

بسم الله الرحمن الرحيم

آموزش برآورد اثرات ثابت یکطرفه و دوطرفه در نرم افزار STATA

حسین خاندانی

مدرس داده کاوی و اقتصادسنجی

سخن مدیر

در این فایل آموزشی نحوه برآورد مدل‌های پانل اثرات ثابت یک و دوطرفه در نرم افزار استاتا آموزش داده می‌شود. /

(حسین خاندانی) <http://econometrics.blog.ir>

بعد از وارد نمودن داده‌ها در نرم افزار استاتا باید به نرم افزار متغیر مقطع و زمان را معرفی نمایید تا نرم افزار ساختار پانل را تشخیص دهد. از این‌رو از دستور اسلاید بعد استفاده نمایید.

البته قبل از هر اقدام ابتدا چک نمایید که نرم افزار همه متغیرهای شما را تشخیص داده باشد. اگر داده‌های متغیری به رنگ قرمز بود نرم افزار آن را به عنوان یک رشته متن در نظر گرفته است. برای رفع این مشکل هم می‌توانید از دستور `destring variable, replace` استفاده نمایید.

معرفی ساختار داده‌های پانل به نرم افزار استاتا

□ برای معرفی ساختار داده‌های پانل لازم است داده‌های ورودی شما علاوه بر متغیرهای وابسته و توضیحی شامل دو ستون مقاطع و زمان نیز باشد.

دستور در برنامه: `xtset cross year`

```
. xtset cross year
      panel variable:  cross (strongly balanced)
      time variable:  year, 2002 to 2012
                   delta:  1 unit
```

Command

```
xtset cross year
```

دستور

نکته:

`cross` در اینجا نام متغیر مقاطع و یک اسم دلخواه است.

`year` در اینجا نام متغیر زمان و یک اسم دلخواه است.

تخمین مدل اثرات ثابت نسخه متغیر مجازی به تفکیک یکطرفه و دو طرفه

□ اثرات ثابت یکطرفه مقطعی

در صورتی که در مدل پانل، فقط عرض از مبدأ مقاطع متفاوت فرض شود به این مدل اثرات ثابت مقطعی گویند.

□ اثرات ثابت یکطرفه زمانی

در صورتی که در مدل پانل فقط عرض از مبدأ در طی زمان متفاوت فرض شود به این مدل اثرات ثابت زمانی گویند.

□ اثرات ثابت دو طرفه (مکانی و زمانی)

در صورتی که در مدل پانل علاوه عرض از مبدأ متفاوت به ازای مقاطع، این تفاوت و ناهمگنی در طی زمان نیز فرض شود به این مدل اثرات ثابت دو طرفه گویند.

تخمین مدل اثرات ثابت یکطرفه مقطعی نسخه متغیر مجازی

□ اثرات ثابت یکطرفه مقطعی

برای تخمین مدل اثرات ثابت یکطرفه مقطعی در نرم افزار استاتا از دستور زیر استفاده کنید:

$$\text{Reg } y \text{ } x_1 \text{ } x_2 \text{ } i.\text{crossName}$$

نام متغیر مقطع

```
. reg pgdp tb inf i.cross
```

Source	SS	df	MS	Number of obs	=	55
Model	7.7586e+09	6	1.2931e+09	F(6, 48)	=	15.17
Residual	4.0913e+09	48	85234880	Prob > F	=	0.0000
Total	1.1850e+10	54	219442352	R-squared	=	0.6547
				Adj R-squared	=	0.6116
				Root MSE	=	9232.3

pgdp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
tb	-543.1979	999.5871	-0.54	0.589	-2553.003 1466.607
inf	1899.234	1266.087	1.50	0.140	-646.4056 4444.873
cross					
2	-342.7899	4577.937	-0.07	0.941	-9547.348 8861.769
3	-30041.37	9575.937	-3.14	0.003	-49295.08 -10787.66
4	-8800.11	12385.97	-0.71	0.481	-33703.77 16103.55
5	16053.28	10516.54	1.53	0.133	-5091.635 37198.19
_cons	48955.94	23171.61	2.11	0.040	2366.296 95545.59

عرض از مبداهای مختلف به ازای هر مقطع

Command

```
reg pgdp tb inf i.cross
```

بررسی اثرات ثابت یکطرفه مقطعی

برای بررسی اثر وجود اثرات مقطعی و تست یکسان بودن (pool) تمامی مقاطع می‌توانید از دستور زیر استفاده نمایید:

> testparm i.cross

```
. testparm i.cross  
  
( 1)  2.cross = 0  
( 2)  3.cross = 0  
( 3)  4.cross = 0  
( 4)  5.cross = 0  
  
      F(  4,    48) =    11.61  
      Prob > F =    0.0000
```

تفسیر نتایج آزمون (مدل یکطرفه مقطعی):

با توجه به اینکه احتمال آماره آزمون زیر ۵ صدم است، در نتیجه فرضیه صفر مبنی برابری و صفر بودن تمامی مقاطع با هم رد می‌گردد و در نتیجه وجود اثرات مقطعی را نمی‌توان رد نمود.

Command

```
testparm i.cross
```

تخمین مدل اثرات ثابت یکطرفه زمانی نسخه متغیر مجازی

```
. reg pgdp tb inf i.year
```

Source	SS	df	MS	Number of obs	=	55
Model	7.7622e+09	12	646853273	F(12, 42)	=	6.65
Residual	4.0876e+09	42	97324946.2	Prob > F	=	0.0000
Total	1.1850e+10	54	219442352	R-squared	=	0.6550
				Adj R-squared	=	0.5565
				Root MSE	=	9865.3

pgdp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
tb	1131.431	172.9723	6.54	0.000	782.3588 1480.503
inf	-4644.458	1895.976	-2.45	0.019	-8470.694 -818.2231
year					
2003	1895.764	6322.443	0.30	0.766	-10863.44 14654.97
2004	8586.943	6240.208	1.38	0.176	-4006.307 21180.19
2005	10388.87	6242.142	1.66	0.103	-2208.28 22986.03
2006	13583.34	6251.635	2.17	0.035	967.0343 26199.65
2007	17439.58	6248.817	2.79	0.008	4828.954 30050.2
2008	32306.52	7477.943	4.32	0.000	17215.42 47397.62
2009	13135.65	6619.54	1.98	0.054	-223.12 26494.43
2010	20258.93	6241.498	3.25	0.002	7663.077 32854.78
2011	28181.71	6372.713	4.42	0.000	15321.06 41042.37
2012	25296.4	6286.154	4.02	0.000	12610.43 37982.37
_cons	7576.92	6546.472	1.16	0.254	-5634.395 20788.24

□ اثرات ثابت یکطرفه زمانی

برای تخمین مدل اثرات ثابت یکطرفه زمانی در نرم افزار استاتا از دستور زیر استفاده کنید:

Reg y x1 x2 i.TimeName

نام متغیر زمان

عرض از مبداهای مختلف به ازای هر سال

بررسی اثرات ثابت یکطرفه زمانی

برای بررسی اثر وجود اثرات زمانی و تست یکسان بودن (pool) تمامی مقاطع از دستور زیر استفاده نمایید:

> testparm i.year

```
. testparm i.year

( 1) 2003.year = 0
( 2) 2004.year = 0
( 3) 2005.year = 0
( 4) 2006.year = 0
( 5) 2007.year = 0
( 6) 2008.year = 0
( 7) 2009.year = 0
( 8) 2010.year = 0
( 9) 2011.year = 0
(10) 2012.year = 0

F( 10, 42) = 4.07
Prob > F = 0.0006
```

تفسیر نتایج آزمون (مدل یکطرفه زمانی):

با توجه به اینکه احتمال آماره آزمون زیر ۵ صدم است، در نتیجه فرضیه صفر مبنی همگن و یکسان بودن تمامی سالها با هم رد می‌گردد و در نتیجه وجود اثرات زمانی را نمی‌توان رد نمود.

Command

```
testparm i.year
```

تخمین مدل اثرات ثابت دوطرفه نسخه متغیر مجازی

□ اثرات ثابت دوطرفه

برای تخمین مدل اثرات ثابت دوطرفه در نرم افزار استاتا از دستور زیر استفاده کنید:

$$\text{Reg } y \text{ } x_1 \text{ } x_2 \text{ } i.\text{crossName} \text{ } i.\text{TimeName}$$

`reg pgdp tb inf i.cross i.year` → دستور مدل دوطرفه

Source	SS	df	MS	Number of obs	=	55
Model	1.1159e+10	16	697468246	F(16, 38)	=	38.39
Residual	690395075	38	18168291.4	Prob > F	=	0.0000
Total	1.1850e+10	54	219442352	R-squared	=	0.9417
				Adj R-squared	=	0.9172
				Root MSE	=	4262.4

pgdp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
tb	-894.8062	493.4216	-1.81	0.078	-1893.686 104.0736
inf	-2523.103	961.6055	-2.62	0.012	-4469.771 -576.4343
cross					
2	-2040.081	2158.46	-0.95	0.351	-6409.655 2329.492
3	-35043	4734.406	-7.40	0.000	-44627.31 -25458.7
4	-18244.26	6270.612	-2.91	0.006	-30938.45 -5550.072
5	16402.81	5094.864	3.22	0.003	6088.801 26716.83
year					
2003	3328.331	2744.915	1.21	0.233	-2228.459 8885.122
2004	9369.433	2701.344	3.47	0.001	3900.848 14838.02
2005	11934.04	2729.132	4.37	0.000	6409.203 17458.88
2006	13953.29	2711.341	5.15	0.000	8464.467 19442.11
2007	19565.04	2768.455	7.07	0.000	13960.6 25169.49
2008	29769.11	3474.301	8.57	0.000	22735.76 36802.47
2009	15713.69	2919.01	5.38	0.000	9804.466 21622.92
2010	19920.5	2700.274	7.38	0.000	14454.08 25386.92
2011	26531.29	2774.32	9.56	0.000	20914.97 32147.61
2012	25359.95	2735.419	9.27	0.000	19822.38 30897.51
_cons	53533.1	11845.86	4.52	0.000	29552.41 77513.78

عرض از مبدهای مختلف به ازای هر مقطع و هر سال

بررسی اثرات ثابت دوطرفه

```
. testparm i.cross i.year
```

```
( 1)  2.cross = 0  
( 2)  3.cross = 0  
( 3)  4.cross = 0  
( 4)  5.cross = 0  
( 5)  2003.year = 0  
( 6)  2004.year = 0  
( 7)  2005.year = 0  
( 8)  2006.year = 0  
( 9)  2007.year = 0  
(10)  2008.year = 0  
(11)  2009.year = 0  
(12)  2010.year = 0  
(13)  2011.year = 0  
(14)  2012.year = 0
```

```
      F( 14,    38) =    28.93  
      Prob > F =    0.0000
```

برای بررسی اثر وجود اثرات زمانی و تست یکسان بودن (pool) تمامی مقاطع و سالها از دستور زیر استفاده نمایید:

```
> testparm i.cross i.year
```

تفسیر نتایج آزمون (مدل طرفه):

با توجه به اینکه احتمال آماره آزمون زیر ۵ صدم است، در نتیجه فرضیه صفر مبنی یکسان و صفر بودن تمامی مقاطع و سالها با هم (POOLED) رد می‌گردد و در نتیجه وجود اثرات دوطرفه (مقطع و زمان) را نمی‌توان رد نمود.

ملاحظات پایانی

۱- لازم به ذکر است در نسخه متغیرمجازی اثرات ثابت نیز می‌توانید از طریق R^2 یا RSS مقید و نامقید آزمون F لیمر را بصورت دستی انجام و محاسبه نمایید.

۲- نکته دومی که باید بدان اشاره نمود این مسأله است که در تخمین مدل اثرات ثابت در نرم افزار استاتا از طریق **نسخه میانگین درون گروهی (Within)**:

`xtreg y x1 x2, fe`

نرم افزار استاتا بصورت **پیش فرض** اثرات ثابت **مقطعی** را تخمین می‌زند که این نکته را حتماً در پژوهش‌های خود در نظر داشته باشید.

Poolability Test (between Pooled Regression and FE Model)

Compared with our old pooled regression model, the new LSDV fixed effects model has a higher R^2 value. Hence the question comes up: Which model is better? The pooled regression with constant slope and constant intercept or the LSDV fixed effects model with constant slope and variable intercept for companies? The question can be reframed also as: Can we assume that there is neither significant cross sectional nor significant temporal effect, and pool the data and run an OLS regression model with an intercept α and slope coefficients β s constant across companies and time? This is the poolability test.

Note that compared with the second (FE) model, the first one (pooled regression) is a restricted model; it imposes a common intercept on all companies: $\mu_2 = \mu_3 = \mu_4 = \mu$. Hence we have to do the restricted F test given by

$$F = \frac{(R_{UR}^2 - R_R^2)/J}{(1 - R_{UR}^2)/(n - k)}$$

where $R_{UR}^2 = R^2$ of the unrestricted regression (second model)

$R_R^2 = R^2$ of the restricted regression (first model)

J = number of linear restrictions on the first model

k = number of parameters in the unrestricted regression

n = NT = number of observations

This poolability test we can do in Stata after the regression with the factor variable i.ind, by typing

```
. testparm i.ind
```

پایان بخش **چهارم** از مجموعه آموزش های نرم افزار Stata

زکات علم نشر آن است

Econometrics.blog.ir

مدیریت وب و تهیه کننده:

حسین خاندانی