

The empirical method of economic research and stata code

在做任何分析之前都要做协变量平衡分析，防止由于对照组和控制组变量分布造成的误差。

Instrument Variable

我们在使用工具变量时，需要进行检验，最常见的就是排除性和相关性。

进行IV时我们需要讲故事，并且数据检验其合理性：

弱工具变量检验

1. F检验

```
reg y x ,robust // OLS回归估计
ivregress 2sls y (x=z1,z2),robust // 2SLS回归估计
reg x z1 z2,robust // 第一阶段回归估计
test z1 z2 //查看是否有弱工具变量问题·F检验 大于10即可 F估计与弱IV的关系来自于
causal inference
```

可以通过以上的第一阶段回归查看第一阶段的参数从而判断工具变量的相关性

也可以比较OLS和2SLS的结果，看看是否有差异

2. Cragg-Donald检验

一般条件是同方差，无自相关 假设扰动项为iid，则看CDW检验统计量 CDW检验一般过15%，10%的临界值就可以，过了5%的临界值更好

```
ivreg2 y (x1 x2 = z1 z2), robust //Cragg-Donald检验,要大于 10
```

3. Kleibergen-Paap检验 无iid假设 不对扰动项作iid的假设，则看KP W rk F统计量

```
ivreg2 y (x1 x2 = z1 z2), robust //Kleibergen-Paap检验,要大于 10
```

外生性（排除性）检验

1. Hausman检验

```
//豪斯曼检验 这是在同方差条件下的检验
reg y x1 x2
estimates store ols
ivregress 2sls y (x1 = z) x2
estimates store iv
hausman iv ols, constant sigmamore
//chi - squared和p - value。p 小于0.05 · 拒原 · 认为变量是内生变量,p最好大一点
```

2. DWH检验

用上一个检验的结果就行，也会输出DWH检验的结果。这是在异方差条件下的检验

3. GMM估计 用于异方差条件下的检验工具变量

```
ivregress gmm y (x1 = z1 z2), twostep robust
estat overid //使用结果与2SLS的结果进行比较，其假设是工具变量是外生的，所以结果
相似就可以认为是外生的，也可以通过hassen J值 其P值只要大于0.05就可以认为是外生的
```

过度识别检验

1. Sargan检验 用于线性模型中的工具变量过度识别检验 要求p大于0.1 iid时用Sargan统计量

```
ivregress 2sls y (x1 = z1 z2) 其P值只要大于0.05就可以认为是外生的
```

加选项robust时汇报Hansen J统计量

2. Anderson - Rubin 检验 用于非线性模型或联立方程模型中的工具变量过度识别检验 以联立方程模型为例

```
sysreg (eq1: y1 = x1 x2 (y2 = z1 z2)) (eq2: y2 = x3 x4 (y1 = z3 z4))
test [eq1_y2] [eq2_y1] // 原假设是不存在过度识别问题
```

3. Hansen J统计量 非iid时用Hansen J统计量 和Sargon检验类似 非iid时用Hassen统计量