What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. Clément de Chaisemartin, Xavier D'Haultfoeuille（2022）. Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey. The Econometrics Journal, forthcoming.

## （一）双向固定效应估计量

在线性回归的一般性假设和无预期效应假设下，传统的向双固定效应（TWFE）模型声明如下：

$$Y_{i,t} = \alpha_i + \alpha_t + \beta D_{i,t} + \epsilon_{i,t} \quad (1)$$

其中， $Y_{i,t}$ 表示结果变量， $\alpha_i$ 表示个体固定效应， $\alpha_t$ 表示时间固定效应， $D_{i,t}$ 表示处理变量， $\epsilon_{i,t}$ 表示随机扰动项，且与处理变量均值独立，下标  $i$  和  $t$  分别表示截面和时间识别码。需要注意的是，在面板数据环境下，处理变量并不一定是二值型，最可能是一个连续性或多值离散变量。

TWFE 的方程（1）隐含着两个假设：① 控制组的潜在结果满足  $Y_{i,t}(0) = \alpha_i + \alpha_t + \epsilon_{i,t}$ ；② 处理组的潜在结果满足  $Y_{i,t}(1) = Y_{i,t}(0) + \beta$ ，这个假设意味着无动态处理效应，且处理效应恒定。因此，结合这两个假设可以得到结果变量实现值的函数方程：

$$Y_{i,t} = D_{i,t}Y_{i,t}(1) + (1 - D_{i,t})Y_{i,t}(0) \quad (2)$$

当获得了方程（2）的数据生成过程（DGP）信息后，研究者就可以用最小二乘（LS）方法来估计 TWFE 方程（1）的参数：

$$(\hat{\beta}^{TWFE}, \hat{\alpha}^i, \hat{\alpha}^t) = \arg \min_{\beta, \alpha_i, \alpha_t} \sum_{i=1}^N \sum_{t=1}^T (Y_{i,t} - \alpha_i - \alpha_t - \beta D_{i,t})^2 \quad (3)$$

需要注意的是，在面板数据环境下，处理变量  $D_{i,t}$  并不一定是二值型，最常见的形式是一个连续性或多值离散变量。

## （二）双重差分估计量

处理变量为 0-1 二值型时，研究者通常采用更加灵活的回归方法来获得 DID 估计量，例如， $2 \times 2$  DID 可以当做是一种固定效应估计量，而且很容易在回归里加入个体和时间虚拟变量，以及其它协变量（Angrist 和 Pischke, 2009）。因此，在实践中，大部分学者都会采用式（3）形式的双向固定效应（TWFE）模型来获得 DID 估计量：

$$Y_{i,t} = \alpha_i + \alpha_t + \beta^{DID} D_{i,t} + \epsilon_{i,t} \quad (4)$$

其中， $Y_{i,t}$ 表示结果变量， $\alpha_i$ 表示个体固定效应， $\alpha_t$ 表示时间固定效应， $D_{i,t}$ 表示示性变量，即处理变量——处理期的处理个体的值，在二值虚拟变量的情形下， $D_{i,t} = I_i * T_t$ ，其中，如果个体接受处理， $I_i=1$ ，否则为 0；如果在处理期后， $T_t=1$ ，否则为 0。 $\beta^{DID}$ 就是研究者感兴趣的平均处理效应 ATT（或者 ATE）。基于双向固定效应回归的 DID 有许多优势：（1）为平均处理效应  $\beta^{DID}$  提供了点估计值和标准误；（2）可以加入协变量、时间趋势、动态处理效应，变换处理变量形式等来适用于形式更加一般化、多样化的 DID 研究设计，例如连续型 DID（Nunn and Qian, 2011）、混合截面 DID（Huang and Liu, 2023）、队列 DID（Chen et al., 2020）等。

正因为研究者用 TWFE 模型来估计 DID 估计量，所有 TWFE 的识别假设也适用于 TWFE DID

估计量。且同样使用最小二乘法得到的回归参数为：

$$(\hat{\beta}^{DID}, \hat{\alpha}^i, \hat{\alpha}^t) = \arg \min_{\beta, \alpha_i, \alpha_t} \sum_{i=1}^N \sum_{t=1}^T (Y_{i,t} - \alpha_i - \alpha_t - \beta D_{i,t})^2 \quad (5)$$

用潜在结果的条件期望 $E[Y_z|post]$ 、 $E[Y_z|pre]$ 、 $E[Y_j|post]$ 和 $E[Y_j|pre]$ 来分别替代处理组和控制组的结果变量均值。那么，DID 估计量可以表示为：

$$\begin{aligned} \hat{\beta}^{DID} = & (E[Y_z^1|post] - E[Y_z^0|pre]) - (E[Y_j^0|post] - E[Y_j^0|pre]) + (E[Y_z^0|post] - E[Y_z^0|pre]) \\ & = \underbrace{(E[Y_z^1|post] - E[Y_z^0|post])}_{\text{平均处理效应ATT}} + \underbrace{[(E[Y_j^0|post] - E[Y_j^0|pre]) - (E[Y_z^0|post] - E[Y_z^0|pre])]}_{\text{平行趋势偏差}} \end{aligned} \quad (6)$$

DID 研究设计实质上是估计处理组的平均处理效应（ATT）。而式（6）表明，只有当第二个等号右边的第二项中括号内容等于 0 的情况下，才可能干净地识别出 ATT。第二项内容为 0 有另一个熟知的叫法——平行趋势假设。即只有假设处理组和对照组满足平行趋势，才能干净识别处理组的平均处理效应 ATT。

平行趋势假设成立可以推导出： $E[Y_z^0|pre] - E[Y_j^0|pre] = E[Y_z^0|post] - E[Y_j^0|post]$ 。这表明平行趋势假设可以由两个部分组成：① 处理前，处理组 $E[Y_z^0|pre]$ 和控制组 $E[Y_j^0|pre]$ 存在一定的差异（趋势）；② 处理后，如果没有发生处理，处理前的差异趋势会延续到处理之后。这也就意味着，从经验角度来看，平行趋势假设不可检验。因为第二部分是“没有发生处理”的反事实，处理组的反事实结果观测不到，因此平行趋势假设不可能用现实数据来证明是成立的，但是我们可以为它提供坚实的经验证据。

虽然平行趋势假设不可检验，但在实践中，许多研究者都将平行趋势假设分解成两步：第一步，处理组和控制组处理前的变化趋势，这种变化趋势可以通过数据来检验，被称为“处理前趋势检验（pretrends test）”（Roth, 2022）；第二步，在处理后期，外生政策冲击以相同的方式、程度影响处理组和控制组，被称为“共同冲击（common shocks）”（Dimick and Ryan, 2014; Ryan et al., 2015）。

应用研究者对“处理前趋势”较为熟悉，通常使用事件研究法来实现，可以参见张子尧和黄炜（2023）的综述。而“共同冲击”则有两个核心的要点：① 外生的政策冲击，也就是说政策干预 $D_{i,t}$ 是外生的。一般来说，研究者在评估政策效应时，都是借助“自然实验”或者“准实验”框架（The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel, 2021），其核心在于政策实施具有一定的随机性，满足一定的外生性；② 以相同的方式、相同的程度影响处理组和控制组，这意味着政策干预 $D_{i,t}$ 的效应具有同质性（许文立，2023）。从理论来说，真正地检验“平行趋势假设”应该是“处理前趋势”和“共同冲击”二者检验的结合。

### （三）异质性处理效应稳健双重差分估计量

在交叠处理的情形下，(5)式得到的 TWFE DID 估计量难以理解（Goodman-Bacon, 2021）。下面，采用 Arkhangelsky and Imbens（2023）的分析框架来分解 TWFE DID 估计量。定义

$t$  为时间,  $a$  为处理时点。对于两个时间  $t$  和  $t'$ , 以及两个处理时点  $a$  和  $a'$ , 满足  $t' < a \leq t$ , 且  $a' \leq t'$  或者  $t \leq a'$ 。定义  $2 \times 2$  DID 估计量:

$$\hat{\tau}_{t,t'}^{a,a'} = (Y_{t|a} - Y_{t'|a}) - (Y_{t|a'} - Y_{t'|a'}) \quad (7)$$

如果  $t \leq a'$ , 也就是说, 处理时点为  $a'$  的个体在  $t$  期还未受到处理, (7) 式的第二个差分就是“干净的”控制组, (7) 式得到的  $\hat{\tau}_{t,t'}^{a,a'}$  是标准的  $2 \times 2$  DID 估计量。但是, 当  $a' \leq t'$  时, (7) 式就使用了“先处理组”作为控制组。当存在异质性处理效应时, (7) 式的期望为:

$$E[\hat{\tau}_{t,t'}^{a,a'}] = \tau_{t|a} - \tau_{t|a'} + \tau_{t'|a'} \quad (8)$$

(8) 式意味着,  $E[\hat{\tau}_{t,t'}^{a,a'}]$  是处理效应的加权平均, 且权重之和为 1, 但是有一些权重可能为负 (Goodman-Bacon, 2021; de Chaisemartin and D' Haultfoeuille, 2020)。

为了纠正负权重所带来 TWFE DID 估计量偏误, 近几年的 DID 理论文献提出了许多异质性处理效应稳健估计量。对于这些估计量的介绍和比较, 已经有许多文献综述 (de Chaisemartin and D' haultfoeuille, 2022; Roth et al., 2023; 刘冲等, 2022; 张子尧和黄炜, 2023; 许文立, 2023)。但需要强调的是, Callaway and Sant Anna (2021) 估计量只要求处理前后一个时点的平行趋势延续即可, de Chaisemartin and D' Haultfoeuille (2020, 2021, 2022) 估计量则要求处理后一期延续处理前多期的平均趋势, Borusyak et al. (2023) 估计量则要求处理前后多个对应时点满足平行趋势, 最为严格。

#### (四) 异质性处理效应下连续型双重差分估计量

在二值型处理变量的情形下, 政策干预仅仅表示“有无”的粗略处理。但在许多情形下, 政策有强弱, 甚至连续变化 (Callaway et al., 2021), 例如, 受教育年限 (Angrist and Krueger, 1994)、最低工资 (Card, 1994; 夏璋煦等, 2024)、农作物种植面积 (Qian, 2008; Nunn and Qian, 2011) 等等。这些因果效应估计方法称为“连续/强度 DID”。

传统的连续型 DID 回归方程声明为:

$$Y_{i,t} = W_i + Post_t + \beta^{CDID} Post_t \times W_i + \epsilon_{i,t} \quad (9)$$

其中,  $W_i$  表示连续处理变量,  $Post_t$  表示政策干预时点, 处理后一直为 1, 否则为 0。这种 TWFE 模型本质上还是区分处理前后, 以及处理对不同个体的影响程度不同来划分处理组和控制组。这种连续变量与时间虚拟变量交乘项的方式来识别连续型处理变量的因果效应在应用研究中比较常见, 例如, Finkelstein (2007, QJE) 的模型包含医疗补助政策时间指标与私人保险可及性交乘项, Dube and Vargas (2013, RES) 的模型包含世界油价与地区石油生产强度的交乘项, 以及 Zhang and Zhu (2011, AER)、Dafny et al. (2012, AER)、Pierce and Schott (2016, AER) 等。

在交叠处理情形下, 异质性处理效应同样会使得上述连续型 DID 估计量产生偏误 (Callaway et al., 2021; de Chaisemartin et al., 2022; D' Haultfoeuille et al., 2023)。

假设有两期  $t-1$  和  $t$  期，在  $t-1$  期所有的个体都没有受到处理，在  $t$  期有一些个体受到处理。且处理变量  $D_{i,t}$  是连续的计量（does）。处理组的潜在结果变量  $Y_{i,t} = Y_{i,t}(D_{i,t})$ ，控制组的结果变量  $Y_{i,t}(0)$ 。Callaway et al. (2021) 定义了连续处理变量情形下的平均处理效应：

$$ATT(d'|d) = E[Y(d') - Y(0)|D = d] = E[Y(d') - Y(0)|D = d] + E[Y(d') - Y(0)|D = 0] - E[Y(d') - Y(0)|D = 0] \quad (10)$$

其中， $ATT(d'|d)$  表示处理强度为  $d$  的组群，实际发生处理  $d'$  的平均处理效应。(10) 式也表明，在连续型处理变量下，只要平行趋势假设成立  $E[Y - Y(0)|D = 0] = E[Y - Y(0)|D = d]$ ，就可以识别处理效应。但是在连续处理变量情形中，研究者根据可观测数据得到的可能是下列处理效应：

$$ATT(d'_2|a) - ATT(d'_1|c) = \underbrace{ATT(d'_2|a) - ATT(d'_1|a)}_{\text{平均处理效应ATT}} + \underbrace{ATT(d'_1|a) - ATT(d'_1|c)}_{\text{异质性处理效应偏误}} \quad (11)$$

(11) 式意味着不同处理强度的组群之间如果存在异质性处理效应，那么，就会存在偏误。

Callaway et al. (2021) 和 de Chaisemartin et al. (2022) 所讨论的连续型处理变量的估计量更适合于一般形式的面板数据模型：

$$Y_{i,t} = \alpha_i + \alpha_t + \beta^{CDID} D_{i,t} + \epsilon_{i,t} \quad (12)$$

其中， $D_{i,t}$  为时变的连续型处理变量。这与传统的连续型 DID 的交乘项回归系数估计存在本质区别。

为了矫正异质性处理效应对连续 DID 估计量带来的偏误，de Chaisemartin and d' Haultfoeuille (2018, 2019) 提出了“模糊双重差分估计量”(Fuzzy DID)，可以使用 stat 命令 `fuzzydid` 和 `did_multipligt` 来实现；Cerulli (2015) 提出了“计量响应函数”(Does Response Function, DRF)，可以使用 stata 命令 `ctreatreg` 来实现；Hegland (2023) 扩展 Wooldridge (2021) 估计量，可以使用 stata 命令 `wooldid` 实现。

### (五) 多处理变量的双重差分估计量

在回归方程包含多个处理变量是一种常见的回归方程设置，例如，Krueger et al. (1999, QJE) 小班教学和配备助教是否可以提高初级教育的水平。由于存在两个处理组——小班和有助教班，作者引入两个虚拟变量 `small` 和 `aide`，前者表示学生在小班为 1，否则为 0，后者表示学生在有助教班级为 1，否则为 0。回归方程可以设定为：

$$score_i = \alpha + \beta_1 small_i + \beta_2 aide_i + X_i \Gamma' + u_i \quad (13)$$

其中， $score_i$  表示学生测试成绩；如果学生在小班  $small_i = 1$ ，否则为 0；如果学生在有助教班级， $aide_i = 1$ ，否则为 0。 $X_i$  表示控制变量向量。相对于正常班级来说，小班教学的效应是  $\beta_1$ ，而有助教的效应是  $\beta_2$ 。由于上述教学实验属于随机实验，研究者可以使用 OLS 来估计两个效应的差分估计量 (Stock and Watson, 2019)。

还有一种更常见的多处理变量 DID 回归，那就是在排除其他政策干扰时，研究者通常会

在 TWFE 回归方程中引入干扰政策的虚拟变量作为控制变量（de Chaisemartin and D'haultfœuille, 2023）。此外，面板事件研究中通常设置每个时点一个虚拟变量，从而存在多个处理变量（Sun and Abraham, 2021）。

由此，在 DID 设计中，多处理变量主要存在以下几种类别，以最简单的 2 政策为例，可以划分为 3 种情形，如表 1 所示：①（1）表示同一政策，在不同时点上多个处理变量，例如事件研究设计；②（1 0）和（0 1）表示政策 A 和 B 是相互排斥的，即个体只能接受其中一种政策，例如，Krueger et al.（1999, QJE）的小班教学和助教教学的效应；③（1 1）表示个体可以同时接受两种政策的处理。对于第一种情形——事件研究设计，目前张子尧和黄炜（2023）已经较好地综述。

表 1 多处理变量的分类

	(政策 A 政策 B)
个体	(1) (1 0) (0 1) (1 1)

对于第二种多处理变量的情形，研究者常见的政策评估情形是单政策/处理，例如，只评估小班教学的效应，上述 OLS 回归中的  $small$  就是核心解释变量，而  $aide$  和  $X$  就是控制变量，表示当控制  $aide$  和  $X$  不变时， $small$  的效应，也就相当于分组回归，相同的  $aide$  和  $X$  水平的样本进行比较。此时，研究者只关心 OLS 估计量  $\beta_1$ ：

$$\beta_1 = \underbrace{\frac{Cov(\widehat{small}_i, score_i(0))}{var(\widehat{small}_i)}}_{=0} + \frac{Cov(\widehat{small}_i, \widehat{small}_i \times \tau_i)}{var(\widehat{small}_i)} = \frac{var(\widehat{small}_i) \times \tau_i}{var(\widehat{small}_i)} = \tau_i \quad (14)$$

其中， $score_i(0)$ 表示未处理的结果， $\tau_i = score_i(1) - score_i(0)$ 表示处理效应。上式等号右边第一项等于 0，因为未处理时，基于控制变量的条件下， $E(\widehat{small}_i) = 0$ 。

在多处理/政策情形下，例如小班教学和助教教学两类改革。此时，小班教学的效应  $\beta_1$ ，助教的效应  $\beta_2$ 。且定义  $\tau_{1i}$  和  $\tau_{2i}$  分别为小班和助教潜在结果变量差分。小班教学的效应  $\beta_1$  可以分解为：

$$\beta_1 = \underbrace{\frac{Cov(\widehat{small}_i, score_i(0))}{var(\widehat{small}_i)}}_{=0} + \frac{Cov(\widehat{small}_i, \widehat{small}_i \times \tau_{1i})}{var(\widehat{small}_i)} + \underbrace{\frac{Cov(\widehat{small}_i, aide_i \times \tau_{2i})}{var(\widehat{small}_i)}}_{\neq 0} \quad (15)$$

因为此时有两个处理变量  $small_i$ 、 $aide_i$ ，所以上式出现了三项求和。前两项与单处理变量一致，多出了第三项。也就是说，小班教学的处理效应还依赖于是否进入助教班级和助教的效应。第三项不等于 0 是因为多个处理之间是相互依赖的，即选择了小班，就不能选择助教班级，即当  $small_i = 1$ ， $aide_i = 0$ ，反之亦然。 $small_i$  和  $aide_i$  的条件协方差并不等于 0： $Cov(\widehat{small}_i, aide_i) \neq 0$ 。因此，估计互斥的多个处理/政策效应时，如果使用 OLS 估计量，

就会存在偏误，这个偏误是由于其它处理/政策效应“传染”所致。

为了矫正相互排斥的多处理变量带来的偏误，Goldsmith-Pinkham et al. (2022) 给出了三种稳健估计量：① 无加权(等权重)ATE，即最简化估计量(simple estimator)，由 Imbens and Wooldridge (2009) 提出；② 特定处理的方差加权 ATE，由 Angrist (1998) 提出；③ 可比有效加权估计量。这些稳健估计量均可以使用 stata 命令 `multte` 来实现。

对于第三种多处理变量的情形，假设有四个个体、两期，个体 1 从未处理，个体 2 在第二期开始第一个处理，个体 3 在第二期开始第二个处理，个体 4 在第二期同时开始两个处理。此时，第一个处理的效应：

$$\hat{\beta}_{fe} = \frac{1}{2} (Y_{2,2} - Y_{2,1} - (Y_{1,2} - Y_{1,1})) + \frac{1}{2} (Y_{4,2} - Y_{4,1} - (Y_{3,2} - Y_{3,1}))$$

上式的第二个 DID 比较的是个体 4 处理前后和个体 3 处理前后结果的变化，将其变换为：

$$\begin{aligned} & E(Y_{4,2} - Y_{4,1} - (Y_{3,2} - Y_{3,1})) \\ &= E(Y_{4,2}(1, 1) - Y_{4,1}(0, 0) - (Y_{3,2}(0, 1) - Y_{3,1}(0, 0))) \\ &= E(Y_{4,2}(1, 1) - Y_{4,2}(0, 1)) + E(Y_{4,2}(0, 1) - Y_{4,2}(0, 0)) - E(Y_{3,2}(0, 1) - Y_{3,2}(0, 0)) \\ & \quad + E(Y_{4,2}(0, 0) - Y_{4,1}(0, 0) - (Y_{3,2}(0, 0) - Y_{3,1}(0, 0))) \\ &= E(Y_{4,2}(1, 1) - Y_{4,2}(0, 1)) + E(Y_{4,2}(0, 1) - Y_{4,2}(0, 0)) - E(Y_{3,2}(0, 1) - Y_{3,2}(0, 0)) \end{aligned}$$

由此，得到第一个处理的效应估计量：

$$\begin{aligned} \beta_{fe} &= \frac{1}{2} E(Y_{2,2}(1, 0) - Y_{2,2}(0, 0)) + \frac{1}{2} E(Y_{4,2}(1, 1) - Y_{4,2}(0, 1)) \\ & \quad + \frac{1}{2} E(Y_{4,2}(0, 1) - Y_{4,2}(0, 0)) - \frac{1}{2} E(Y_{3,2}(0, 1) - Y_{3,2}(0, 0)) \end{aligned}$$

上式表明，第一个处理的 TWFE 估计量为两项的加权和：① 第一项是等号右边的第一行，即保持处理二的状态不变，个体 2 和个体 4 在第二期的处理一的效应；② 第二项是等号右边第二行，即保持处理一状态不变，处理二从 0 到 1，个体 4 和个体 3 的处理二的效应之差。这也就意味着，如果存在异质性处理效应，也就是个体 4 和个体 3 的处理二的效应不相同，那么，处理二的异质性处理效应会导致处理一的处理效应产生偏误。

为了矫正多个处理变量带来的偏误，de Chaisemartin and D’Haultfœuille (2023) 提出了稳健估计量，可以使用 stata 命令 `did_multiplegt_dyn` 实现。目前，赵仁杰等 (2023)、Rico-Straffon et al. (2023) 应用了多处理变量偏误的稳健估计量。

### 三、因子模型和合成控制法



### （一）因子模型

因子模型在面板数据因果效应评估中也使用较为广泛（Bai, 2009）。最简单的因子模型可以表示如下：

$$y_{it} = f_t' \gamma_i + u_{it} \quad (17)$$

其中， $f_t$  是  $p \times 1$  的不可观测因子向量， $\gamma_i$  是  $p \times 1$  的不可观测因子载荷向量。

在应用研究中，研究者可以将因子  $f_t$  理解成宏观经济冲击（例如，对所有地区，所有行业，所有个体都产生影响的冲击），而  $\gamma_i$  是个体暴露于冲击的程度（例如，有些行业受到宏观经济冲击较小，而有些行业影响较大）。二者乘积描述的是个体如何受到宏观经济冲击的影响。另一种理解是  $\gamma_i$  表示时间不变的特征，而其对结果的效应  $f_t$  随时间变化。。

TWFE 实际上是因子模型的一种特例，即取  $\gamma_i = (1, \alpha_i)$  和  $f_t = (\alpha_t, 1)$ 。TWFE 模型变成  $y_{it} = \alpha_i + \alpha_t + u_{it}$ 。而因子模型允许个体有差异化的趋势（主要是因为不同个体暴露在共同冲击中的程度不同）。因子模型最大的问题在于，使用因子模型需要长面板（长时期）来估计因子和因子载荷。而我们使用 DID 的大部分情形都是较短的时期（短面板）。

### （二）合成控制法

合成控制法由 Abadie and Gardeazabal（2003）提出，Abadie et al. (2010, 2015) 讨论更详细的技术要点。这个方法与传统的 TWFE 和 DID 独立发展，有很大的不同，是过去十几年来，面板数据因果推断领域最重要的进展。合成控制法已经广泛应用于评估法律、政策等的效应（Donohue et al., 2017; Cunningham and Shah, 2018; Bohn et al., 2014; Acemoglu et al., 2016），劳动经济学领域的研究（Borjas, 2017; Peri and Yasenov, 2017; Allegretto et al., 2017; Jardim et al., 2017; Neumark and Wascher, 2017; Reich et al., 2017），以及经济学之外的领域（Heersink et al., 2017; Pieters et al., 2017）。Abadie（2021）对合成控制法的最新理论进展做了较好的综述，但本文的综述角度稍有不同，以便与 TWFE/DID 进行比较。

传统的合成控制法文献主要关注于利用控制组的凸组合来构造处理组的合成控制组，从而插补（impute）出处理组的反事实潜在结果。合成控制组的关键在于利用数据来计算最优的凸组合权重，因此，它更多基于算法——主要源自于因子模型的算法（Arkhangelsky and Imbens, 2023）。

假设未处理的个体的潜在结果  $Y_{it}(0) = f_t' \gamma_i + u_{it}$ ，处理组的潜在结果  $Y_{it}(1) = Y_{it}(0) + \beta^{SC} D_{it} = \beta^{SC} D_{it} + f_t' \gamma_i + u_{it}$ 。因为传统合成控制法仅仅适用于单一/少量处理组，因此，时间不变的个体特征退化为  $\gamma_i = (1, 1)$ ，即  $Y_{it}(1) = Y_{it}(0) + \beta^{SC} D_{it} = \beta^{SC} D_{it} + f_t' + u_{it}$ ，令  $f_t = (\alpha_t, 1)$ ，由此，可以得到 Doudchenko and Imbens（2016）所刻画的合成控制法最小二乘估计量：

$$(\hat{\beta}^{SC}, \hat{\alpha}^t) = \arg \min_{\beta, \alpha_t} \sum_{i=1}^N \sum_{t=1}^T \hat{\omega}_i^{SC} (Y_{it} - \alpha_t - \beta^{SC} D_{it})^2 \quad (18)$$

非负权重 $\hat{\omega}_i^{sc}$ 定义了处理个体  $i$  的合成控制组。需要注意的是 Abadie and Gardeazabal (2003) 和 Abadie et al. (2010, 2015) 的初始结论是非负权重是稀疏的，也就是说只有少量的控制组个体会获得正的权重。同时，由于存在不可观测的时变因子 $f_t'$ ，合成控制法也要求处理前有较长的时期。

Athey et al. (2021) 指出，(20) 式中的合成控制法的权重估计量 $\hat{\omega}_i^{sc}$ 依赖于依赖于  $N-1$  个未处理个体的  $T-2$  期的未处理时期的观测值的垂直回归，而处理效应估计量 $\hat{\beta}^{sc}$ 依赖于  $N-1$  个未处理个体的 2 期处理时期的观测值的水平回归，如图 1 所示。需要注意的是，最小二乘算法得到的权重可能并不唯一，Spiess et al. (2023) 提出了一种更优的解法。

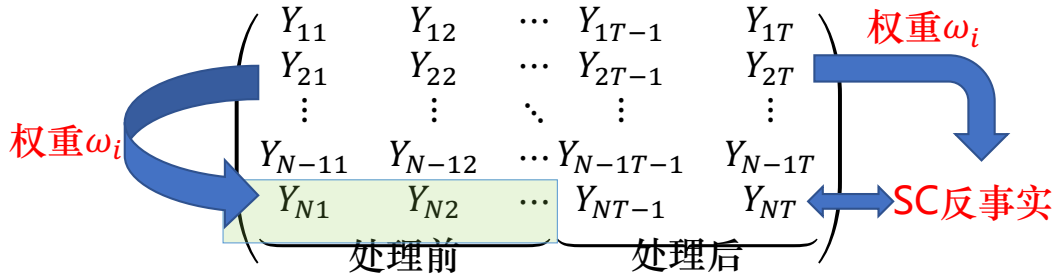


图 1 合成控制法的回归特征示意图

传统合成控制法仅仅加权控制组。从理论上说，研究者也可以加权处理组，类似于加权处理组处理前时期的个体 (Kuusmanen et al., 2021)。除此之外，还有学者建议将匹配方法和合成控制法结合起来，以避免合成控制法的外推偏误 (Kellogg et al., 2021)。

### (三) 合成控制法的推断

合成控制法的推断一直以来都极具挑战，学术界关于估计方差-协方差矩阵，进而构建置信区间的最佳方式并没有形成一致 (Arkhangelsky and Imbens, 2023)。目前，应用研究中最常用的方法是利用安慰剂方法来检验无处理效应的原假设 (Cattaneo et al., 2022a)。为此，Abadie et al. (2010) 提出了“重新排序检验” (permuted test)，即假设控制组随机接受处理，检验控制组的安慰剂处理的效应和处理组的处理效应差异。如果估计的处理组处理效应明显异于安慰剂效应，那么，推断出处理效应显著。但是，这个推断过程假设安慰剂处理随机配置，这在观测数据中很难满足 (Abadie, 2021)。

近年来，一类被称为“共形推断” (conformal inference) 的文献已经作为合成控制法的推断框架 (Chernozhukov et al., 2021; Cattaneo et al., 2021; Cattaneo et al., 2022b)。共形推断方法就是将处理效应估计问题当做反事实预测问题。

检验合成控制法中处理后的处理效应的时间路径 $\theta = (\theta_a, \dots, \theta_T)$ 的原假设：

$$H_0: \theta = \theta^0$$

例如，如果研究者希望检验无处理效应，则设置 $\theta^0 = (0, \dots, 0)$ 。注意，这个假设完全决定了反事实结果：

$$Y_t(0) = Y_t(1) - \theta_t = Y_t - \theta_t$$

共形推断的核心思想是从原假设中生成数据，用生成的数据来检验反事实结果 $Y_t(0)$ 模型的余值。如果这些余值非常极端，就意味着数据不可能来自于原假设。基于此，定义共形推断的统计量  $S$ ：

$$S(\hat{u})_q = (\sum_{t=a}^T |u_t|^q)^{1/q} \quad (19)$$

其中， $u_t$ 是反事实结果 $Y_t(0)$ 模型的余值。注意这个统计量仅仅用处理后时期来计算。这个统计量的值越高意味着处理后的拟合越差，因此，拒绝原假设。但是在实践中，即使原假设为真，研究者可能会遇到一个很大  $S$  统计量，意味着反事实结果拟合较差。因此，在实践中通常利用处理后的  $S$  统计量和处理前的余值的相对大小来对合成控制法的处理效应进行推断。同时，还可以重排余值，计算每个重排的  $S$  统计量。一旦进行  $\Pi$  次重排，就可定义共形推断的  $p$  值：

$$P\text{-value} = \frac{1}{|\Pi|} \sum_{\pi \in \Pi} 1 \{S(\widehat{u}_{\pi_0}) \leq S(\widehat{u}_{\pi})\} \quad (20)$$

其中， $\widehat{u}_{\pi_0}$ 表示原始余值向量。

实践中，研究者可以使用 python 代码来实现 Chernozhukov et al. (2021) 的处理后余值排序推断方法，而 Cattaneo et al. (2021) 和 Cattaneo et al. (2022a,b) 的随机潜在结果方法则可以使用 stata 命令 `scpi` 命令来实现。

## 四、结合双重差分法与合成控制法的统一框架

Cavallo et al. (2013, RES) 将单处理个体的合成控制方法扩展至多处理个体的合成控制方法，且允许交叠处理时点。应用多处理个体的合成控制方法也有许多文献 (Dube & Zipperer, 2015; Acemoglu et al., 2016; Abadie & L'Hour, 2020; Abadie, 2021; Ben-Michael et al., 2021b; Wiltshire, 2021a)。这些理论与应用也为合成控制法与双重差分法的结合提供了基础。

最近有三种方法结合了合成控制法和双向固定效应 DID 的优势，形成了一种统一的面板数据因果推断框架。

### (一) 广义合成控制法

Xu (2017) 提出了一种交互固定效应反事实估计量，这种方法将合成控制法和线性固定效应结合在一起，用基于线性交互固定效应模型的控制组信息来插补每个处理个体的反事实结果。作者将反事实潜在结果定义为 $Y_{it}(0) = \alpha_i + \alpha_t + f_t' \gamma_i + u_{it}$ ，那么，广义合成控制法就可以表示为下列因子模型：

$$Y_{it} = \beta^{GSC} D_{it} + Y_{it}^N = \beta^{GSC} D_{it} + (\alpha_i + \alpha_t + f_t' \gamma_i + u_{it}) \quad (21)$$

由此可以得到“交互固定效应反事实估计量” (Liu et al., 2022)：

$$(\hat{\beta}^{GSC}, \hat{\alpha}^i, \hat{\alpha}^t) = \arg \min_{\beta, \alpha_t} \sum_{i=1}^N \sum_{t=1}^T \hat{\omega}_i^{sc} (Y_{i,t} - \alpha_i - \alpha_t - f_t' \gamma_i - \beta^{GSC} D_{i,t})^2 \quad (22)$$

广义合成控制法的优势在于因子结构可以解决时变混淆因子和异质性处理效应引起的

平行趋势不成立偏误，并调和合成控制法的外推偏误。但是，由于估计因子结构，它也需要较长的处理前时期。

## （二）矩阵完成法

广义合成控制法假设不可观测的时变混淆因子的函数形式已知，且采用交互固定效应形式  $f_t' \gamma_i$ 。但是在现实中，不可观测的时变混淆因子的函数形式通常未知。对于未知的函数形式，Athey et al. (2021) 提出了一种估计方法——矩阵完成法 (matrix completion)。作者假设要估计的未知函数形式的混淆因子  $L_{it}$  的估计值为：

$$\hat{L} = \arg \min_L \left[ \sum_{(i,t) \in \mathcal{O}} \frac{(Y_{it} - L_{it})^2}{|\mathcal{O}|} + \lambda_L \|L\| \right] \quad (23)$$

其中， $\mathcal{O} = \{(i, t) \mid D = 0\}$ ， $\|L\|$  表示选择矩阵  $L$  的模， $\lambda_L$  表示调节参数或者惩罚参数。作者用机器学习算法来迭代得到高维矩阵  $L$  的估计值  $\hat{L}$ 。只要获得这个估计量，就可以计算出每个处理后处理个体的反事实结果：

$$Y_{it}(0) = \alpha_i + \alpha_t + L_{it} + u_{it} \quad (24)$$

由此，即可计算出处理组的平均处理效应。实践中，广义合成控制法和矩阵完成法可以通过 stata 命令 `fect` 来实现。但这些方法都包含不可观测时变混淆因子，因此都要求处理前时期较长，而且要声明调节参数。这也使得它们偏离了传统的 TWFE 方法，应用不够灵活。例如，在矩阵完成法中，调节参数刻画了模惩罚的重要性，而在交互固定效应反事实估计量中，研究者需要声明因子模型的阶 (rank)。理论上，研究者可以使用机器学习的算法来选择合适的调节参数，例如样本交叉差验证法。但是截面维度为长时序的交叉验证法带来的其它挑战。

## （三）合成双重差分法

一般来说，未处理个体中有一些与处理个体更为相似，而更靠近处理的时期与处理时期也会更为相似。而从 TWFE DID 估计量和合成控制法估计量来看，双重差分的最小二乘估计量隐含着未处理个体和时间的权重相等（都为 1），即将所有未处理个体和时间都同等待待。而合成控制法虽然包含了未处理个体权重，但对于时间层面并未区分重要程度。

通常来说，DID 更适合于有大量处理个体的政策评估环境，且它严重依赖于平行趋势假设，但平行趋势又不可检验，因此，从本质上来说，我们难以评判控制组选择是否合适。而合成控制法更适用于单一或少量处理个体的政策评估，且并不需要依赖于处理组和控制组间满足平行趋势，而是利用数据来合成一个控制组，从而匹配处理组。Arkhangelsky et al. (2021) 提出了一种新的估计量——合成双重差分 (Synthetic Difference in Differences, SDID)，这个估计量综合了双重差分和合成控制法各自的优势。SDID 估计量的作者们指出，该估计量有两个方面的优势：第一，重新加权回归的过程极大地削弱了估计量对平行趋势或者处理外生性假设的依赖性；第二，将个体加权和权重引入双重差分中也可以使得该估计量适用于任意数量的处理个体。这就使得 SDID 估计量既适用于 DID 环境，又适用于 SC 环境。SDID 估计量为处理前时期和截面个体分别计算权重来构造一个合成反事实结果，计算实际处理结

果和合成反事实结果之间的差异，获得单一的平均处理效应，从而进行因果推断。

### 1、静态 SDID 估计量

SDID 则结合了 DID 和 SC 两者的优势，即在保留了 DID 的双向固定效应 $\alpha_i$ 、 $\beta_t$ ，也保留了 SC 的个体权重 $\hat{\omega}_i^{sdid}$ 。与此同时，还增加了时间权重 $\hat{\lambda}_t^{sdid}$ ：

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \quad (25)$$

与 DID 估计量相比，SDID 估计量中的个体权重强调与处理前时期处理个体越相似的潜在控制组就越重要，即处理个体的平均结果与控制个体的加权平均近似平行；时间权重则强调与处理后时期越相似的时期就越重要，即每个控制组个体的处理后结果的均值与同一个体处理前结果的加权平均值相差一个常数。而且，SDID 还解决了 Roth (2022) 指出的 DID 实践中处理前趋势检验的问题。

与 SC 估计量相比，SDID 的时间权重可以通过移除那些与处理后时期差异较大的时期来消除偏误，并提高精度。而包含个体固定效应则使得模型更加的灵活，且个体固定效应可以解释相当大比例的结果变动，并提高精度。

假设数据产生过程如下：

$$Y = L + W_{it}\tau_{it} + E \quad (26)$$

其中， $\tau_{it}$ 表示处理效应，且允许个体和时间上存在异质性处理效应。正如 Arkhangelsky et al. (2021) 指出，SDID 的目标是要估计处理后处理组的平均处理效应，即

$$\tau = \frac{1}{N_{tr}T_{post}} \sum_{i=N_{co}+1}^N \sum_{t=T_{pre}+1}^T \tau_{it} \quad (27)$$

我们可以将结果变量表示成如下矩阵的形式

$$Y = \begin{pmatrix} Y_{co,pre} & Y_{co,post} \\ Y_{tr,pre} & Y_{tr,post} \end{pmatrix}$$

方差为：

$$\Xi = \begin{pmatrix} \Xi_{pre,pre} & \Xi_{pre,post} \\ \Xi_{post,pre} & \Xi_{post,post} \end{pmatrix}$$

因此，总的估计量可以表示为

$$\hat{\tau} = \left[ \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left( \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (28)$$

其中，个体权重 $\hat{\omega}_i^{sdid}$ 和时间权重 $\hat{\lambda}_t^{sdid}$ 的定义与最优化过程请参考 Arkhangelsky et al. (2021) 的正文和附录。

下面用矩阵形式表示个体和时间加权后的结果变量和方差：

$$\begin{pmatrix} \mathbf{Y}_{::} & \mathbf{Y}_{:T} \\ \mathbf{Y}_{N:} & \mathbf{Y}_{NT} \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_{co,pre} & \mathbf{Y}_{co,post} \hat{\lambda}_{post}^{sdid} \\ \hat{\omega}_{tr}^{sdid} \mathbf{Y}_{tr,pre} & \hat{\omega}_{tr}^{sdid} \mathbf{Y}_{tr,post} \hat{\lambda}_{post}^{sdid} \end{pmatrix}$$

其中，对于单一平均处理效应估计量来说， $\hat{\lambda}_{post}^{sdid}$  和  $\hat{\omega}_{tr}^{sdid}$  时间和个体权重向量，分别等于  $\frac{1}{T_{post}}$  和  $\frac{1}{N_{tr}}$ 。在这种紧缩形式中，行  $\varepsilon_{it}$  表示均值为 0，方差为  $\Xi$  的独立高斯向量，方差矩阵  $\Xi$  满足下列关系：

$$\Xi = \begin{pmatrix} \Xi_{pre,pre} & \Xi_{pre,post} \hat{\lambda}_{post} \\ \hat{\lambda}_{post} \Xi_{post,pre} & \hat{\lambda}_{post} \Xi_{post,post} \hat{\lambda}_{post} \end{pmatrix}$$

Kranz (2022) 在 Arkhangelsky et al. (2021) 的基础上，提出了另一种引入协变量的方式来改善静态估计量的精度。Porreca (2022) 则进一步将该估计量推广到交叠处理的情形，从而获得静态平均处理效应。

正如 (5) 式所示，SDID 估计量可以允许个体和时间的异质性处理效应。但是 Arkhangelsky et al. (2021) 在论文中并未展开异质性处理效应分析，也没有给出异质性处理效应的正式估计量。今年来，因果效应推断计量理论取得了极大的进展，尤其关注时间层面的异质性处理效应，即处理效应随时间变化而变化 (Goodman-Bacon, 2021; Sun and Abraham, 2021)，提出了估计处理后每一期的处理效应 (Clarke and Tapia Schythe, 2021)。因此，Liu et al.(2022)、Clarke et al. (2023) 也考虑用 SDID 估计量来估计处理后每一期的处理效应，并给出正式的估计量形式。

## 2、动态 SDID 估计量：统一处理时点

利用 SDID 估计处理后每一期的处理效应仅仅只需要对处理后的每一期分别运行一次 SDID 估计过程。首先，我们来考察只有单一处理个体的估计量。假设处理后有两期，那么，结果矩阵就可以表示为：

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{co,pre} & \mathbf{Y}_{co,post1} & \mathbf{Y}_{co,post2} \\ \mathbf{Y}_{tr,pre} & \mathbf{Y}_{tr,post1} & \mathbf{Y}_{tr,post2} \end{pmatrix}$$

其中， $post1$  和  $post2$  分别表示处理后第一期和处理后第二期。我们可以将上述结果矩阵分解为两个子矩阵：

$$\begin{aligned} \mathbf{Y}^1 &= \begin{pmatrix} \mathbf{Y}_{co,pre} & \mathbf{Y}_{co,post1} \\ \mathbf{Y}_{tr,pre} & \mathbf{Y}_{tr,post1} \end{pmatrix} \\ \mathbf{Y}^2 &= \begin{pmatrix} \mathbf{Y}_{co,pre} & \mathbf{Y}_{co,post2} \\ \mathbf{Y}_{tr,pre} & \mathbf{Y}_{tr,post2} \end{pmatrix} \end{aligned}$$

基于上述两个结果子矩阵，分别应用 SDID 估计处理后的处理效应。也就是说，处理后的每一期处理效应估计量可以表示为：

$$\hat{\tau}_{t'} = \left[ \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (29)$$

其中， $t'$  表示处理后的时期。与 (28) 式相比，在 (29) 式的处理后每期处效应估计量

中，由于只有一个处理组，因此， $N_{tr} = 1$ ，而处理组个体之和 $\sum_{i=N_{co}+1}^N = 1$ ，因此，处理组结果差分只需要使用单一处理组的结果减去合成结果 $Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it}$ 。同理，在估计处理后的每一期处理效应时，将每一期的分解，并分别估计，因此，对于处理后每一期估计量来说，处理后时期也只是一期，即 $T_{post} = 1$ ，且处理后的时期总数为 $\sum_{t=T_{pre}+1}^T = 1$ 。

下面，假设有多个处理个体 $N_{tr} > 1$ ，有多个处理后时期 $T_{post}$ ，且初次处理时点相同。此时，我们仍然可以将结果矩阵表示为

$$Y = \begin{pmatrix} Y_{co,pre} & Y_{co,post1} & Y_{co,post2} \cdots & Y_{co,postT} \\ Y_{tr,pre} & Y_{tr,post1} & Y_{tr,post2} \cdots & Y_{tr,postT} \end{pmatrix}$$

其中，处理组结果变量矩阵为 $Y_{tr:} = \begin{pmatrix} Y_{tr1:} \\ \vdots \\ Y_{trN:} \end{pmatrix}$ 。因为本文并不关注处理效应在个体之间的

差异，因此，将不同处理组合并为一个矩阵。根据上文的处理后每一期分解，我们仍然可以将上述多期结果矩阵分解为每一期对应的子矩阵：

$$\begin{aligned} Y^1 &= \begin{pmatrix} Y_{co,pre} & Y_{co,post1} \\ Y_{tr,pre} & Y_{tr,post1} \end{pmatrix} \\ Y^2 &= \begin{pmatrix} Y_{co,pre} & Y_{co,post2} \\ Y_{tr,pre} & Y_{tr,post2} \end{pmatrix} \\ &\vdots \\ Y^T &= \begin{pmatrix} Y_{co,pre} & Y_{co,postT} \\ Y_{tr,pre} & Y_{tr,postT} \end{pmatrix} \end{aligned}$$

每一期的估计量为

$$\hat{\tau}_{t'} = \left[ \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (30)$$

其中， $t' \in \{post1, post2, \dots, postT\}$ 表示处理后的单一时期。与 (29) 式相比，(30) 中多了处理个体的加权平均项 $\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right)$ 。因为 SDID 估计程序就是对每个处理个体合成一个控制个体，然后比较处理个体与合成个体的结果差异，那么，当样本中有 $N_{tr}$ 个处理个体时，就计算所有单一个体差异的加权平均，权重为 $\frac{1}{N_{tr}}$ 。

### 3、动态 SDID 估计量：交叠处理

首先，我们来考察两类处理个体（它们的处理时点不同）的估计量。假设先处理的 cohort 处理后有两期，后处理的 cohort 处理后只有一期，那么，结果矩阵就可以表示为：

$$Y = \begin{pmatrix} Y_{co,pre} & Y_{co,post1} & Y_{co,post2} \\ Y_{tr1,pre} & Y_{tr1,post1} & Y_{tr1,post2} \\ Y_{tr2,pre1} & Y_{tr2,pre2} & Y_{tr2,post1} \end{pmatrix}$$

其中， $tr1$ 和 $tr2$ 表示两类处理组， $post1$ 和 $post2$ 分别表示处理后第一期和处理后第二期。我们可以将上述结果矩阵分解为两个子矩阵：

#### (1) 2×2 的分块

$$\begin{aligned}
Y^{11} &= \begin{pmatrix} Y_{co,pre} & Y_{co,post1} \\ Y_{tr1,pre} & Y_{tr,post1} \end{pmatrix} \\
Y^{12} &= \begin{pmatrix} Y_{co,pre} & Y_{co,post2} \\ Y_{tr1,pre} & Y_{tr,post2} \end{pmatrix} \\
Y^{21} &= \begin{pmatrix} Y_{co,pre} & Y_{co,post2} \\ Y_{tr2,pre1} & Y_{tr2,post1} \end{pmatrix}
\end{aligned}$$

(2) N×T 矩阵分块

$$\begin{aligned}
Y^{11} &= \begin{pmatrix} Y_{co,pre} & Y_{co,post1} \\ Y_{tr1,pre} & Y_{tr1,post1} \\ Y_{tr2,pre1} & Y_{tr2,pre2} \end{pmatrix} \\
Y^{21} &= \begin{pmatrix} Y_{co,pre} & Y_{co,post1} & Y_{co,post2} \\ Y_{tr2,pre1} & Y_{tr2,pre2} & Y_{tr2,post1} \end{pmatrix} \\
Y^{12} &= \begin{pmatrix} Y_{co,pre} & Y_{co,post2} \\ Y_{tr1,pre} & Y_{tr1,post2} \end{pmatrix}
\end{aligned}$$

下面的 N×T 矩阵分解与上面的 2×2 是等价的，因为个体权重和时间权重会根据数据来选择最优的合成控制组。

同一处理时点的处理组类别 j 每一期的估计量为

$$\hat{\tau}_{t'}^j = \left[ \frac{1}{N_{tr}^j} \sum_{i=N_{co}+1}^N \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (31)$$

其中， $t' \in \{post1, post2, \dots, postT\}$  表示处理后的单一时期。与 (30) 式相比，(31) 中多了处理个体的加权平均项  $\frac{1}{N_{tr}^j} \sum_{i=N_{co}+1}^N \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right)$ 。因为 SDID 估计程序就是对每个处理个体合成一个控制个体，然后比较处理个体与合成个体的结果差异，那么，当样本中有  $N_{tr}$  个处理个体时，就计算所有单一个体差异的加权平均，权重为  $\frac{1}{N_{tr}^j}$ ， $N_{tr}^j$  表示处理类别 j 中处理个体的数量。

那么，相对事件时间的 ATE 为

$$\hat{\tau}_{t'} = \sum_j^J \theta_{t'} \hat{\tau}_{t'}^j \quad (32)$$

其中， $\hat{\tau}_{t'}^j$  在相对事件时间  $t'$  处第 j 个矩阵块的 ATE； $\theta_{t'} = \frac{N_{tr,t'}^j}{\sum_{j=1}^J N_{tr,t'}^j}$  是权重。

在 SDID 的应用方面，主要有两种应用的方式：① 将 SDID 作为主要的识别方法，例如，Campos et al. (2022) 应用 SDID 研究了加入欧盟对挪威经济和产业的影响。Liu et al.(2022) 应用 SDID 研究了中国草原生态系统保护政策的效应。② 将 SDID 作为稳健性检验，Borebly (2022) 利用 SDID 作为双重差分的稳健性检验分析了英国住房补贴对劳动市场的影响。许文立和孙磊 (2023) 利用 SDID 作为双重差分的稳健性检验分析了中国碳排放权交易对能源消费结构的影响。但总体上来讲，对于该估计量的应用仍然非常少。



在应用研究中使用 SDID 需要满足一些要求：第一，也是最重要的，所用数据是强平衡面板，结果变量、处理变量、协变量都不能有任何缺失值，如果存在缺失值，要么把这些样本个体删除，要不补齐；第二，如果存在总是处理的个体，即从样本第一期开始就接受处理，那么，这样的个体应该删除。因为对于总是处理的个体，没有处理前时期作为计算时间权重的样本观测数据。如果实在要保留总是处理的个体，那么，至少要扩充样本时期至更早前；第三，样本中一定要包括从未处理组，如果没有从未处理个体，那么，SDID 无法产生合成控制组；第四，sdid 的推断方法各有各的要求，bootstrap 和 jackknife 推断要求处理个体大于 1，如果只有一个处理个体，那么应该使用 placebo 推断，且安慰剂推断要求控制个体数量超过处理个体的数量；第五，其它一些限制条件，见 Arkhangelsky et al. (2021)的假设和 Clarke et al. (2023)。

## 五、面板数据因果推断的应用

房地产税改革是我国在高质量发展过程中促进共同富裕的关键性举措。2011 年，重庆和上海两地试行房产税，我国房地产税改革正式登上舞台。2021 年 10 月 16 日出版的第 20 期《求是》杂志发表习近平总书记重要文章《扎实推动共同富裕》，明确指出“要积极稳妥推进房地产税立法和改革，做好试点工作”。同年 10 月 23 日，全国人大常委会授权国务院在部分地区开展房地产税改革试点工作，这意味着未来房地产税将在一些地区试点、推广。

但不得不承认，房地产税改革难点众多，还需从长计议，行稳致远（白彦锋和张丹昱，2022）。房地产税改革最大的难点就在于政策改革的效应难以预估，这也是学者和政策制定者们最为关心的问题，即房地产税改革的政策效果。过去十年来，国内学者对此进行了大量的评估，大部分文献都集中在重庆和上海两地的房产税试点的效应上。这些文献使用的评估方法主要有两种：第一种是双重差分法，评估房产税的房价效应（Bai et al., 2014；王家庭和曹清峰，2014）、收入分配效应（范子英和刘甲炎，2015）、土地价格效应（杨超等，2021）；第二种是合成控制法，评估房产税改革对房价的影响（刘甲炎和范子英，2013）、产业转移的影响（刘友金和曾小明，2018）、企业投资的影响（徐鹏杰等，2020）。不过因为样本和研究方法的不同，上述文献对房产税改革产生的经济效应方向和程度也存在较大差异（杨超等，2021）。

正如前文所述，无论是双重差分法，还是合成控制法都有严格的假设和适用条件，例如，双重差分法对平行趋势假设的严重依赖，而重庆与上海两地房产税试点可能使得平行趋势更加不可信，而重庆和上海分别作为人口最多和经济最强的直辖市又可能不满足合成控制法的凸集假设。样本数据与识别方法之间的不匹配可能造成实证结果不精确，甚至得到相反的结论。

为此，我们使用 Arkhangelsky et al. (2021) 提出的合成双重差分估计量和上文扩展的动态合成双重差分估计量来评估重庆和上海房产税试点的经济效应。数据来源于刘友金和曾

小明（2018）<sup>2</sup>，作者们研究了房产税对产业转移的效应，主要评估方法是合成控制法，也使用了双重差分法作为稳健性检验。之所以使用他们的数据是因为该数据集包含产业产值和产业就业两个大方向的经济数据<sup>3</sup>，可以比较综合地反映经济状况，而不仅仅只是关注于经济的某一方面（例如，房价、地价等），进而综合评估房产税的经济效应。房产税的开征，理论上主要通过降低企业的生产成本和改变企业的投资取向来影响产业转移（刘友金和曾小明，2018）。

### （一）静态结果

表 2 的面板（Panel）A-C 呈现了房产税试点对经济的总效应、重庆和上海分地区效应的 SDID 估计量，面板（Panel）D 呈现的是复制刘友金和曾小明（2018）得出的上海效应的双重差分结果<sup>4</sup>。

表 2 房产税改革的经济效应

	工业相对产值	工业相对就业率	服务业相对产值	服务业相对就业率
Panel A: 总的效应				
	-0.17**	0.81***	0.02	1.27***
	(0.075)	(0.20)	(0.04)	(0.15)
Panel B: 重庆的效应				
	0.32***	1.82***	0.32***	2.11***
	(0.12)	(0.26)	(0.07)	(0.24)
Panel C: 上海的效应				
	-0.72***	0.14	-0.28**	0.43
	(0.11)	(0.27)	(0.07)	(0.29)
Panel D: 上海的效应--刘友金和曾小明（2018）的双重差分结果				
	-1.40***	-0.40**	-0.52***	0.36*
	(0.11)	(0.18)	(0.05)	(0.19)

注：括号显示了标准误，是安慰剂方差估计方法得到；\*\*\*，\*\*，\*分别表示 1%、5%和 10%置信水平。

<sup>2</sup> 原始数据从中国知网下载，

[https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2018&filename=GGYY201811007&uniplatform=NZKPT&v=ZYH4GJEAQjMeqfETIFnCquYqfV6RPIJO\\_vrkAlxSXXoKx3FzS98dyTkU9215pnUN](https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2018&filename=GGYY201811007&uniplatform=NZKPT&v=ZYH4GJEAQjMeqfETIFnCquYqfV6RPIJO_vrkAlxSXXoKx3FzS98dyTkU9215pnUN)。

<sup>3</sup> 样本数据和指标的详细解释与描述，请参见刘友金和曾小明（2018）的原文。

<sup>4</sup> 需要注意的是，刘友金和曾小明（2018）提供的双重差分估计程序（do 文件）中只包含了时间固定效应，并没有包含个体固定效应。因此，本文使用高维固定效应命令 reghdfe 来复制没有协变量的双重差分结果。我们的复制结果与原作者并不完全一致，暂不清楚不一致的原因。且我们还增加了上海房产税试点对工业相对就业率、服务业相对产值、服务业相对就业率的双重差分结果。

表 2 的结果显示，重庆和上海试点的房产税对两地工业相对产值的平均效应为-0.17，且在 95%的置信水平上显著，对两地的工业和服务业相对就业率的平均效应均在 99%置信水平上显著为正，效应大小分别为 0.81 和 1.27，而没有统计证据显示对服务业相对产值有影响。之所以产生这样的结果，是因为（1）重庆的房产税试点对工业相对产值的影响显著为正（0.32），而上海的房产税试点对工业相对产值的影响显著为负（-0.72），且房产税对上海工业相对产值的负效应超过了重庆，所以从平均意义上来看，房产税显著抑制了工业相对产值；（2）房产税试点对重庆工业和服务业相对就业率的影响都显著为正，而对上海两个产业的相对就业率影响虽然也为正，但在统计意义上不显著，因此，两个产业相对就业率的平均效应均低于重庆；（3）服务业相对产值的平均效应不显著主要是由于房产税对重庆和上海两地的服务业相对产值的影响程度较为接近，但影响方向正好相反。

从上述结果可知，房产税对重庆的工业相对产值有“促增作用”，这与刘友金和曾小明（2018）的结论一致。但刘友金和曾小明（2018）在评估房产税对上海的工业相对产值影响时，由于合成控制法的限制，没有找到上海合适的合成控制组，因此，没有发现明确的上海工业相对产值下降证据，而本文的 SDID 估计量则给出了明确的结果，上海确实由于房产税试点导致了工业相对产值的下降。采用刘友金和曾小明（2018）对上海房产税试点效应评估的补充做法，结果显示，房产税试点对上海经济的双重差分估计量与 SDID 估计量存在明显差异。刘友金和曾小明（2018）指出上海产业相对产值与相对就业率出现背离现象，并对此做出了一定的解释。根据表 1 的结果，更可能是房产税效应识别方法本身的问题，或者说合成控制法和双重差分法并不适合识别上海房产税试点的经济效应，因为在只有上海一个试点地区的情形下使用 DID 估计的效应并不可信，例如上海工业相对产值和相对就业率的 DID 估计量显示它们并没有出现背离现象，而数据显示上海服务业相对产值虽然下降，但相对就业率却在波动。更确切的说，没有证据显示，房产税试点会影响上海工业和服务业的相对就业率，如面板 C 的 SDID 估计量所示。

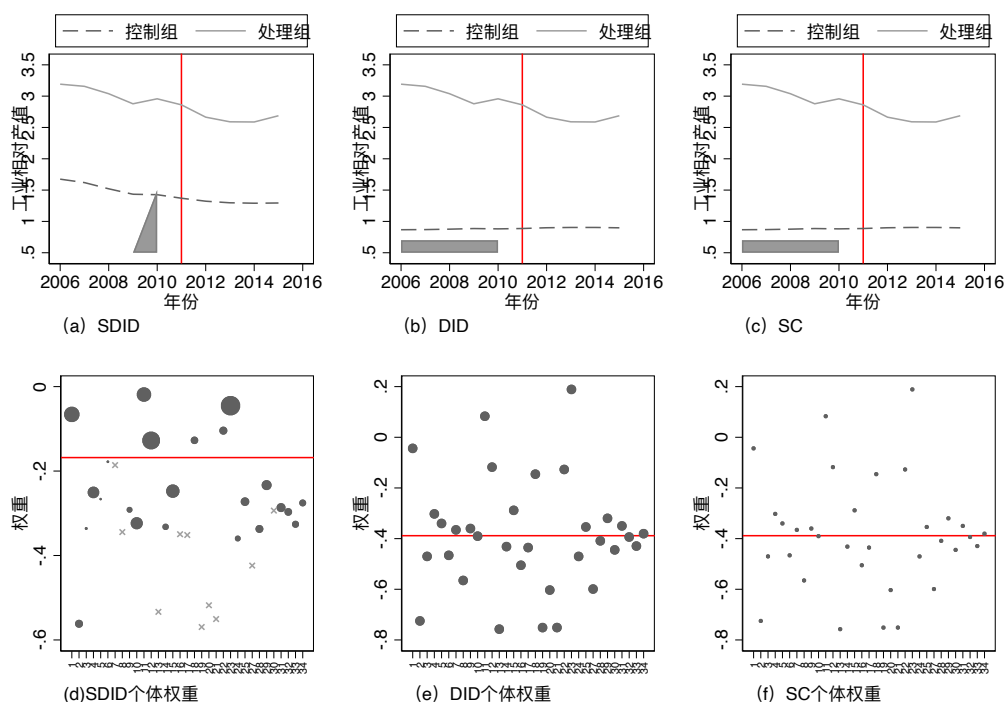


图 2 房产税效应的 DID、SC 和 SDID 估计量比较

房产税试点经济效应 SDID 估计量意味着房产税对工业的平均效应比服务业要显著，且对两次产业的就业的平均效应都有显著的促进作用。但房产税的经济效应存在明显的地区异质性：房产税对重庆的产值和就业都有显著的促进作用，对上海的产值有显著的抑制作用，但没有证据显示对就业有影响。此外，SDID 可以弥补合成控制法和 DID 估计量的不足。

## (二) 动态结果

下面，我们来考察房产税试点的动态经济效应，动态效应的评估能够展示更多政策效果的实施细节和变化趋势。我们利用上文扩展的动态 SDID 估计量识别房产税试点后每一期的经济效应。两个地区的动态平均效应如图 3 所示。它展示了房产税试点后每一年(2011-2015)的动态 SDID 估计量及其对应的置信区间。对于工业相对产值来说，房产税的负效应随着时间的推移先逐渐变大，后向“0”刻度线趋近，呈现“U”型变化趋势，区别于首尾两期，其余时期均在 95%的置信水平下显著，而工业相对就业率除第一期不显著外，其它时期均在 95%置信水平上显著为正。房产税对服务业相对产值在初期显著为负，但随后几年效应出现反转，且没有显著的影响，而对于服务业相对就业率来说，房产税的效应一直显著为正。开征房产税对地区产业结构的影响随时间变化有高级化趋势，具体表现为服务业相对产值的趋于升高（虽然不显著）和工业相对产值的总体下降，以及相对就业率的显著提高。

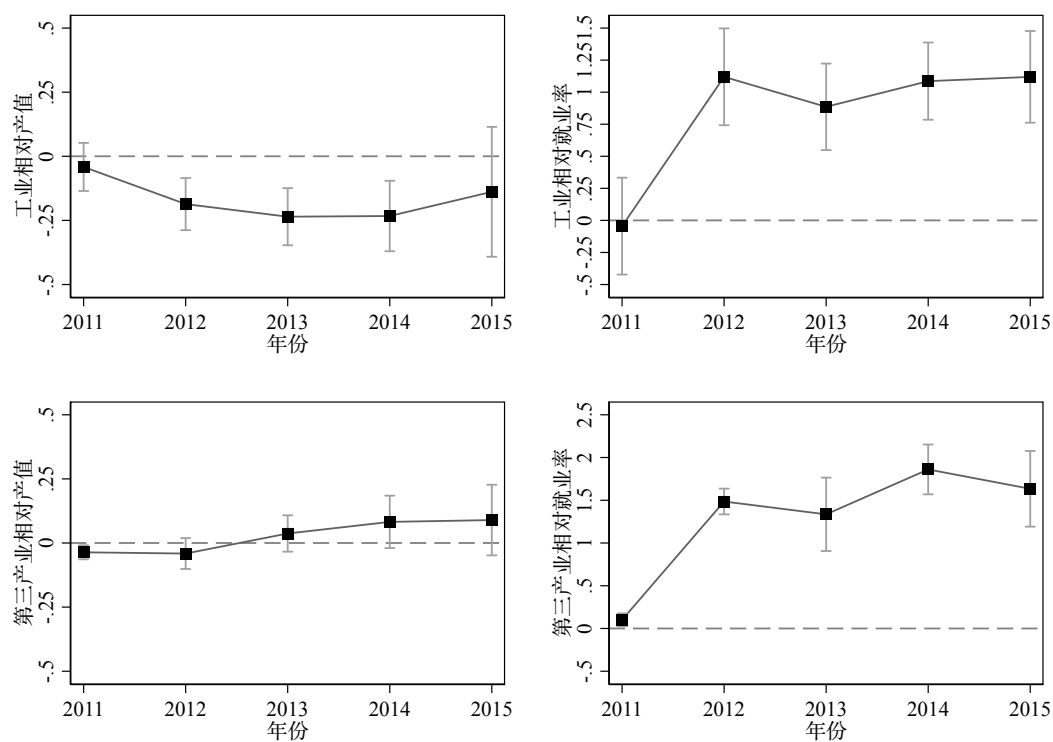


图 3 房产税试点对经济的时间异质性效应

房产税对重庆经济的动态影响如图 4 所示。2011 年试点房产税后，2011-2012 年的工业相对产值效应并不显著，而是在试点后第三年才逐渐表现出显著的促进作用，且效应逐渐增大。房产税对重庆服务业相对产值的促进作用从 2011 年开始就一直保持显著，且在第三年后保持稳定。从重庆就业的效应来看，房产税的实施除在第一年不显著外，2012-2015 年的就业效应均显著为正。

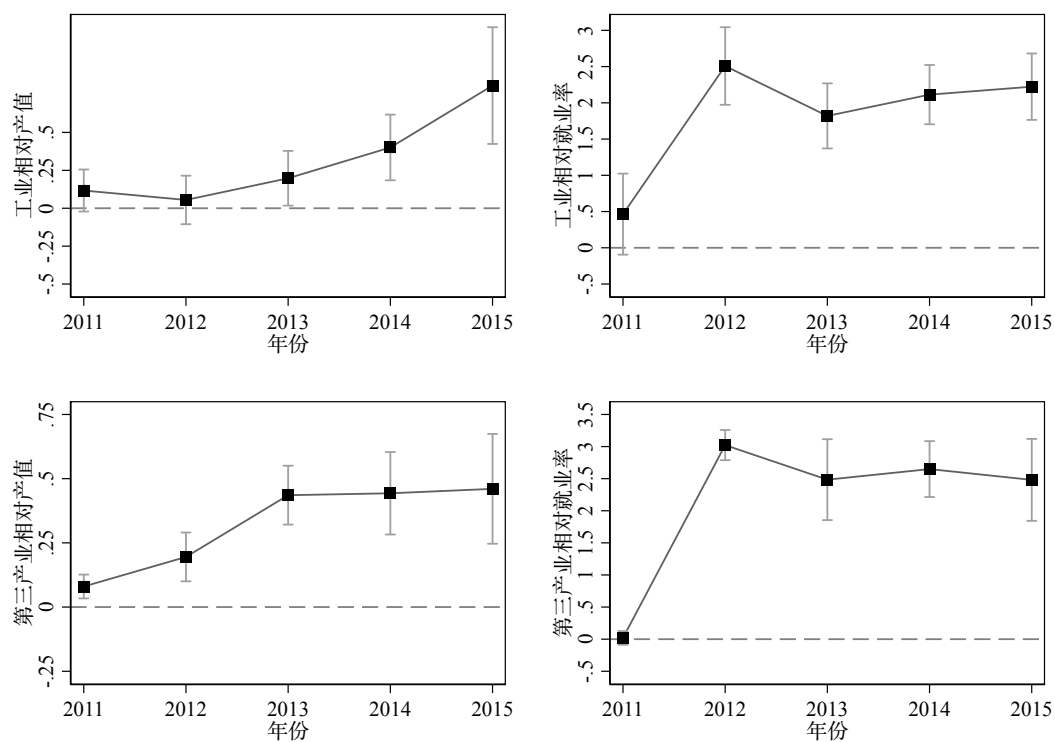


图 4 重庆房产税试点对经济的时间异质性效应

房产税对上海经济的影响如图 5 所示。房产税对上海工业相对产值的负效应从 2011-2015 年逐渐增大，且均在 95% 的置信水平下保持显著，而对服务业相对产值的负效应也一直保持显著，但与工业相对产值的变化模式不同，它呈现出“V”型变化趋势。而房产税对上海工业相对就业率的动态影响虽然为正，但所有时期的效应在 95% 置信水平下并不显著，对上海服务业相对就业率的动态效应在大部分时期也显著为正。

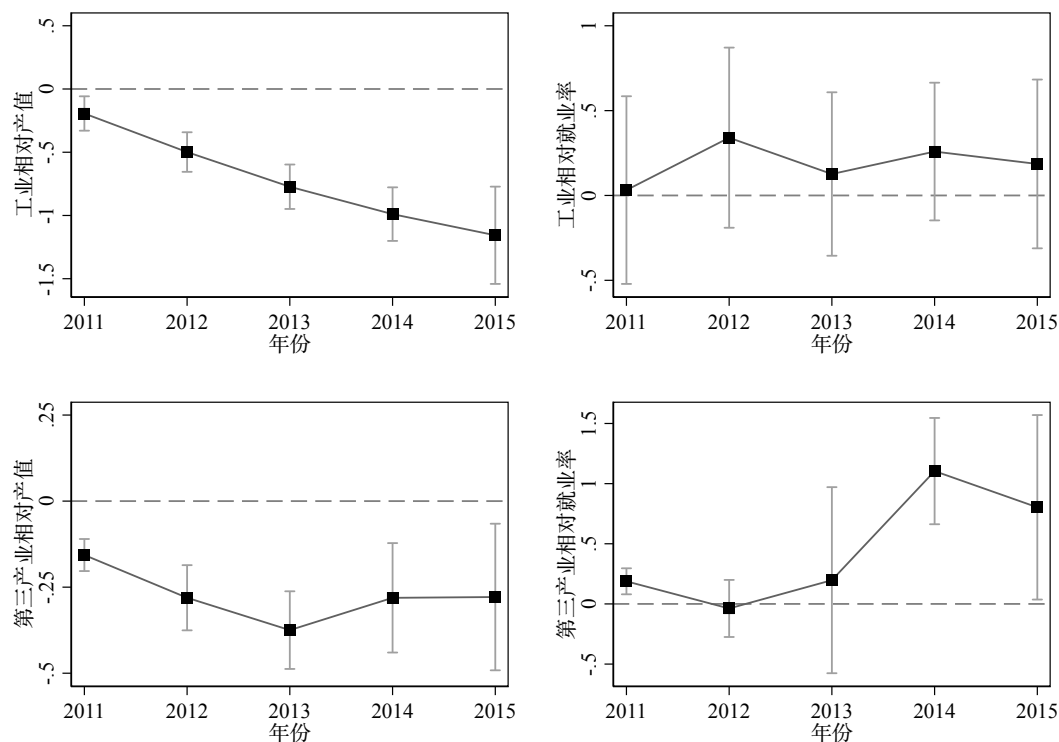


图 5 上海房产税试点对经济的时间异质性效应

本文的结论在一定程度上丰富了刘友金和曾小明 (2018) 对房产税开征政策评估的研究成果，特别是在动态研究方面。一方面，我们发现房产税的开征总体上加速了工业的外移和相对就业率的提升，但是对服务业却没有显著影响，不过，房产税政策的实施对地区产业发展“腾笼换鸟”的作用随着时间变化作用趋势逐渐明显。另一方面，从区域异质性来看，与刘友金和曾小明 (2018) 的结论相似，重庆作为房价水平较低、处于工业化中期的典型区域，房产税的开征不仅加速劳动力的集聚，工业、服务业相对产值也显著提高，承接产业转移迹象明显，对当地产业规模增长和产业结构优化具有积极作用；而上海经济发达、并已处于工业化后期，较低强度房产税政策对当地已然处于高位的房价上涨几乎没有起到抑制作用（杨励和欧嘉丽，2020），对域外工业企业和服务业企业进入成本的降低“杯水车薪”，并没有转变产业既定向外转移的走势，同时具有“限制外来人口”导向的排外性房产税政策让很多劳动力认为这座城市“难以企及”、进而“望而生畏”，在短期难以拉升当地的相对就业率，但是，在 2014 年前后对上海服务业相对就业率有显著提升作用，且远大于服务业相对产值上升比例，这种现象说明房产税政策在一段时间后对服务业劳动效率产生了不利影响，上海的产业转变更多依赖于低端服务业的发展，这与杨励和欧嘉丽（2020）的研究结论相契合。

## 六、总结性评论

本文系统归纳总结了面板数据因果推断主要方法的基本思想、最新进展和实际应用事项。一直以来,大量的国内社会科学,尤其是自大规模微观调查数据可用后,经济学和管理学领域的经验研究使用面板数据方法来进行因果效推断和政策评估。近年来,面板数据领域最常用的三类方法——双向固定效应、双重差分和合成控制法——已经取得了极大的进展。最新的理论文献表明,在交叠处理下,异质性处理效应会使得双向固定效应双重差分估计量产生负权重问题,进而产生偏误,许多学者已经提出了异质性处理效应稳健估计量。不仅如此,异质性处理效应还会使得连续型 DID 和多处理变量 DID 估计量也产生偏误,应用新的稳健估计量可能矫正这些偏误。传统文献均忽略了因子模型结构在合成控制法中的作用,以及对合成控制法和双重差分法之间的联结作用,且合成控制法的推断框架也开始转向最前沿的“共形推断”。正因为 DID 和合成控制法存在这些联结,并取得了许多新的进展,所以结合二者优势的新因果效应评估框架成为最前沿的发展趋势,即双重差分法和合成控制法的统一框架。

本文试图较为详细地介绍面板数据方法的最新进展和实现方法,期望能为中国社会科学领域的应用研究者们更好地使用面板数据评估政策因果效应提供一些有价值的参考。建议经验研究者可以根据数据结构(面板数据、重复截面数据等)、处理配置(统一处理时点、交叠处理等)、截面与时间层面的相对规模(长面板、短面板),平行趋势是否事前可信等几个方面来选择最适合的因果效应识别方法。本文的分析和建议是基于笔者对当前的前沿理论计量和实证研究文献的理解和自身研究经验归纳总结而来,作为一家之言必定存在谬误和不足。期待本文的“抛砖引玉”能引发更广泛地讨论,以促成经济学、管理学、政治学、历史学、法学、社会学等中国社会科学各领域的共同进步。

## 参考文献

1. de Chaisemartin C, D'haultfœuille X. Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey[J]. *Econometrics Journal*, 2022.
2. Roth J, Sant'Anna P H C, Bilinski A, et al. What's trending in difference-in-differences? A synthesis of the recent econometrics literature[J]. *Journal of Econometrics*, 2023.
3. 刘冲,沙学康,张妍.交错双重差分:处理效应异质性与估计方法选择[J].数量经济技术经济研究,2022,39(09):177-204.
4. 张子尧,黄炜.事件研究法的实现、问题和拓展[J].数量经济技术经济研究,2023,40(09):71-92.
5. 许文立.双重差分法的最新理论进展与经验研究新趋势[J].广东社会科学



学,2023,(05):51-62.

6. Arkhangelsky D, Imbens G. Causal Models for Longitudinal and Panel Data: A Survey[J]. arxiv preprint arxiv:2311.15458, 2023.
7. Athey S, Imbens G W. The state of applied econometrics: Causality and policy evaluation[J]. Journal of Economic perspectives, 2017, 31(2): 3-32.
8. Abadie A. Using synthetic controls: Feasibility, data requirements, and methodological aspects[J]. Journal of Economic Literature, 2021, 59(2): 391-425.
9. Athey S, Bayati M, Doudchenko N, et al. Matrix completion methods for causal panel data models[J]. Journal of the American Statistical Association, 2021, 116(536): 1716-1730.
10. Xu Y. Generalized synthetic control method: Causal inference with interactive fixed effects models[J]. Political Analysis, 2017, 25(1): 57-76.
11. Arkhangelsky D, Athey S, Hirshberg D A, et al. Synthetic difference-in-differences[J]. American Economic Review, 2021, 111(12): 4088-4118.
12. Huang W, Liu H. Early childhood exposure to health insurance and adolescent outcomes: Evidence from rural China[J]. Journal of Development Economics, 2023, 160: 102925.
13. Nunn N, Qian N. The potato's contribution to population and urbanization: evidence from a historical experiment[J]. The quarterly journal of economics, 2011, 126(2): 593-650.
14. Chen Y, Fan Z, Gu X, et al. Arrival of young talent: The send-down movement and rural education in China[J]. American Economic Review, 2020, 110(11): 3393-3430.
15. Angrist J D, Pischke J S. Mostly harmless econometrics: An empiricist's companion[M]. Princeton university press, 2009.
16. Roth J. Pretest with caution: Event-study estimates after testing for parallel trends[J]. American Economic Review: Insights, 2022, 4(3): 305-322.
17. Goodman-Bacon A. Difference-in-differences with variation in treatment timing[J]. Journal of Econometrics, 2021, 225(2): 254-277.
18. de Chaisemartin C, d'Haultfoeuille X, Pasquier F, et al. Difference-in-differences estimators for treatments continuously distributed at every period[J]. arxiv preprint arxiv:2201.06898, 2022.
19. D'Haultfoeuille X, Hoderlein S, Sasaki Y. Nonparametric difference-in-differences in repeated cross-sections with continuous treatments[J]. Journal of Econometrics, 2023, 234(2): 664-690.
20. Callaway B, Goodman-Bacon A, Sant'Anna P H C. Difference-in-differences with a continuous treatment[J]. arxiv preprint arxiv:2107.02637, 2021.
21. De Chaisemartin C, d'Haultfoeuille X. Fuzzy differences-in-differences[J]. The Review of

- Economic Studies, 2018, 85(2): 999-1028.
22. de Chaisemartin C, D'Haultfœuille X, Guyonvarch Y. Fuzzy differences-in-differences with Stata[J]. The Stata Journal, 2019, 19(2): 435-458.
  23. Cerulli G. ctreatreg: Command for fitting dose–response models under exogenous and endogenous treatment[J]. The Stata Journal, 2015, 15(4): 1019-1045.
  24. Wooldridge J M. Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators[J]. Available at SSRN 3906345, 2021.
  25. Hegland, Thomas A. wooldid: Estimation of Difference-in-Differences Treatment Effects with Staggered Treatment Onset Using Heterogeneity-Robust Two-Way Fixed Effects Regressions. Statistical Software Components s459238,2023.
  26. de Chaisemartin C, D'haultfœuille X. Two-way fixed effects and differences-in-differences estimators with several treatments[J]. Journal of Econometrics, 2023, 236(2): 105480.
  27. 赵仁杰,赵欣仪,钟世虎等.征管模式、征管技术与税收治理——来自取消税管员固定管户的证据[J].经济研究,2023,58(09):133-151.
  28. Rico-Straffon J, Wang Z, Panlasigui S, et al. Forest concessions and eco-certifications in the Peruvian Amazon: Deforestation impacts of logging rights and logging restrictions[J]. Journal of Environmental Economics and Management, 2023, 118: 102780.
  29. Bai Jushan. Panel data models with interactive fixed effects. Econometrica, 77(4):1229–1279, 2009.
  30. Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california’s tobacco control program. Journal of the American statistical Association, 105(490):493–505, 2010.
  31. Liu M, Xu W, Zhang H, et al. Livestock production, greenhouse gas emissions, air pollution, and grassland conservation: Quasi-natural experimental evidence[J]. Greenhouse Gas Emissions, Air Pollution, and Grassland Conservation: Quasi-Natural Experimental Evidence (December 19, 2022), 2022.
  32. Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Comparative politics and the synthetic control method. American Journal of Political Science, pages 495–510, 2015.
  33. Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. American Economic Review, 93(-):113–132, 2003.
  34. Doudchenko N, Imbens G W. Balancing, regression, difference-in-differences and synthetic control methods: A synthesis[R]. National Bureau of Economic Research, 2016.
  35. Susan Athey, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. Matrix completion methods for causal panel data models. Journal of the

- American Statistical Association, 116(536):1716–1730, 2021.
36. Jann Spiess, Guido Imbens, and Amar Venugopal. Double and single descent in causal inference with an application to high-dimensional synthetic control. arXiv preprint arXiv:2305.00700, 2023.
  37. Kuosmanen T, Zhou X, Eskelinen J, et al. Design flaw of the synthetic control method[J]. working paper, 2021.
  38. Maxwell Kellogg, Magne Mogstad, Guillaume A Pouliot, and Alexander Torgovitsky. Combining matching and synthetic control to tradeoff biases from extrapolation and interpolation. Journal of the American statistical association, 116(536):1804–1816, 2021.
  39. Abadie A. Using synthetic controls: Feasibility, data requirements, and methodological aspects[J]. Journal of Economic Literature, 2021, 59(2): 391-425.
  40. Cattaneo M D, Feng Y, Palomba F, et al. scpi: Uncertainty Quantification for Synthetic Control Methods[J]. arxiv preprint arxiv:2202.05984, 2022a.
  41. Chernozhukov V, Wüthrich K, Zhu Y. An exact and robust conformal inference method for counterfactual and synthetic controls[J]. Journal of the American Statistical Association, 2021, 116(536): 1849-1864.
  42. Cattaneo M D, Feng Y, Palomba F, et al. Uncertainty Quantification in Synthetic Controls with Staggered Treatment Adoption[J]. arxiv preprint arxiv:2202.05984, 2022b.
  43. Cattaneo, Feng and Titiunik (2021): Prediction Intervals for Synthetic Control Methods. Journal of the American Statistical Association 116(536): 1865-1880.
  44. Clarke D, Pailańir D, Athey S, et al. Synthetic difference in differences estimation[J]. arxiv preprint arxiv:2301.11859, 2023.
  45. Porreca Z. Synthetic difference-in-differences estimation with staggered treatment timing[J]. Economics Letters, 2022, 220: 110874.