双重差分法的安慰剂检验：一个实践的指南

代码说明

代码说明 1 贝特朗等（2004）的混合安慰剂检验

在文献中，贝特朗等（2004）或许最早使用 DID 的混合安慰剂检验，主要关注在面板数据存在序列相关的情况下[①]，对于 DID 估计量标准误的稳健估计。由于此特殊目的，故该文使用了特别的混合安慰剂检验，与本文正文所述有所不同。该文所用面板数据源于 1979-1999 年美国"当前人口调查"，由 25-50 岁的女性所构成，结果变量为"周工资对数"[②]。然而，在样本期间，并未实施影响女性工资的任何新政策，故真实处理效应为 0。进一步，既然无政策冲击，故也没有真正的处理组或处理时间。贝特朗等（2004）随机抽取美国一半的州（即 25 个州），这些州在样本内的女性构成"伪处理组"；而其余 25 州的女性则为控制组。进一步，作者从区间$[1985, 1995]$上的均匀分布随机抽取一个伪处理时间，由此得到一个安慰剂样本，进行 DID 估计，并检验安慰剂效应的显著性。

由于真实的安慰剂效应为 0[③]，且双向固定效应估计量（TWFE）为一致估计（即使扰动项存在序列相关），故若 DID 标准误正确，则在 5%的显著性水平下，在大样本中拒绝原假设"安慰剂效应为 0"的频率也应接近 5%。贝特朗等 (2004)重复此过程 200 次，即随机抽取 200 个安慰剂样本，并进行 200 次检验。结果发现，若使用忽略序列相关的普通标准误，拒绝原假设"安慰剂效应为 0"的频率高达 67.5%。具体而言，在这 200 次检验中，67.5%的 t 统计量绝对值均大于 1.96，导致拒绝原假设。由于 t 统计量等于 DID 估计量除以其标准误，而 DID 估计量为一致估计，故问题只能出在标准误。这意味着，未考虑序列相关的普通标准误大大低估了真实的标准误（因为序列相关多为正相关）。自从贝特朗等（2004） 以来，有关 DID 的实证研究一般均转而使用聚类稳健标准误，包括"聚类自助标准误"及"野聚类自助标准误"。

代码说明 2 标准 DID 的安慰剂检验 Stata 案例（曹和陈, 2022）

在此通过"漕粮海运"的案例（曹和陈, 2022），使用本团队开发的 Stata 命令 `didplacebo`,

[①] 这种样本数据一般称为"聚类数据"。其中，每位个体不同时期的观测值构成聚类，在每个聚类内部存在序列相关，但不同聚类之间不相关，且不同聚类允许存在异方差。
[②] 个体进入样本的准则之一为，其周工资须为正数，故可取对数。
[③] 在本例中，安慰剂效应与处理效应均为 0。

演示标准 DID 的安慰剂检验操作。首先，运行以下命令即可从 SSC 下载安装 didplacebo：

```
. ssc install didplacebo, all replace
```

其中，选择项"all"将安装此命令所附带的两个演示数据集（cao_chen.dta 与 bbb.dta），而选择项"replace"表示可覆盖此命令的旧版本。加载数据集 cao_chen.dta，并设为面板数据集：

```
. use cao_chen.dta, clear

. xtset county year

Panel variable: county (strongly balanced)
 Time variable: year, 1650 to 1911
         Delta: 1 unit
```

遵照曹和陈（2022），使用非官方 reghdfe 命令进行双向固定效应估计[①]：

```
. reghdfe rebel canal_post, absorb(i.county i.year)
cluster(county)
```

其中，选择项"absorb(i.county i.year)"将所有的县与年度的虚拟变量均"吸收"，而选择项"cluster(county)"指定以县为聚类的聚类稳健标准误。

```
(MWFE estimator converged in 2 iterations)
HDFE Linear regression                        Number of obs   =    140,432
Absorbing 2 HDFE groups                       F(  1,    535) =       5.23
Statistics robust to heteroskedasticity       Prob > F        =     0.0226
                                              R-squared       =     0.0308
                                              Adj R-squared   =     0.0253
                                              Within R-sq.    =     0.0002
Number of clusters (county)  =        536     Root MSE        =     0.3848
                                  (Std. err. adjusted for 536 clusters in county)
```

		Robust				
rebel	Coefficient	std. err.	t	P>|t|	[95% conf. interval]	
canal_post	.0380143	.016621	2.29	0.023	.0053639	.0706647
_cons	.0313251	.0007227	43.35	0.000	.0299054	.0327447

```
Absorbed degrees of freedom:
```

Absorbed FE	Categories	- Redundant	= Num. Coefs	
county	536	536	0	*
year	262	0	262	

```
* = FE nested within cluster; treated as redundant for DoF computation
```

然后，将回归结果存为"did_cao_chen"，以便后续调用：

[①] 安装 reghdfe 的命令为"ssc install reghdfe, all replace"。此命令也可等价地通过官方命令"xtreg rebel canal_post i.year,fe cluster(county)"来实现，但由于时间虚拟变量很多，故使用命令 reghdfe 更为方便。

. estimates store did_cao_chen

其次，进行时间安慰剂检验：

. didplacebo did_cao_chen, treatvar(canal_post) pbotime(1(1)10)

其中，"did_cao_chen" 即为上述所存储的 TWFE 回归结果，而必选项 "treatvar(canal_post)" 指定处理变量为 canal_post。选择项 "pbotime(1(1)10)" 表示将处理时间分别滞后 1 至 10 期，进行时间安慰剂检验。

```
Implementing in-time placebo test using fake treatment time shifted back by 1, 2
> , 3, 4, 5, 6, 7, 8, 9, 10 periods respectively.
Results of in-time placebo test using fake treatment times:
```

	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
canal_post						
L1.	.0237488	.0323933	0.73	0.463	-.039741	.0872386
L2.	-.0018263	.0177129	-0.10	0.918	-.0365429	.0328903
L3.	-.0072098	.0123804	-0.58	0.560	-.0314749	.0170552
L4.	.0365355	.0313244	1.17	0.243	-.0248592	.0979302
L5.	.0193318	.0257651	0.75	0.453	-.0311668	.0698304
L6.	.0085124	.0216505	0.39	0.694	-.0339218	.0509467
L7.	.0156334	.0202899	0.77	0.441	-.0241341	.055401
L8.	.0412773	.0259976	1.59	0.112	-.0096771	.0922318
L9.	.0348408	.0231507	1.50	0.132	-.0105337	.0802153
L10.	.0283695	.020891	1.36	0.174	-.0125762	.0693151

```
Note: The standard errors are computed using the same method as specified by
      the Stata command previously used for DID estimation. For example, if
      "xtreg, r" or "reghdfe, cluster(clustvar)" is used, then cluster-robust
      standard errors are reported.
Finished.
```


In-time Placebo Test
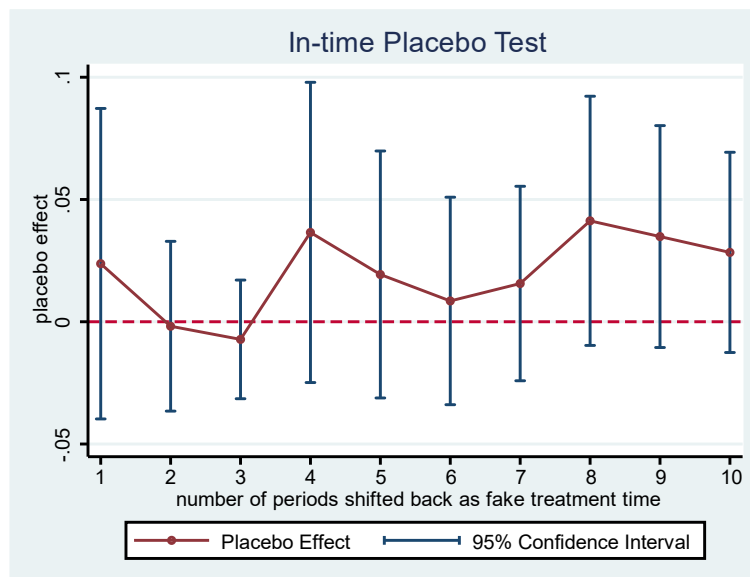
其中，正如回归结果的注释所示，在进行安慰剂检验时，将使用与原命令同样的方法计算标准误。例如，若原命令 reghdfe 使用聚类稳健标准误，则 didplacebo 也将使用同样的聚类稳健标准误。

再次，进行空间安慰剂检验：

```
. didplacebo did_cao_chen, treatvar(canal_post) pbounit rep(500)
seed(1)
```

其中，选择项"pbounit"表示进行空间安慰剂检验，选择项"rep(500)"表示重复500次（这是默认选项，故可省略），而选择项"seed(1)"指定随机种子为1（这是默认选项，故可省略；使用相同随机种子可保证结果可复现）。

```
Implementing in-space placebo test using fake treatment units:
Simulations (500):.........10.........20.........30.........40.........50.......
> ..60.........70.........80.........90.........100.........110.........120.....
> ....130.........140.........150.........160.........170.........180.........19
> 0.........200.........210.........220.........230.........240.........250.....
> ....260.........270.........280.........290.........300.........310.........32
> 0.........330.........340.........350.........360.........370.........380.....
> ....390.........400.........410.........420.........430.........440.........45
> 0.........460.........470.........480.........490.........500
Results of in-space placebo test results using fake treatment units:
```

	Coefficient		P-value	
		Two-sided	Left-sided	right-sided
canal_post	0.038014	0.0080	0.9920	0.0080

```
Note: (1) The two-sided p-value is the frequency that the absolute values of
      the placebo effects are greater than or equal to the absolute value of
      estimated treatment effect.
      (2) The left-sided (right-sided) p-value is the frequency that the
      placebo effects are smaller (greater) than or equal to the estimated
      treatment effect.
Finished.
```
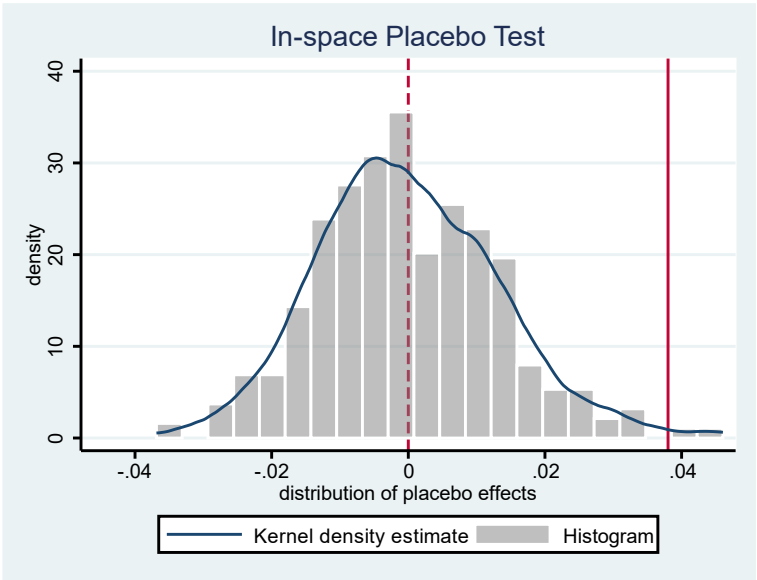

结果显示，双边与右边 p 值均为0.8%，故平均处理效应在1%水平上显著。最后，进行混合安慰剂检验：

```
.   didplacebo   did_cao_chen,   treatvar(canal_post)   pbomix(1)
```

```
seed(1)
```

其中，选择项"`pbomix(1)`"表示进行适合于标准 DID 的混合安慰剂检验。

```
Implementing mixed placebo test for standard DID (version 1) using both fake tre
> atment units and times:
```

The number of units randomly selected as fake treatment units	The range within which fake treatment times are randomly selected
73	[1651, 1911]

```
Simulations (500):.........10.........20.........30.........40.........50.......
> ..60.........70.........80.........90.........100........110........120.....
> ....130........140........150........160........170........180........19
> 0........200........210........220........230........240........250.....
> ....260........270........280........290........300........310........32
> 0........330........340........350........360........370........380.....
> ....390........400........410........420........430........440........45
> 0........460........470........480........490........500
Results of mixed placebo test for standard DID (version 1) using both fake treat
> ment units and times:
```

	Coefficient	P-value Two-sided	Left-sided	right-sided
canal_post	0.038014	0.0060	0.9960	0.0040

```
Note: (1) The two-sided p-value is the frequency that the absolute values of
         the placebo effects are greater than or equal to the absolute value of
         estimated treatment effect.
      (2) The left-sided (right-sided) p-value is the frequency that the
         placebo effects are smaller (greater) than or equal to the estimated
         treatment effect.
Finished.
```
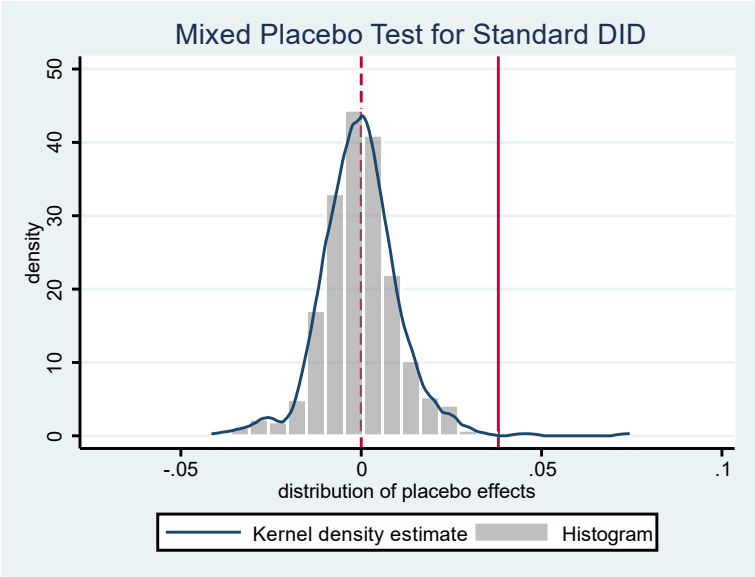

以上结果显示，双边与右边 p 值均小于 1%，故平均处理效应在 1%水平上显著。若想同时进行时间、空间与混合安慰剂检验，则可输入如下命令（结果从略）：

```
. didplacebo did_cao_chen, treatvar(canal_post) pbotime(1(1)10)
```

pbounit pbomix(1) seed(1)

代码说明 3 交叠 DID 的安慰剂检验 Stata 案例（贝克等, 2010）

在此通过"美国银行放松管制对收入分配的影响"的案例（贝克等, 2010），使用本团队开发的 Stata 命令 didplacebo，演示交叠 DID 的安慰剂检验操作。首先，运行以下命令即可从 SSC 下载安装 didplacebo：

```
. ssc install didplacebo, all replace
```

其中，选择项"all"将安装此命令所附带的两个演示数据集（cao_chen.dta 与 bbb.dta），而选择项"replace"表示可覆盖此命令的旧版本。加载数据集 bbb.dta，并设为面板数据集：

```
. use bbb.dta, clear
. xtset statefip wrkyr
```

其中，"statefip"表示州，而"wrkyr"表示年。

```
Panel variable: statefip (strongly balanced)
 Time variable: wrkyr, 1976 to 2006
        Delta: 1 unit
```

由于该文所用协变量名称较繁琐, 故通过命令 global 设定"全局暂元"（global macro），并以 cov 作为所有协变量的简写：

```
. global cov gsp_pc_growth prop_blacks prop_dropouts
prop_female_headed unemploymentrate
```

如此定义之后，在后续命令中，只要用"$cov"即可指代所有协变量"gsp_pc_growth prop_blacks prop_dropouts prop_female_headed unemploymentrate"。接着，进行双向固定效应估计：

```
. xtreg log_gini _intra $cov i.wrkyr, fe r
```

其中，"i.wrkyr"为时间虚拟变量，选择项"fe"指定固定效应模型，而选择项"r"指定使用聚类稳健标准误。

```
Fixed-effects (within) regression              Number of obs     =      1,519
Group variable: statefip                       Number of groups  =         49

R-squared:                                     Obs per group:
     Within  = 0.3876                                        min =         31
     Between = 0.0396                                        avg =       31.0
     Overall = 0.2339                                        max =         31

                                               F(36, 48)         =      98.01
corr(u_i, Xb) = -0.2892                         Prob > F          =     0.0000

                            (Std. err. adjusted for 49 clusters in statefip)
```

log_gini	Coefficient	Robust std. err.	t	P>|t|	[95% conf. interval]	
_intra	-.0177239	.0064033	-2.77	0.008	-.0305986	-.0048493
gsp_pc_growth	-.0288419	.0412498	-0.70	0.488	-.1117802	.0540965
prop_blacks	-.2128497	.1584963	-1.34	0.186	-.5315277	.1058284
prop_dropouts	.1641627	.0713148	2.30	0.026	.0207748	.3075507
prop_female~d	.0190403	.056497	0.34	0.738	-.0945545	.1326352
unemploymen~e	.0063327	.0012649	5.01	0.000	.0037895	.0088759
wrkyr						
1977	.0076353	.0068442	1.12	0.270	-.0061259	.0213964
1978	.0630635	.0072286	8.72	0.000	.0485295	.0775976
1979	.0892775	.0090404	9.88	0.000	.0711006	.1074543
1980	.0661498	.0099681	6.64	0.000	.0461077	.0861919
1981	.0805363	.0099324	8.11	0.000	.0605659	.1005067
1982	.1508562	.0123343	12.23	0.000	.1260565	.1756559
1983	.0835548	.0121092	6.90	0.000	.0592076	.1079021
1984	.0920258	.0130212	7.07	0.000	.065845	.1182066
1985	.0901488	.0121423	7.42	0.000	.0657351	.1145626
1986	.0947377	.0127701	7.42	0.000	.0690616	.1204138
1987	.0852106	.0132274	6.44	0.000	.0586152	.1118061
1988	.0803313	.0140429	5.72	0.000	.0520961	.1085664
1989	.0898124	.0152262	5.90	0.000	.0591981	.1204267
1990	.0771691	.015829	4.88	0.000	.0453429	.1089954
1991	.0734357	.0170896	4.30	0.000	.0390747	.1077966
1992	.0781385	.018361	4.26	0.000	.0412212	.1150558
1993	.0873723	.0196049	4.46	0.000	.047954	.1267905
1994	.0905129	.0200469	4.52	0.000	.0502059	.1308199
1995	.1251676	.0205097	6.10	0.000	.0839301	.1664052
1996	.1072482	.0232226	4.62	0.000	.060556	.1539403
1997	.1224012	.0223157	5.48	0.000	.0775325	.1672699
1998	.1099866	.0224811	4.89	0.000	.0647853	.1551878
1999	.1060401	.0243267	4.36	0.000	.0571279	.1549522
2000	.1332583	.0248512	5.36	0.000	.0832915	.183225
2001	.1175197	.0254826	4.61	0.000	.0662835	.1687559
2002	.1116299	.0262021	4.26	0.000	.0589471	.1643127
2003	.1192856	.0273775	4.36	0.000	.0642394	.1743319
2004	.1139813	.0255872	4.45	0.000	.0625348	.1654277
2005	.1348639	.0266757	5.06	0.000	.0812288	.1884991
2006	.1269117	.0277339	4.58	0.000	.0711488	.1826745
_cons	-.9676361	.0266693	-36.28	0.000	-1.021258	-.914014

```
    sigma_u |  .03262543
    sigma_e |  .03651163
        rho |  .44396696   (fraction of variance due to u_i)
```

然后，将回归结果存为"did_bbb"，以便后续调用：

. estimates store did_bbb

其次，进行时间安慰剂检验：

. didplacebo did_bbb, treatvar(_intra) pbotime(1(1)10)

8

其中，必选项"`treatvar(_intra)`"指定处理变量为 _intra；而选择项"`pbotime(1(1)10)`"表示，将处理时间分别滞后 1 至 10 期，进行时间安慰剂检验。

```
Implementing in-time placebo test using fake treatment time shifted back by 1, 2
> , 3, 4, 5, 6, 7, 8, 9, 10 periods respectively.
Results of in-time placebo test using fake treatment times:
```

	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
_intra						
L1.	.0068046	.0065114	1.05	0.296	-.0059575	.0195667
L2.	.0134762	.0058096	2.32	0.020	.0020895	.0248629
L3.	.0112213	.0055885	2.01	0.045	.0002681	.0221744
L4.	.0053115	.0064456	0.82	0.410	-.0073216	.0179446
L5.	.0071231	.006268	1.14	0.256	-.0051619	.0194082
L6.	.0113446	.0073201	1.55	0.121	-.0030025	.0256917
L7.	.002451	.00702	0.35	0.727	-.0113078	.0162099
L8.	-.0009252	.0053834	-0.17	0.864	-.0114765	.0096262
L9.	.0040415	.0058517	0.69	0.490	-.0074277	.0155106
L10.	.0071297	.0081124	0.88	0.379	-.0087703	.0230296

```
Note: The standard errors are computed using the same method as specified by
      the Stata command previously used for DID estimation. For example, if
      "xtreg, r" or "reghdfe, cluster(clustvar)" is used, then cluster-robust
      standard errors are reported.
Finished.
```
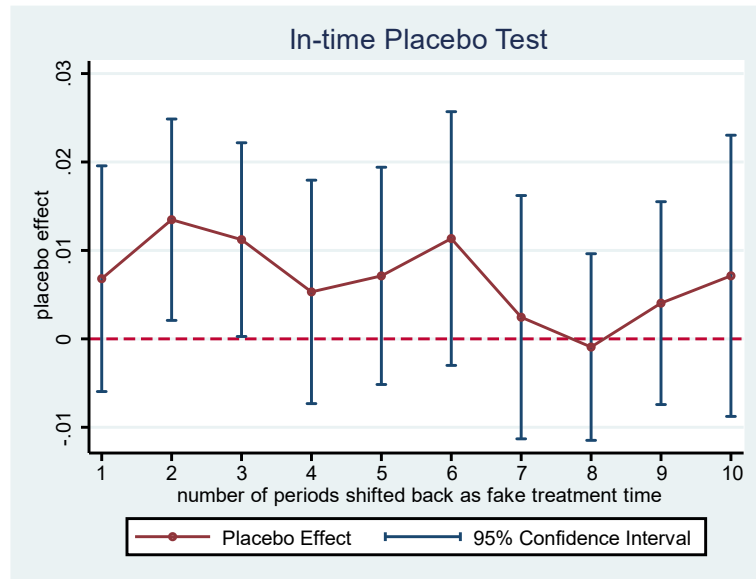

结果显示，滞后 2 期与 3 期的安慰剂效应置信区间均在横轴上方（未包括 0），故这两期的安慰剂效应显著为正；而其余各期的安慰剂效应则不显著。这说明，平行趋势假定可能并不完全满足。再次，进行空间安慰剂检验：

. didplacebo did_bbb, treatvar(_intra) pbounit rep(500) seed(1)

其中，选择项"`pbounit`"表示进行空间安慰剂检验，选择项"`rep(500)`"表示重复 500 次（这是默认选项，故可省略），而选择项"`seed(1)`"指定随机种子为 1（这是默认选项，故可省略；使用相同随机种子可保证结果可复现）。

```
Implementing in-space placebo test using fake treatment units:
Simulations (500):.........10.........20.........30.........40.........50.......
> ..60.........70.........80.........90.........100.........110.........120.....
> ....130.........140.........150.........160.........170.........180.........19
> 0.........200.........210.........220.........230.........240.........250.....
> ....260.........270.........280.........290.........300.........310.........32
> 0.........330.........340.........350.........360.........370.........380.....
> ....390.........400.........410.........420.........430.........440.........45
> 0.........460.........470.........480.........490.........500
Results of in-space placebo test results using fake treatment units:
```

	Coefficient	Two-sided	P-value Left-sided	right-sided
_intra	-0.017724	0.0120	0.0060	0.9940

```
Note: (1) The two-sided p-value is the frequency that the absolute values of
          the placebo effects are greater than or equal to the absolute value of
          estimated treatment effect.
      (2) The left-sided (right-sided) p-value is the frequency that the
          placebo effects are smaller (greater) than or equal to the estimated
          treatment effect.
Finished.
```
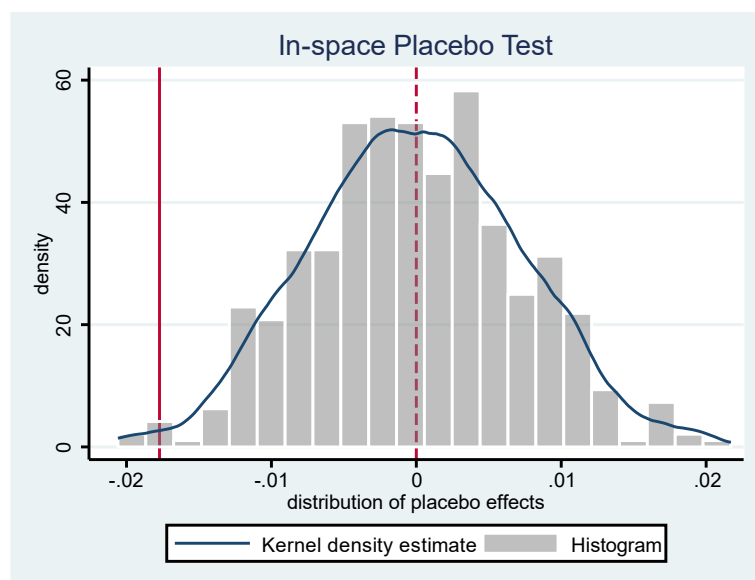

结果显示,平均处理效应的取值位于安慰剂效应分布的左端。进一步,双边 p 值为 1.2%,而左边 p 值为 0.6%,分别在 5% 与 1% 水平上显著,故平均处理效应显著。最后,进行混合安慰剂检验:

```
. didplacebo did_bbb, treatvar(_intra) pbomix(2) seed(1)
```

其中,选择项"pbomix(2)"表示以无约束(unrestricted)的方式(version 2),进行适合于交叠 DID 的混合安慰剂检验。

```
Implementing unrestricted mixed placebo test for staggered DID (version 2) using
>  both fake treatment units and times:
```
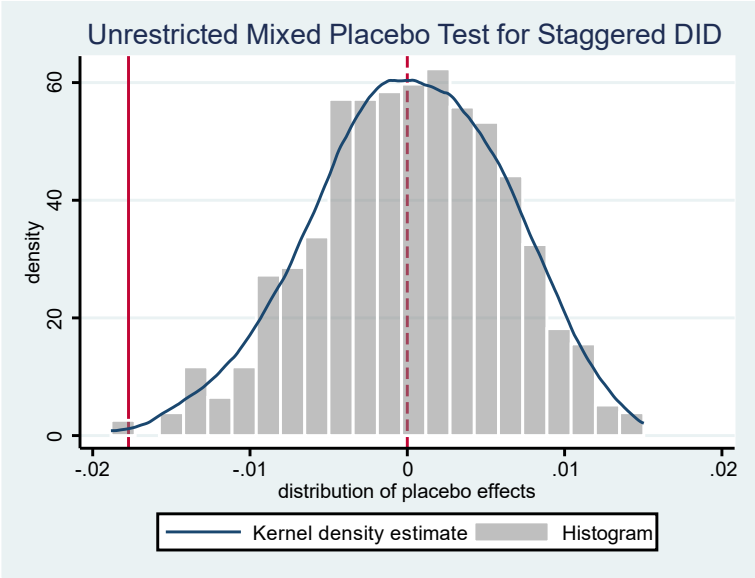
The number of units randomly selected as fake treatment units	The range within which fake treatment times are randomly selected
49	[1976, 2000]

```
Simulations (500):.........10.........20.........30.........40.........50.......
> ..60.........70.........80.........90.........100.........110.........120.....
> ....130.........140.........150.........160.........170.........180.........19
> 0.........200.........210.........220.........230.........240.........250.....
> ....260.........270.........280.........290.........300.........310.........32
> 0.........330.........340.........350.........360.........370.........380.....
> ....390.........400.........410.........420.........430.........440.........45
> 0.........460.........470.........480.........490.........500
Results of unrestricted mixed placebo test for staggered DID (version 2) using b
> oth fake treatment units and times:
```

	Coefficient	P-value Two-sided	Left-sided	right-sided
_intra	-0.017724	0.0040	0.0040	0.9960

```
Note: (1) The two-sided p-value is the frequency that the absolute values of
          the placebo effects are greater than or equal to the absolute value of
          estimated treatment effect.
      (2) The left-sided (right-sided) p-value is the frequency that the
          placebo effects are smaller (greater) than or equal to the estimated
          treatment effect.
Finished.
```


结果显示，平均处理效应的取值位于安慰剂效应分布的左端，且双边与左边 p 值均为 0，故平均处理效应高度显著。接着，进行有约束（restricted）的混合安慰剂检验，可保持交叠 DID 的组群结构。

. didplacebo did_bbb, treatvar(_intra) pbomix(3) seed(1)

其中，选择项"pbomix(3)"表示针对交叠 DID 模型，进行有约束的混合安慰剂检验（version 3），以保持组群结构。

```
Implementing restricted mixed placebo test for staggered DID (version 3) using b
> oth fake treatment units and times:

      The range within which fake
  treatment times are randomly selected
 ────────────────────────────────────────
              [1976, 2000]
 ────────────────────────────────────────

Simulations (500):.........10.........20.........30.........40.........50.......
> ..60.........70.........80.........90.........100.........110.........120.....
> ....130.........140.........150.........160.........170.........180.........19
> 0.........200.........210.........220.........230.........240.........250.....
> ....260.........270.........280.........290.........300.........310.........32
> 0.........330.........340.........350.........360.........370.........380.....
> ....390.........400.........410.........420.........430.........440.........45
> 0.........460.........470.........480.........490.........500
Results of restricted mixed placebo test for staggered DID (version 3) using bot
> h fake treatment units and times:
```

		P-value		
	Coefficient	Two-sided	Left-sided	right-sided
_intra	-0.017724	0.0000	0.0000	1.0000

```
Note: (1) The two-sided p-value is the frequency that the absolute values of
          the placebo effects are greater than or equal to the absolute value of
          estimated treatment effect.
      (2) The left-sided (right-sided) p-value is the frequency that the
          placebo effects are smaller (greater) than or equal to the estimated
          treatment effect.

Finished.
```
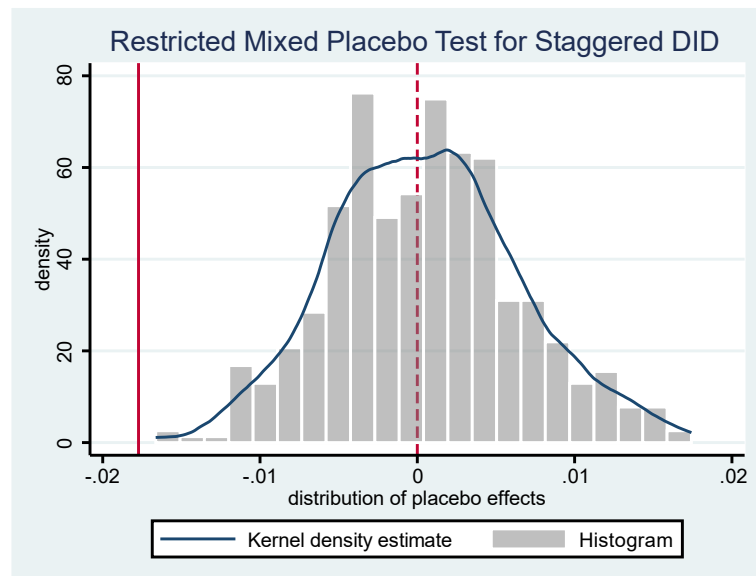

结果显示，平均处理效应的取值位于安慰剂效应分布的左端，且双边与左边 p 值均在 1%

水平上显著，故平均处理效应显著。

然而，若处理效应随时间而变，则以 TWFE 估计交叠 DID 模型将带来偏差。但由于异

质性稳健的交叠 DID 方法有多种，故不便整合进命令 didplacebo 中。尽管如此，交叠

DID 安慰剂检验的原理仍基本相同，只是将 TWFE 估计换为异质性稳健的交叠 DID 估计方

法。为此，下面使用异质性稳健的 CSDID（Callaway and Sant, 2021），进一步演示交叠 DID

的安慰剂检验。

在 Stata 中，可通过非官方命令 csdid 进行 CSDID 的估计①，其下载安装命令为 "ssc install csdid,all replace"。安装后，可运行如下命令：

. csdid log_gini $cov, ivar(statefip) time(wrkyr) gvar(branch_reform) method(dripw) wboot rseed(1) agg(simple)

其中，选择项 "ivar(statefip)" 指定个体变量为 "statefip"，选择项 "time(wrkyr)" 指定时间变量为 "wrkyr"，选择项 "gvar(branch_reform)" 指定变量 "branch_reform" 为个体开始受到处理的时间（以此将样本中个体分为若干组群），选择项 "method(dripw)" 表示进行双稳健估计（也是默认选项，故可省略），选择项 "wboot" 表示使用 "野自助法"（wild bootstrap）估计标准误（在小样本下表现更佳），选择项 "rseed(1)" 设定野自助法的随机种子，而选择项 "agg(simple)" 则指定汇报各组群与各期的总平均处理效应。

```
No never treated observations found. Using Not yet treated data
Units always treated found. These will be ignored
xxxxxxxxxxxxxxxxxxxxxxxxxxx..........xxxxxxxxx
xxxxxxxxx.........xxxxxxxxxxxxxxxxxx...x......
.xxxxxxxxxxxxxxxxxx.............xxxxxxxxxxxxxxx
............xxxxxxxxxxxxxx.....xxxxxxxxxxxxxx
xxxxxxxxxxx.x...x...xxxxxxxxxxxxxxxxxx.........
......xxxxxxxxxxxx.............xxxxxxxxxxxxxxx
................xxxxxxxxx................
..xxxxxxxx..xxxx...x...xxxxxxxxxxxxxxx.........
......xxxxxxxxxxxx.x.xx.xxx.xxxx...xxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
Difference-in-difference with Multiple Time Periods

                                             Number of obs = 636

Outcome model  : least squares
Treatment model: inverse probability
```

	Coefficient	Std. err.	t	[95% conf. interval]	
ATT	-.0007548	.0085642	-0.09	-.0174154	.0159058

```
Control: Not yet Treated
See Callaway and Sant´Anna (2021) for details
```

结果显示，平均处理效应的点估计为-0.00075，其符号与 TWFE 一致。然而，95%的置信区间为[-0.0174, 0.0159]，包含 0，故并不显著。然后，将回归结果存为 "csdid_bbb"，并将总平均处理效应的估计值记为全局宏 "tr_eff"（便于后续调用）：

. estimates store csdid_bbb

. global tr_eff = _b[ATT]

---

① 也可通过 Stata 18 的官方命令 xthdidregress 来实现 CSDID。

```
. dis $tr_eff
```

−.0007548

如上所示，使用命令"dis $tr_eff"即可展示全局宏 tr_eff 的取值。下面，进行 CSDID 的时间安慰剂检验。考虑最简单的情形，将处理时间滞后 1 期，可定义伪处理时间如下：

```
. gen branch_reform_1 = branch_reform - 1
```

需要指出，若个体从第 1 期开始即受到处理，则滞后 1 期后便超出样本区间。尽管如此，该个体在样本中仍从一开始就接受处理，故不影响最终结果。使用伪处理时间 branch _reform_1 进行 CSDID 的估计：

```
. csdid log_gini $cov if wrkyr < branch_reform, ivar(statefip) time(wrkyr) gvar(branch_reform_1) wboot rseed(1) agg(simple)
```

其中，"if wrkyr < branch_reform"表示仅使用处理前的数据进行时间安慰剂检验。

```
No never treated observations found. Using Not yet treated data
Units always treated found. These will be ignored
  (output omitted)
Will use observations with Pair balanced (observed at t0 and t1)
.xxxxxxxxxxxxxxxxxxxxx..xxxxxxxxxxxxxxxxxxxx...xxx
xxxxxxxxxxxxxxxx....xxxxxxxxxxxxxxxx.....xxxxxx
xxxxxxxxxx......xxxxxxxxxxxxxxx.x...xxxxxxxxxxx
xxxx........xxxxxxxxxxxx.........xxxxxxxxxxxx..
........xxxxxxxxxx..........xxxxxxxxxx..xxxx.
..x.xxxxxxxxx...........xxxxxxxxx.x.xx.xxx.xxxx
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxxxxxxxxxxxx
Difference-in-difference with Multiple Time Periods

                                              Number of obs = 361
Outcome model  : least squares
Treatment model: inverse probability
```

	Coefficient	Std. err.	t	[95% conf. interval]	
ATT	−.0060173	.0132436	−0.45	−.0333741	.0213396

```
Control: Not yet Treated
See Callaway and Sant´Anna (2021) for details
```

结果显示，滞后 1 期的时间安慰剂检验并不显著，因为 95%的置信区间包含 0。如果要进行滞后 2-10 期的时间安慰剂检验，也可类似地进行，只要定义相应的伪处理时间即可。

更正式地，可通过如下 for 循环程序来批量实现，并实现自动的图表输出。其中，定义取值为 10 的全局暂元变量 K，表示迭代次数。每次迭代时，创建变量 branch_reform _new 作为伪处理时间并运行 csdid 命令，将平均处理效应和方差分别存储于 att_b 矩阵和 att_V 矩阵，再使用命令 ereturn post 和 ereturn display 显示返回结果。具体

14

程序如下：

```
. global K = 10

. matrix att_b = J(1, $K, 0)

. matrix att_V = J($K, $K, 0)

. forvalues i = 1(1)$K{

    cap drop branch_reform_new

    qui gen branch_reform_new = branch_reform - `i'

    qui csdid log_gini $cov if wrkyr < branch_reform, ivar(state
fip) time(wrkyr) gvar(branch_reform_new) wboot rseed(1) agg(simple)

    matrix att_b[1, `i'] = e(b)[., "ATT"]

    matrix att_V[`i', `i'] = e(V)["ATT", "ATT"]

}

. mata: st_local("names", invtokens("L":+strofreal(1..$K):+".AT
T"))

. matrix colnames att_b = `names'

. matrix colnames att_V = `names'

. matrix rownames att_V = `names'

. ereturn post att_b att_V

. ereturn display
```

|	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
ATT						
L1.	-.0060173	.0132436	-0.45	0.650	-.0319743	.0199398
L2.	-.0137552	.0191677	-0.72	0.473	-.0513233	.0238128
L3.	-.0033001	.0095432	-0.35	0.729	-.0220045	.0154043
L4.	.0008408	.0113579	0.07	0.941	-.0214203	.0231018
L5.	.0237718	.0130823	1.82	0.069	-.0018691	.0494127
L6.	-.0066938	.014051	-0.48	0.634	-.0342333	.0208456
L7.	.0075708	.0176524	0.43	0.668	-.0270272	.0421688
L8.	-.0013085	.0106612	-0.12	0.902	-.0222041	.0195871
L9.	-.0077957	.0128759	-0.61	0.545	-.033032	.0174405
L10.	.0363283	.0193316	1.88	0.060	-.0015608	.0742175

结果显示，滞后 1-10 年的安慰剂效应均不在 5%水平上显著，尽管滞后 5 年与 10 年的安慰剂效应都在 10%水平上显著。然后，使用命令 coefplot 画将以上置信区间可视化：

```
. ssc install coefplot,replace
```

```
.  coefplot,vertical msymbol(smcircle_hollow) yline(0, lp(dash))
 xtitle("number of periods shifted back as fake treatment time") yt
itle("placebo effect") title("In-time Placebo Test") legend(order
(2 "Placebo Effect" 1 "95% Confidence Interval")) ciopts(recast(rca
p)) addplot(line @b @at) coeflabels(L.ATT=1 L2.ATT=2 L3.ATT=3  L4.A
TT=4 L5.ATT=5 L6.ATT=6 L7.ATT=7 L8.ATT=8 L9.ATT=9 L10.ATT=10)
```
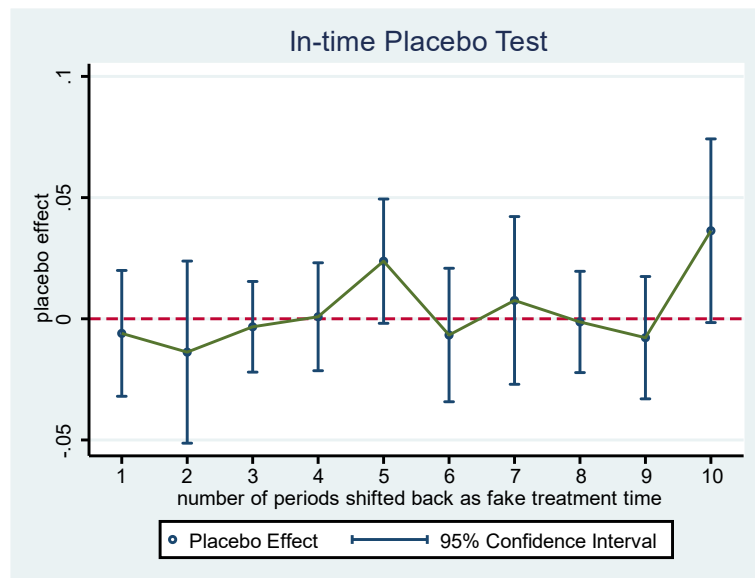

下面，进行 CSDID 的空间安慰剂检验。为此，定义一个名为
"InSpacePlaceboTest"的程序估计安慰剂效应：

```
.  capture drop branch_reform_new

.  capture program drop InSpacePlaceboTest

.  program def InSpacePlaceboTest, rclass

    preserve

    xtshuffle branch_reform, gen(branch_reform_new)

    qui csdid log_gini $cov, ivar(statefip) time(wrkyr) gvar(bra
nch_reform_new) agg(simple)

    return scalar pbo_eff = _b[ATT]

.  end
```

其中，使用命令 didplacebo 的附带命令 xtshuffle，可快捷地实现针对面板数据
的随机置换，详见 help xtshuffle。该命令将处理时间 branch_reform 进行随机置换
（每位个体不同时期的 branch_reform 变量取值作为整体进行置换），并将置换所得的伪

处理时间记为"branch_reform_new",而面板中其余数据不变。命令 xtshuffle 是 didplacebo 的基础命令,为 didplacebo 所调用,且在安装 didplacebo 时已同时安装。

然后,使用命令 simulate 运行"InSpacePlaceboTest"程序 500 次(由于每次运行 csdid 较费时,故此命令将花较长时间):

. simulate pbo_eff = r(pbo_eff), seed(1) reps(500): InSpacePlaceboTest

接着,将内存中的所有安慰剂效应估计值存为数据集 bbb_InSpacePbo.dta(以便后续使用),并画安慰剂效应分布的直方图与核密度图:

. save bbb_InSpacePbo.dta, replace

. graph twoway (kdensity pbo_eff) (histogram pbo_eff, fcolor(gs8%50) lcolor(white) lalign(center) below), xline(0, lp(dash)) xline($tr_eff) xtitle("distribution of placebo effect") ytitle("density") title("In-space Placebo Test") legend(order(1 "Kernel density estimate" 2 "Histogram") rows(1)) name(pbounit, replace)
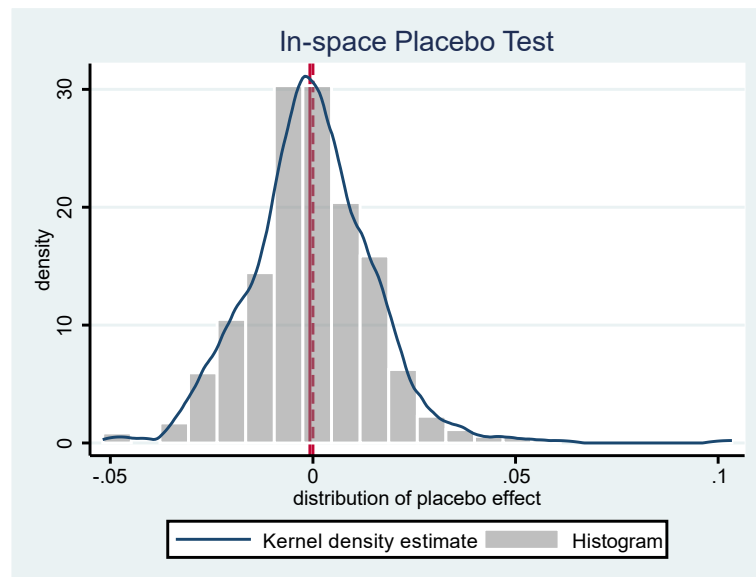
从上图可知,处理效应的取值位于安慰剂效应分布的中部,并非极端值。进一步,计算双边 p 值:

. gen extreme_abs = (abs(pbo_eff)>=abs($tr_eff))

. sum extreme_abs

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_abs	500	.956	.2053005	0	1

由上表可知，双边 p 值为 0.956。计算左边 p 值：

. gen extreme_left = (pbo_eff<=$tr_eff)

. sum extreme_left

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_left	500	.492	.5004367	0	1

由上表可知，左边 p 值为 0.492。计算右边 p 值：

. gen extreme_right = (pbo_eff>=$tr_eff)

. sum extreme_right

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_ri~t	500	.508	.5004367	0	1

由上表可知，右边 p 值为 0.508。总之，无论双边、左边或右边 p 值，均大幅超过常用的显著性水平（比如 5% 或 10%）。

下面，进行 CSDID 的混合安慰剂检验。首先进行准备工作。

. use bbb.dta, clear

. xtset statefip wrkyr

. global cov gsp_pc_growth prop_blacks prop_dropouts prop_female_headed unemploymentrate

. csdid log_gini $cov, ivar(statefip) time(wrkyr) gvar(branch_reform) wboot rseed(1) agg(simple)

. global tr_eff = _b[ATT]

其次，定义一个名为"MixedPlaceboTest2"的程序，以估计无约束的交叠 DID 安慰剂效应（version 2）：

. capture program drop MixedPlaceboTest2

. prog def MixedPlaceboTest2, rclass

    preserve

    xtrantreat _intra, method(2) gen(_intra_new)

    tofirsttreat _intra_new, gen(branch_reform_new)

```
    qui csdid log_gini $cov, ivar(statefip) time(wrkyr) gvar(bra
nch_reform_new) agg(simple)

    return scalar pbo_eff = _b[ATT]

  . end
```

其中，使用命令 didplacebo 的附带命令"xtrantreat,method(2)"针对交叠 D
ID 以无约束（unrestricted）的方式随机生成伪处理变量_intra_new。这意味，样本中的每
位个体均随机抽取一个伪处理时间，故一般无法保持组群结构（组群内的个体数目一般与原
始样本不同），详见 help xtrantreat。若希望保持组群结构，可使用命令"xtrantre
at,method(3)"（详见下文）。然后，在用命令 didplacebo 的另一附带命令"tofirst
treat"将伪处理变量_intra_new 变为个体首次受处理的时间 branch_reform_new。

之后，使用命令 simulate 运行程序"MixedPlaceboTest2"500 次（由于每次运
行 csdid 较费时，故此命令将花较长时间）：

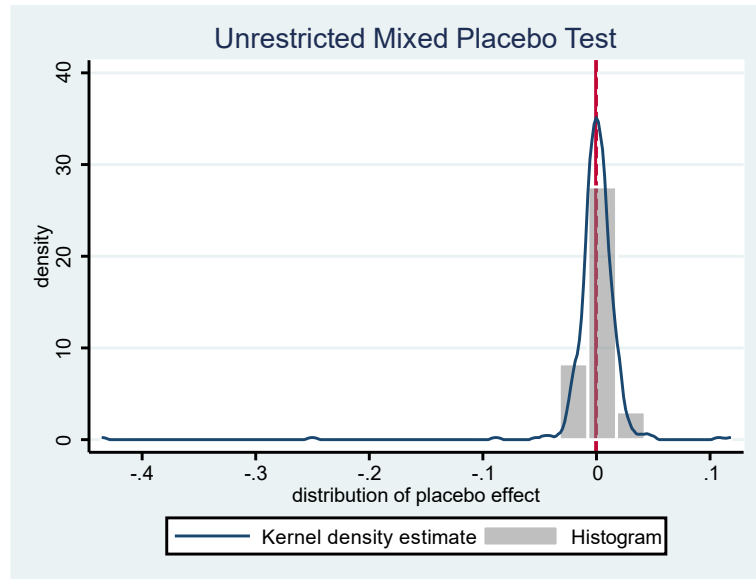 . simulate pbo_eff = r(pbo_eff), seed(1) reps(500): MixedPlaceb
oTest2

将内存中所有的安慰剂效应估计值存为数据集"bbb_MixedPbo2.dta"（以便后续使
用），并画安慰剂效应分布的直方图与核密度图：

  . save bbb_MixedPbo2.dta, replace

  . graph twoway (kdensity pbo_eff) (histogram pbo_eff,
fcolor(gs8%50) lcolor(white) lalign(center) below), xline(0,
lp(dash)) xline($tr_eff) xtitle("distribution of placebo effect")
ytitle("density") title("Unrestricted Mixed Placebo Test")
legend(order(1 "Kernel density estimate" 2 "Histogram") rows(1))
name(pbomix, replace)

从上图可知，处理效应的取值位于安慰剂效应分布的中部，并非极端值。进一步，计算双边 p 值：

```
. gen extreme_abs = (abs(pbo_eff)>=abs($tr_eff))
. sum extreme_abs
```

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_abs	500	.952	.2139803	0	1

由上表可知，双边 p 值高达 0.952。计算左边 p 值：

```
. gen extreme_left = (pbo_eff<=$tr_eff)
. sum extreme_left
```

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_left	500	.454	.4983781	0	1

由上表可知，左边 p 值为 0.454。计算右边 p 值：

```
. gen extreme_right = (pbo_eff>=$tr_eff)
. sum extreme_right
```

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_ri~t	500	.546	.4983781	0	1

由上表可知，右边 p 值为 0.546。总之，无约束混合安慰剂检验的双边、左边与右边 p 值均远大于常规的显著性水平（比如 5%或 10%），故平均处理效应并不显著。

最后，进行有约束（restricted）的混合安慰剂检验，可保持交叠 DID 模型的组群结构不变。首先进行准备工作。

```
. use bbb.dta, clear

. xtset statefip wrkyr

. global cov gsp_pc_growth prop_blacks prop_dropouts prop_femal
```

20

```
e_headed unemploymentrate
```

```
. csdid log_gini $cov, ivar(statefip) time(wrkyr) gvar(branch_r
eform) wboot rseed(1) agg(simple)
```

```
. global tr_eff = _b[ATT]
```

其次，定义一个名为"MixedPlaceboTest3"的程序，以估计有约束的交叠 DID 安慰剂效应（version 3）：

```
. capture program drop MixedPlaceboTest3
```

```
. prog def MixedPlaceboTest3, rclass
    preserve
    xtrantreat _intra, method(3) gen(_intra_new)
    tofirsttreat _intra_new, gen(branch_reform_new)
    qui csdid log_gini $cov, ivar(statefip) time(wrkyr) gvar(bra
nch_reform_new) agg(simple)
    return scalar pbo_eff = _b[ATT]
```

```
. end
```

其中，使用命令 didplacebo 的附带命令"xtrantreat,method(3)"针对交叠 DID 以有约束的方式随机生成伪处理变量（version 3）。具体而言，记样本中的处理时间分别为 $t_1, \cdots, t_G$，而相应组群内所包含的个体数目分别为 $n_1, \cdots, n_G$。在生成伪处理变量时，首先将样本中个体随机地划分为 G 个组群，分别包含 $n_1, \cdots, n_G$ 位个体。其次，随机地给这 G 个组群分配 G 个不同的伪处理时间。

然后，使用命令 simulate 运行程序"MixedPlaceboTest3"500 次（由于每次运行 csdid 较费时，故此命令将花较长时间）：
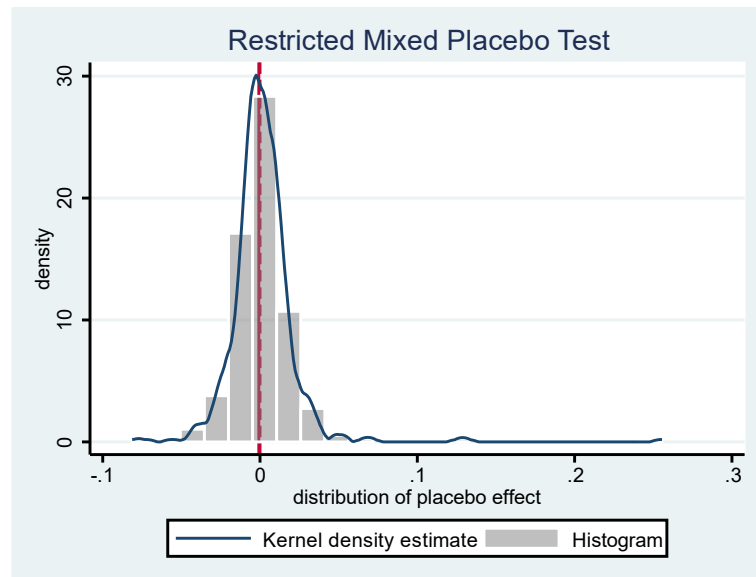
```
. simulate pbo_eff = r(pbo_eff), seed(1) reps(500): MixedPlaceb
oTest3
```

将内存中所有的安慰剂效应估计值存为数据集"bbb_MixedPbo3.dta"（以便后续使用），并画安慰剂效应分布的直方图与核密度图：

```
. save bbb_MixedPbo3.dta, replace
```

```
. graph twoway (kdensity pbo_eff) (histogram pbo_eff, fcolor(gs
8%50) lcolor(white) lalign(center) below), xline(0, lp(dash)) xlin
```

```
e($tr_eff) xtitle("distribution of placebo effect") ytitle("densit
y") title("Restricted Mixed Placebo Test") legend(order(1 "Kernel d
ensity estimate" 2 "Histogram") rows(1)) name(pbomix, replace)
```


从上图可知，处理效应的取值位于安慰剂效应分布的中部，并非极端值。进一步，计算双边 p 值：

```
. gen extreme_abs = (abs(pbo_eff)>=abs($tr_eff))
. sum extreme_abs
```

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_abs	500	.966	.1814106	0	1

由上表可知，双边 p 值为 0.966。计算左边 p 值：

```
. gen extreme_left = (pbo_eff<=$tr_eff)
. sum extreme_left
```

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_left	500	.474	.4998236	0	1

由上表可知，左边 p 值为 0.474。计算右边 p 值：

```
. gen extreme_right = (pbo_eff>=$tr_eff)
. sum extreme_right
```

Variable	Obs	Mean	Std. dev.	Min	Max
extreme_ri~t	500	.526	.4998236	0	1

由上表可知，右边 p 值为 0.526。总之，有约束混合安慰剂检验的双边、左边与右边 p 值均远大于常规的显著性水平（比如 5% 或 10%），故平均处理效应并不显著。