

# 动态合成双重差分估计量：以房产税试点的经济效应为例

许文立 孙 磊

xuweny87@hotmail.com

(安徽大学经济学院, 合肥, 230601)

初稿：2022-5-6

最新稿：2022-12-7

**摘 要：**因果推断已经成为经济学经验研究的主要目的，其中，合成控制法和双重差分法是最流行的两类因果推断方法，它们各具优势，也有不足。Arkhangelsky et al. (2021, AER) 提出了一种新的因果推断估计量——合成双重差分估计量，兼具了合成控制法和双重差分法的优势。本文将合成双重差分估计量扩展到动态环境，得到处理后每一期的处理效应，类似于事件研究。为了说明动态合成双重差分估计量的用处和重要性，本文基于刘友金和曾小明 (2018, 中国工业经济, 11:98-116) 的数据来重新分析我国房产税试点对产业转移的效应。结果显示：(1) 动态合成双重差分估计量与合成控制和双重差分估计量存在明显差异，且对两个传统估计量的效应识别起到了补充作用；(2) 进一步完善原文结论，房产税的开征总体上加快了地区产业结构“腾笼换鸟”的动态趋势，特别是对重庆的工业、服务业的相对产值和相对就业率均有“促增”作用，但是抑制了上海的产业结构升级，产业结构的调整更多依赖于低端服务业的发展。最后，给出了一些实践应用的建议。

**关键词：**动态合成双重差分估计量；房产税；因果推断

## 一、引言与文献回顾

应用经济学研究者和政策制定者通常最关心政策干预的因果效应，但是回答因果效应并不容易，因为我们很难知晓政策不干预时会产生什么样的结果 (The Royal Swedish Academy of Sciences, 2021)。2021 年诺贝尔经济学奖三位得主 Card、Angrist、Imbens 和著名经济学家 Krueger 在 20 世纪 90 年代提出利用随机试验的方式来评估政策效应，即将政策随机地配置给个体——一部分个体受到政策影响，另一部分个体不受政策影响，然后来比较两个

组群结果的差异,得到政策效应。但是,政策干预的配置通常并不随机 (Arkhangelsky et al., 2021),甚至给定可观测协变量的条件下,无混淆因子假设也可能并不可信 (Imbens and Rubin, 2015)。因此,在缺少随机试验的情况下,经济学研究者们利用统计模型来模拟(预测)反事实结果。

在这些模型中,双重差分 (Difference in Difference, 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。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)、许文立等 (2022)。<sup>①</sup>

除此之外,过去十几年来,因果推断方法中最大的进展是合成控制法 (Synthetic Control, SC) (Athey and Imbens, 2017)。Abadie and Gardeazabal (2003)、Abadie et al. (2010, 2015) 和 Abadie and L'Hour (2016) 提出和发展的合成控制法已经成为案例研究领域最重要的定量分析工具。合成控制法利用数据来获得计算出一组最优个体权重,并将这些权重应用于处理前时期潜在控制组,来为少量处理个体合成一个最优的控制组。在国内外,也有非常多的研究使用合成控制法来评估政策效应,例如, Cunningham (2020)、Huntington-Klein (2021)、刘甲炎和范子英 (2013)、刘友金和曾小明 (2018)。<sup>②</sup>

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,将个体加权和时间加权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

---

<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. Callaway, B. (2022). Difference-in-Differences for Policy Evaluation. Forthcoming at Handbook of Labor, Human Resources and Population Economics.

<sup>②</sup> 关于 SC 最新进展的文献回顾,可以参见 Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature, 59(2), 391-425.

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。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单一估计量的精度。Porreca (2022) 则进一步将该估计量推广到交叠处理的情形, 从而获得单一的平均处理效应。

在 SDID 的应用方面, 据我们所知, 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。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。但在经济学研究领域, 对于该估计量的应用总体较少。在国内, SDID 估计量的应用更是处于空白。

基于此, 本文考虑在政策评估过程中, 处理后处理效应可能存在时间异质性问题, 而上述 SDID 估计量均只给出了单一的平均处理效应估计量的正式结果, 因此, 将 Arkhangelsky et al. (2021) 的 SDID 估计量推广到动态处理效应, 即处理后每一期都估计一个处理效应, 并给出正式的估计量<sup>①</sup>, 这也是本文的第一个贡献。为了说明动态合成双重差分估计量的重要性和实践应用, 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, 这也是本文的第二个贡献。此外, 在应用动态 SDID 估计量来分析我国房产税试点的产业转移效应时, 进一步补充和完善了刘友金和曾小明 (2018) 的结论, 这也丰富了我国房产税效应的研究, 并为我国未来房产税的开征提供了一些新的政策启示, 这可能是本文的第三个贡献。

本文的结构如下: 第二部分回顾 Arkhangelsky et al. (2021) 的合成双重差分估计量的单一平均处理效应估计量; 第三部分将单一 SDID 估计量扩展为动态估计量, 并给出正式的处理后每一期处理效应估计量; 为了说明 SDID 估计量和动态 SDID 估计量的应用, 第四部分重新评估我国房产税试点的产业转移效应; 第五部分总结, 并给出一些应用建议。

## 二、合成双重差分估计量

我们在这里简要回顾一下 SDID 的估计量和方差, 更多的技术细节可以详见 Arkhangelsky et al. (2021) 的正文和附录。考察  $N$  个个体和  $T$  期的平衡面板数据,  $Y_{it}$  表示第  $i$  个个体  $t$  期的结果变量,  $W_{it}$  表示二值型虚拟变量。假设有  $N_{co}$  个个体从未接受处理, 接受处理的个体数  $N_{tr} = N - N_{co}$ , 处理前时期数  $T_{pre}$ 。我们先呈现出 DID 估计量和 SC 估计量。DID 估计量是依靠解向双固定效应回归得到:

---

<sup>①</sup> Matheus Facure Alves (2022) 提出了应对时间异质性处理效应的一种非正式估计过程。Campos et al. (2022) 也在研究中估计了处理后每一期的处理效应。但我们有三点需要说明, 第一, 本文在写作的时候, 这两个研究还未公开发表, 因此, 我们还未看到他们的结果; 第二, 他们并没有给出处理后每一期处理效应的正式估计量; 第三, 前一个文献使用的 python 程序得到处理后每个时期的处理效应估计量, 后一个文献则使用 R 语言, 而本文则使用经济学领域最常用的 stata 程序。在编写 stata 程序期间, 我们得到了 stata 包 sdid 的开发者, 智利大学 Daniel Pailařir 的极大帮助, 在此一并感谢。

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

其中， $\mu$ 、 $\alpha_i$ 、 $\beta_t$ 分别表示常数项、个体固定效应和时间固定效应； $\tau$ 表示处理效应。这些变量加上“帽子”表示估计值。

与 DID 估计量相比，SC 估计量忽略了个体固定效应，且增加了个体权重 $\hat{\omega}_i^{sc}$ ：

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

而 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 $\hat{\lambda}_t^{sdid}$ ：

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg \min_{\tau, \mu, \alpha, \beta} \left\{ \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \quad (3)$$

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：

$$Y = L + W_{it}\tau_{it} + E \quad (4)$$

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

$$\tau = \frac{1}{N_{tr}T_{post}} \sum_{i=N_{co}+1}^N \sum_{t=T_{pre}+1}^T \tau_{it} \quad (5)$$

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

$$\hat{\tau} = \left[ \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left( \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^T Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (6)$$

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 $\varepsilon_i$ 表示均值为 0，方差为 $\Sigma$ 的独立高斯向量，方差矩阵 $\Sigma$ 满足下列关系：

$$\Sigma = \begin{pmatrix} \Sigma_{pre,pre} & \Sigma_{pre,post} \hat{\lambda}_{post} \\ \hat{\lambda}_{post} \Sigma_{post,pre} & \hat{\lambda}_{post} \Sigma_{post,post} \hat{\lambda}_{post} \end{pmatrix}$$

### 三、动态合成双重差分估计量

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。近年来，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，本文也考虑用 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。

#### （一）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：

$$\hat{\tau}_{t'} = \left[ \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (6)$$

其中， $t'$ 表示处理后的时期。与（5）式相比，在（6）式的处理后每期效应估计量中，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

$$\hat{\tau}_{t'} = \left[ \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^N \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} \left( Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it} \right) \right] \quad (7)$$

其中， $t' \in \{post1, post2, \dots, postT\}$ 表示处理后的单一时期。与（6）式相比，（7）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 $\frac{1}{N_{tr}}$ 。

## （二）Staggered Treatment Timing

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：

2×2 的分块

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

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

$$\hat{\tau}_{t'}^j = \left[ \frac{1}{N_{tr}^j} \sum_{i=N_{co}+1}^N (Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it}) \right] - \left[ \sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} (Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it}) \right] \quad (7)$$

其中,  $t' \in \{post1, post2, \dots, postT\}$  表示处理后的单一时期。与 (6) 式相比, (7) 中多了处理个体的加权平均项  $\frac{1}{N_{tr}^j} \sum_{i=N_{co}+1}^N (Y_{it} - \sum_1^{T_{pre}} \hat{\lambda}_t^{sdid} Y_{it})$ 。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

$$\hat{\tau}_{t'} = \sum_j \theta_{t'} \hat{\tau}_{t'}^j$$

其中,  $\hat{\tau}_{t'}^j$  表示在相对事件时间  $t'$  处, 第  $j$  个矩阵块的 ATT,  $\theta_{t'} = \frac{N_{tr,t'}^j}{\sum_{j=1}^J N_{tr,t'}^j}$  表示权重。

(三) 推断

参照 Arkhangelsky et al. (2021)，在一定的条件下，上述处理后每一期的估计量的渐进方差 $\Xi$ 仍然存在一致的方差-协方差估计量。传统的置信区间可以表示为：

$$\tau_{t'} \in \hat{\tau}_{t'} \pm z_{\alpha/2} \sqrt{\hat{\Xi}_{\tau_{t'}}} \quad (8)$$

Arkhangelsky et al. (2021) 提供了三种方差-协方差估计方法：(1) Bootstrap 方差估计；(2) Jackknife 方差估计；(3) 安慰剂 (Placebo) 方差估计。

(1) Bootstrap 方差估计。首先从样本数据  $(\mathbf{Y}, \mathbf{W})$  中重复抽样，获得一个重复抽样数据样本  $(\mathbf{Y}^B, \mathbf{W}^B)$ ，如果重复抽样样本中没有处理组或者没有潜在控制组，就放弃样本，并再次重复抽样，直到获得合适的重复抽样样本。然后，基于重复抽样数据样本  $(\mathbf{Y}^B, \mathbf{W}^B)$  计算每一期的 SDID 估计量  $\hat{\tau}_{t'}^B$ 。最后，计算出每一期的方差估计量：

$$\hat{\Xi}_{\tau_{t'}}^B = \frac{1}{M} \sum_{m=1}^M \left( \hat{\tau}_{t'}^B - \frac{1}{M} \sum_{m=1}^M \hat{\tau}_{t'}^B \right)^2$$

其中，M 表示重复抽样次数。

(2) Jackknife 方差估计。计算  $\hat{\tau}_{t'}^{(-i)}$ ：  $\arg \min_{\hat{\tau}_{t'}, \{\alpha_j, \beta_t\}_{j \neq i, t'}} \sum_{j \neq i, t'} (Y_{jt} - \alpha_j - \beta_t - \tau W_{it})^2 \hat{\omega}_j \hat{\lambda}_t$

每一期的方差估计量为：

$$\hat{\Xi}_{\tau_{t'}}^{Jack} = \frac{N-1}{N} \sum_{i=1}^N (\hat{\tau}_{t'}^{(-i)} - \hat{\tau}_{t'})^2$$

(3) 安慰剂 (Placebo) 方差估计。安慰剂方差估计过程就是在潜在控制组中随机抽取 M 安慰剂处理组，构造安慰剂处理变量矩阵  $\mathbf{W}^P$ ，基于数据样本  $(\mathbf{Y}^P, \mathbf{W}^P)$  计算每一期的 SDID 估计量  $\hat{\tau}_{t'}^P$ 。最后，计算出每一期的方差估计量：

$$\hat{\Xi}_{\tau_{t'}}^P = \frac{1}{M} \sum_{m=1}^M \left( \hat{\tau}_{t'}^P - \frac{1}{M} \sum_{m=1}^M \hat{\tau}_{t'}^P \right)^2$$

## 四、房产税对产业转移的效应

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>①</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>②</sup>，可以比较综合地反映经济状况，而不仅仅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 1 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>③</sup>。

表 1 房产税改革的经济效应

<sup>①</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=ZYH4GJEAQjMeqfETIFnQcuYqfV6RPIJO\\_vrkAlxSXXoKx3FzS98dyTkU9215pnUN](https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2018&filename=GGYY201811007&uniplatform=NZKPT&v=ZYH4GJEAQjMeqfETIFnQcuYqfV6RPIJO_vrkAlxSXXoKx3FzS98dyTkU9215pnUN)。

<sup>②</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>③</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。

	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。

表 1 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。

我们的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 1 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。

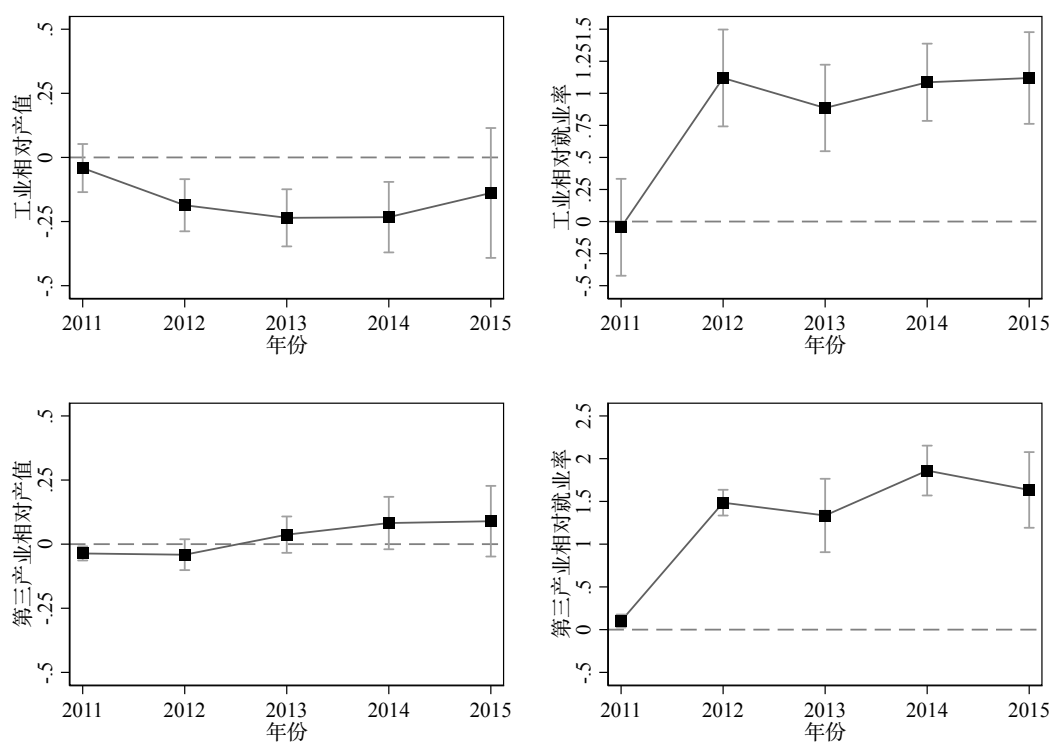


图 1 房产税试点对经济的时间异质性效应

房产税对重庆经济的动态影响如图 2 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。

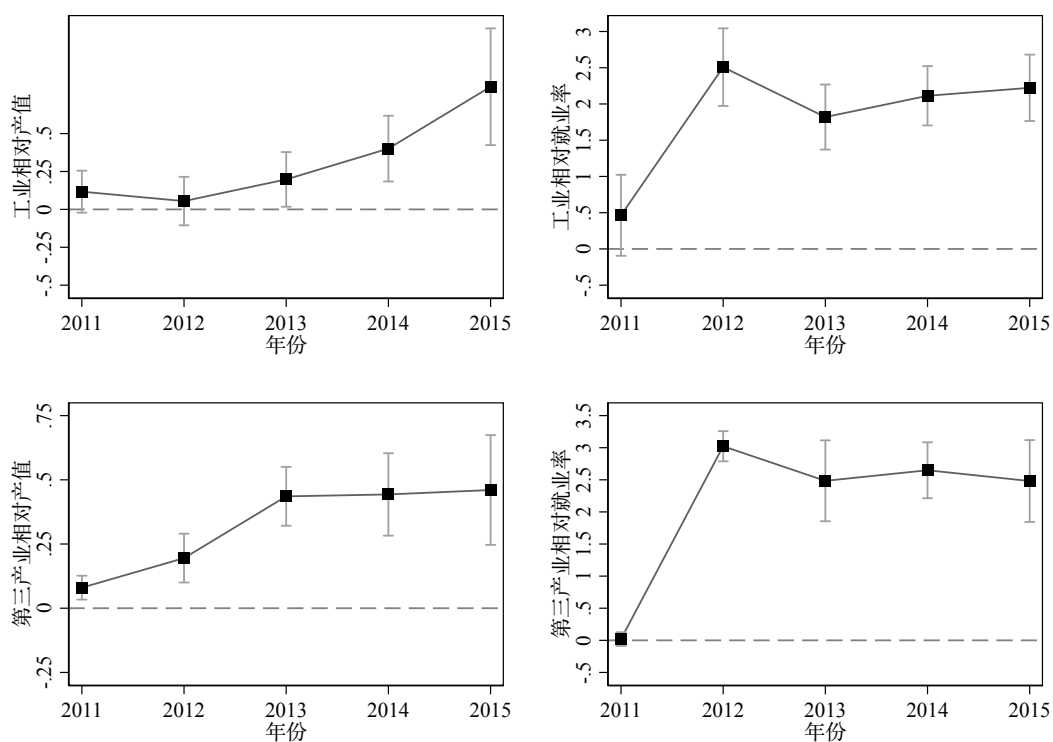


图 2 重庆房产税试点对经济的时间异质性效应

房产税对上海经济的影响如图 3 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。

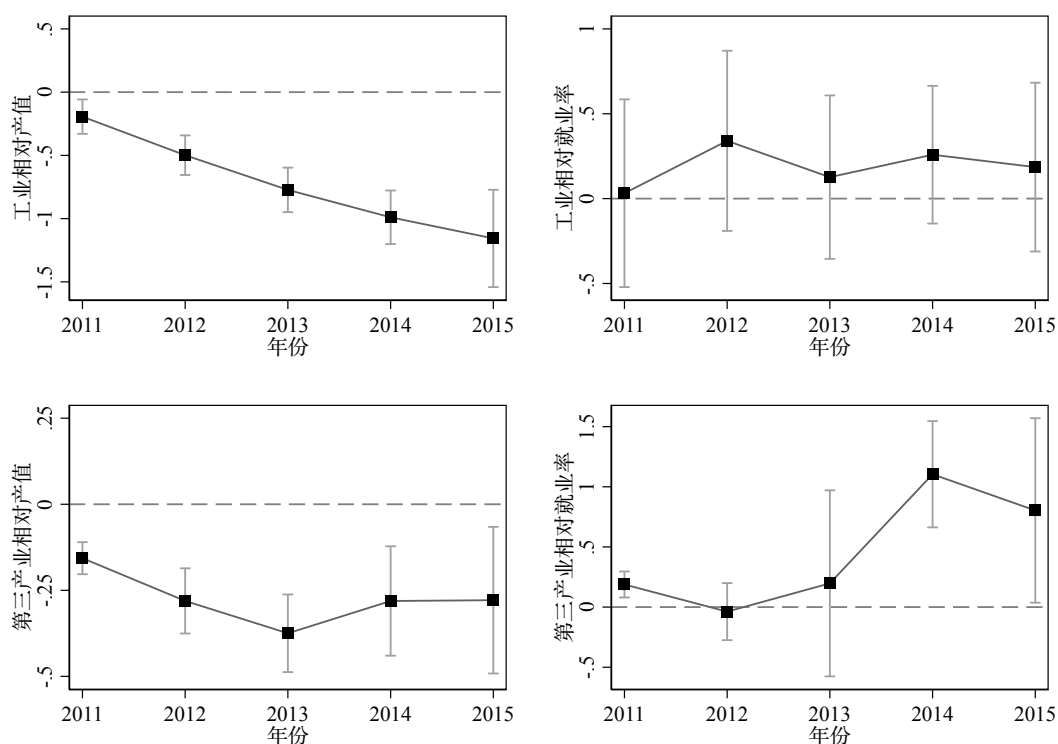


图 3 上海房产税试点对经济的时间异质性效应

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。

## 五、启示与实践建议

合成控制法和双重差分法是应用经济研究领域最重要、最常用的因果推断方法，但其核心识别假设一直颇受质疑，且适用环境受到很大的限制。为了减弱对这些识别假设的依赖，并拓宽政策评估的适用范围，Arkhangelsky et al. (2021) 提出了一种新的估计量——合成双重差分估计量，它综合了合成控制法加权优势和双重差分固定效应的灵活性，从而可以提高因果效应识别的精度 (Porreca, 2022)。本文将合成双重差分估计量扩展成动态合成双重差分估计量——政策干预后每一期的平均处理效应及其置信区间，并给出每一期估计量正式公式和方差公式，同时运用动态合成双重差分估计量来评估 2011 年在重庆和上海试点的房产税的经济效应。本文的研究将为政策效应评估手段的丰富和房产税政策的设计与实施提供有益的启示。研究结果显示：(1) 合成双重差分估计量可以补充合成控制法和双重差分法的不足之处，并为政策效应的识别提供更清晰的统计证据；(2) 房产税政策的设计与实施需要兼顾全国的整体性和区域的差异性，一方面，要加快房产税立法的进程，遏制房价上涨对实体部门的挤出，引导产业合理转移和结构优化，推动建立适应产业协调发展的房产税政策制度；另一方面，要从制度配套上把握房产税在产业间、地区间和时间维度上的差异化特征，对于像重庆房价较低、处于工业化中期的城市，应鼓励其积极承接产业转移，并与现有产业结构相匹配，加速产业结构升级与优化，而对于上海房价较高、处于工业化后期的城市，要着力化解高房价对人力资本的阻拦作用和产业流失对城市经济的消极影响，出台人才补贴和产业扶持政策留住并吸引与城市产业结构相适应的人才和新兴产业，促进城市产业与经济的可持续发展。

因果效应识别的最大困难在于构造反事实结果，合成控制法和双重差分法均可以在一定程度上解决这个难点，但前者依赖于凸性假设，后者依赖于平行趋势假设，这些假设的不可证实使得因果效应估计和推断受到很大的质疑。而合成双重差分法克服了上述两个关键假设，可以有效补充合成控制法和双重差分方法。在因果效应评估的实践中，既可以将合成双重差分估计量作为合成控制法和双重差分法的稳健性检验，也可以单独使用合成双重差分作为效应评估主要方法来使用。

## 参考文献

- [1] The Royal Swedish Academy of Sciences, 2021. Natural experiments help answer important questions. <https://www.nobelprize.org/uploads/2021/10/popular-economicsciencesprize2021-3.pdf>.
- [2] Imbens G W, Rubin D B. Causal inference in statistics, social, and biomedical sciences[M]. Cambridge University Press, 2015.
- [3] Currie J, Henrik Kleven, and Esmee Zwiers. Technology and big data are changing economics: Mining text to track methods. In AEA Papers and Proceedings, 2020,

volume 110, pages 42-48.

- [4] Angrist J D, Pischke J S. Mostly harmless econometrics: An empiricist's companion[M]. Princeton university press, 2009.
- [5] Ashenfelter O, David Card. Using the longitudinal structure of earnings to estimate the effect of training programs. *The Review of Economics and Statistics*, 1985, 67(4):648-660.
- [6] Bertrand M, Esther Duflo, and Sendhil Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 2004, 119(1):249-275.
- [7] Angrist, JD, J-S Pischke (2015), *Mastering Metrics: The Path from Cause to Effect*, Princeton University Press.
- [8] Abadie, A. Diamond and J. Hainmueller. 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the Statistical*.
- [9] D. Arkhangelsky, S. Athey, D. Hirshberg, G. Imbens and S. Wager. 2021. Synthetic Difference in Differences. *American Economic Review*.
- [10] S. Kranz. 2022. Synthetic Difference-in-Differences with Time-Varying Covariates. Working paper.
- [11] Campos, N. F., Coricelli, F., & Franceschi, E. (2022). Institutional integration and productivity growth: Evidence from the 1995 enlargement of the European Union. *European Economic Review*, 142, 1040-14.
- [12] Borebly, 2022. "The impact of housing subsidy cuts on the labour market outcomes of claimants: evidence from England," *Journal of Housing Economics*, Volume 57: 1018-59.
- [13] Porreca Z. Synthetic Difference-In-Differences Estimation With Staggered Treatment Timing. *Economics letter*. 2022, forthcoming.
- [14] 白彦锋,张丹昱.房地产税改革:从长计议,行稳致远[J].*中国经济评论*,2022(03):44-48.
- [15] 范子英,刘甲炎.为买房而储蓄——兼论房产税改革的收入分配效应[J].*管理世界*,2015(05):18-27+187.
- [16] 王家庭,曹清峰.房产税能够降低房价吗——基于 DID 方法对我国房产税试点的评估[J].*当代财经*,2014(05):34-44.
- [17] Bai, ChongEn, Qi Li, and Min Ouyang. "Property taxes and home prices: A tale of two cities." *Journal of Econometrics* 180.1 (2014): 1-15.
- [18] 刘甲炎,范子英.中国房产税试点的效果评估:基于合成控制法的研究[J].*世界经*



济,2013,36(11):117-135.

- [19] 杨超,肖璐,凌爱凡.房产税抑制了土地价格吗?——基于微观住宅用地出让数据的实证研究[J].经济评论,2021,(06):19-33.
- [20] 徐鹏杰,匡萍,王宁.房产税对企业投资行为的影响——来自上海和重庆的经验分析[J].经济体制改革,2020(02):136-143.
- [21] 刘友金,曾小明.房产税对产业转移的影响:来自重庆和上海的经验证据[J].中国工业经济,2018(11):98-116.
- [22] 彭飞,许文立,吕鹏,吴华清.未预期的非税负担冲击:基于“营改增”的研究[J].经济研究,2020,55(11):67-83.
- [23] 杨励,欧嘉丽.房产税促进地区产业结构优化的路径与异质性表现——产业转移和企业创新视角的沪渝比较[J].西部论坛,2020,30(06):13-26.
- [24] 田淑英,孙磊,许文立,范子英.绿色低碳发展目标下财政政策促进企业转型升级研究[J].财政研究,2022,第8期.