Winter 12-17-2012

Effect of Post-Secondary Education on Longevity

Pavlo Buryi
Southern Illinois University Carbondale, pasha@siu.edu

EFFECTS OF POST-SECONDARY EDUCATION ON LONGEVITY

by

Pavlo Buryi

B.A., Southern Illinois University, 2010

A Research Paper

Submitted in Partial Fulfillment of the Requirements for the

Master of Science in Economics

Department of Economics

Southern Illinois University Carbondale

December 2012

RESEARCH PAPER APPROVAL

EFFECT OF POST-SECONDARY EDUCATION ON LONGEVITY

By

Pavlo Buryi

A Research Paper Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Masters of Science

in the field of Economics

Approved by:

Dr. Subhash C. Sharma, Chair

Dr. Scott Gilbert

Graduate School

Southern Illinois University Carbondale

12/13/2012

TABLE OF CONTENTS

CHAPTER **PAGE**

CHAPTER 1

INTRODUCTION

In this paper I investigate the effect of education on longevity, or simply speaking if higher education prolongs life. To be more precise, I examine the effect of post secondary education on longevity. Most of the studies find that education is strongly associated with higher levels of longevity. However, the causal effect was poorly studied until in 2005 Adriana Lleral-Muney came up with the paper, "The Relationship between Education and Adult Mortality in the United States". Her paper used a unique quasi-natural experiment that happened between 1915 and 1939, at least 30 states changed their compulsory schooling laws and child labor laws. If compulsory schooling laws force people to get more schooling than they normally would, then individuals who spend their youth in the affected states would live longer. She finds that more education is highly associated with additional years of life. Her paper caused a lot of attention, and on January the 3rd of 2007 New York Times published an article with a title, "A Surprising Secret to a Long Life: Stay in School" that told her story and shed the light of her findings to the public. After her paper a few other scientists came up with similar studies which also use quasi-natural experiments to confirm positive effect of education on longevity. Hans van Kippersluis and colleagues (2010) use Dutch compulsory laws to find that education has a positive and significant effect on longevity. Another study by Damon Clark and Heather Royer (2010) uses British compulsory schooling laws to find the causal effect of education on longevity, but find little evidence. Despite the strong correlation between education and health, it is difficult to determine whether education has a causal effect on longevity because of the two econometric problems: reverse causality (expected longevity may affect education), and confounding variables. Savelyev (2010) suggests that an unobserved factor such as non-cognitive trails may

affect both education and longevity. Previous studies are limited to the effect of compulsory education on longevity, but the effects of post-secondary education are less studied. In the diverse environment such as the United States it would be useful to know if some groups benefit more from education than others do. Research by Jay Olshansky and colleagues (2012) find that education and race are important factors in determining life expectancy. The paper finds that "in 2008 for people with less than twelve years of education life expectancy was not much better than for all adults in 1950 and 1960" (Olshanskiy, 2012). However, this paper does not find the effects of education on longevity by groups it simply suggests that people with a higher education are more likely to live longer.

It is natural to think that more education would be associated with higher levels of longevity, because there are so many ways in which education may affect health. First of all, more education is highly correlated with greater income. Secondly, higher education leads to less risky choices of occupation. Thirdly, education improves the ability to understand the use of sanitation and hygiene, as well providing motivation for people to use health care and practice a healthier lifestyle and in many other ways (Case, 2004). Leon Feinstein et al. (2006) suggests that education can improve self-esteem and provide protection against some adverse health outcomes through fostering resilience. They also find evidence from a randomized clinical trial that people with higher education show faster recovery from depression and social malfunctioning. In this paper I investigate the effects of education on longevity and compare it among different groups. My hypothesis is that education has positive effect for all groups regardless of gender, race, and marital status.

CHAPTER 2

DATA

In this study I use data on mortality from the National Vital Statistics System for the year 2006. The data contains records on all the known deaths that occurred in the United States of America during 2006. Each record in the data includes of the following information: age of the person, when he or she died, cause of death, sex, marital status, number of children, race, and level of education. The data is coded, so the manual of how to interpret the information is needed. The Mortality Multiple Cause-of-Death Public Use Record contains a detailed description of how to use the data.

I restrict the data to the people who died at age 18 and above to ignore infants and children who died for reasons that could not be affected by education. The dependent variable is longevity, which is the age of the person when he or she died. The independent variable is education. There are two different records of education: educ1989 and educ2003. I use educ2003 because it has more observations. This record contains information about the level of formal education. Numbers are assigned is the following manner: 1- 8th grade or less, 2 – 9 to 12th grade and no diploma, 3-high school graduate or GED, 4 - completed some college credit, but no degree, 5- Associate degree, 6- Bachelor's degree, 7- Master's, 8- degree Doctorate or professional degree, 9- Unknown (2006 Mortality Multiple Cause-of-Death Public Use Record). I restrict the number of records, to those that satisfy criteria 3<educ2003<9, because the area of my interest is solely post-secondary education, which is anything greater than high school. It is worth mentioning that my data set does not allow for income levels, and enables me to study the causal effect, which is not a subject of my interest in this research. The goal of this paper is to evaluate the overall effect of education on longevity, and the information I have is satisfactory

for this purpose. In the following section I am going to discuss the methodology.

CHAPTER 3

METHODS

To decide which method to use in the regression of education on longevity, I construct a series of graphs with fitted lines to see what kind of relationship longevity and education have. Graph 1 shows a scatter plot of longevity and education and two fitted lines: linear fitted line and quadratic fitted line. It is clear that both fitted lines suggest that longevity and education have linear relationship. To make sure that the relationship is linear I use STATA software to run kernel regression and the product of this regression can be found in Graph 2. Graph 2 also suggests the linear relationship between the dependent and the independent variables. According the graphs, it is clear that the use of OLS is appropriate in estimating the effect of education on longevity in a regression model. Nevertheless, OLS is a useful tool in estimating a linear model, but the problem of possible heteroscedasticity needs to be addressed. So I tested for heteroscedasticity, and the results are found in Appendix C. There is no evidence of heteroscedasticity, but to avoid possible inaccuracies I use WLS as well[1]. To find the effect of education on longevity in a specific group I generate additional variables that help me restrict the sample. The first variable that I generate is k, it restricts the total sample of observations to only those that satisfy the following conditions: 3<educ2003<9 and 18<longevity<150. Then I create "dummy" variables: male, female, white, black, single, married, divorced. For example, the "dummy" variable "male" turns 1 if observed individual is male and 0 otherwise. To select observations that satisfy group characteristics I generate a specific variable "grm" (gender, race, marital status) for each specific group of interest. For example, if the group of interest is white single males I generate "mws" variable that satisfies the following conditions:

[1] Results are reported in Table 3

mws = (male=1)&(white=1)&(single=1). (1)

Additional "dummy" variable "activity" is created to find by how many years higher education could prolong life of those who died at work or earning income. Then I create "act" variable that satisfies the following condition act= (activity==4). It turns 1 if person died at work or earning income and 0 other wise. Now that all variable are well defined, I proceed to use OLS to estimate the effect of education on longevity, conditional that group characteristics are taken into account. The model takes the following form

$$Y = \alpha+\beta_1 X+\varepsilon \quad \text{if } k*grm=1 \tag{2}$$

Where Y is longevity, X is education, k is a variable that restricts the total sample to observations that satisfy conditions described above, and grm is a variable that accounts for group specific characteristics. Results of all regressions can be found in Table 1 and 4. The coefficients in Table 1 explain how education affects longevity, and by comparing them one can find the difference in effects for different groups. To test if the difference is valid I use a model with a dummy variable:

$$Y = \beta0 + \beta1D+ \beta2X + \beta3DX +\varepsilon \tag{3}$$

where D is a dummy variable, "gender", that takes value 1 if it is a women, and 0 if it is a man. DX is a product of gender and education. In the case when D is zero we get coefficient for man, to get the coefficient for women add the coefficient of the product. Coefficients that I get are: 2.16531 for a male individual, and (2.16531-1.593115)= 0.5721954 for a female. By doing this exercise I find that education has different effects on longevity for a man than for a woman, and the result of the original regression show the same. In order to check the robustness of my findings, I perform the bootstrapping procedure that generates random subsamples from the

original data, and runs the regression using generated subsample. The results of this operation are reported in Table 2. Coefficients for newly generated subsamples are identical to those in the original samples; this suggests that coefficients are significant and are reliable source of explaining the effect of education in improving longevity.

CHAPTER 4

RESULTS

For all categories[2] more education is associated with longer life. All coefficients are significant at 1% level and suggest strong effects of education on longevity. As can be found in Table 4 education has the strongest effect for white single males. Also, it suggests that receiving one additional degree, or making one more "step" in education after high-school on average prolongs life by almost 6 years. For example, if a white single man decides to get some college credit after high school it will add 6 years to his life. For men, regardless of race and marital status, one more "step" in education prolongs life by an average of 2 years. Results for women are a little weaker. However, black females, whether they are single, married, or divorced, would still enjoy the significant addition to their life span from acquiring an additional "step" in their education, which is on average 2 more years. For all single females additional education would get 5 more years of life, but married and divorced women would get significantly less. The only category that does not show to be affected by more education is the white married women. For them more education does not mean longer life. It could be due to the fact that having children or being a housewife contributes more to longevity than education does. It is not possible to interpret the cause of these findings from the information used in this study. On average, black people gain more years of life from education than white people do. It is 2.5 years for blacks and a little over a year of additional life for whites. I also construct a kernel density plot that shows probability on y axis and age on x axis. Graph 3 suggests that people with higher levels of education have better chances to live longer. People with PhD or professional degrees, such as doctor or lawyer, have the highest probability of dying in older ages, and people who earned

[2] But white married females

some college credit have the lowest chances among observed groups. These results are consistent with the work of Dustin C. Brown et. al. (2012) who finds that educational attainment in the United States maximizes life chances by delaying the biological aging process. My results confirm Brown's finding and prove that for whatever reason people who have more education are more likely to live longer.

CHAPTER 5

CONCLUSION

This paper confirms the hypothesis that post-secondary education has a positive effect on longevity. Use of the classical OLS method is justified by graphs that suggest that there is a linear relationship between longevity and education. I find strong and robust evidence that more education prolongs life for almost all categories, at most by 6 years for single white males. The causal effect is not studied so any inferences regarding the causality cannot be made. The results are simply suggesting that for whatever reason higher education is associated with longer life, whether it is due to the higher income or due to some other fact is not studied in this paper. Higher education leads to a safer choice of occupation. In fact, people who died at work or earning income, could have gained on average 14 more years of life for each additional step made in post-secondary education. Implications of this paper to the policy making are very limited, because of the econometric problems such as inverse causality and confounding variable. With the help of the studies made previously (Lleral-Muney, 2004), we can assume that education does affect longevity, and now from the results that I derive we know that this effect is not the same across different groups. One particular result is that post-secondary education is more effective in prolonging life among black population, and this can have implications for public policy.

REFERENCES

Brown, B., Hayward, M., Montez, J., Hummer, R., Chiu, C., Hidajat, M. (2012) "*The significance of Education for Mortality Compression in the United States",* Demography(2012) 49:819-840

Clark, D., Royer, H., (2010) "*The Effect of Education on Adult Health and Mortality: Evidence from Britain*" (NBER, Working Paper No. 16013

Feinstein, L., Sabates, R., Anderson,T., Sorhaindo, A., Hammond, C. (2006) "*What are the effects of education on health?"* Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium-OECD

Kippersluis, H., O'Donnell,O., and Doorslaer, E.(2011) *"Long-Run Returns to Education: Does Schooling Lead to an Extended Old Age?*", J.Human Resources September 21, 2011 46:695-721

Kolata, G. (2007) "*A surprizeing Secret to a Long Life: Stay in School*" The New York Times, January 3,2007

Lleras-Muney, A. (2005) *"The Relationship Between Education and Adult Mortality in the United States",* Review of Economic Studies 72(1): 189-221 doi:10.1111/0034-6527.00329

Olshansky, J., Antonucci, T., Berkman, L., Binstock, R., Boersch-Supan, A. Cacioppo, J., Carnes, B., Carstensen, L., Fried, L., Goldman, D., Jackson, J., Kohli, M., Rother, J., Zheng, Y., and Rowe, J. (2012) *"Differences In Life Expectancy Due To Race And Educational Differences Are Widening, And Many May Not Catch Up*", Health Affairs, 31, no.8 (2012):1803-1813

Sanchez, Y.,(2010) *The Longevity Gains of Education,* University of Chicago

(2006) Mortality Multiple Cause-of-Death Public Use Record, http://www.nber.org/data/multicause.html

APENDICIES

APPENDIX A. TABLES

TABLE 1. COEFFIENTS ON EDUCATION WHEN ORDINARY LEAST SQUERED IS USED

Race\Genger	Men	Women
All	2.16531 (77.49)** *195601*	.5721954 (15.71)** *162079*
White	1.971186(68.41)** *174497*	.4214288(11.16)** *142415*
Black	2.815528(22.96)** *13159*	1.961042(15.45)** *14682*

Italicized numbers are numbers of observations. All represent general population and includes all races.
** significant at 1% level.

TABLE 2. COEFFICIENTS AFTER BOOTSTRAPPING

Race \ Gender	Men	Women
All	2.16531(84.84)**	.5721954(16.63)**
White	1.971186(71.09)**	.4214288(11.63)**
Black	2.815528(22.78)**	1.961042(16.63)**

Z statistics are reported in parenthesis
** significant at 1% level.

TABLE 3. COEFFICIENTS WHEN WHEIGHTED LEAST SQUERED IS USED

Race \ Gender	Men	Women
All	2.213589 (83.54)**	0.5415235(15.55)**
White	2.02505(73.99)**	.3829716 (10.62)**
Black	2.92233(26.12)**	1.968973(16.31)**

** significant at 1% level.

TABLE 4. COEFFICIENTS FOR SPECIFIC GROUPS

Race\Gender	Male		
White	5.912317(50.54)** *17049*	1.137224(36.45)** *107375*	2.013074(27.97)** *21694*
Black	3.639316(12.99)** *2883*	1.725526(12.12) *6650*	1.63949(7.32)** *2320*
All	5.851574(55.44)** *20850*	1.215677(40.18)** *119449*	2.016202(29.87)** *24755*
Marstat[3]	**Single**	**Married**	**Divorced**

Race\Gender	Female		
White	4.798847(29.33)** *12290*	-.0535877(-0.86) *43792*	1.003196(11.11)** *20602*
Black	3.716468(12.21)** *2583*	1.198608(5.79)** *3843*	1.162509(5.51)** *3274*
All	5.188163(35.86)** *15440*	.0149512(0.25) *49852*	1.021496(12.36)** *24469*
Marstat	**Single**	**Married**	**Divorced**

Number in parenthesis is a t statistics. Italicized numbers are numbers of observations. ** Significant at 1%

[3] Martstat- marital status

APPENDIX B. FIGURES

FIGURE 1

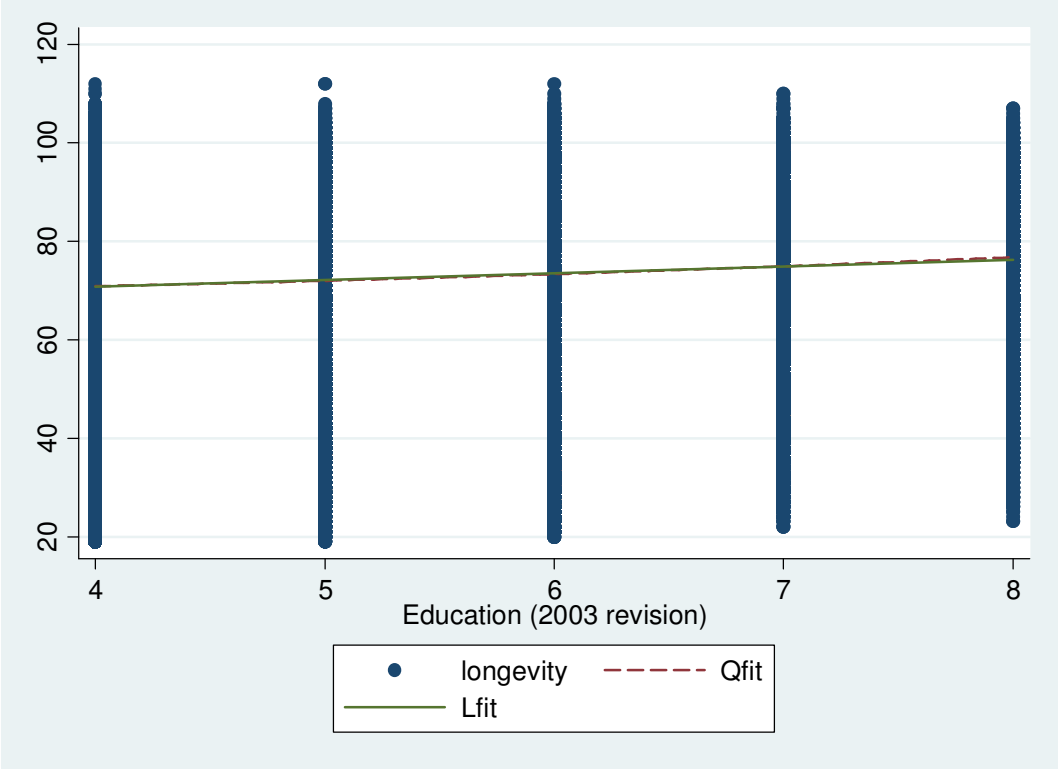


FIGURE 2

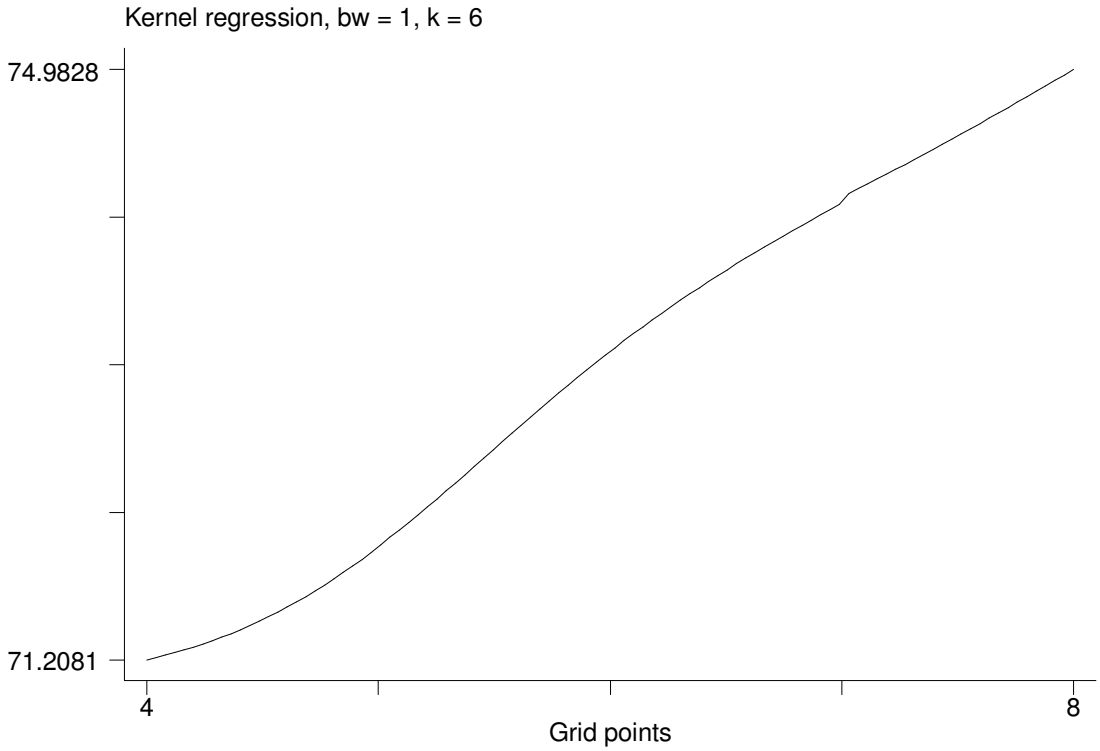
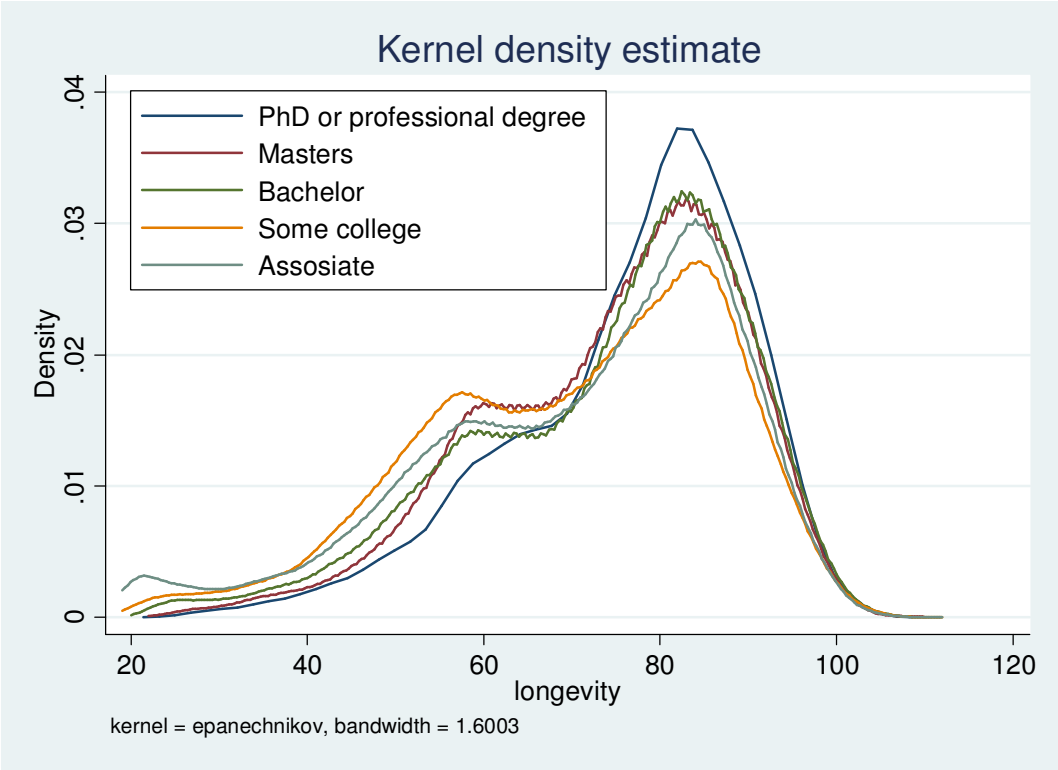


FIGURE 3


Kernel density estimate
PhD or professional degree
Masters
Bachelor
Some college
Assosiate
Density
0
.01
.02
.03
.04
20
40
60
80
100
120
longevity
kernel = epanechnikov, bandwidth = 1.6003

APPENDIX C

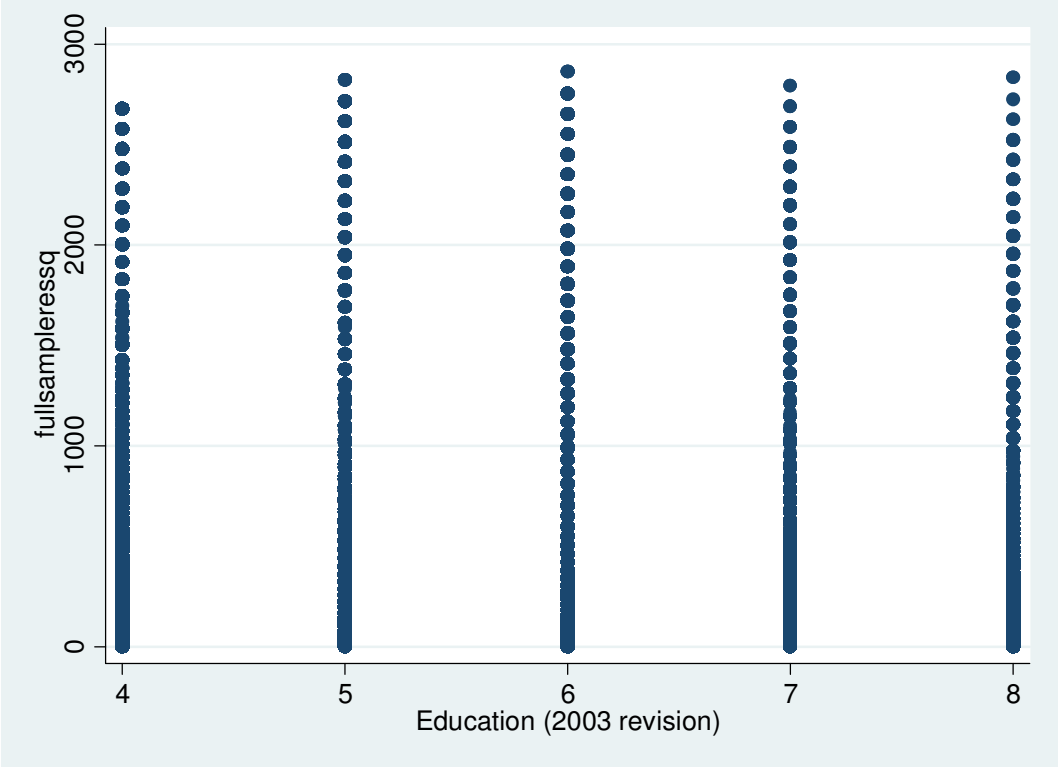
Regress longevity educ2003 if k==1, *first I regress longevity on education for full sample*

Predict fullsampleres, residual , *then I get residuals*

Generate fullsamplesq=fullsample^2 , *generate new variable that is equal to residuals squared*

Regress fullsampleressq educ2003 educ2003sq, *regress new variable on education and education squared*

Plot fullsampleressq educ2003*, create the graph to see how residuals are distributed. It is obvious from the graph given below that residuals are evenly distributed so there is no problem with heteroscedasticity*



R-squared is 0.0110, which means independent variable is not very good at predicting residuals. So there is no problem with heteroscedasticity.

VITA

Graduate School

Southern Illinois University

Pavlo Buryi

pasha@siu.edu

Southern Illinois University Carbondale

Bachelor of Arts, Economics, December 2010

Research Paper Title:

EFFECT OF POST-SECONDARY EDUCATION ON LONGEVITY

Major Professor: Scott Gilbert