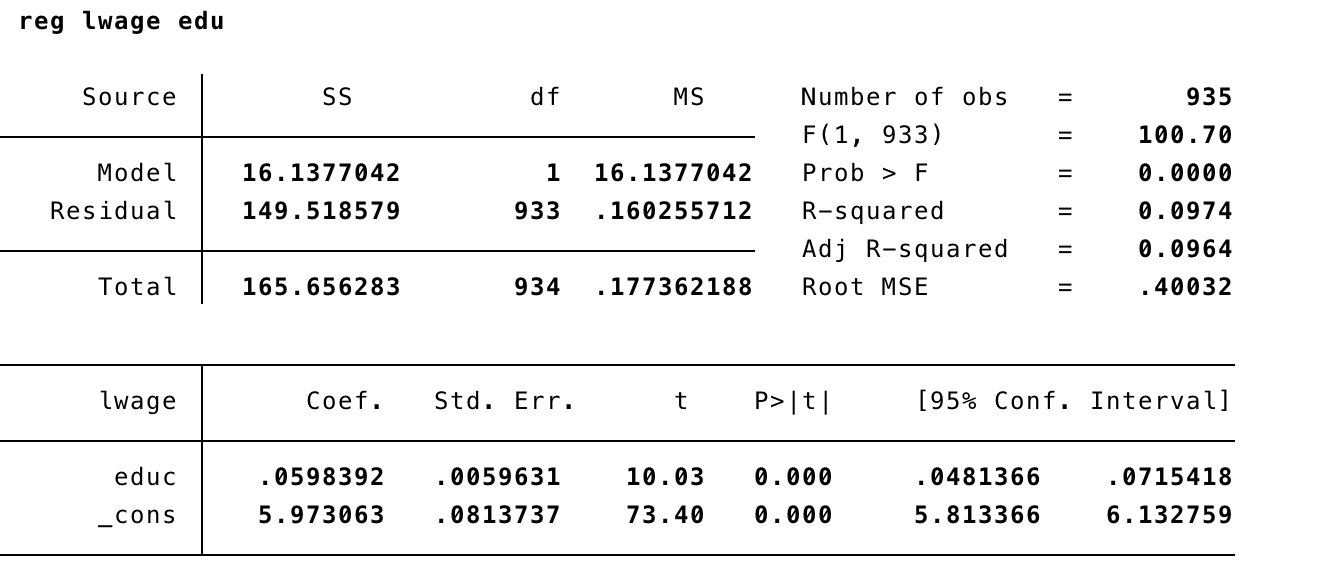
ECON 6511: Advanced Applied Econometrics

Homework 6

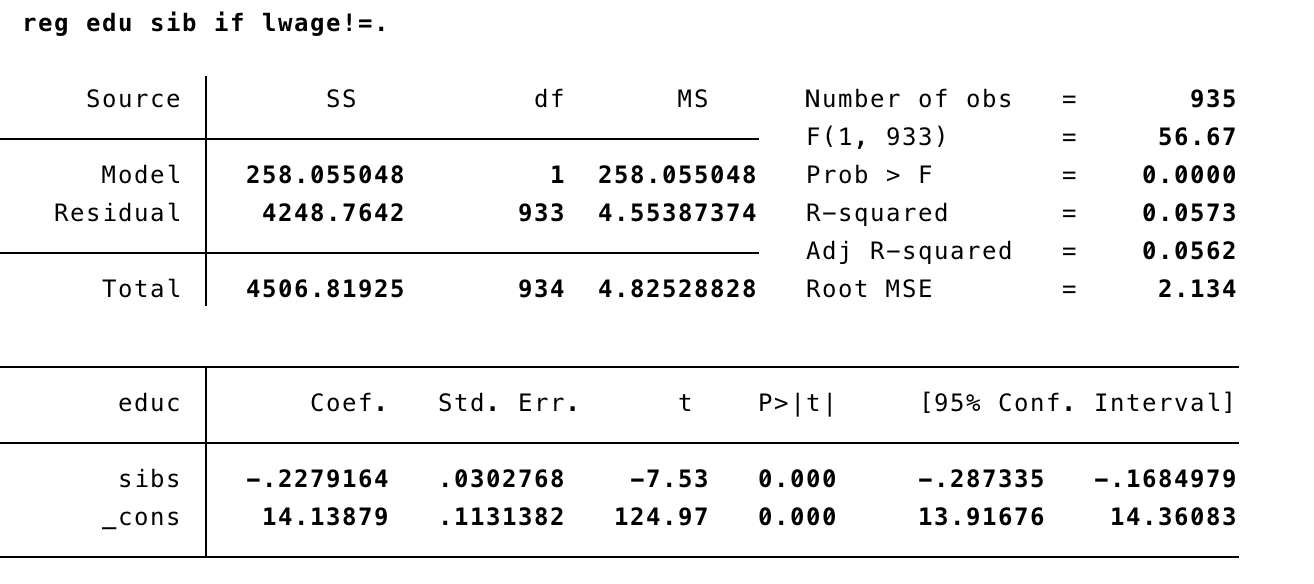
Surabhi Asati

1. (Wooldridge, Chapter 15, Problem 1) Use the data in WAGE2.dta for this exercise.

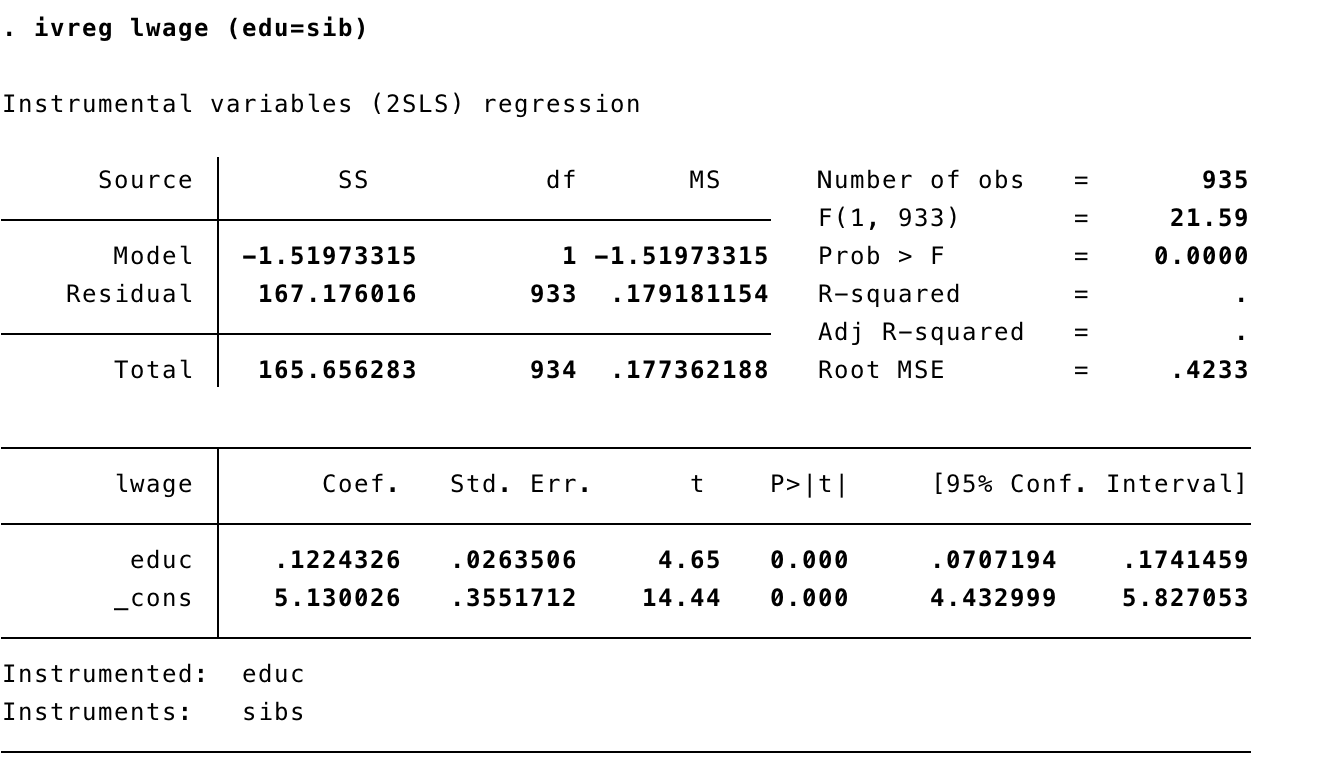
(a)  In class we considered using sibs as a instrument for educ (which would yield an IV estimate of the return to education of 0.122). To convince yourself that using sibs as an IV for educ is not the same as just plugging sibs in for educ and running an OLS regression, run the regression of log(wage) on sibs and explain your findings.



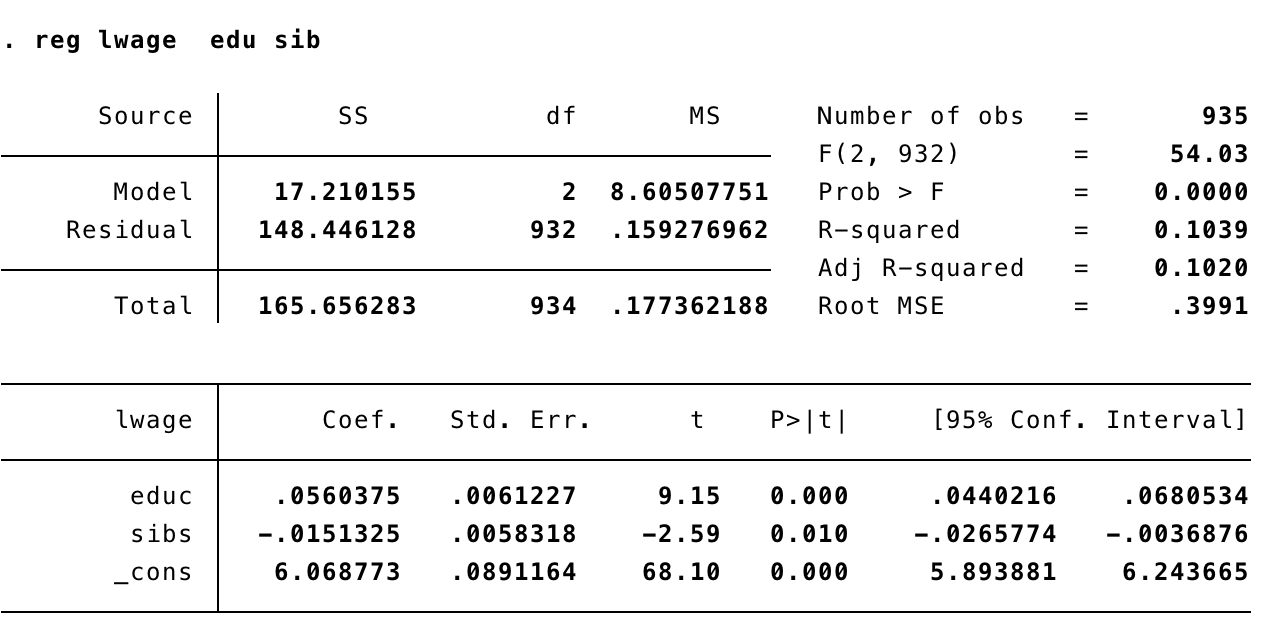
5.9% return for another year of education. Very significant



Years of education and number of siblings are correlated

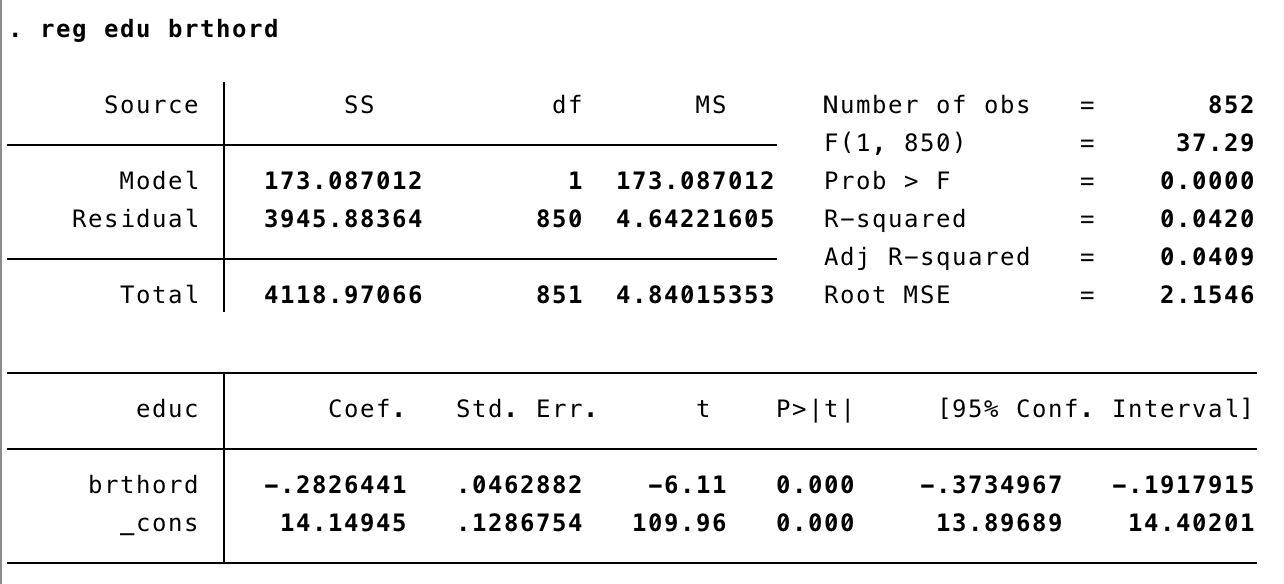


IV estimate of return to education is 12.24%



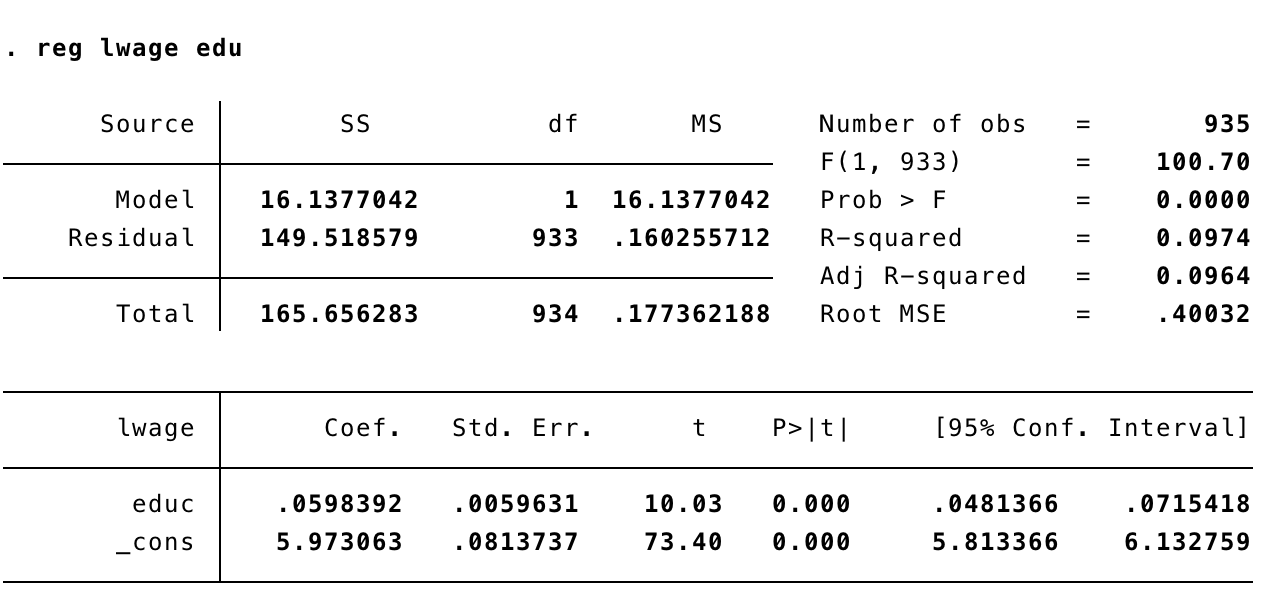
return drops by 1.51% for each sibling. As oppose to the IV estimate of return to education which leads to increase in returns by 12.24%

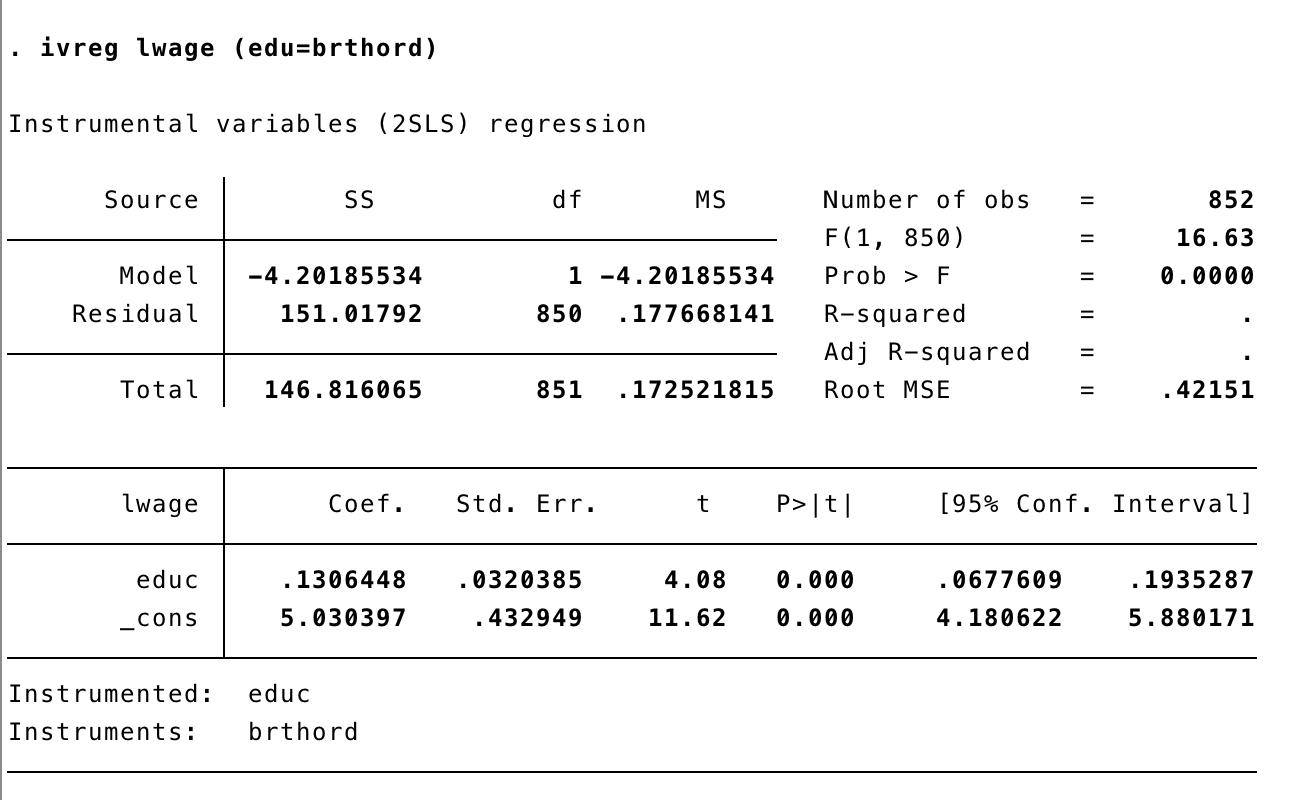
(b)  The variable brthord is birth order (brthord is one for a first-born child, two for a second- born child, and so on.) Explain why educ and brthord might be negatively correlated. Regress educ on brthord to determine whether there is a statistically significant negative correlation.



with t statistics = -6.11 and p-value = 0.00 < 5%, birth order has a very statistically significant negative correlation with education. Higher birth order means smaller age i.e. less years of education as compared to lower birth order. Hence, negative correlation.

(c)  Use brthord as an IV for educ in the equation: log(wage) = β0 + β1educ + u.  Report and interpret the results.





IV estimate of return to education is 13.06%. more than double the OLS estimate (5.9%). with t statistics = 4.08 and p-value = 0.00 < 5%, it is statistically significant and has larger confidence interval

(d)  Now, suppose that we include number of siblings as an explanatory variable in the wage equation; this controls for family background, to some extent:  log(wage) = β0 + β1educ + β2sibs + u. Suppose that we want to use brthord as an IV for educ, assuming that sibs is exogenous.  The reduced form for educ is: educ = π0 + π1sibs + π2brthord + v  State and test the identification assumption.

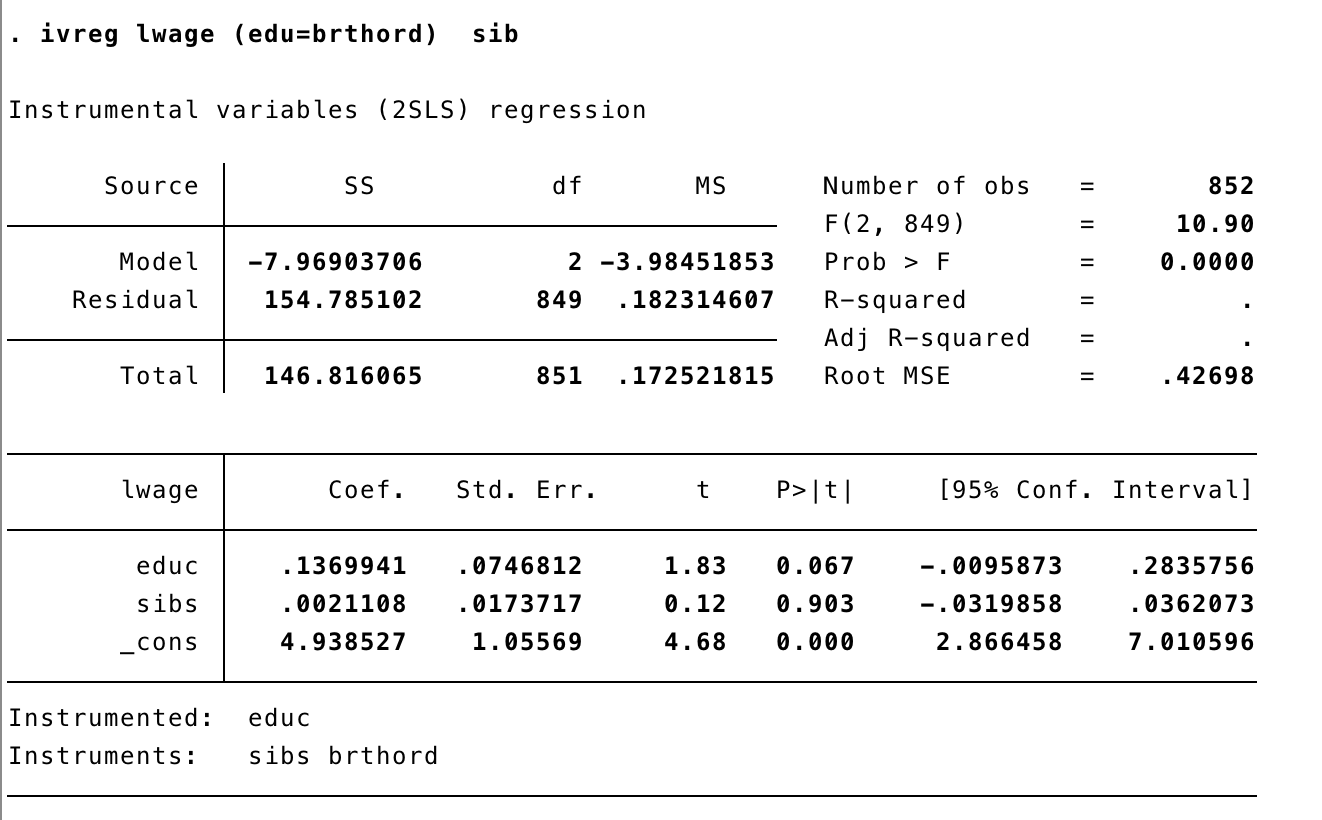
educ = π0 + π1sibs + π2brthord + v

Assumption: For log(wage) equation, an instrument birth order for educ must be uncorrelated with ability and any other unobserved factors affecting wage in the error term and correlated with education.

Given, sibs is exogenous. birth order satisfies the second requirement i.e. it is correlated with education and satisfies the first requirement to uncorrelate with ability(intuition)

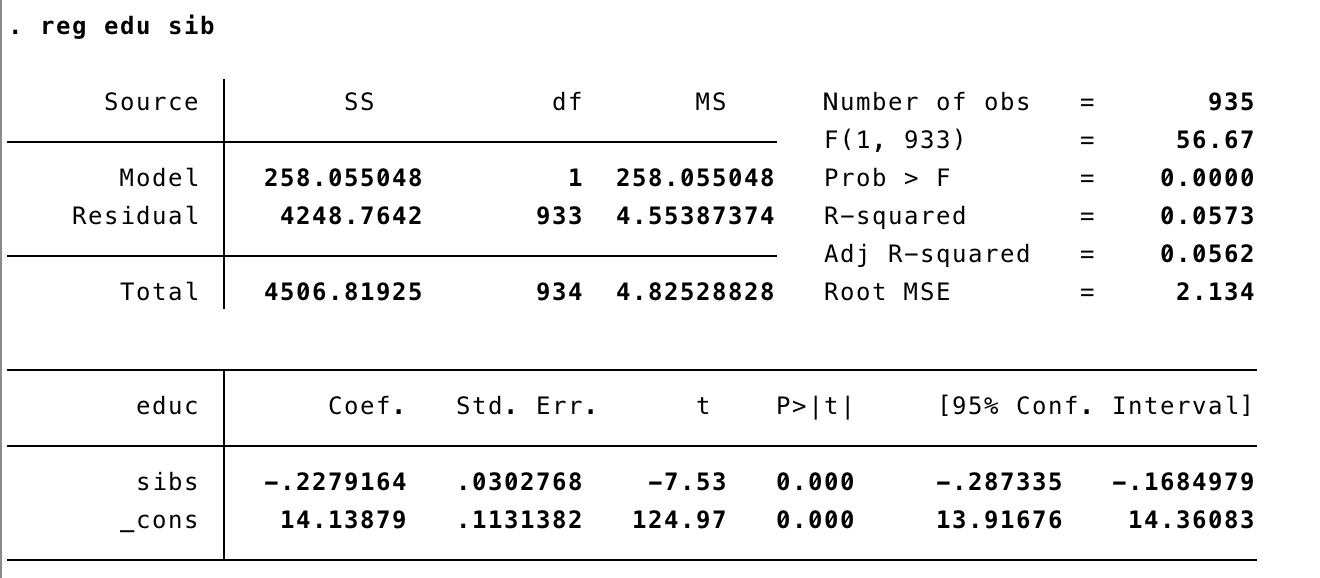
So, birth order satisfies the two assumptions: 1 Cov(z,u)=0 2 Cov(z,x)!=0

(e)  Estimate the equation from part (d) using brthord an an IV for educ (and sibs as its own IV). Commend on the standard errors for βˆeduc and βˆsibs.

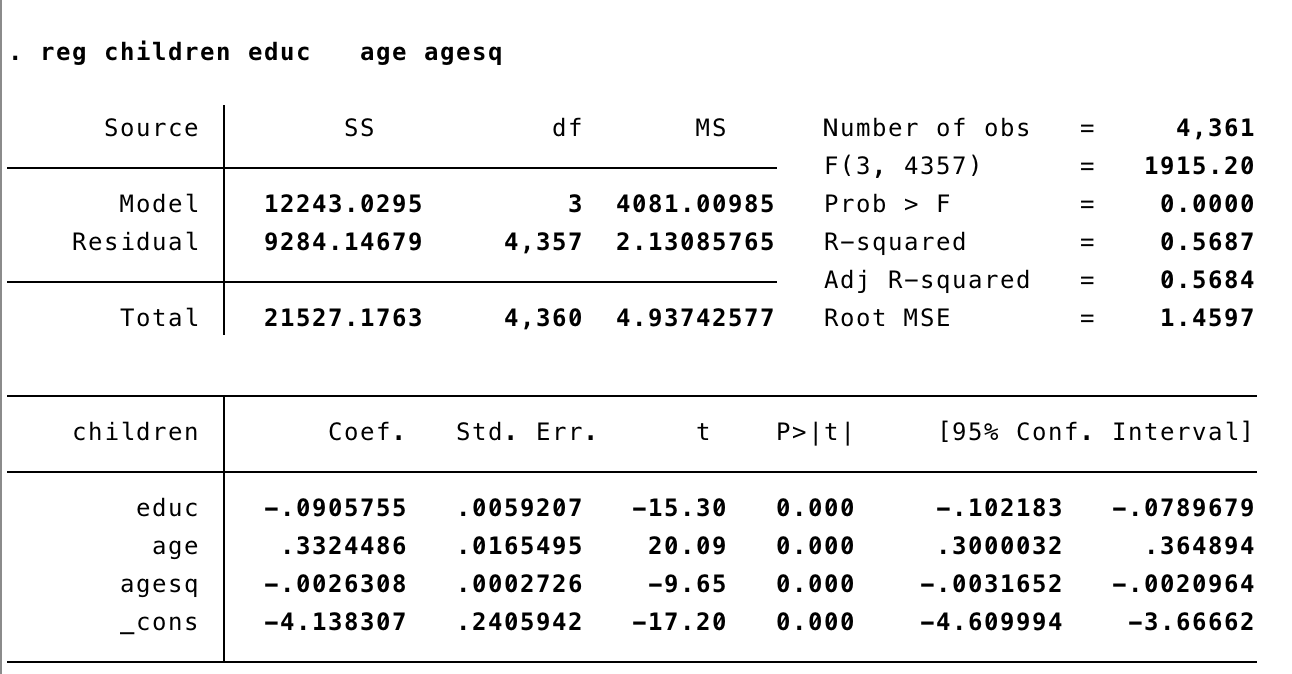




(f) Using the fitted values from part (d), educ, compute the correlation between educ and sibs. Use this result to explain your findings from part (e).



1. (Wooldridge, Chapter 15, Problem 2) The data in FERTIL2.dta include, for women in Botswana during 1988, information on number of children, years of education, age, and religious and economic status variables. Estimate the model children = β0 + β1educ + β2age + β3age2 + u by OLS, and interpre the estimates. In particular, holding age fixed, what is the esti- mated effect of another year of education on fertility? If 100 women receive another year of education, how many fewer children are they expected to have?



Estimated equation:

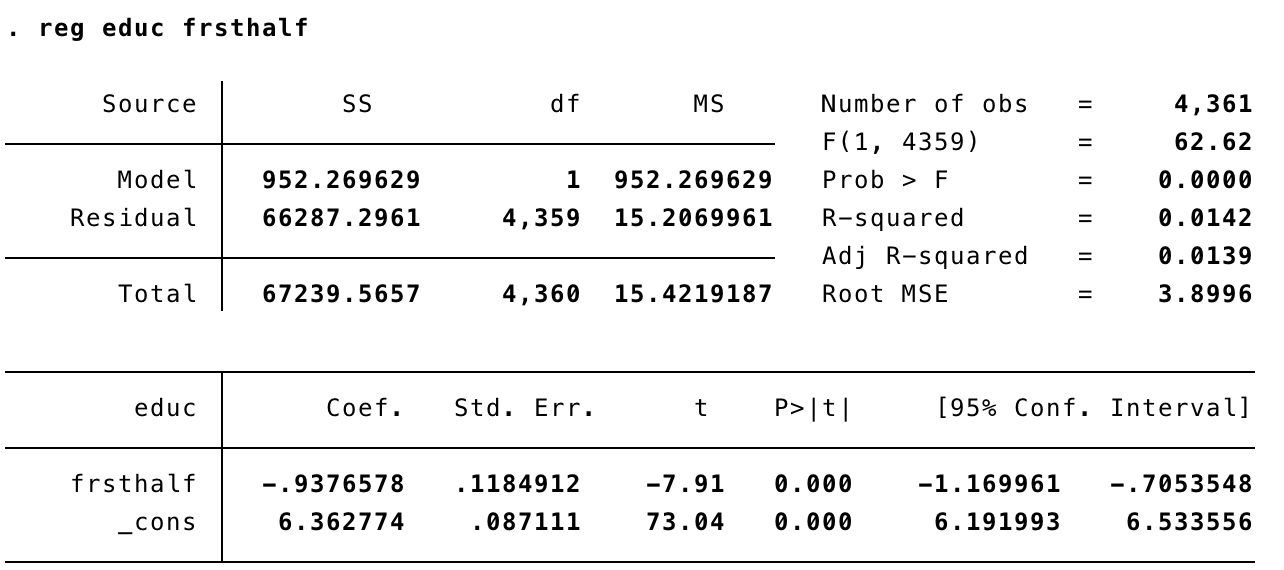
children = -4.14 – 0.09educ + 0.33age – 0.01age2

the t statistics and p-value are very significant for edu, age and agesq. That means all three variables are statistically significant. Though educ and agesq are negatively correlated with fertility.

holding age fixed, 1 year of education drops fertility by 0.09.

If 100 women receive another year of education, they are expected to have 9.05 less children

(b)  The variable frsthalf is a dummy variable equal to one if the woman was born during the first six months of the year. Assuming that frsthalf is uncorrelated with the error term from part (a), show that frsthalf is a reasonable IV candidate for educ. (Hint: You need to do a regression.)



An instrumental variable satisfies two assumptions:

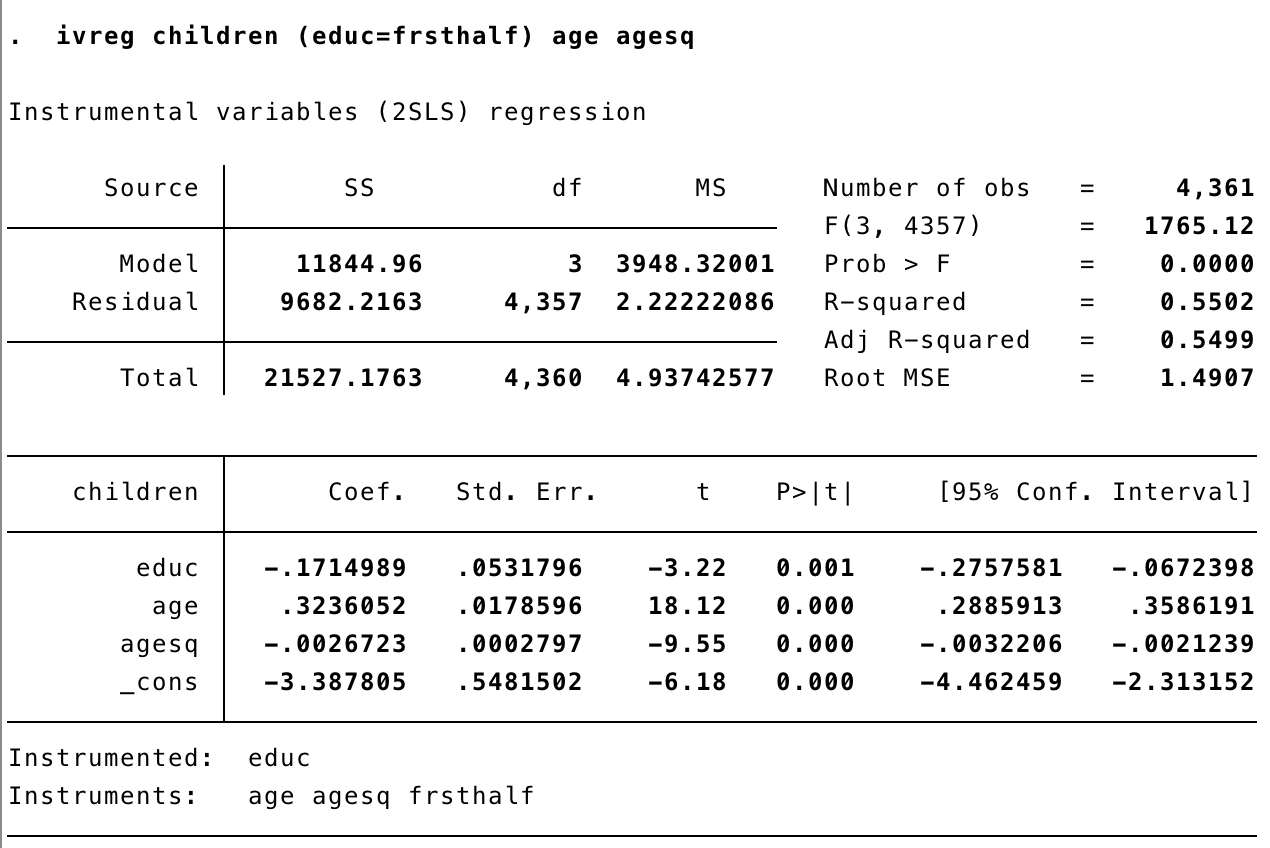
1: Cov(z,u)=0

2: Cov(z,x)̸=0

z = frsthalf

Given assumption: frsthalf is uncorrelated with the error term so Cov(z,u)=0 (“instrumental exogeneity”)  and frsthalf is related to educ negatively (“instrumental  relevance”) with t statistics = -7.91 and p-value = 0.00. This satistfies both the requirements so, frsthalf is a reasonable IV candidate for educ.

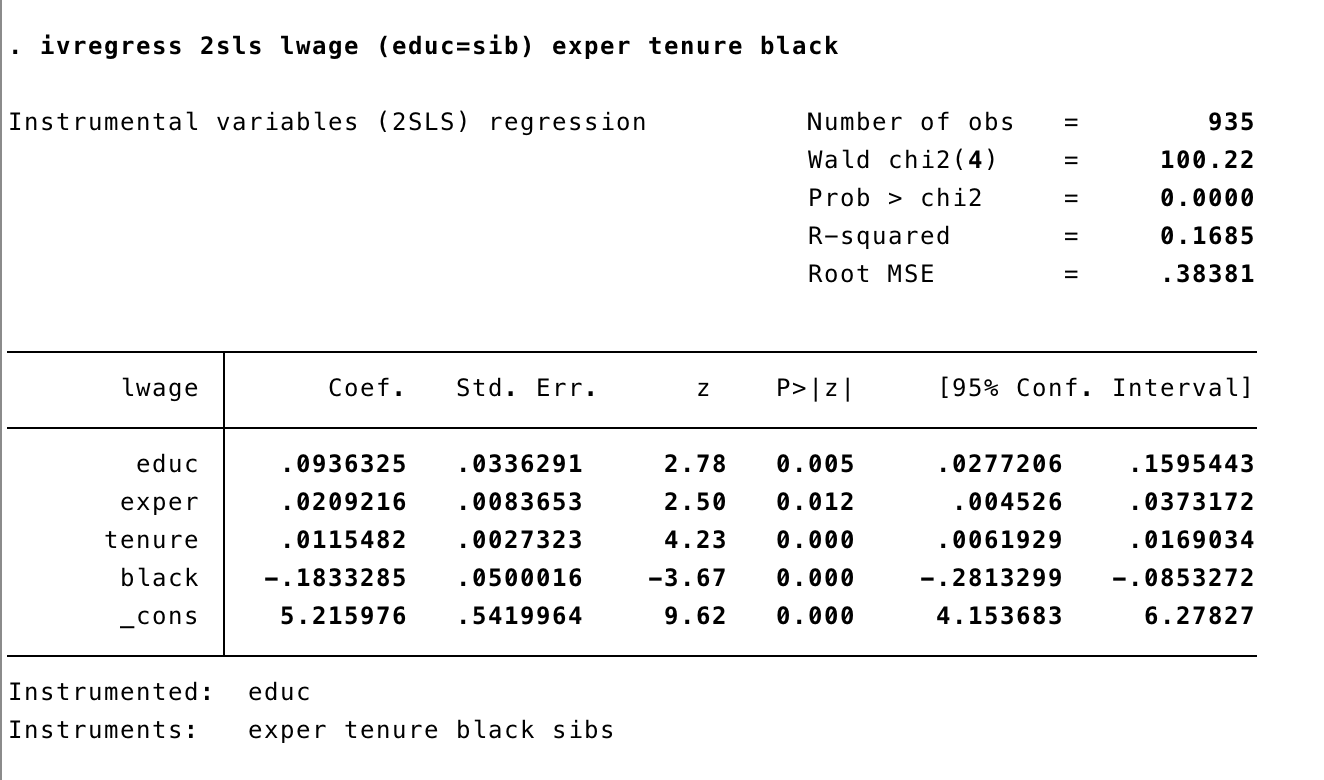
(c)  Estimate the model from part (a) using frsthalf as an IV for educ. Compare the estimated effect of education with the OLS estimate from part (a).



IV estimate of fertility is -0.17 which is almost double the OLS estimate. coefficient of IV estimate is negative and standard error is small. With with t statistics = -3.22 and p-value = 0.01this is significant. With every year of education, fertility drops by 0.17 though confidence interval is larger than OLS estimates’

4. The purpose of this exercise is to compare the estimates and standard errors obtained by correctly using 2SLS with those obtained using inappropriate procedures. Use the data file WAGE2.dta.

(a)  Use a 2SLS routine to estimate the equation: log(wage) = β0 + β1educ + β2exper + β3tenure + β4black + u,  where sibs is the IV for educ. Report the results in the usual form.

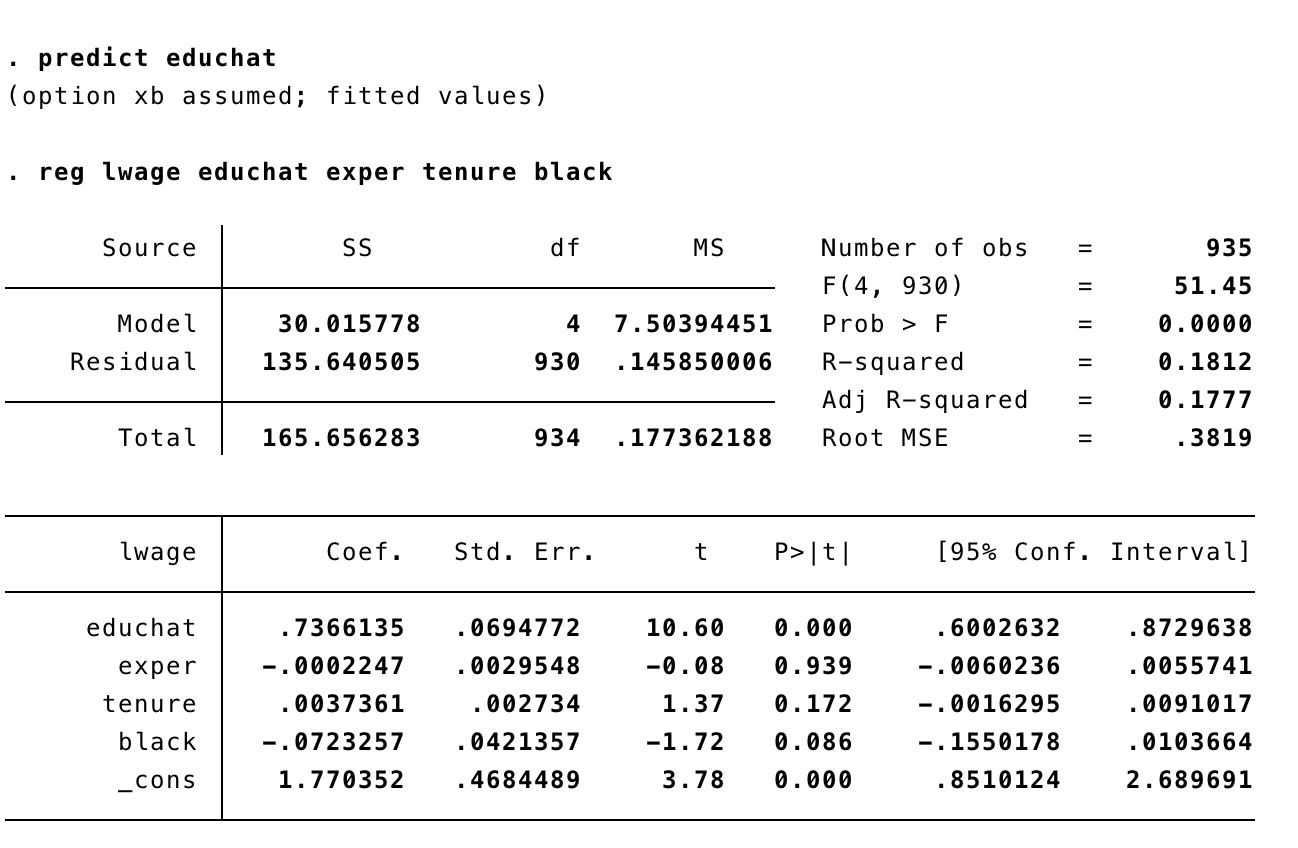


(b)  Now, manually carry out 2SLS. That is, first regress educi on sibsi, experi, tenurei,

and blacki and obtain the fitted values, educi, i = 1, . . . , n. Then, run the second stage

regression log(wagei) on educi, experi, tenurei, and blacki, i = 1, . . . , n. Verify that the βˆj are identical to those obtained from part (a), but that the standard errors are somewhat

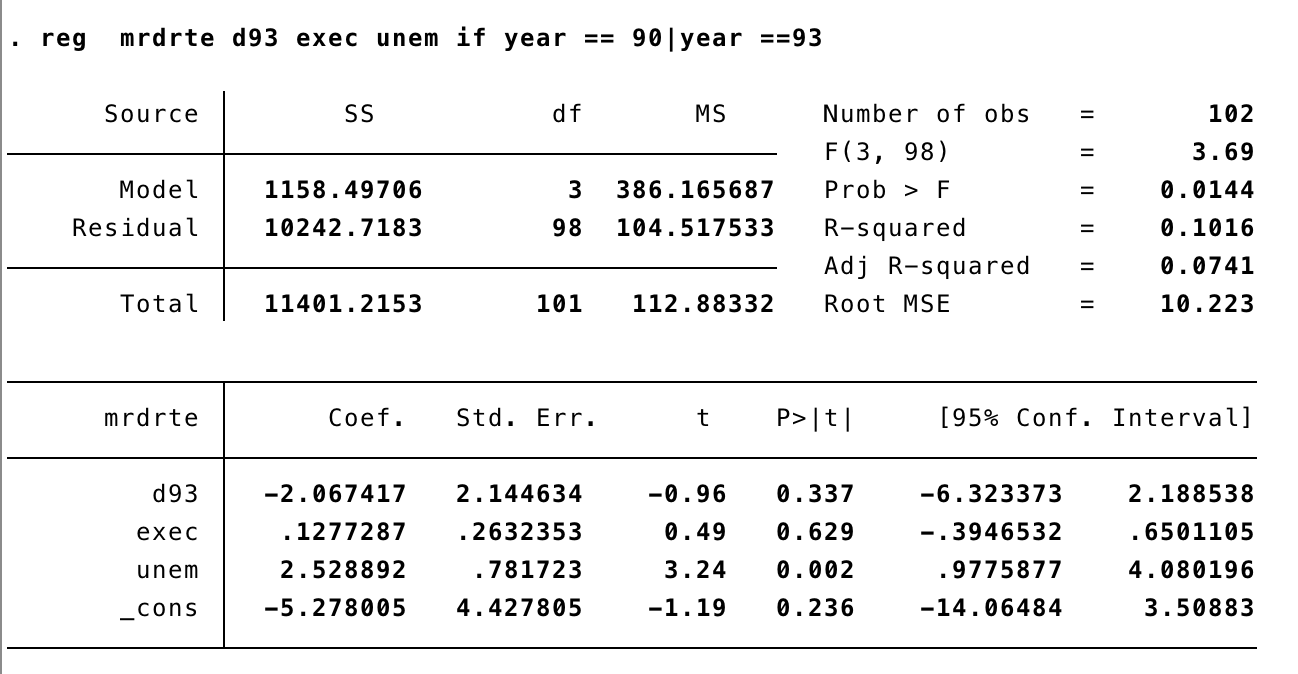
different. The standard errors obtained from the second stage regression when manually carrying out 2SLS are generally inappropriate.



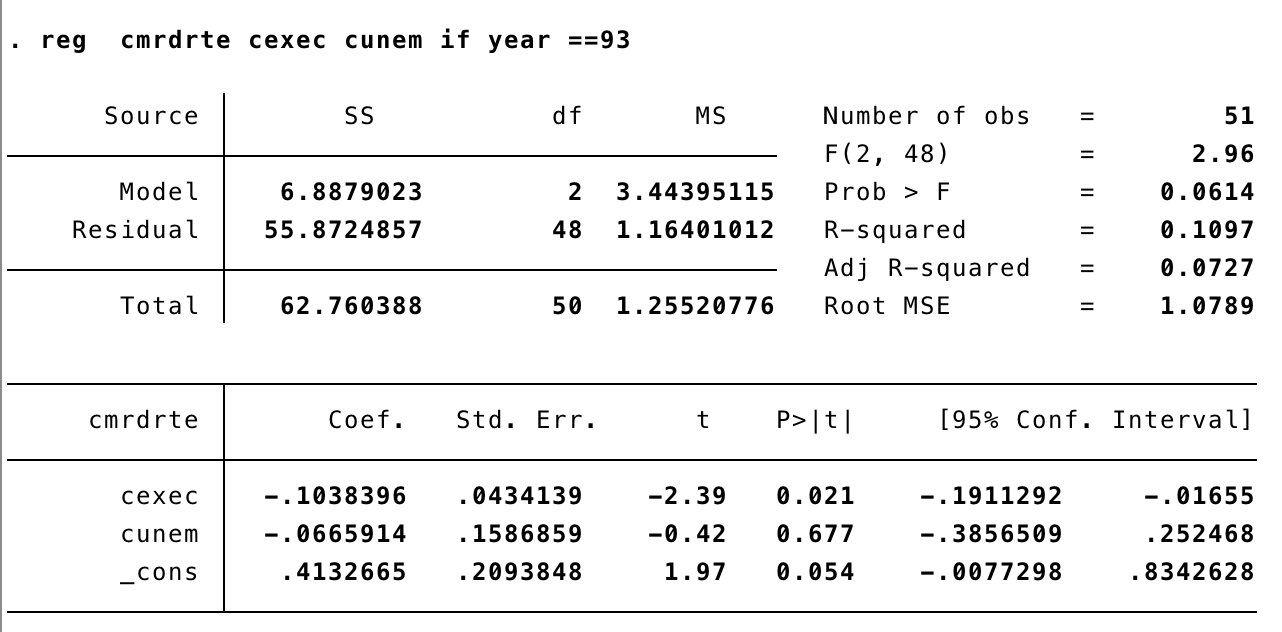
1. (Wooldridge, Chapter 15, Problem 6) Use the data in MURDER.dta for this exercise. The variable mrdrte is the muder rate, that is, the number of murders per 100,000 people. The variable exec is the total number of prisoners executed for the current and prior two years; unem is the state unemployment rate.
2. How many states executed at least one prisoner in 1991, 1992, or 1993? Which state had the most executions?

16 states executed at least one prisoner in 1991, 1992, or 1993. Texas has the most executions.

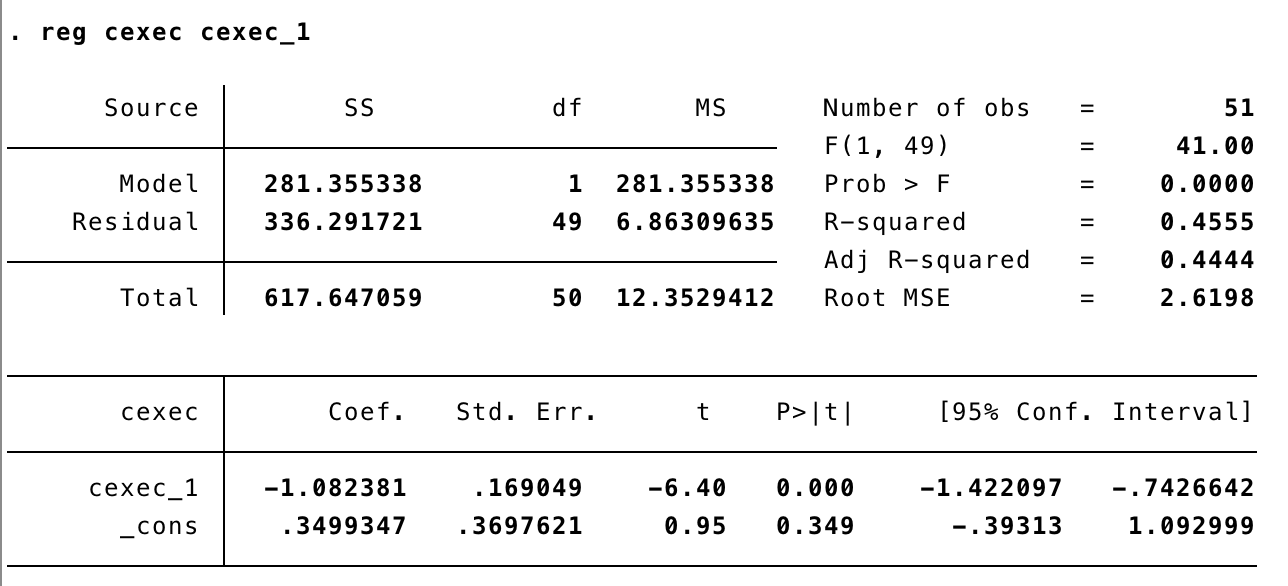
1. Using the two years 1990 and 1993, do a pooled regression of mrdrte on d93, exec, and unem. What do you make of the coefficient on exec?



1. Using the changes from 1990 to 1993 only (for a total of 51 observations), estimate the equation  ∆mrdrte = β0 + β1∆exec + β2∆unem + ∆u by OLS and report the results in the usual form. Now, does capital punishment appear  to have a deterrent effect?



1. The change in executions may be at least partly related to changes in the expected murder rate, so that ∆exec is correlated with ∆u in part (c). It might be reasonable to assume that ∆exec−1 is uncorrelated with ∆u. (After all, ∆exec−1 depends on executions that occurred three ore more years ago!) Regress ∆exec on ∆exec−1 to see if they are sufficiently correlated; interpret the coefficient on ∆exec−1.



1. Reestimate the equation in part (c), using ∆exec−1 as an IV for ∆exec. Assume that ∆unem is exogenous. How do your conclusions change from part (c)?

Q3 b) Using the two years 1990 and 1993, do a pooled regression of mrdrte on d93, exec, and unem. What do you make of the coefficient on exec?

