

Roadshow effectiveness analysis

1. 数据和模型

数据为 balanced panel **平衡面板** (每个时期 t 每个省份 n , 数据都没有缺失) ; 数据为 **长面板** (时期数 $t >$ 个体数 n) ;

模型为 :

$$sales_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}$$

其中 \mathbf{x}_{it}' 可以包括 *roadshow* , 常数项、时间趋势项、个体虚拟变量和不随时间变化的解释变量 z_i , 随机扰动项 ε_{it} 可写作 :

$$\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + v_{it}$$

其中 $|\rho_i| < 1$, v_{it} 为独立同分布且期望为 0.

Trends of sales and roadshow of provinces:



Graphs by province

ifroadshow=roadshow*2500

进行回归前考虑以下三种情形：

1. **组间异方差**：记 $Var(\varepsilon_{it}) = \sigma_i^2$, $Var(\varepsilon_{jt}) = \sigma_j^2$, 则 $\sigma_i^2 \neq \sigma_j^2$.即每个省份，影响 sales 的除 roadshow 以外的因素的波动不相同（认为这种波动在不同时期对每个省份是不变的）
2. **组内自相关**：.即不同时期，每个省份影响 sales 的除 roadshow 以外的因素之间存在相关关系
3. **组间同期相关**： $Cov(\varepsilon_{it}, \varepsilon_{jt}) \neq 0 (i \neq j, for \forall t)$.即不同省份之间，在同一时间，除 roadshow 以外影响 sales 的其他因素之间存在相关关系，比如相邻省份之间的影响。

2. 三种关系的检验

2.1 检验组间异方差

Greene 沃尔德检验（xtreg,fe 和 xtglsls 两种方法）

原假设 H_0 ：组间同方差

xtreg,fe 检验：

```
Modified Wald test for groupwise heteroskedasticity  
in fixed effect regression model
```

```
H0: sigma(i)^2 = sigma^2 for all i
```

```
chi2 (12)   =      5462.36  
Prob>chi2  =      0.0000
```

xtglsls 检验:

```
Modified Wald test for groupwise heteroskedasticity  
in cross-sectional time-series FGLS regression model
```

```
H0: sigma(i)^2 = sigma^2 for all i
```

```
chi2 (12)   =      5462.36  
Prob>chi2  =      0.0000
```

结果相同，P 值非常小，强烈拒绝原假设，**认为组间存在异方差**

2.2 检验组内自相关

Wooldridge 沃尔德检验

原假设 H_0 : 无一阶的组内自相关

```
Wooldridge test for autocorrelation in panel data
H0: no first order autocorrelation
F( 1, 11) = 408.099
Prob > F = 0.0000
```

P 值非常小，强烈拒绝原假设，**认为每一组内存在一阶自相关**

1. 检验组间同期相关

Greene: Breusch-Pagan LM 检验

原假设 H_0 : 无组间同期相关

Correlation matrix of residuals:

	__e1	__e2	__e3	__e4	__e5	__e6	__e7	__e8	__e9	__e10	__e11	__e12
__e1	1.0000											
__e2	0.8771	1.0000										
__e3	0.9566	0.8595	1.0000									
__e4	0.9397	0.8943	0.9665	1.0000								
__e5	0.9445	0.9492	0.9448	0.9736	1.0000							
__e6	0.9160	0.7879	0.9695	0.9526	0.9115	1.0000						
__e7	0.8462	0.6903	0.9269	0.8972	0.8042	0.9253	1.0000					
__e8	0.8664	0.7149	0.9459	0.9047	0.8324	0.9447	0.9619	1.0000				
__e9	0.8364	0.6334	0.9129	0.8673	0.7778	0.9500	0.9639	0.9646	1.0000			
__e10	0.6566	0.6468	0.6891	0.5997	0.5986	0.6382	0.6464	0.6633	0.6414	1.0000		
__e11	0.7023	0.4467	0.8122	0.7382	0.6234	0.8549	0.9138	0.9058	0.9510	0.5429	1.0000	
__e12	0.7058	0.4799	0.8059	0.7478	0.6299	0.8362	0.9279	0.8899	0.9461	0.6036	0.9620	1.0000

Breusch-Pagan LM test of independence: chi2(66) = 1355.144, Pr = 0.0000
Based on 30 complete observations over panel units

P 值非常小，强烈拒绝原假设，**认为组间存在同期相关**；从残差相关系数矩阵中也可看出，组间同期相关系数靠近 1，表现出很高的相关性

3. 回归结果

Model .1

聚类稳健标准误的 OLS 回归

Model .2

面板校正标准误 PCSE 回归 (Panel-Corrected Standard Error) (**考虑了组间异方差和组间同期相关**)

Model .3

服从 **AR(1)**过程的组内自相关 **FGLS** (Feasible Generalized Least Square) (**只考虑了组内自相关**) (认为 $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + v_{it}$ 中的 ρ_i **自回归系数，对所有省份都相同**，即 $\rho_i = \rho$)

Model .4

服从 **PSAR(1)**过程的组内自相关 **FGLS** (Feasible Generalized Least Square) (**只考虑了组内自相关**) (认为 $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + v_{it}$ 中，**每个省份可以有自己的 ρ_i 自回归系数**)

结果分析：

1. roadshow: 显著性上，四种模型中 **roadshow** 系数在 95%置信度下都**不显著**；取值上：**增加一次 roadshow 对 sales 的提升在 50 左右**，但有很大的标准误（足以让提升变为负）
2. 各省份销量：除广东外，其他各省份销量在 99%置信度小明显低于上海，从北京到湖南销量越来越少；广东省的销量跟上海最为接近，在仅考虑组内自相关的 AR(1) 和 PSAR(1)模型中，与上海的销量没有显著差异
3. t: 四种模型均显示出了 **sale 随时间增长**的趋势,每月的增长在 **20** 左右，但在只考虑组内自相关的 AR(1)和 PSAR(1)两模型中不显著
4. constant: **常数项值在 4000 左右，且均显著**，AR(1)和 PSAR(1)中常数项的标准误超过 500，OLS 和 PCSE 的标准误相对较小

```
. esttab OLS PCSE AR1 PSAR1,r2 se mtitles star(* 0.1 ** 0.05 *** 0.01)
```

	(1) OLS	(2) PCSE	(3) AR1	(4) PSAR1
roadshow	58.04 (140.6)	58.04 (231.5)	50.40 (136.8)	58.03 (139.1)
1.province	0 (.)	0 (.)	0 (.)	0 (.)
2.province	-208.2*** (4.687)	-208.2* (109.9)	-276.5 (211.2)	-281.2 (228.2)
3.province	-806.8*** (4.687)	-806.8*** (76.23)	-834.2*** (180.9)	-829.0*** (154.1)
4.province	-1716.0*** (9.374)	-1716.0*** (106.8)	-1735.8*** (231.8)	-1733.4*** (210.8)
5.province	-2401.5*** (4.687)	-2401.5*** (129.2)	-2419.0*** (283.8)	-2414.5*** (242.3)
6.province	-2979.6*** (4.687)	-2979.6*** (164.1)	-3007.1*** (342.0)	-3004.2*** (321.4)
7.province	-3167.2*** (4.687)	-3167.2*** (175.2)	-3211.5*** (375.3)	-3207.5*** (307.5)
8.province	-3299.3*** (4.687)	-3299.3*** (182.7)	-3344.5*** (387.2)	-3337.8*** (342.1)
9.province	-3400.4*** (6.66e-12)	-3400.4*** (183.6)	-3439.5*** (384.8)	-3433.9*** (339.0)
10.province	-3426.4*** (4.687)	-3426.4*** (191.6)	-3483.2*** (413.5)	-3472.7*** (372.3)
11.province	-3667.4*** (4.687)	-3667.4*** (197.0)	-3704.6*** (413.5)	-3699.1*** (347.5)
12.province	-3728.5*** (4.687)	-3728.5*** (197.1)	-3770.5*** (419.1)	-3768.4*** (344.2)
t	25.36*** (5.803)	25.36** (10.96)	14.33 (20.33)	14.20 (20.27)
_cons	3882.3*** (89.99)	3882.3*** (279.8)	4090.8*** (569.8)	4086.3*** (525.8)
N	360	360	360	360
R-sq	0.801	0.801	0.511	0.530

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Model .5

全面 FGLS (考虑组间异方差 , 组间同期相关 , 组内自相关 (ρ_i 相同))

Model .6

全面 FGLS (考虑组间异方差 , 组间同期相关 , 组内自相关 (ρ_i 不同))

结果分析：

1. roadshow: 显著性上 , 两种模型中 **roadshow 系数**同样都**不显著** ; 取值上 : **增加一次 roadshow 对 sales 的提升不足 20** , 但标准误也相对较小
2. 各省份销量 : 除广东外其他各省份销量仍然在 99% 置信度小明显低于上海 , 各省份销量大致情况与前 4 种模型类似 ; 广东省的销量跟上海最为接近 , 且与上海的销量没有显著差异
3. t: 两模型种 **sales 随时间都显著增长**, 每月的增长为 **13~14**
4. **constant:** 常数项值均**超过 4000 , 且均显著** , 标准误在 400 左右

	(1)	(2)
	FGLS_AR1	FGLS_PSAR1
roadshow	19.12 (12.93)	14.51 (13.98)
1.province	0 (.)	0 (.)
2.province	-277.4 (211.2)	-340.7 (223.8)
3.province	-835.2*** (180.8)	-820.0*** (152.1)
4.province	-1733.9*** (231.7)	-1701.3*** (203.8)
5.province	-2418.1*** (283.8)	-2382.3*** (231.7)
6.province	-3008.0*** (342.0)	-2945.2*** (304.4)
7.province	-3212.4*** (375.2)	-3168.0*** (291.6)
8.province	-3345.4*** (387.2)	-3284.4*** (323.7)
9.province	-3437.7*** (384.7)	-3374.9*** (320.8)
10.province	-3484.1*** (413.5)	-3421.6*** (352.7)
11.province	-3703.8*** (413.5)	-3645.6*** (328.5)
12.province	-3769.7*** (419.1)	-3721.7*** (325.3)
t	13.00*** (1.990)	14.13*** (2.383)
_cons	4113.2*** (476.0)	4018.3*** (398.8)
N	360	360
R-sq		

Standard errors in parentheses

* p<0.1, ** p<0.05, *** p<0.01

Model.7

变系数模型——每个省份的 roadshow 的系数不同，并视为常数

```
. reg sales roadshow i.province i.province#c.roadshow t,vce(cluster province)
```

```
Linear regression               Number of obs   =           360
                               F(0, 11)         =             .
                               Prob > F          =             .
                               R-squared          =           0.8071
                               Root MSE       =           678.41
```

(Std. Err. adjusted for 12 clusters in province)

sales	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
roadshow	-964.6773	57.86603	-16.67	0.000	-1092.04	-837.315
province						
2	-313.0912	2.71963	-115.12	0.000	-319.0771	-307.1054
3	-891.6429	2.71963	-327.85	0.000	-897.6288	-885.6571
4	-1813.712	2.728642	-664.69	0.000	-1819.718	-1807.706
5	-2491.625	.4195173	-5939.27	0.000	-2492.548	-2490.702
6	-3020.879	2.926558	-1032.23	0.000	-3027.32	-3014.438
7	-3241.234	2.512701	-1289.94	0.000	-3246.765	-3235.704
8	-3365.781	2.71963	-1237.59	0.000	-3371.767	-3359.795
9	-3451.225	.4286373	-8051.62	0.000	-3452.169	-3450.282
10	-3496.712	2.71963	-1285.73	0.000	-3502.698	-3490.726
11	-3721.449	.8572746	-4341.02	0.000	-3723.336	-3719.562
12	-3780.311	.8572746	-4409.69	0.000	-3782.198	-3778.424
province#c.roadshow						
2	2124.059	23.72286	89.54	0.000	2071.846	2176.273
3	1523.611	23.72286	64.23	0.000	1471.397	1575.825
4	1244.428	8.4682	146.95	0.000	1225.79	1263.067
5	1242.144	15.0935	82.30	0.000	1208.923	1275.364
6	214.6952	29.93071	7.17	0.000	148.8181	280.5722
7	1198.354	17.51501	68.42	0.000	1159.804	1236.905
8	970.7489	23.72286	40.92	0.000	918.5353	1022.963
9	762.3771	6.429559	118.57	0.000	748.2258	776.5285
10	1086.68	23.72286	45.81	0.000	1034.466	1138.894
11	599.7846	32.1478	18.66	0.000	529.0278	670.5415
12	530.6467	32.1478	16.51	0.000	459.8899	601.4035
t	25.84802	6.000922	4.31	0.001	12.64008	39.05596
_cons	3942.901	89.15655	44.22	0.000	3746.669	4139.133

分析：

所有解释变量均非常显著，回归效果较好；

t：month 每增加 1，sales 显著增加约 25；

roadshow:

$$\frac{\partial sales}{\partial roadshow} = \beta_{roadshow} + \beta_{roadshowprovince} \cdot province$$

此时**增加一次 roadshow 对 sales 的影响**可由以上两个系数相加得到，如下表：

Province	Code	Influence
Shanghai	1	Base group
Guangdong	2	1159.382
Beijing	3	558.9337
Jiangsu	4	279.7507
Zhejiang	5	277.4667
Sichuan	6	-749.982
Liaoning	7	233.6767
Shandong	8	6.0716
Hubei	9	-202.3
Chongqing	10	122.0027
Henan	11	-364.893
Hunan	12	-434.031

以 Guangdong 为例，1159.382 代表在广东省，增加 1 次 roadshow 会使 sales 相比上海市来讲增加 1159.382；以此类推，roadshow 在**广东，北京，江苏，浙江，辽宁，重庆**对 sales 具有明显的**促进作用**，在**四川，湖北，河南，湖南**具有有**明显副作用**，在**山东**对 sales 的**促进作用较小**。

该模型可能的缺陷：

命令中 i.province#c.roadshow 中的 c 是 continuous 的简称，即将 roadshow 视为连续性变量，但因为大部分省份 roadshow 取值为 0/1，只有一个省份出现了 2 次 roadshow，所以将**roadshow 看成连续性变量，而不是虚拟变量也许有一些不合适**。

Model.8

变系数模型——随机系数模型：每个省份的 roadshow 的系数不同，并视为随机变量

```
. xtrc sales roadshow t, betas
```

Random-coefficients regression
Group variable: province

Number of obs = 360
Number of groups = 12

Obs per group:

min = 30
avg = 30.0
max = 30

Wald chi2(2) = 16.72
Prob > chi2 = 0.0002

sales	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
roadshow	-3.08234	194.2364	-0.02	0.987	-383.7787	377.614
t	25.77578	6.372877	4.04	0.000	13.28517	38.26639
_cons	1465.264	344.3722	4.25	0.000	790.3072	2140.221

Test of parameter constancy: chi2(33) = 1137.43 Prob > chi2 = 0.0000

Group-specific coefficients

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Group 1						
roadshow	-236.4486	300.634	-0.79	0.432	-825.6804	352.7831
t	45.33567	7.598206	5.97	0.000	30.44346	60.22788
_cons	3517.694	236.6877	14.86	0.000	3053.794	3981.593
Group 2						
roadshow	301.5206	282.5097	1.07	0.286	-252.1883	855.2295
t	56.2777	7.74595	7.27	0.000	41.09591	71.45948
_cons	3155.701	222.7433	14.17	0.000	2719.132	3592.27
Group 3						
roadshow	102.978	293.6085	0.35	0.726	-472.4841	678.4401
t	35.09253	7.894538	4.45	0.000	19.61952	50.56554
_cons	2882.41	206.1619	13.98	0.000	2478.34	3286.48
Group 4						
roadshow	173.366	269.766	0.64	0.520	-355.3657	702.0976
t	31.36618	7.617468	4.12	0.000	16.43622	46.29614
_cons	2048.367	182.7157	11.21	0.000	1690.25	2406.483

Group 5							
roadshow	192.853	242.3359	0.80	0.426	-282.1166	667.8226	
t	32.53247	7.61195	4.27	0.000	17.61332	47.45162	
_cons	1360.128	158.0093	8.61	0.000	1050.435	1669.82	
Group 6							
roadshow	-499.4087	281.0252	-1.78	0.076	-1050.208	51.39065	
t	19.20264	6.450599	2.98	0.003	6.559696	31.84558	
_cons	1018.134	120.2153	8.47	0.000	782.5161	1253.752	
Group 7							
roadshow	203.1049	270.5067	0.75	0.453	-327.0785	733.2883	
t	13.99322	6.143908	2.28	0.023	1.951386	26.03506	
_cons	886.8191	113.3535	7.82	0.000	664.6503	1108.988	
Group 8							
roadshow	36.86219	245.2774	0.15	0.881	-443.8727	517.597	
t	16.80048	5.449927	3.08	0.002	6.118816	27.48214	
_cons	717.4381	99.00038	7.25	0.000	523.4009	911.4752	
Group 9							
roadshow	-73.60059	193.4401	-0.38	0.704	-452.7362	305.5351	
t	13.95878	5.388749	2.59	0.010	3.39703	24.52054	
_cons	668.3246	95.75711	6.98	0.000	480.6441	856.0051	
Group 10							
roadshow	110.6208	267.5003	0.41	0.679	-413.6701	634.9116	
t	21.2234	6.055127	3.51	0.000	9.355571	33.09123	
_cons	521.0646	111.5091	4.67	0.000	302.5108	739.6183	
Group 11							
roadshow	-135.0732	189.7639	-0.71	0.477	-507.0035	236.8571	
t	11.045	4.086701	2.70	0.007	3.035214	19.05479	
_cons	443.8546	71.34152	6.22	0.000	304.0278	583.6815	
Group 12							
roadshow	-213.7624	190.2742	-1.12	0.261	-586.693	159.1681	
t	12.4813	4.09858	3.05	0.002	4.448233	20.51437	
_cons	363.2379	71.56177	5.08	0.000	222.9794	503.4964	

分析：

roadshow:

12 个省份中 roadshow 系数在 95%置信度上均不显著；取值上，roadshow 对 sales **促进**作用从大到小为**广东(G2)> 辽宁(G7)> 浙江(G5)> 江苏(G4)>重庆(G10)>北京(G3)>山东(G8)>湖北(9)>河南(11)**,**抑制**作用从大到小为**四川(G6)>上海(G1)>湖南(G12)>河南(G11)>湖北(G9)**

t:

所有 **t 值**均在 99%置信水平**显著**，t 值从大到小排列为：

广东 56>上海 45>北京 35>浙江 32>江苏 31>重庆 21>四川 19>山东 16>辽宁 13.99322>湖北 13.95878>湖南 12.4813>河南 11.045