ECN 102, Summer 2020

Week 4 Recap
OLS Tests

RESET Test 1

- Suppose we want to explain wage with education, IQ, number of siblings, and birthorder

$$\widehat{wage} = b_1 + b_2 educ + b_3 IQ + b_4 sibs + b_5 brthord$$

- Are we missing any nonlinear variables? I dunno, let's test it.

- Key insight: $\widehat{wage}$ is a function of $educ, IQ, sibs, brthord$

- Therefore $\widehat{wage}^2$ contains a lot of nonlinear functions of these variables (as does $\widehat{wage}^3$ and so on)

- Try the regression

$$\widehat{wage} = b_1 + b_2 educ + b_3 IQ + b_4 sibs + b_5 brthord + a_1 \widehat{wage}^2 + a_2 \widehat{wage}^3$$

- If a_1 and a_2 are jointly zero, then all of that nonlinear stuff is insignificant, so we feel more justified using the first regression

RESET Test in R

- Null $H_0 : \alpha_1 = \alpha_2 = 0$, alternative H_1 : at least one of $\alpha_1, \alpha_2 \neq 0$

```
ols1 = lm(wage ~ educ+iq+sibs+brthord, data = wages)
resettest(ols1)

        RESET test

data:  ols1
RESET = 0.63518, df1 = 2, df2 = 845, p-value = 0.5301
```

- p-value is large, so we fail to reject the null hypothesis: we seem to be okay doing the regression without the nonlinear terms
- By default `resettest()` tests for significance of $\hat{y}^2$ and $\hat{y}^3$
- Can specify powers, e.g. `resettest(ols1, power = 2:4)` will test second through fourth powers of $\hat{y}$

Jarque-Bera Test for Normality of Disturbances

- The null hypothesis is that disturbances are normal; the alternative hypothesis is that disturbances are not normal
- What do we know about normal distributions? Have zero skew and zero excess kurtosis
- We use the test statistic

$$\mathrm{JB} \equiv n \left[\frac{\widehat{\mathrm{skew}}^2}{6} + \frac{(\widehat{\mathrm{kurt}} - 3)^2}{24} \right] \sim \chi^2(2)$$

- If non-normal, then at least one of $\widehat{\mathrm{skew}}^2$ and $(\widehat{\mathrm{kurt}} - 3)^2$ will be positive, in which case JB is positive
- Therefore reject the null when JB is sufficiently far from zero in the positive direction

Jarque-Bera Test in R

- The null hypothesis is that disturbances are normal; the alternative hypothesis is that disturbances are not normal

```
jarque.bera.test(ols1$residuals)

        Jarque Bera Test

data:  ols1$residuals
X-squared = 532.92, df = 2, p-value < 2.2e-16
```

- *p*-value is essentially zero, so reject the null hypothesis and conclude that disturbances are not normally distributed

Breusch-Pagan Heteroskedasticity Test

- Can be shown that $\text{Var}(\epsilon) = E[\epsilon^2]$ when OLS assumptions 1-2 hold
- If homoskedastic, then $E[\epsilon^2]$ should not depend on $x_2, \ldots, x_k$
- If homoskedastic, then

$$\epsilon^2 = \alpha_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k + \eta$$

  should have no overall significance, that is, $H_0 : \alpha_2 = \ldots = \alpha_k = 0$
- Alternative is that ϵ^2 does depend on regressors, so
  H_1 : at least one of $\alpha_2, \ldots, \alpha_k \neq 0$

Breusch-Pagan Heteroskedasticity Test in R

- $wage = \beta_1 + \beta_2 educ + \beta_3 IQ + \beta_4 sibs + \beta_5 brthord + \epsilon$
- $\epsilon^2 = \alpha_1 + \alpha_2 educ + \alpha_3 IQ + \alpha_4 sibs + \alpha_5 brthord + \eta$
- $H_0 : \alpha_2 = \ldots = \alpha_5 = 0$ against H_1 : at least one of $\alpha_2, \ldots, \alpha_5 \neq 0$

```
ols1 = lm(wage ~ educ+iq+sibs+brthord, data = wages)
esq = (ols1$residuals)^2

olsbp = lm(esq ~ educ+iq+sibs+brthord, data = wages)
stargazer(olsbp, type = "text")

-----------------------------------------------
Observations                     852
R2                              0.021
Adjusted R2                     0.016
Residual Std. Error   316,474.600 (df = 847)
F Statistic           4.560*** (df = 4; 847)
===============================================
Note:                 *p<0.1; **p<0.05; ***p<0.01
```

- F-statistic significant, reject the null: conclude heteroskedasticity