Educational Attainment and Life Satisfaction

Hongyuan Zhang and Yuejun Chen

Abstract:

Previous research has shown contradictory results of the relationship between educational attainment and life satisfaction. In this study, we examine whether educational attainment can cause a change in life satisfaction in the U.S. by using instrumental variable regression with time fixed effects on county-level data. Our results indicate that there may not exist a causal relationship between educational attainment and life satisfaction. These results can provide policy makers with additional considerations when it comes to enacting policies which aim to improve public welfare.

JEL: C I

Keywords: Life Satisfaction, Education, Well-being, Welfare

Introduction and Literature Review

Maximizing society's utility has always been a focal point of economics; one of the direct interpretations of utility is life satisfaction. Much research centers on the determinants of life satisfaction and there are mixed results regarding whether higher-level education can improve people's life satisfaction. Investigating how education impacts happiness in the context of East Asia, Chen (2012) found that "by enhancing one's ability and propensity to connect with the wider social world, education may improve an individual's subjective well-being." Trostel (2010) also reported a significant positive relationship between educational attainment and happiness in his U.S.-based research. However, after conducting surveys in Britain, Stewart-Brown et al. (2015) found that people with the highest educational attainment tend to have higher levels of anxiety, as do the people in higher-level managerial or professional positions. Similarly, after analyzing the data from the Longitudinal Surveys of Australian Youth, Dockery (2010) concluded that students' happiness declines following the completion of their college degrees.

These contradictory results motivate us to investigate the relationship between educational attainment and life satisfaction. In this study, we examine the question of whether educational attainment affects life satisfaction in the United States and whether that relationship is causal. In addition to the educational attainment and life satisfaction variables, we collect 4 economic indicators, 1 general health indicator and 3 demographic indicators as control variables. These variables allow us to conclude a causal relationship between educational attainment and life satisfaction more confidently. The 4 economic indicators are GINI index, per capita income, poverty rate, and unemployment rate. We

include per capita income as a variable because Easterlin (2010) et al. found that "in the short-term happiness and income go together for a country, i.e., happiness tends to fall in economic contractions and rise in expansions." We include unemployment rate as well because Clark and Oswald (1994) found that "the effect of being jobless is, at any conventional level, statistically significant and is negatively correlated with well-being." It has also been found that inequality does matter to individual welfare in Latin America (Graham and Felton, 2005), so GINI index is included. Well-being differences have also been found among the extremely poor, the moderately poor and the not poor in Mexico (Lever 2004), inspired by which we include poverty rate in our study. We also include the general health variable. According to Subramanian, Kim, and Kawachi (2005), "poor health and unhappiness are highly positively correlated within individuals, and communities that are healthier tend to be happier and vice versa."

Unlike previous literature, we address the potential problem of double causality between educational attainment and life satisfaction by factoring in the number of institutions per capita in each state as an instrumental variable in our analysis. This helps us determine the one-way causal relationship between educational attainment and life satisfaction, namely education causing satisfaction and the extent to which educational attainment impacts life satisfaction. Our study contributes to the field by identifying county-level causal relationships and can help induce policy changes on access to higher-level education.

In this study, we fail to conclude that there exists such a causal relationship in all our models, including a model with an instrumental variable, a model without Per Capita

Income, a model focusing on Top 10 educated states, a model focusing on Bottom 10 educated states and a model without an instrumental variable. Our paper proceeds as follows. In the Data and Empirical Model section, we describe data selection, provide data overview and explain and justify our full model. In the Results section, we describe our primary results and secondary results, check model assumptions and discuss our study's limitations. In the Conclusion section, we discuss potential applications of our results.

Data and Empirical Model

We collect 11 variables (Table 1 in Appendix). One of them is our response variable, Life Satisfaction, one is an instrumental variable for Educational Attainment, the Number of Institutions Per Capita, and the other 9 are our explanatory variables. The data sources are the Behavioral Risk Factor Surveillance System (BRFSS), a national survey conducted by the Centers for Disease Control and Prevention (CDC), the American Community Survey (ACS), U.S. Department of Education and U.S. Census Bureau. According to the BRFSS documentation, a higher value for Life Satisfaction indicates less satisfaction and a higher value for General Health indicates less health. The observational unit is one county per year. We focus on the 2006-2010 period because BRFSS has stopped asking subjects to rate their life satisfaction since 2011. Therefore, we use the most recent available data in our study so our conclusion is more generalizable to the present. After we download the raw data containing every county in the U.S., we mark the counties that have values for all the variables and then remove the counties that have missing value(s). 172 counties from 47

states and Washington DC remain in our dataset. We then reshape the dataset to ensure that our observational unit is one county per year.

To address the potential problem of double causality between educational attainment (measured by rates of college completion) and life satisfaction, we use the number of institutions per capita of a state as the instrumental variable for educational attainment. It is a reasonable instrumental variable since we believe that 1) a county's rate of college completion should be higher if there are more institutions in this county's state because most students choose to attend colleges in their home states due to cheaper tuition, and 2) the number of institutions should not affect life satisfaction except through college attendance and completion. We recognize that the presence of colleges may bring in additional facilities and forms of entertainment such as concerts, art exhibitions, and sports events, which may increase the general life satisfaction of locals living in the area. However, the presence of colleges also causes noise in nearby neighborhoods, which can decrease the locals' life satisfaction. Thus, we believe that the combined effects of the factors mentioned above barely impact life satisfaction; to a large extent, the number of institutions per capita cannot directly affect life satisfaction unless they actually attend one institution. Through this instrumental variable, we hope to isolate the causal effect of life satisfaction on educational attainment and obtain the actual one-way causal effect.

After we plot each variable in boxplot graphs, we do not detect any outliers which are clearly due to measurement errors. We check the correlations of all possible pairs of variables, and no correlation's absolute value exceeds 0.6, and thus there is no severe multicollinearity among our explanatory variables. Since we have 8 control variables and since we believe that each can be correlated with both Educational Attainment and Life

Satisfaction and measure a different social dimension, we minimize omitted variable bias and can conclude causality more confidently. Even though we drop part of the data due to incomplete information, the variable averages of the remaining dataset are still close to the national averages (Table 2 in Appendix). Therefore, we are confident that our sample does not suffer much from selection bias and can be representative of all counties in the U.S. Table 3 shows descriptive statistics for Life Satisfaction and Educational Attainment.

	Mean	Std. Dev.
Life Satisfaction	.313889	.0921791
Educational Attainment	1.671784	.1425054

N = 860

Table 3 Summary Statistics for Life Satisfaction and Educational Attainment

$$Education_{it} = \beta_0 + \beta_1 NumberOfInstitution_{it} + \beta_2 Unemployment_{it} + \beta_3 Income_{it} + \beta_4 GINI_{it} + \beta_5 PovertyRate_{it} + \beta_6 Health_{it} + \beta_7 \%Male_{it} + \beta_8 MedianAge_{it} + \beta_9 \%White_{it} + \delta_t + \mu_{it}$$

(First stage model)

$$LifeSatisfaction_{it} = \beta_0 + \beta_1 Education_{it} + \beta_2 Unemployment_{it} + \beta_3 Income_{it} + \beta_4 GINI_{it} + \beta_5 PovertyRate_{it} + \beta_6 Health_{it} + \beta_7 \%Male_{it} + \beta_8 MedianAge_{it} + \beta_9 \%White_{it} + \delta_t + \mu_{it}$$

(Second stage model)

Model 1 Full Model

Model 1 is our instrumental variable regression model with time fixed effects where i=1, ..., 172 and t=1, …, 5. For example, Life Satisfaction$_{12}$ is the value of life satisfaction for the first

county in our dataset in the year 2007. Each β represents a regression coefficient that we will estimate. δ is a time fixed effects coefficient. μ is the error term. We fix the variation that comes from time since we want to find the actual effect of educational attainment on life satisfaction. We are unable to have county fixed effects since our instrumental variable only varies from county to county, not from year to year. We also use robust standard errors to fix the potential problem of heteroskedasticity.

Results

Life Satisfaction	(1) Full	(2) Omitted Income	(3) Top 10 States	(4) Bottom 10 Stets	(5) No Instrument
Education	0.151	0.269	-0.173	-0.201	-0.064
	(0.398)	(0.870)	(0.339)	(0.604)	(0.040)
Unemployment	1.711***	1.763**	0.051	0.539	1.599***
	(0.480)	(0.626)	(0.478)	(0.312)	(0.414)
Income	-0.000000811		0.00000240	-0.000000726	0.000000880
	(0.00000316)		(0.00000244)	(0.00000579)	(0.000000964)
GINI	0.360*	0.184	0.091	0.045	0.375*
	(0.173)	(0.748)	(0.131)	(0.314)	(0.165)
Poverty Rate	-0.373	-0.235	0.846**	0.305	-0.241
	(0.327)	(0.310)	(0.310)	(0.368)	(0.212)
Health	0.215**	0.249	0.250***	0.246***	0.184***
	(0.0718)	(0.199)	(0.0546)	(0.0581)	(0.0365)
%Male	0.929	1.051	0.156	0.433	0.784
	(0.593)	(0.919)	(0.529)	(1.264)	(0.558)
Median Age	-0.00258	-0.00229	0.00540*	0.00107	-0.00335*
	(0.00224)	(0.00312)	(0.00252)	(0.00511)	(0.00167)
%White	-0.074*	-0.061	-0.025	0.037	-0.078*
	(0.037)	(0.060)	(0.032)	(0.057)	(0.037)
_cons	0.558	0.394	0.616	0.736	0.741*
	(0.426)	(1.020)	(0.354)	(0.842)	(0.289)
N	860	860	250	90	860
F stat	6.119	1.159	6.353	4.195	

Robust standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 4 Regression Results

Table 4 shows our regression results. The first column is the result of our Model 1 (full model). The second to fifth columns in Table 4 show secondary results. For the full model, the results reveal insignificance (p-value = 0.379) of Educational Attainment. Thus we do not have enough evidence to conclude that educational attainment has an effect on life satisfaction and we may not conclude that a change in educational attainment will cause a change in life satisfaction. The coefficient of 0.151 indicates that a 1% increase in Educational Attainment will lead to a 0.151% increase in Life Satisfaction when holding other variables constant. However, since a higher value of Life Satisfaction means less satisfaction, our result means that more education can make you less satisfied about your life. However, the standard error of 0.398, a relatively large value compared to the coefficient, indicates that the coefficient is not precise. Thus, we should be careful when making use of our interpretation of the coefficient. The 95% confidence interval for Educational Attainment is (-0.63, 0.93), which is a wide range considering the range of Life Satisfaction (1-5). The wideness of this confidence interval indicates that educational attainment may make people feel very bad or good about their life. It is also possible that educational attainment has little or no effect on people's life satisfaction. Thus our results inform people that they should be conservative when thinking about whether there exists a relationship between educational attainment and life satisfaction.

Unemployment Rate, General Health, GINI Index and Percent White are statistically significant at significance levels of 0.001, 0.01, 0.05 and 0.05, respectively. This means these four variables each have an effect on Life Satisfaction, which confirms previous researchers' results. The sign of these variables' coefficients is reasonable considering a higher value of Life Satisfaction means less satisfaction. Unemployment Rate, especially,

has a large coefficient of 1.71 and is significant at the 0.001 significance level. This result indicates that status of economy can have a large effect on life quality.

A scatterplot (Figure 1) of Life Satisfaction against Educational Attainment confirms our previous insignificant result. As we can see, generally a county's residents are satisfactory about their life, regardless of how high the rate of college completion is. After examining the counties which have Life Satisfaction greater than or equal to 2, we find that most of these counties are in California; this phenomenon may be due to California's strong economy and thus strong competition for job positions. We run the full model again without these outliers and still find insignificance of Educational Attainment.

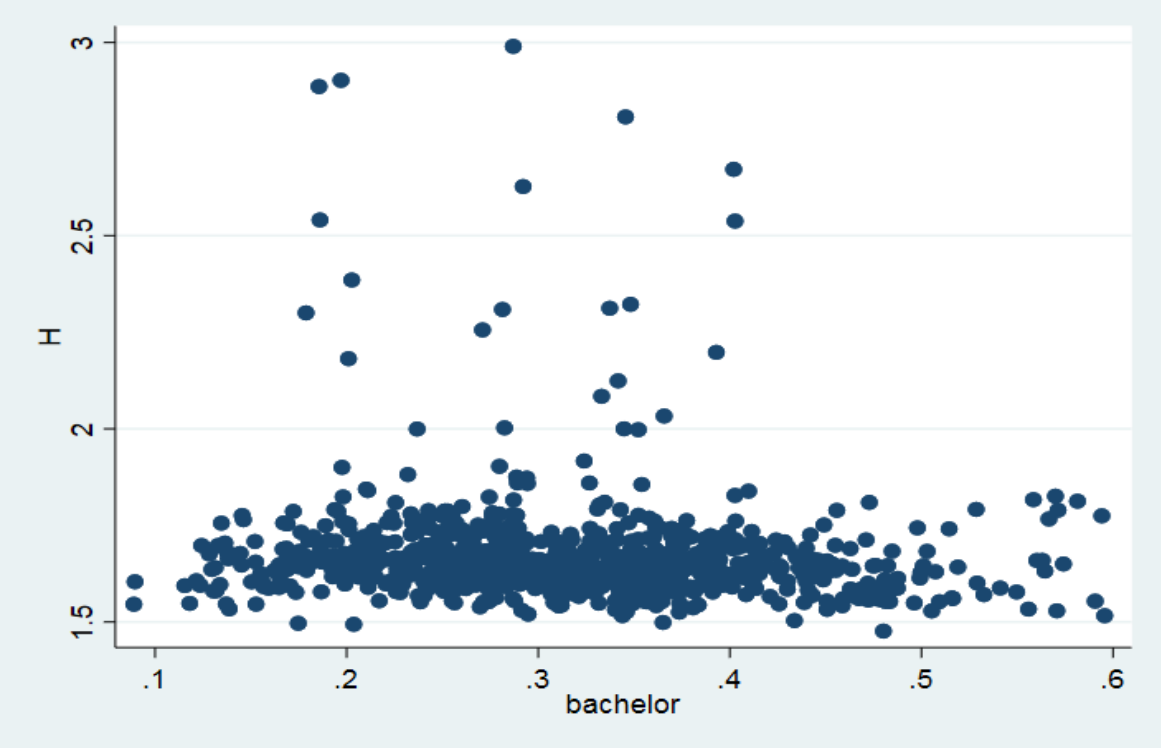

Figure 1 Scatterplot of Life Satisfaction and Educational Attainment (measured by rate of college completion)

After checking Residuals vs Observation Order plot and Residual Histogram (Figure 2 and Figure 3 in Appendix), we find that the residuals do not have a clear pattern of autocorrelation and are nearly normally distributed except for some moderate outliers. Although we use robust standard errors, we still check for heteroskedasticity. According to Residuals versus Fitted Values Plot (Figure 4 in Appendix), we do not have heteroskedasticity problem except for the outliers. The outliers in these plots are the same outliers in Figure 1. These residual plots suggest that we do not completely break model assumptions.

The second to fifth columns in Table 4 show our secondary results. The second time we run our model, we eliminate the variable Per Capita Income because we suspect that the correlation between Per Capita Income and Poverty Rate (0.60) may have induced multicollinearity in our model and biased our results. Hence, we drop Per Capita Income to check if the bias is severe. We also suspect the significance of Educational Attainment may depend on the level of Educational Attainment for each state, so we focus on the most 10 educated and least 10 educated states in our dataset, whose results correspond to the third and fourth columns in Table 4, respectively. Lastly, since our instrumental variable's F-statistic is 6.11 (< 10), we take out our instrumental variable to see whether our result may differ.

In general, our secondary results are consistent with our primary results. According to Table 4, Educational Attainment is insignificant at the significance level of 0.05 in all models. However, the sign of the coefficient changes for the last three models. Thus, the direction of the relationship may be reversed in Top 10 educated states and in Bottom 10

educated states. Unemployment Rate is significant in two of the four models. Additionally, General Health is significant at the significance level of 0.001 in all columns but the second one, where we drop Per Capita Income. However, the coefficients for General Health in all models are generally much smaller than those for Unemployment Rate. This indicates that being jobless can affect life satisfaction more than being unhealthy. Looking at the results in the second column in Table 4, we find that the coefficients do not change much from the first column (full model). Moreover, the standard error for each variable but Poverty Rate increases, and F-statistic becomes only 1.159. Thus, this indicates that the correlation of 0.60 between Per Capita Income and Poverty Rate does not bias our full model results much. Looking at the last column in Table 4, we find that the standard error for each variable but Percent White decreases, which is a sign of the weakness of our instrumental variable. In this model, the magnitude of the coefficient (0.0643) of Educational Attainment is largely different from that in the other models. This small coefficient suggests that Educational Attainment may have little effect on Life Satisfaction in fact.

It is worth noting that all five models may suffer from bias. For our full model, the weakness of our instrumental variable may induce additional bias that we cannot control for. The results in the second column may be affected by omitted variable bias since we take out Per Capita Income, a control variable that can be correlated with both Life Satisfaction and Educational Attainment. The third column and fourth column of results have selection bias since we arbitrarily select a subset of the data to analyze. The last column has an endogeneity problem because we do not use an instrumental variable to control for double causality. Although they have different biases, the results are consistent, especially the insignificance of Educational Attainment and the significance of General

Health. These results suggest that the biases may not be severe and in general there may not be a relationship between educational attainment and life satisfaction. They also suggest that in general there is a relationship between general health and life satisfaction.

Our insignificant results may be due to our study's limitations. First, the number of counties in our sample (172) is relatively small, considering the number of counties in the U.S. We drop many counties due to the limited availability of data. This limits our results' precision. Second, we do not have county fixed effects in our model. This is because our instrumental variable varies only from county to county, not from year to year. Hence, we are unable to control for unknown factors that vary from county to county. Third, all F-statistics are less than 10, which indicates that our instrumental variable is not strong enough for any model. The decreases in standard errors in the model without our instrumental variable also allude to the weakness of our instrumental variable. Finally, we only have one instrumental variable, and that is for Educational Attainment, the independent variable of our main interest. Nonetheless, there may exist double-causality problems between life satisfaction and other variables, for example, general health, and we do not have instrumental variables for each potentially endogenous variable. This may have biased our results.

Conclusion

This paper analyzes the impact of educational attainment on life satisfaction. By using instrumental variable – the Number of Institutions Per Capita – we hope to extract a causal relationship between educational attainment and life satisfaction. However, the

small sample size, lack of county fixed effects, weak instrumental variable and existence of other potentially endogenous variables reduce our level of confidence in making the conclusion. Our findings suggest that there is no significant relationship and thus no significant causal relationship between educational attainment and life satisfaction. The insignificance from our results, however, well deserve attention from the public. For many years, education has been given a lot of credit for solving various social issues; increasing governmental spending on education has also been a unanimous focus across the political spectrum. This paper contributes to the knowledge of social welfare by offering a quantitative approach to evaluate the effect of educational attainment on subjective well-being. The paper fails to draw a conclusion that is consistent with other evidence that supports the significant effect of education in society. Due to the limitations of our study, we are not confident in saying that the importance of education should be downplayed. Rather we offer policy makers additional considerations when it comes to making policies to improve the overall social welfare; they should be more cautious when concluding that subsidizing educational programs can improve the general well-being of the public. For example, programs that aim to improve the public's mental well-being may now consider subsidizing health programs and reducing unemployment rate, as both the general health and unemployment rate are shown to be significantly related to life satisfaction in most of our models.

In summary, there may not exist a causal relationship between educational attainment and life satisfaction. College education is neither an automatic money maker nor a magical antidote. While it provides us with extensive social networks, advanced technical skills, and unforgettable experiences, it also comes with financial burdens, higher

expectations for the future, and more social responsibilities. More research should be conducted to obtain more precise results if more data become available, a better instrumental variable that varies both from county to county and from year to year can be identified or additional instrumental variables for the potentially endogenous variables can be identified. Given the consistent significance of General Health in our models, more research should be conducted to investigate the effect of health status on life satisfaction, provided the researcher can find a reasonable instrumental variable for general health.

References

Chen, W. (2012, March). How Education Enhances Happiness: Comparison of Mediating Factors in Four East Asian Countries. Social Indicators Research, 106(1), 117-131.

Clark, A., & Oswald, A. (1994). Unhappiness and Unemployment. The Economic Journal, 104(424), 648-659. doi:10.2307/2234639

Dockery, A. M. (2010, April 19). Education and happiness in the school-to-work transition. Retrieved March 07, 2018, from https://www.ncver.edu.au/publications/publications/all- publications/education-and-happiness-in-the-school-to-work-transition#

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. Nations and Households in Economic Growth, 89-125. doi:10.1016/b978-0-12-205050-3.50008-7

Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness–income paradox revisited. Proceedings of the National Academy of Sciences of USA. doi:https://doi.org/10.1073/pnas.1015962107

Graham, C., & Felton, A. (2005). Inequality and happiness: Insights from Latin America. The Journal of Economic Inequality, 4(1), 107-122. doi:10.1007/s10888-005-9009-1

Lever, J. P. (2004). Poverty and Subjective Well-being in Mexico. Social Indicators Research, 68(1), 1-33. doi:10.1023/b:soci.0000025567.04153.46

Stewart-Brown, S., Samaraweera, P. C., Taggart, F., Kandala, N., & Stranges, S. (2015, March

19). Socioeconomic gradients and mental health: Implications for public health. The British Journal of Psychiatry. doi:10.1192/bjp.bp.114.147280

Subramanian, S., Kim, D., & Kawachi, I. (2005). Covariation in the socioeconomic determinants of self-rated health and happiness: a multivariate multilevel analysis of individuals and communities in the USA. Journal of Epidemiology and Community Health, 59(8), 664–669. http://doi.org/10.1136/jech.2004.025742

Trostel, P. (2010). It's Not Just the Money: The Benefits of College Education to Individuals and to Society. Retrieved March 7, 2018, from https://www.luminafoundation.org/files/resources/its-not-just-the-money.pdf

Appendix

Variable	Source
Life Satisfaction (Scale: 1-5)	CDC BRFSS City and County Data
Educational Attainment (Rate of college completion)	American Community Survey
Unemployment Rate	American Community Survey
Income per Capita	American Community Survey
GINI	American Community Survey
Poverty Rate	American Community Survey
General Health (Scale: 1-5)	CDC BRFSS City and County Data
Gender (Ratio of Male)	American Community Survey
Median Age	American Community Survey
Race (Ratio of White)	American Community Survey
Number of Institutions	US Department of Education
State Population	US Census Bureau

Table 1 Variables and Sources

Variable	Sample Average	National Average
Life Satisfaction	1.6717	1.63
Educational Attainment	0.3138	0.32
Unemployment Rate	0.0763	0.079
Per Capita Income	28561.84	27334
General Health	2.5197	2.52
%Male	0.4902	0.4927
Median Age	36.826	36.4
%White	0.7317	0.8065

N = 860

Table 2 Summary Statistics for Sample Averages Compared to the National Averages

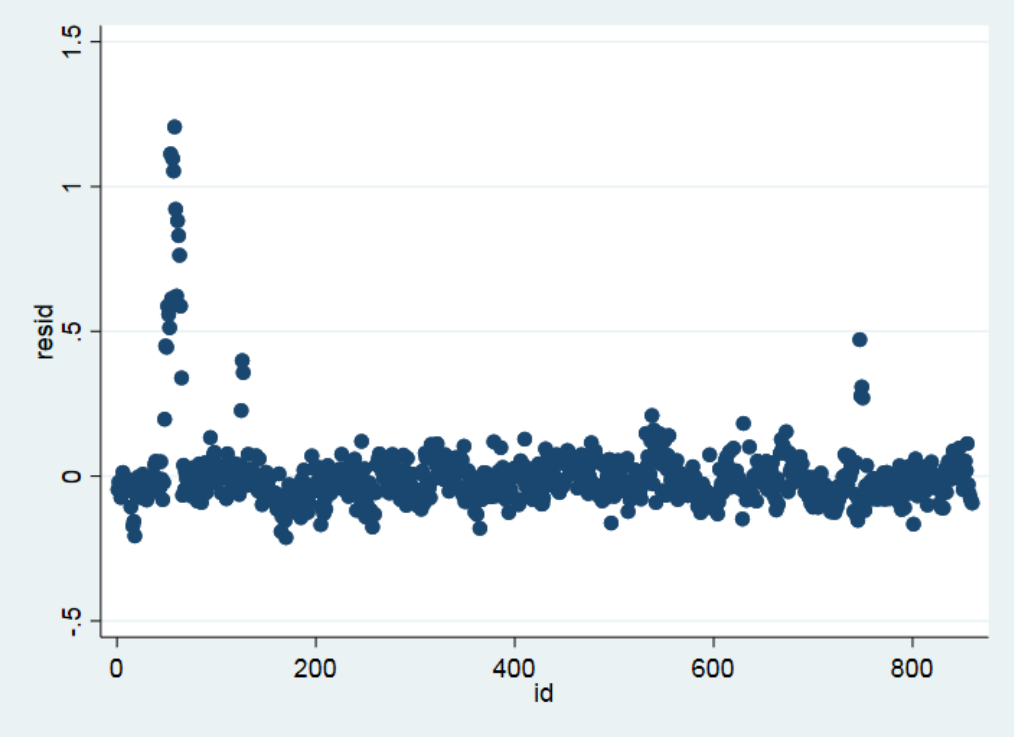

Figure 2 Residuals vs Observation Order Plot

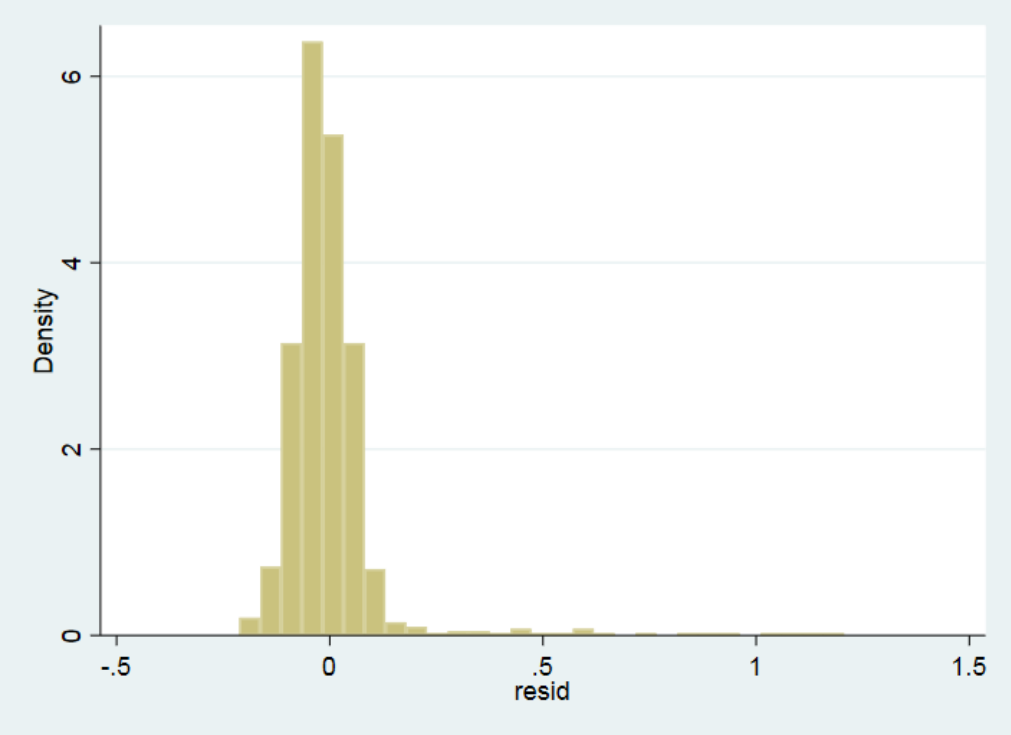

Figure 3 Residual Histogram

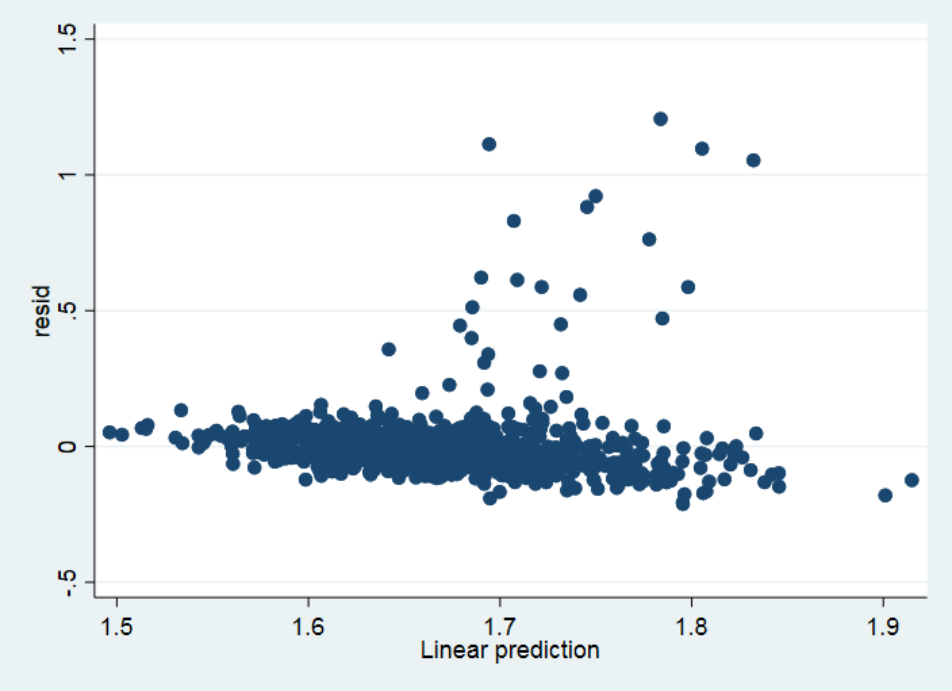

Figure 4 Residuals versus Fitted Values Plot