# Economics 471: Introductory Econometrics

Department of Economics, Finance and Legal Studies

University of Alabama

Fall 2019

Final Exam - Key

The exam consists of four questions on six pages. Each question is of equal value.

1. Consider a regression model without a regressor: $y_i = \alpha + u_i,\ i = 1, 2, \ldots, n$. Given this information, answer the following:

   (a) Derive the method of moments estimator of $\alpha$.

   (b) Derive the variance of the estimator achieved in part (a).

   (c) For this model, prove that the goodness-of-fit measure, $R^2 = 0$.

   (d) Suppose that there were two groups in the population (group 0 and group 1) and we wished to account for this in our model by using a dummy variable ($D$) via $y_i = \alpha + \delta D_i + u_i$. Derive the ordinary least-squares estimators of $\alpha$ and $\delta$?

   (e) For the model in part (d), how would you test the null hypothesis that there is no difference between the two groups (state the null, test statistic, distribution of the test statistic, and decision rule)?

(a) $E(u)=0 \Rightarrow \frac{1}{n}\sum_{i=1}^{n}\hat{u}_i = 0 \Rightarrow \frac{1}{n}\sum_{i=1}^{n} y_i - \hat{\alpha} = 0 \Rightarrow \hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} y_i$

(b) $V(\hat{\alpha}) = V\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) \overset{iid}{=} \frac{1}{n^2}\sum_{i=1}^{n} V(y_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$

(c) $R^2 = \frac{SSE}{SST} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}(\bar{y} - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 0$

(d) $\min_{\alpha,\delta} \sum_{i=1}^{n}\hat{u}_i^2 \Rightarrow \frac{\partial}{\partial \hat{\alpha}} = -2\sum_{i=1}^{n}(y_i - \hat{\alpha} - \hat{\delta} D_i) = 0$

$\frac{\partial}{\partial \hat{\beta}} = -2\sum_{i=1}^{n}(y_i - \hat{\alpha} - \hat{\delta} D_i) D_i = 0$

$\hat{\alpha} = \bar{y} - \hat{\delta}\bar{D} \qquad \hat{\delta} = \sum_{i=1}^{n}(y_i - \bar{y})(D_i - \bar{D}) \Big/ \sum_{i=1}^{n}(D_i - \bar{D})^2$

(e) $H_0: \delta = 0$



$t = \frac{\hat{\delta} - 0}{se(\hat{\delta})} \sim t_{n-2}$ if $|t| > t_{n-2,\alpha} \Rightarrow$ reject $H_0$

2. Consider the $F-$statistic discussed in class

$$F = \frac{(SSR_R - SSR_U)/q}{SSR_U/(n-k-1)}$$

(a) Define each component on the right hand side of the equation.

(b) What is the distribution of this test statistic (be sure to list the degrees of freedom)? Draw this distribution.

(c) What is the range of the test statistic? Why?

(d) Using this $F-$statistic, derive the $F-$statistic in terms of the $R^2$ formulation. Show your work.

(e) Define $SST$. Does $SST$ need a subscript for $R$ or $U$? If so, why? If not, why not?

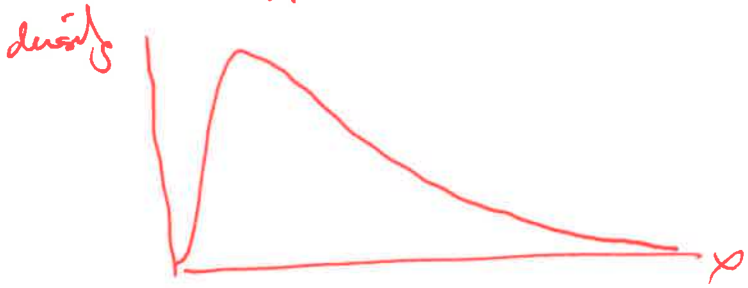
(a) $SSR_R$ - SSR from restricted model (null)

$SSR_u$ - SSR from unrestricted model (alternate)

$q$ - numerator df (# of restrictions in null)

$(n-k-1)$ - denominator df (# of obs minus # of estimated parameters)

(b) $F \sim F_{q,n-k-1}$




(c) 0 to $\infty$ because $SSR_R \geq SSR_u$

(d) $$F = \frac{\left(\frac{SSR_R}{SST} - \frac{SSR_u}{SST}\right)/q}{\left(\frac{SSR_u}{SST}\right)/(n-k-1)} = \frac{[(1-R^2_R) - (1-R^2_u)]/q}{[(1-R^2_u)]/(n-k-1)} = \frac{(R^2_u - R^2_R)/q}{(1-R^2_u)/(n-k-1)}$$



(e) $SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$ is independent of the model (no R and u)

3. Consider the population regression function $y = \alpha + \beta x + \gamma x^2 + \delta D + \psi Dx + \chi Dx^2 + u$, where $D$ is a dummy variable. Assuming $\alpha > 0$, $\beta > 0$, $\gamma < 0$, $\delta > 0$, $\psi < 0$, and $\chi > 0$ (where $|\psi| > \beta$ and $|\chi| > \gamma$), in the figure below, perform the following:

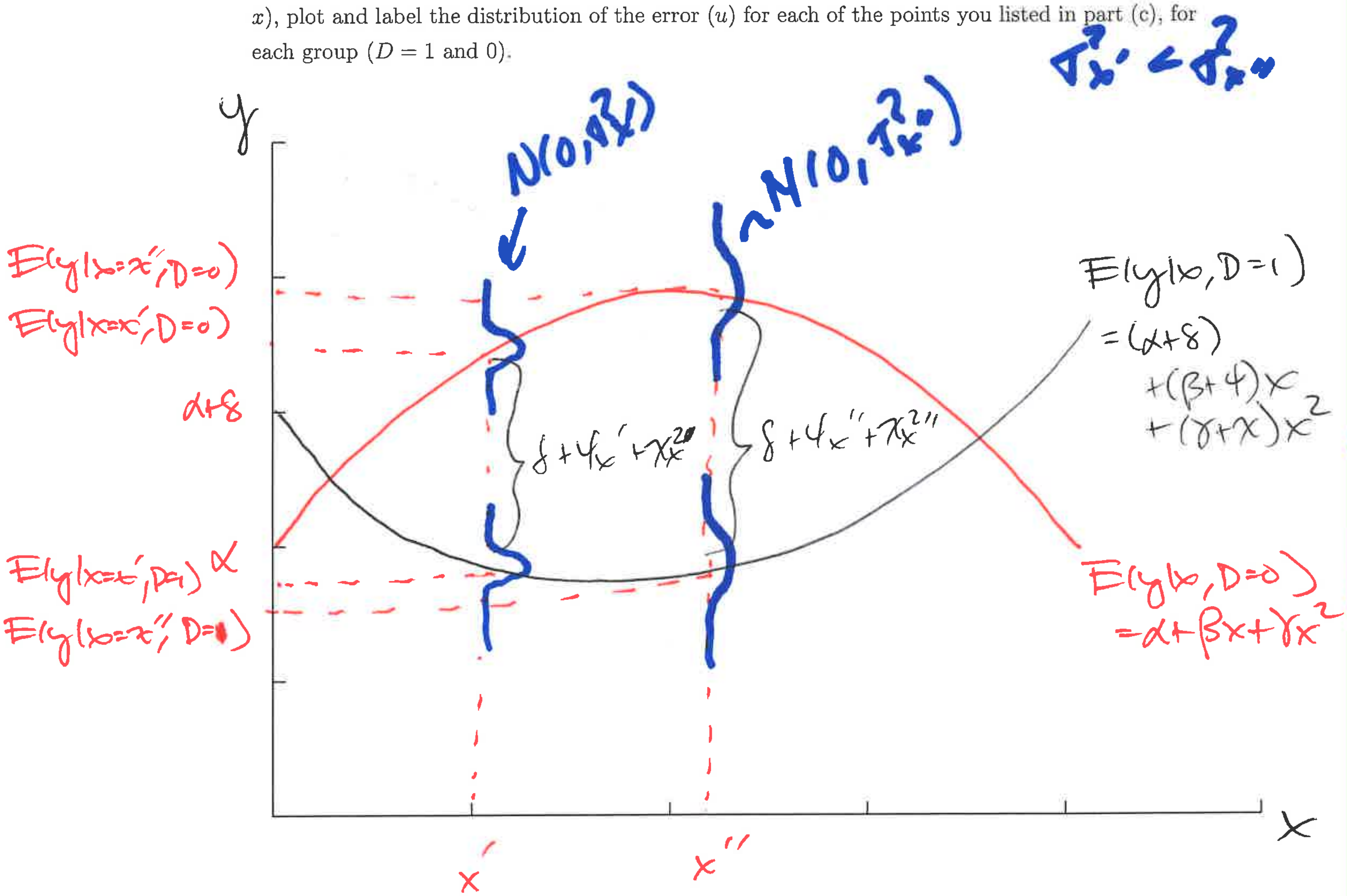
   (a) Label the axes.

   (b) Plot and label the population regression curves for group 1 ($D = 1$) and group 0 ($D = 0$).

   (c) Pick two values for $x$, plot their conditional expectations (i.e., $E(y|x)$) for each group ($D = 1$ and 0).

   (d) For each of those two values of $x$ in part (c), what is the expected impact on $y$ of going from group 0 to group 1?

   (e) Assuming normally distributed, but heteroskedastic errors (assume a variance that increases with $x$), plot and label the distribution of the error ($u$) for each of the points you listed in part (c), for each group ($D = 1$ and 0).



4. List the six Gauss-Markov Assumptions that we discussed in class. Then, consider the four pieces of gretl output below. For each piece of output, list the assumption(s) that is (are) violated. Briefly mention how each violation can be corrected?

(a) (1) $y = \alpha + \beta x + u$

(2) $E(u) = 0$

(3) $E(u|x) = 0$

(4) no perfect collinearity

(5) $V(u|x) = \sigma^2$ . homoskedasticity

(6) $u|x \overset{iid}{\sim} N(0, \sigma^2)$

(b) violated (4)
remove $X_2$

Model 1: OLS, using observations 1–500
Dependent variable: y
Omitted due to exact collinearity: x2

| | Coefficient | Std. Error | $t$-ratio | p-value |
|---|---|---|---|---|
| const | 2.96197 | 0.144189 | 20.54 | 0.0000 |
| x1 | 2.95907 | 0.142094 | 20.82 | 0.0000 |

| | | | |
|---|---|---|---|
| Mean dependent var | 2.929732 | S.D. dependent var | 4.405280 |
| Sum squared resid | 5176.237 | S.E. of regression | 3.223981 |
| $R^2$ | 0.465477 | Adjusted $R^2$ | 0.464403 |
| $F(1,498)$ | 433.6713 | P-value($F$) | 9.59e–70 |
| Log-likelihood | −1293.776 | Akaike criterion | 2591.551 |
| Schwarz criterion | 2599.981 | Hannan–Quinn | 2594.859 |

(c) violate (1) & (3)
add a quadratic term $X_1^2$

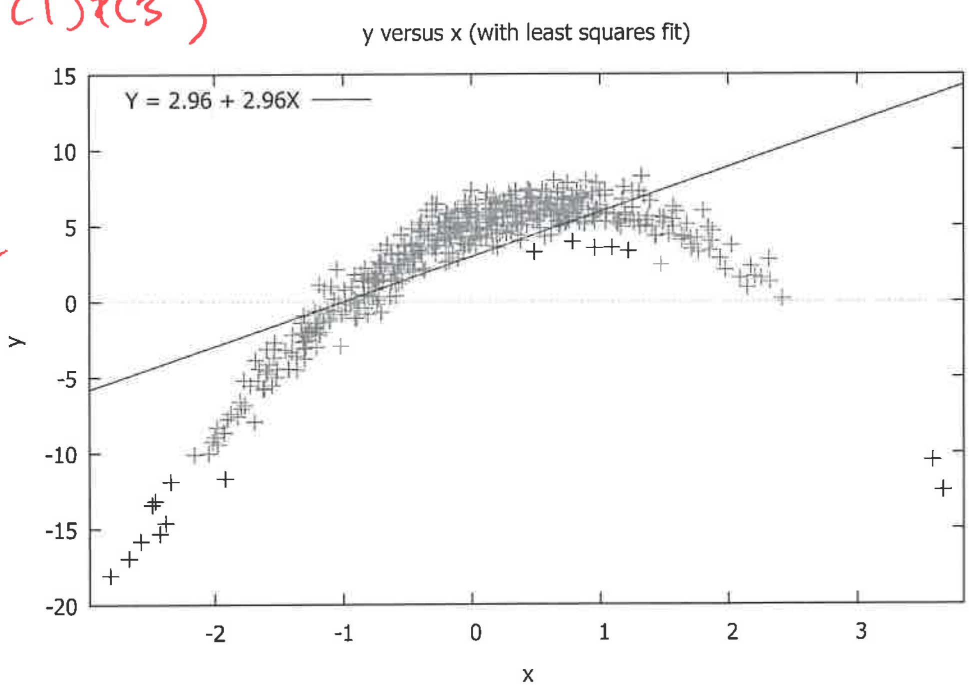



Figure 1: Scatterplot of $y$ versus $x$ with a regression line

(d) violate (5) & (6)

use robust standard errors

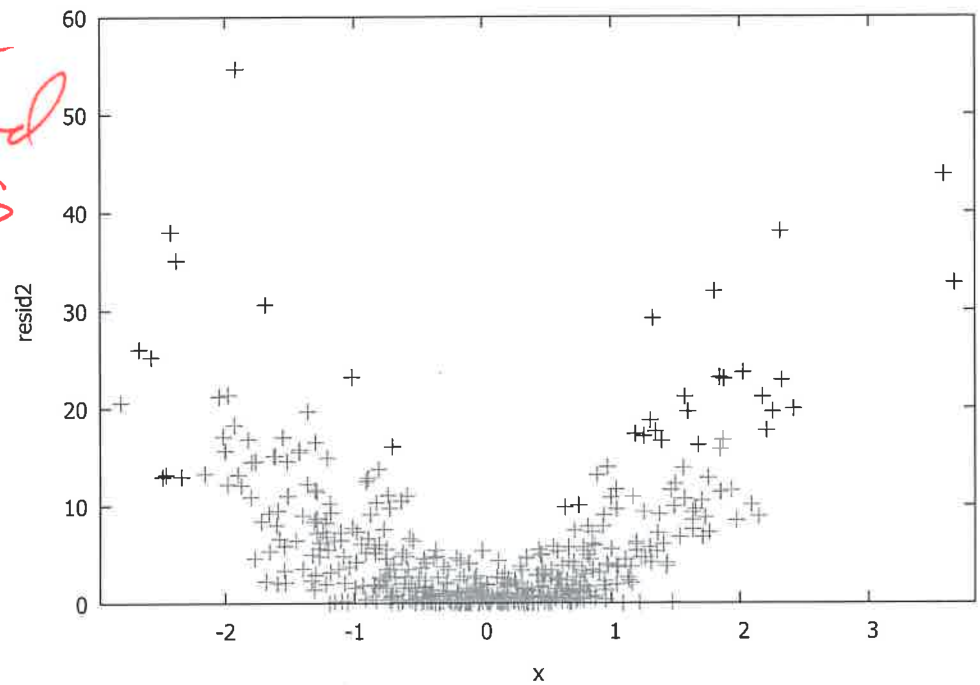



Figure 2: Scatterplot of $\widehat{u}^2$ (squared residuals) versus $x$

(e) violate (6)

bootstrap test stat

(full credit if say violate 6 solely)

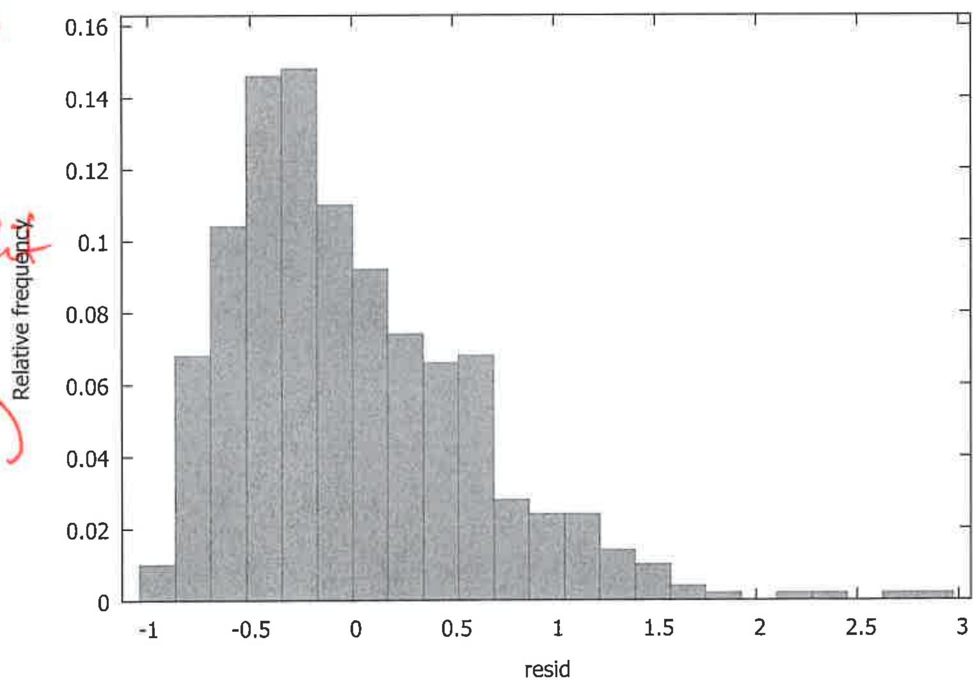



Figure 3: Histogram of $\widehat{u}$ (residuals)