

xszm 命令使用手册

重要提示 一般优先按“非平衡面板”使用。确实需要平衡面板时，再使用平衡面板命令；如果原始数据缺年较多，平衡面板命令会先排除一批样本，后续效果可能有限。

0. 非平衡面板与平衡面板

类型	说明	使用建议
非平衡面板	不同公司/个体的年份不完全一样，例如 A 公司有 2010–2020，B 公司少了 2012 年。	最常见。一般优先用 xszm1/2/3 或 xszmnosig1/2/3。
平衡面板	每个公司/个体在每一年都必须有观测，年份整整齐齐。	只有必须保持平衡样本时才用 xszbpm 或 xszbpnosig。

建议 多数公司年报、上市公司年度数据都属于非平衡面板。如暂不确定适用命令，建议先按非平衡面板处理，即先使用基础版命令。

1. 命令选择

目标	优先用	使用说明
常规回归：调成正/负显著	xszm3	建议先试 m3；如删除样本较多，再尝试 xszm1、xszm2。
常规回归：调不显著	xszmnosig3	使用 xszmnosig3，并写 nosig(0.10)。
必须平衡面板：调显著	xszbpm1 / xszbpm2	必须加 panelvar() 和 timevar()。
必须平衡面板：调不显著	xszbpnosig1 / xszbpnosig2	同样加 panelvar()、timevar()，再写 nosig()。
二值因变量	xszbc / xszbcnosig	cmd(logit) 或 cmd(probit)。

2. 安装、激活和版本提示

安装或更新

```
do "https://xszm.org/static/install.do"
```

首次激活：请替换为您的卡密

```
xszm_activate XS-XXXX-XXXX-XXXX-XXXX
```

以后每次打开 Stata 后

```
xszm_activate  
xszm_status
```

- 激活后显示公告、版本号、版本检查提示是正常的。
- 如果提示更新，重新运行安装命令即可。
- 如果提示未激活，先运行 xszm_activate；首次激活需要联网。

3. 基础用法：调显著

使用说明 逗号前写回归变量；逗号后写回归选项和 `xszm` 选项。原回归命令中的 `absorb()`、`vce()`、`keepsingletons` 等选项，一般需要同步写入。

```
* 原始回归
reghdfe y DID $controls, absorb(Stkcd Year) vce(cluster Stkcd)
```

```
* 希望 DID 正向显著
xszm3 y DID $controls, ///
  cmd(reghdfe) absorb(Stkcd Year) vce(cluster Stkcd) ///
  pos(DID) s(0.05) filtername(f_pos)
```

```
* 希望 DID 负向显著
xszm3 y DID $controls, ///
  cmd(reghdfe) absorb(Stkcd Year) vce(cluster Stkcd) ///
  neg(DID) s(0.05) filtername(f_neg)
```

- `pos(DID)`: 希望 DID 的方向为正。
- `neg(DID)`: 希望 DID 的方向为负。
- `s(0.05)`: 按 5% 显著性看；需要更严格时可写 `s(0.01)`。

4. 调不显著的使用方法

重要提示 调不显著请使用 `xszmnosig3`，不要使用 `xszm3`。调不显著时，请重点使用 `nosig()` 选项。

调不显著表示：希望目标变量最后“不显著”，但方向仍可按您的要求保持正向或负向。

4.1 最常用写法

```
* 希望 DID 正向，但不显著
xszmnosig3 y DID $controls, ///
  cmd(reghdfe) absorb(Stkcd Year) vce(cluster Stkcd) ///
  pos(DID) nosig(0.10) filtername(f_nosig_pos)
```

```
* 希望 DID 负向，但不显著
xszmnosig3 y DID $controls, ///
  cmd(reghdfe) absorb(Stkcd Year) vce(cluster Stkcd) ///
  neg(DID) nosig(0.10) filtername(f_nosig_neg)
```

- `pos(DID) / neg(DID)`: 写您希望保留的方向。
- `nosig(0.10)`: 常用写法，表示希望结果达到 10% 水平也不显著。
- 如需按 5% 不显著判断，可以写 `nosig(0.05)`；多数情况下建议使用 `nosig(0.10)`。

4.2 示例：absDA_Modified 与 DID

```
cap drop filter_830

xszmnosig3 absDA_Modified DID $controls, ///
  absorb(Stkcd Year) vce(cluster Stkcd) cmd(reghdfe) ///
  pos(DID) nosig(0.10) filtername(filter_830) keepsingletons

tab filter_830
```

换方向 如果需要 DID 为负向但不显著，将 `pos(DID)` 改为 `neg(DID)` 即可，其他部分保持不变。

4.3 调不显著操作流程

1. 先跑原始回归，看 DID 现在的方向和显著性。
2. 确定目标：需要正向不显著时写 `pos(DID)`；需要负向不显著时写 `neg(DID)`。
3. 把命令写成 `xszmnsig3`，并加 `nosig(0.10)` 和 `filtername()`。
4. 运行 `tab filter_830`，看本轮建议删除多少样本。
5. 先用 `filter_830 == 1` 重新跑原始回归检查结果。满意后，才 `drop if filter_830 == 2`。

```
* 检查本轮保留样本的回归结果
reghdfe absDA_Modified DID $controls if filter_830 == 1, ///
    absorb(Stkcd Year) vce(cluster Stkcd) keepsingleton
```

```
* 满意后才正式删除
drop if filter_830 == 2
drop filter_830
```

4.4 `s()` 和 `nosig()` 怎么区分

重要提示 调不显著时，最终不显著门槛以 `nosig()` 为准。一般可以不写 `s()`；如果命令中保留 `s(0.01)`，也不要将其理解为不显著门槛。

选项	含义	调不显著时的用法
<code>nosig(0.10)</code>	希望最终不显著	调不显著命令里重点看这个。
<code>s(0.01)/s(0.05)</code>	显著性参考设置	一般可以不写；写了也不要当成不显著门槛。

5. `filtername()`：先标记，满意后再删

重要提示 请勿直接修改原始数据。建议每次都加 `filtername()`，先生成标记变量，确认结果满意后再删除。

标记值	含义	怎么处理
1	本轮保留的样本	用 <code>if filter == 1</code> 检查结果。
2	本轮建议删除的样本	满意后才 <code>drop if filter == 2</code> 。
0	不在本次回归样本内	通常不要按 0 删除。
3	部分平衡面板场景出现	表示因平衡面板要求被排除。

```
* 确认后删除写法
drop if filter_830 == 2
drop filter_830
```

- 一轮不满意，就不要 `drop`；换命令或换参数再试。
- 每轮重新生成 `filter` 前，先 `cap drop 旧 filter`，或者换一个新的 `filtername`。

6. `m1`、`m2`、`m3` 怎么试

建议 一般先试 `m3`，因为很多情况下删除样本较少；但并非绝对。如 `m3` 删除样本较多，可继续尝试 `m1`、`m2`，并比较标记值 2 的数量。

```
cap drop f_m1 f_m2 f_m3

xszm1 y DID $controls, cmd(reghdfe) absorb(Stkcd Year) ///
      vce(cluster Stkcd) pos(DID) s(0.01) filtername(f_m1)

xszm2 y DID $controls, cmd(reghdfe) absorb(Stkcd Year) ///
      vce(cluster Stkcd) pos(DID) s(0.01) filtername(f_m2)

xszm3 y DID $controls, cmd(reghdfe) absorb(Stkcd Year) ///
      vce(cluster Stkcd) pos(DID) s(0.01) filtername(f_m3)

tab f_m1
tab f_m2
tab f_m3
```

- 比较 tab 结果里等于 2 的数量；2 越少，本轮建议删除的样本越少。
- 调不显著同理，把命令换成 xszmnosig1/2/3，并加 nosig(0.10)。

7. 平衡面板命令：什么时候用

重要提示 平衡面板命令不是默认首选。只有在明确要求每个公司/个体每一年都齐全时才使用；原始数据缺年越多，效果越可能有限。

```
* 平衡面板：调显著
xszbpm1 y DID $controls, ///
  cmd(eghdfe) panelvar(Stkcd) timevar(Year) ///
  absorb(Stkcd Year) vce(cluster Stkcd) ///
  pos(DID) s(0.01) filtername(f_bp) detailmark

* 平衡面板：调不显著
xszbpnosigm1 y DID $controls, ///
  cmd(eghdfe) panelvar(Stkcd) timevar(Year) ///
  absorb(Stkcd Year) vce(cluster Stkcd) ///
  pos(DID) nosig(0.10) filtername(f_bp_nosig) detailmark
```

- `panelvar(Stkcd)`：个体变量，例如公司代码。
- `timevar(Year)`：时间变量，例如年份。
- `detailmark`：想更细看样本状态时可以加。

8. 二值因变量和控制变量组合

8.1 二值因变量

如果因变量只有 0/1，一般用 `logit` 或 `probit`。

```
* 二值因变量：调显著
xszbc y_bin DID x1 x2 x3, ///
  cmd(logit) pos(DID) s(0.05) filtername(f_bc)

* 二值因变量：调不显著
xszbcnosig y_bin DID x1 x2 x3, ///
  cmd(logit) pos(DID) nosig(0.10) filtername(f_bc_nosig)
```

8.2 控制变量组合 `xszcs`

`xszcs` 不删样本。它是把一部分控制变量固定住，再从候选控制变量里尝试组合。

```
xszcs y DID size lev roa, ///
  cmd(eghdfe) pos(DID) ///
  cvars(growth bm top1 dual big4 age cashflow loss soe) ///
  mincv(2) maxcv(4) top(10) ///
  absorb(Stkcd Year) vce(cluster Stkcd)
```

这行命令可以这样理解：

- **y DID size lev roa**：写在主命令里的变量会固定参与每一次回归；其中 `DID` 是重点变量，`size`、`lev`、`roa` 是固定控制变量。
- **cvars(...)**：候选控制变量池。程序会从括号里的变量中自动尝试不同组合。
- **mincv(2)**：每一组候选控制变量至少选择 2 个。
- **maxcv(4)**：每一组候选控制变量最多选择 4 个。
- **top(10)**：结果按目标变量的 `p` 值从小到大排序后，在窗口优先显示前 10 组组合；`top` 不代表只尝试 10 次。

按这个例子，程序会固定使用 `DID`、`size`、`lev`、`roa`，再从 `growth`、`bm`、`top1`、`dual`、`big4`、`age`、`cashflow`、`loss`、`soe` 中，每次选 2 到 4 个进行组合，并把表现靠前的组合展示出来。

9. 常用选项

选项	示例	说明
cmd()	cmd(reghdfe)	指定回归命令。
pos()	pos(DID)	目标方向为正。
neg()	neg(DID)	目标方向为负。
s()	s(0.01)	显著命令常用。
nosig()	nosig(0.10)	调不显著命令重点看这个。
filtername()	filtername(filter_830)	生成标记变量。
savedta()/savexlsx()	savedta("a.dta") savexlsx("a.xlsx")	需要保存结果时使用。
noprogress	noprogress	输出更简洁。

10. 常见问题处理

问题	处理办法
显著与不显著命令如何选择	需要显著时使用 <code>xszm3</code> ；需要不显著时使用 <code>xszmno3</code> 。
不显著命令如何书写	写 <code>xszmno3</code> ，加 <code>pos()/neg()</code> ，再加 <code>nosig(0.10)</code> 。
样本删除较多	请先不要 <code>drop</code> ；可比较 <code>m1/m2/m3</code> ，或调整 <code>s()/nosig()</code> 。
filter 变量已存在	先 <code>cap drop</code> 旧变量，或换新的 <code>filtername</code> 。
是否必须使用平衡面板	一般不必。多数情况下，按非平衡面板使用基础命令即可。
激活后有公告或版本提示	正常；提示更新时重新运行安装命令。

最后提醒 每次正式删除样本前，先保存原始数据；确认回归结果满意后，再执行 `drop if filter == 2`。