BS2280 – Econometrics I
Homework 9: Nonlinear Models and Transformation of Variables I

1

Until now we have assumed that our regression model is linear in variables and parameters. Explain what this means.

2

It has often been observed that there is a weak tendency for years of schooling to be inversely related to the number of siblings (brothers and sisters) of an individual. The regression shown below has been fitted on the hypothesis that the adverse effect is nonlinear. Z is defined as the reciprocal of the number of siblings, for individuals with at least one sibling. Sketch the regression relationship and provide an interpretation of the regression results.

```
> EAWE21$Z <- 1 / EAWE21$SIBLINGS
> Sfit <- lm(S~Z,data=subset(EAWE21, SIBLINGS>0))
> summary(Sfit)

Call:
lm(formula = S ~ Z, data = subset(EAWE21, SIBLINGS > 0))

Residuals:
    Min      1Q  Median      3Q     Max
-7.5028 -2.4569  0.4972  1.8917  5.6738

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  13.9340     0.2710   51.41  < 2e-16 ***
Z             1.5688     0.4023    3.90  0.00011 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.727 on 471 degrees of freedom
Multiple R-squared:  0.03128,   Adjusted R-squared:  0.02923
F-statistic: 15.21 on 1 and 471 DF,  p-value: 0.0001102
```

3

The output below shows the result of regressing $LGWT04$, the logarithm of weight in 2004, measured in pounds, on $LGHEIGHT$, the logarithm of height, measured in inches. Provide an interpretation of the slope coefficient and evaluate the regression results.

```
> EAWE21$LGWT04 <- log(EAWE21$WEIGHT04)
> EAWE21$LGHEIGHT <- log(EAWE21$HEIGHT)
> LGWT04fit <- lm(LGWT04~LGHEIGHT,data=EAWE21)
> summary(LGWT04fit)

Call:
lm(formula = LGWT04 ~ LGHEIGHT, data = EAWE21)

Residuals:
     Min       1Q   Median       3Q      Max
-0.40323 -0.13720 -0.03225  0.10760  0.61840

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -3.7883     0.6109  -6.201 1.18e-09 ***
LGHEIGHT      2.1064     0.1449  14.536  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1935 on 498 degrees of freedom
Multiple R-squared:  0.2979,    Adjusted R-squared:  0.2965
F-statistic: 211.3 on 1 and 498 DF,  p-value: < 2.2e-16
```

4

The output below shows the result of regressing $LGWT04$, the logarithm of weight in 2004, measured in pounds, on $HEIGHT$, height measured in inches. Provide an interpretation of the slope coefficient and evaluate the regression results.

```
> LGWT04fit2 <- lm(LGWT04~HEIGHT,data=EAWE21)
> summary(LGWT04fit2)

Call:
lm(formula = LGWT04 ~ HEIGHT, data = EAWE21)

Residuals:
     Min       1Q   Median       3Q      Max
-0.40150 -0.13819 -0.03537  0.10552  0.62194

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.988730   0.145213   20.58   <2e-16 ***
HEIGHT      0.030990   0.002137   14.50   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1936 on 498 degrees of freedom
Multiple R-squared:  0.2969,    Adjusted R-squared:  0.2955
F-statistic: 210.3 on 1 and 498 DF,  p-value: < 2.2e-16
```