Statistics Methods in Finance

Homework 8

0786009 廖家鴻

DUE 2020/12/24 00:00

Outline (HW8 questions)

1.(100%) Run an instrumental regression with Y = wage, X = educ, and IV = sibs.

1. Run an instrumental regression with Y = wage, X = educ, and IV = sibs.

```
from linearmodels import IV2SLS
df['const'] = 1
IVR = IV2SLS(dependent=df.wage,
             exog=df.const,
             endog=df.educ,
             instruments=df.sibs).fit(cov_type='unadjusted')
print(IVR.summary)
```

As shown by the code in the left figure, let wage be the dependent variable, educ be the endogenous variable, and the sibs be the instrumental variable

```
                          IV-2SLS Estimation Summary
==============================================================================
Dep. Variable:                   wage   R-squared:                     -0.0074
Estimator:                    IV-2SLS   Adj. R-squared:                -0.0085
No. Observations:                 935   F-statistic:                    23.525
Date:                Wed, Dec 16 2020   P-value (F-stat)                0.0000
Time:                        22:57:19   Distribution:                  chi2(1)
Cov. Estimator:           unadjusted

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
const        -691.57     340.35    -2.0319     0.0422    -1358.6    -24.498
educ          122.47     25.251     4.8502     0.0000     72.982     171.96
==============================================================================

Endogenous: educ
Instruments: sibs
Unadjusted Covariance (Homoskedastic)
Debiased: False
```

After conducting instrumental regression, both the intercept and beta for educ are significant

Other testing

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   wage   R-squared:                       0.107
Model:                            OLS   Adj. R-squared:                  0.106
Method:                 Least Squares   F-statistic:                     111.8
Date:                Wed, 23 Dec 2020   Prob (F-statistic):           9.35e-25
Time:                        20:07:50   Log-Likelihood:                -6885.5
No. Observations:                 935   AIC:                         1.377e+04
Df Residuals:                     933   BIC:                         1.378e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        146.9524     77.715      1.891      0.059      -5.564     299.469
educ          60.2143      5.695     10.573      0.000      49.038      71.391
==============================================================================
Omnibus:                      197.993   Durbin-Watson:                   1.821
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              530.455
Skew:                           1.083   Prob(JB):                    6.51e-116
Kurtosis:                       5.987   Cond. No.                         85.3
==============================================================================
```

```
import statsmodels.api as sm
X1 = sm.add_constant(df.educ)
LR = sm.OLS(df.wage, X1).fit()
print(LR.summary())
```

Conduct the OLS regression, the intercept seems not significant.

```
Error_term = df.wage - (LR.params.const + LR.params.educ*df.educ)

from scipy import stats
corr, pval = stats.pearsonr(df.educ, Error_term**2)
```

```
In [16]: print(f'correlation={corr:.4f}, p-value={pval:.4f}')
correlation=0.1377, p-value=0.0000
```

By examining the correlation between the educ and the square of error term, I can find that the educ is endogenous variable.