Eli Sakov **Lab Writeup #5**

Econometrics

10.12.2014

**Introduction**

Educational achievement is an outcome of many determinants, including distance to the nearest college or university, as well as socioeconomic status. Regressions were performed to determine how various factors effect educational achievement.

**Data and Methods**

The data is from the *HighSchool and Beyond* Survey, conducted in 1980 and 1986, from the Department of Education. The data was used In Cecilia Rouse's paper “Democratization or Diversification? The Effect of Community Colleges on Educational Attainment,” *Journal of Business and Economic Statistics,* April 1995, Vol. 12, No. 2, pp 217-224. Observations of 3796 students were used. The data is from the United States, but not the Western States. Table 1 defines each of the variables used. Table 2 shows summary statistics of the variables.

For the analysis, years of education were regressed on the independent variables using ordinary least squares regression with heteroskedastic-consistent standard errors for linear, quadratic, and log-linear models.

|  |  |
| --- | --- |
| **Table 1. Variables Used** | |
| **Variable** | **Description** |
| yrsed | Years of Education Completed. 12 For High School Seniors and an additional year for each year of secondary education. |
| female | 1 for Female, 0 for Male |
| black | 1 for Black, 0 for Not-Black |
| hispanic | 1 for Hispanic, 0 for Not-Hispanic |
| bytest | Base Year Composite Test Score. (These are achievement tests given to high school seniors in the sample) |
| dadcoll | 1 if Father is a College Graduate, 0 if Father is not a College Graduate |
| cue80 | County Unemployment rate in 1980 |
| urban | 1 for School in Urban Area, 0 for School not in Urban Area |
| ownhome | 1 for Family Owns Home, 0 for Family Does not Own Home |
| incomehi | 1 for Family Income > $25,000 per year, 0 for Income ≤ $25,000 per year |
| momcoll | 1 if Mother is a College Graduate, 0 if Mother is not a College Graduate |
| tuition | Avg. State 4yr College Tuition in $1000's |
| dist | Distance from 4yr College in 10's of miles |
| stwmfg80 | State Hourly Wage in Manufacturing in 1980 |
| distsquared | Distance from a 4 year college in 10's of miles squared |
| dadmomcoll | The interaction term between momcoll and dadcoll |
| distincomehi | The interaction term between dist and incomehi |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 2. Summary Statistics** | | | | | |
| **Variable** | **Number** | **Mean** | **Std. Dev.** | **Min** | **Max** |
| female | 3796 | .5453109 | .4980083 | 0 | 1 |
| black | 3796 | .1925711 | .394371 | 0 | 1 |
| hispanic | 3796 | .1498946 | .3570151 | 0 | 1 |
| bytest | 3796 | 51.00193 | 8.819251 | 28.95 | 71.36 |
| dadcoll | 3796 | .2020548 | .4015858 | 0 | 1 |
| momcoll | 3796 | .1393572 | .3463645 | 0 | 1 |
| ownhome | 3796 | .8192835 | .3848338 | 0 | 1 |
| urban | 3796 | .243941 | .4295141 | 0 | 1 |
| cue80 | 3796 | 7.654874 | 2.86577 | 1.4 | 24.9 |
| stwmfg80 | 3796 | 9.556499 | 1.364411 | 6.59 | 12.15 |
| dist | 3796 | 1.724921 | 2.133836 | 0 | 16 |
| tuition | 3796 | .9131396 | .2835778 | .43418 | 1.40416 |
| yrsed | 3796 | 13.82929 | 1.813969 | 12 | 18 |
| incomehi | 3796 | .2863541 | .4521164 | 0 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 3. Linear and Nonlinear Regressions on Ed** | | | | | |
| **Model** | **(1) Linear** | **(2) Log-linear** | **(3) Quadratic** | **(4) Quadratic** | **(5) Quadratic** |
| **Dependant Variable** | **yrsed** | **Ln(yrsed)** | **yrsed** | **yrsed** | **yrsed** |
| dist | -.0366613\*\*  (.0127154) | -.0026072\*\*  (.0009027) | -.0811732\*\*  (.0255486) | -.0810001\*\*  (.0255346) | -.0995324\*\*\*  (.0274353) |
| female | .1429742\*\* | .0103059\*\* | .1433144\*\* | .1406184\*\* | .1442333\*\* |
|  | (.0504367) | (.0035808) | (.0504168) | (.0504031) | (.0504036) |
| bytest | .0930377\*\*\* | .0066561\*\*\* | .0926367\*\*\* | .0925664\*\*\* | .0927563\*\*\* |
|  | (.0031816) | (.0002259) | (.0031866) | (.003185) | (.0031863) |
| tuition | -.1910519 | -.0139382 | -.1928193 | -.1939714 | -.2038559\* |
|  | (.1008799) | (.0071621) | (.1008433) | (.1007892) | (.1009917) |
| black | .3506095\*\*\* | .0261676\*\*\* | .3339309\*\*\* | .3305619\*\*\* | .3346437\*\*\* |
|  | (.0712302) | (.0050571) | (.0716843) | (.0716603) | (.071663) |
| hispanic | .3617649\*\*\* | .0259986\*\*\* | .3333104\*\*\* | .3297465\*\*\* | .3296601\*\*\* |
|  | (.0772702) | (.0054859) | (.0785278) | (.0785003) | (.0785286) |
| incomehi | .3718305\*\*\* | .0265197\*\*\* | .3694975\*\*\* | .3623156\*\*\* | .2802678\*\*\* |
|  | (.0607215) | (.004311) | (.0607082) | (.0607572) | (.0778088) |
| ownhome | .1385475\* | .0098332\* | .14327\* | .1412131\* | .1449606\* |
|  | (.0667233) | (.0047371) | (.066738) | (.0667075) | (.0667236) |
| dadcoll | .5709712\*\*\* | .0405374\*\*\* | .5611581\*\*\* | .6538031\*\*\* | .5701874\*\*\* |
|  | (.0736951) | (.0052321) | (.0738274) | (.0843096) | (.0739687) |
| momcoll | .3778102\*\*\* | .0266016\*\*\* | .3777022\*\*\* | .5693549\*\*\* | .3798051\*\*\* |
|  | (.0815252) | (.005788) | (.0814925) | (.1172733) | (.0814752) |
| cue80 | .0286753\*\* | .0020357\*\* | .0259537\*\* | .0257697\*\* | .0262992\*\* |
|  | (.0098657) | (.0007004) | (.0099544) | (.0099493) | (.0099531) |
| stwmfg80 | -.0425003\* | -.0028642\* | -.0425539\* | -.0415432\* | -.0429006\* |
|  | (.0202082) | (.0014347) | (.0202001) | (.0201939) | (.0201947) |
| distsquared |  |  | .0046413\* | .0046773\* | .0054327\* |
|  |  |  | (.0023109) | (.0023097) | (.0023502) |
| dadmomcoll |  |  |  | -.3664802\* |  |
|  |  |  |  | (.1613445) |  |
| distincomehi |  |  |  |  | .0566549 |
|  |  |  |  |  | (.0309169) |
| Constant | 8.920823\*\*\* | 2.265819\*\*\* | 9.012167\*\*\* | 9.00197\*\*\* | 9.041118\*\*\* |
|  | (.2515861) | (.0178617) | (.2555647) | (.2554637) | (.255973) |
| **Regression Statistics** | | | | | |
| SER | 1.5378 | 0.10918 | 1.5372 | 1.5363 | 1.5367 |
| N | 3796 | 3796 | 3796 | 3796 | 3796 |
| R2 | 0.284 | 0.285 | 0.284 | 0.285 | 0.285 |
| adj. R2 | 0.281 | 0.283 | 0.282 | 0.283 | 0.282 |
| F | 124.8 | 125.9 | 115.6 | 107.8 | 107.7 |

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001 Standard Errors in parenthesis

All regressions have an adjusted R2 of about 0.282, meaning that the regressions explain just more than 28% of the variance in the data.

The greatest determinants in the first model are those of parent's educational achievement, race, income, and home ownership.

From model 2, we can say with 99% confidence that years of education decreases by 0.26% for every 10 miles away from a four year school a student is. We can also say with 99% confidence that being female increases years of education by 1.03%. With 99.9% significance, we can say that being black increases years of education by 2.616%, and being Hispanic increases it by 2.599%. Having a family income higher than $25,000 and a father who graduated college increases years of education by 2.65% and 4.05%, respectively.

In the quadratic models, distance from a four year school accounts for much more of the variance. The significance of the coefficient of the quadratic term is greater than 95%. It is necessary to include the interaction term for both parents having graduated college, and that is accounted for in model 4 with 95% confidence on the coefficient. In Model 5 the interaction between distance from a four year college and income greater than $25,000 is accounted for, but it is not significant at 95% confidence.

Figure 1 shows effects of distance on years of education in the quadratic and linear models. Distsquared has a significantly greater effect on educational achievement than the linear model at greater distances.

**Conclusions**

With 95% significance, distance from a four year school decreases the number of years in post-secondary education quadratically. Race, income, home-ownership, and whether a student's parents graduated college account for almost 28% of the amount of post-secondary education a student will achieve.

Figure 1.

