1. Introduction

Compulsory schooling laws are one of the most popular ways for economists to estimate the returns to education. A large number of empirical literatures have explored this methodology and concluded that additional schooling associated with compulsory school laws increases earnings in many countries including twelve European countries (Brunello et al, 2009), the US, Canada, and the UK (Oreopoulos, 2006). In contrast, a relatively small number of papers have otherwise focused on the effect of compulsory schooling laws on income inequality. Income inequality is an important issue that can affect society socially, politically and economically. If additional schooling can reduce the dispersion of earnings and inequality, then compulsory schooling laws can be seen as a powerful policy.

To answer "how education affects income inequality", we examine the effect of compulsory schooling laws on education and income inequality based on Stephens & Yang (2014). Using the 1960-1980 U.S. Censuses, we use the ordinary least squares (OLS) and two-stage least squares (2SLS) to estimate the impact of compulsory schooling laws on earnings and income inequality. We also use quantile regression to examine whether or not the effect of compulsory schooling laws is more significant on individuals who have a lower income. We find that compulsory schooling laws have positive impacts on the returns to education, which is consistent with the

previous empirical literature. Moreover, additional education associated with these compulsory schooling laws is efficient in reducing income inequality.

The paper is organized as follows: a brief review of the empirical literatures will be presented in Section 2. Data and methodology will be depicted in Section 3 and 4 respectively. Finally, results will be presented in Section 5 and discussed in Section 6, followed by conclusions.

2. Review of the Literature

The earlier study conducted by Angrist & Krueger (1991) examined the returns of schooling based on compulsory schooling laws and the season of birth. Angrist & Krueger (1991) use both the OLS and 2SLS methods, which estimate log weekly earnings based on years of education. Quarter of birth, an instrument for years of education, is a dummy variable indicating whether the individual was born in the first three quarters of a year. By using the U.S. 1970 and 1980 censuses for men who were born from 1920 to 1929, 1930 to 1939 and 1940 to 1949, the result showed additional education attainment associated with compulsory schooling laws was the cause of students earning higher wages.

To address the possible bias resulting from the weak correlation between education attainment and quarter of birth, Oreopoulos (2006) uses a dummy variable as an instrument to

indicate whether a cohort faced a minimum school leaving age of 15 when they were 14 years old. The OLS and 2SLS methods which regress log annual earnings on age-finished full-time school, age controls, and birth cohort polynomial controls, are both clustered by birth cohort and region. The estimated returns to schooling is around 15%, by using the United Kingdom's data that combines 15 U.K. General Household Surveys (GHS) from 1983 to 1998 with 14 Northern Ireland Continuous Household Surveys from 1985 to 1998 (Oreopoulos, 2006). A cross-country comparison was made by applying the same methodology for the U.S. and Canadian dataset, showing that effects on educational attainment caused by changes in the compulsory schooling laws in the United States and Canada were much smaller than in the United Kingdom.

Another study done by Brunello et al. (2009) references Oreopoulos's paper and uses the data from 12 European countries (the U.K. is not included). Based on the regression in the previous study, Brunello et al. added the individual's ability as one of the explanatory variables in a quantile regression framework. Furthermore, as compulsory schooling laws are different among these 12 countries, the instrumental variable is the number of years of compulsory education rather than the quarter of birth. Similar to prior results, they also found that the change in compulsory schooling laws has a greater impact on people with fewer skill sets. Additional schooling also has a positive effect on income inequality, reducing the disparity of income distributions.

On the other hand, Devereux & Hart (2010) reanalyze Oreopoulos's dataset by adding a gender dummy and using the law variable as an instrument, indicating whether the individual born before, after, or exactly in 1933[1]. Based on this, they find that the returns to schooling are only around 3% on average, which is much smaller than Oreopoulos's 15%. Similar results were shown by applying the same methodology on the New Earnings Survey Panel Data-set (NESPD) which has more earning information than GHS.

In the United States' context, Stephens & Yang (2014) use the 1960-1980 U.S. Censuses for their analysis. They conduct multiple tests with different compulsory schooling measures as their instruments, including required schooling (RS), child labor (CL) and corrected compulsory attendance (CCA). The results of the effect of additional schooling on log weekly wages are significant for both the OLS and 2SLS but become either different signs or statistically insignificant after adding region by year of birth indicators. Similar results are found by examining the impact of additional schooling on other outcomes, such as log occupational score, unemployed and divorced. For the explanation of these results, high discount rates and low private returns are the main reasons pointed out by the authors.

[1] Devereux & Hart (2010) use 1933 as breakpoint because the cohorts impacted by the law change were born after April 1, 1933. They also find that there are 75% of the 1933 cohort get affected, so the law variable is equal to 0.75 if an individual is born in 1933.

All the studies cited above use compulsory schooling laws as a tool to estimate local average treatment effects (LATE) because it has a powerful and immediate impact on a group of the population. However, most of them do not discuss the effect of compulsory schooling laws on income inequality. Building upon the previous studies, this paper aims to further examine how additional education associated with compulsory schooling laws affects income inequality in the United States.

3. Data

Similar to many studies, this paper uses the 1960-1980 U.S. Censuses to examine the benefits of schooling for the United States. The U.S. Census is a decennial national survey of individuals who live in the United States, including both citizens and non-citizens. The first U.S. Census was taken in 1790 and its microdata was available for research purposes since 1850, being a part of the Integrated Public Use Microdata Series (IPUMS). IPUMS is now the largest individual-level population database in the world, providing rich information for evaluating related topics in the United States[2], such as educational and healthcare programs.

[2] IPUMS consists of microdata samples from the United States (IPUMS-USA) and international (IPUMS-International) census records. As our analysis focus on the United States, IPUMS is referring to the IPUMS-USA in this paper.

Similar to the study performed by Stephens & Yang (2014), we use 1% of IMPUS 1960 Census, both 1% of IMPUS 1970 Census, and 5% of IMPUS 1980 Census sample microdata. We also narrow down our analysis on individuals between the ages of 25 to 55 across these three Censuses, which includes the birth cohorts from 1905 to 1954. Although the majority of our sample are Whites, we do not exclude other races. In Stephens & Yang (2014)'s analysis, they only focus on Whites as they find that the first stage estimates for Blacks are wrongly signed after adding differential trends across regions. However, this specification is not included in our analysis, so we use a dummy to indicate whether the individual is white as one of our control variables.

To increase the precision of our model, we use age, gender, race, quarter of birth, state of birth and year of birth from the 1960-1980 Census as our demographic data. For dependent and explanatory variables, we follow Stephens & Yang (2014) to use the highest grade completed as a measurement for years of schooling and divide annual earnings by weeks worked for log weekly earnings. Descriptive statistics[3] for the sample and the distributions of log weekly

[3] The table 3 presents the descriptive statistic for our data, which removes people whose income is zero, and excludes people who born in Alaska and Hawaii from the original dataset.

earnings, year of schooling and mandated year of schooling[4] are provided in Table 3, Figure 1, Figure 2 and Figure 3.

4. Methodology

Following Stephens & Yang (2014), we analyze the effect of compulsory schooling laws on returns to education and then examine the effect of compulsory schooling laws on income inequality by revising their equations.

We use equation (1) and (2) to estimate ordinary least squares (OLS) and two-stage least squares (2SLS) respectively.

$$Log\ Y_{st,i} = \alpha + \beta Educ_{st,i} + \chi_s + \delta_t + \gamma X_{st,i} + \epsilon_{st,i} \quad (1)$$

$$Log\ Y_{st,i} = \alpha + \beta \widehat{Educ_{st,i}} + \chi_s + \delta_t + \gamma X_{st,i} + \epsilon_{st,i} \quad (2)$$

where $Log\ Y_{st,i}$ is log weekly earnings, $Educ_{st,i}$ is the years of schooling of individual i born in state s in year t, and $\widehat{Educ_{st,i}}$ is the predicted years of schooling of individual i born in state s in year t gained from the first stage (equation 3). χ_s and δ_t represent fixed effects for state of birth and year of birth respectively. $X_{st,i}$ stands for all the other control variables, including a

[4] The figure 1 to 3 displays the unweighted distribution of log weekly earnings, year of schooling and mandated year of schooling. For figure 2, the number of observation with 12 years of schooling is much more than others as the sample consists of both 1% 1970 Census and the majority of the population for that finished 12 years of schooling. We adjust this problem when we run the regressions by using sample weights.

square, cube and quartic in age, Census year indicators, and an indicator for gender and race. For equations (1) and (2), our hypothesis is $\beta = 0$, meaning there is no relationship between years of schooling and log weekly earnings. In other words, if $\beta \neq 0$, then an individual's level of income is affected by his/her highest education level.

In correspondence with equation (2), we use the equation (3) to estimate the first stage

$$Educ_{st,i} = \rho + \pi_1 NY7_{st} + \pi_2 NY8_{st} + \pi_3 NY9_{st} + \lambda_s + \theta_t + \upsilon X_{st,i} + v_{st,i} \quad (3)$$

where λ_s and θ_t represent fixed effects for state of birth and year of birth respectively. $NY7_{st}$, $NY8_{st}$, and $NY9_{st}$ are as three schooling law instruments to indicate whether the requirement of years of schooling is seven, eight, nine or more, respectively[5].

To avoid biases for sample selection, we adjust the sample weight for Census 1970 as there are two 1% samples. We also create year by state of birth variable for clustering because the error terms are usually correlated with each other for those who have the same state of birth or year of birth. Based on these adjustments, we conduct the OLS and IV to examine how compulsory schooling laws affect the returns to education for individuals between 25 to 55 years of age. In addition, we estimate the first stage for the 0.1, 0.5 and 0.9 quantile respectively, testing whether

[5] Based on Figure 3, most of people have 6 to 8 mandated years of schooling, so we choose $NY7_{st}$, $NY8_{st}$, and $NY9_{st}$ as instruments to examine the returns to schooling associated with one more mandated years of schooling.

the impact of compulsory schooling law on average years of schooling is more significant for people who belong to the certain quantile of the distribution of earnings.

For the analysis of the effect of compulsory schooling law on income inequality, we revise equation (1) and (2) to be

$$\widetilde{Log\ Y}_{st,j} = \alpha + \beta\overline{Educ_{st,j}} + \gamma X_{st,j} + \epsilon_{st,j} \quad (4)$$

$$\widetilde{Log\ Y}_{st,j} = \alpha + \beta\widehat{\overline{Educ_{st,j}}} + \gamma X_{st,j} + \epsilon_{st,j} \quad (5)$$

where $\widetilde{Log\ Y}_{st,j}$ indicates the three differences of log weekly earnings of cohort j born in state s in year t: (1) 90th-10th Difference, (2) 50th-10th Difference and (3) 90th-50th Difference. $\overline{Educ_{st,j}}$ is the average years of schooling of cohort j born in state s in year t while $\widehat{\overline{Educ_{st,j}}}$ is the predicted average years of schooling of cohort j born in state s in year t gained from the first stage (equation 6). Similarly, $X_{st,j}$ represents all the other control variables, including a square, cube and quartic in age, Census year indicators, and an indicator for gender and race. Given equations (4) and (5), we are testing whether $\beta = 0$. If $\beta = 0$, there is no relationship between average years of schooling and the divergence of earnings. However, if $\beta \neq 0$ and it is statistically significant, then we can conclude that there is either a positive or negative impact of additional schooling on income inequality.

The corresponding first stage equation for our estimation is

$$\overline{Educ_{st,j}} = \rho + \pi_1 NY7_{st} + \pi_2 NY8_{st} + \pi_3 NY9_{st} + \upsilon X_{st,j} + v_{st,j} \quad (6)$$

where $NY7_{st}$, $NY8_{st}$, and $NY9_{st}$ represent the schooling law instruments, indicating whether the requirement of years of schooling is seven, eight, nine or more, respectively. We do not include fixed effects for state of birth and year of birth in equations (4) to (6) as we group individuals who have the same state of birth and year of birth as a cohort. With respect to that, the standard errors are calculated by heteroskedasticity-consistent (HC) standard errors instead of clustered standard errors and sample weights are adjusted by the sum of every individual's weight in a cohort.

5. Results

5.1 The Impact of Schooling on Wages

The column (1) of Table 1 shows our examination of the causal effect of schooling on log weekly earnings for individuals ages 25 to 55 from the 1960-1980 Censuses of Population. Based on equations (1) and (2), the OLS and 2SLS estimates of the returns to schooling are 7.22% and 12.4% respectively, which are both statistically significant. As the weak instrument threshold of the first stage is 10, our result is well above it with an F-statistic of 45.58. Usually, the more required years of schooling, the higher the levels of schooling. Therefore, the first stage point estimates are all consistent with the monotonicity of the instrument except for $NY7$[6].

[6] However, the estimate of $NY7$ is not statistically significant, so it can be ignored.

5.2 The Impact of Schooling on Income Inequality

To examine the causal effect of schooling on income inequality for individuals between the age of 25 to 55 from the 1960-1980 Censuses of Population, we group the individual by state of birth and year of birth. Using the sample of 2449 cohorts, the OLS and 2SLS estimation shown in columns (2) to (4) of the Table 1. We find that the effect is most significant for 90^{th}-10^{th} difference, followed by 50^{th}-10^{th} and 90^{th}-50^{th} difference under both frameworks. This result shows that the increased educational attainment associated with compulsory schooling laws is useful for reducing the divergence in income. The first stage point estimates increase along with the number of required schooling years, consistent with the monotonicity of the instrument. Besides, these instruments are jointly significant with a F-statistic of 194.06.

To prove those who belong to the 0.1 quantile benefit more from compulsory schooling laws than those of the 0.5 quantile or 0.9 quantile, we conduct another first stage on these three sub-datasets. As shown in Table 2, the F-statistics are 52.1, 9.3, and 0.58 for the 0.1 quantile, 0.5 quantile and 0.9 quantile respectively. This means that the joint significance of NY7, NY8 and NY9 is decreasing with the higher quantile of log weekly earnings. Moreover, the same trend is found for the first stage point estimates of each instrument. Therefore, we conclude that the effect of compulsory schooling laws on additional education attainment is more significant for individuals who have lower income, and it is only statistically significant for the 0.1 quantile.

5.3 The Limitations of Results

Our results show that the returns to schooling using compulsory schooling laws are around 7% and 12% under the OLS and 2SLS frameworks respectively. Much research finds that the returns to additional schooling corresponding to the change of compulsory schooling laws in Europe are not as high as that in the United States. For example, Devereux & Hart (2010) find that the estimated returns due to the 1947 compulsory schooling law change in Britain are only 7% for men and 0% for women. Pischke & von Wachter (2008) find no evidence of any positive returns to additional schooling resulting from changes in compulsory schooling laws in Germany after World War II. If the returns to increased schooling are not significant or even zero, then it is not a useful channel to increase individuals' earnings. In other words, the impact of additional schooling on income inequality will not hold, so our result may not apply across all countries.

Based on Pischke & von Wachter (2008), they suggest researchers to focus on the impact of schooling's quality on earnings rather than the amount of education on earnings. For the discrepancy in results between the U.S. and Germany, they explain that compulsory schooling laws work only if employers use the highest grade completed as the signal and not discriminate between individuals who attend an additional education voluntarily and those who were forced. However, we are not able to determine whether the highest grade completed is the signal based

on our dataset and the results of our analysis may change if we replace years of schooling with schooling quality as the independent variable.

Furthermore, the evidence shows that education and ability are substitutes in terms of generating earnings according to Brunello et al. (2009). Our analysis does not include individuals' ability as it is uncorrelated with the mandated years of schooling under the assumption of 2SLS framework. This implies that our results cannot examine the effect with the different quantiles of education level and ability level. In other words, we only present the causal effect of compulsory schooling laws on income inequality based on the mean of education. These factors may decrease the reliability of the application of our analysis, thus, are the limitations of our results.

6. Discussion

6.1 The Policy implications

There are a few policy implications of our results. As compulsory schooling laws have a significant effect on earnings, especially to those who are at the lower end of the distribution of earnings, policy makers should consider compulsory schooling laws as a tool while setting the policy for addressing income inequality. The effect of additional schooling not only has a direct impact on income inequality, but also boosts a country's economy as income inequality is correlated with many social, political and economic issues. Aside from that, the current

distribution of years of schooling could influence the impact of compulsory schooling laws. According to our results, we find that the more required years of schooling compared to the mean, the more significant the aggregate effect is. Therefore, policy makers should take this into account and decide whether this change in compulsory schooling laws is worth it based on cost-benefit analysis. Considering the aforementioned limitations, policy makers should also examine individuals' skills associated with their incomes, determine the signal in the labor market, and evaluate the reliability of the studies based on their dataset while setting compulsory schooling laws. These can assist policy makers on making education policies, along with the help of cost-benefit analysis.

6.2 The Future Research

As mentioned in the limitations of results, there are a few factors that reduce the reliability of our results. To address this issue, one possible extension could reference Brunello et al. (2009), which computes the returns to schooling with respect to the different quantiles of ability and earnings at the same time. By using a quantile regression, we can also estimate the effect of compulsory schooling laws on earnings while taking the different level of ability into account. To examine the main reason causing the discrepancy in results between North America and Europe, future research can also apply the same methodology to samples from various

countries, and compare the results by using the level of education and schooling quality as the explanatory variable. Based on the current research, we would like to conduct another study with data consisting of countries with different levels of development. These help justify the difference between results and applications across the countries.

7. Conclusion

This paper presents analyses of additional schooling associated with compulsory schooling laws' impact on earnings and income inequality. By using the OLS and 2SLS, we demonstrate that additional education attainment leads to higher earnings. We then conduct a quantile regression for the 0.1, 0.5 and 0.9 quantile, as well as estimate the 90^{th}-10^{th}, 50^{th}-10^{th} and 90^{th}-50^{th} difference. The result shows that people with lower income are more likely to have more schooling due to compulsory schooling laws than those who have a higher income, thus compulsory schooling laws are efficient in decreasing income inequality.

Our research shows the causal effect of additional schooling on income inequality, which should be considered by policy makers while setting the education policy. To sum up, continuous research addressing the limitations can improve the reliability and application of our results.

Bibliography

Angrist, J. D., & Krueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? The Quarterly Journal of Economics, 106(4), 979–1014. doi: 10.2307/2937954

Brunello, G., Fort, M., & Weber, G. (2009). Changes in Compulsory Schooling, Education and the Distribution of Wages in Europe. *The Economic Journal*, *119*(536), 516–539. doi: 10.1111/j.1468-0297.2008.02244.x

Devereux, P. J., & Hart, R. A. (2010). Forced to be Rich? Returns to Compulsory Schooling in Britain. *The Economic Journal*, *120*(549), 1345–1364. doi: 10.1111/j.14680297.2010.02365.x

Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter. *American Economic Review*, *96*(1), 152–175. doi: 10.1257/000282806776157641

Pischke, J.-S., & Wachter, T. V. (2008). Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. *Review of Economics and Statistics*, *90*(3), 592–598. doi: 10.1162/rest.90.3.592

Stephens, M., & Yang, D.-Y. (2014). Compulsory Education and the Benefits of Schooling. *American Economic Review*, *104*(6), 1777–1792. doi: 10.3386/w19369

Table 1

OLS and IV Returns to Compulsory Schooling Estimates for Log Weekly Earnings

(The United States, ages 25-55, 1960-1980)

	Whole Sample	90^{th}-10^{th} Difference	50^{th} -10^{th} Difference	90^{th} -50^{th} Difference
	(1)	(2)	(3)	(4)
OLS	0.0722	-0.1135	-0.0923	-0.0211
	(247.317)***	(-11.898)***	(-10.72)***	(-5.6497)***
2SLS	0.124	-0.1021	-0.0793	-0.0228
	(6.781)***	(-3.5937)***	(-3.0887)***	(-2.0643)**
First Stage:				
NY7	-0.0145		0.0456	
	(-0.25)		(1.2401)	
NY8	0.1527		0.2287	
	(3.329)***		(6.9195)***	
NY9	0.3032		0.3804	
	(6.097)***		(10.297)***	
F-test	45.58		194.06	
N	4170452		2449	

Notes: Figures in parenthesis are z-statistics for column (1) and t-statistics for columns (2) to (4)

***,**,*,denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 2

First Stage Effects of Compulsory Schooling Laws on Education Attainment for the 0.1, 0.5 and 0.9 Quantile

of Log Weekly Earnings

(The United States, ages 25-55, 1960-1980)

Sample	0.1 Quantile (1)	0.5 Quantile (2)	0.9 Quantile (3)
NY7	-0.0121	-0.0789	-0.0334
	(-0.283)	(-1.236)	(-0.448)
NY8	0.231	-0.0402	0.0091
	(5.592)***	(-0.781)	(0.17)
NY9	0.3671	0.0579	0..0273
	(8.259)***	(1.124)	(0.507)
F test	52.1	9.3	0.58

Notes: The dependent variable is years of schooling. Each Regression includes age Figures in parenthesis are z-statistics.

***,**,*,denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3: Descriptive Statistics

Variable	*N*	*Mean*	*Std. Dev.*	*Min*	*Max*
Census Year	4170452	1976.04	6.51	1960	1980
Log (Weekly Earnings)	4170452	5.18	0.91	-2.34	11.23
Years of Schooling	4170452	12.38	2.91	0	17
Mandated Years of Schooling	4170452	8.28	1.29	0	11
Male	4170452	0.58	0.49	0	1
White	4170452	0.88	0.32	0	1
Age	4170452	37.98	8.73	25	55
quarter of birth	4170452	2.52	1.11	1	4
state of birth	4170452	28.65	15.25	1	56
year of birth	4170452	1937.3	11.3	1905	1954
year by state of birth	4170452	3760.7	1129.48	501	5456

Figure 1 : Distributions of Log Weekly Earnings

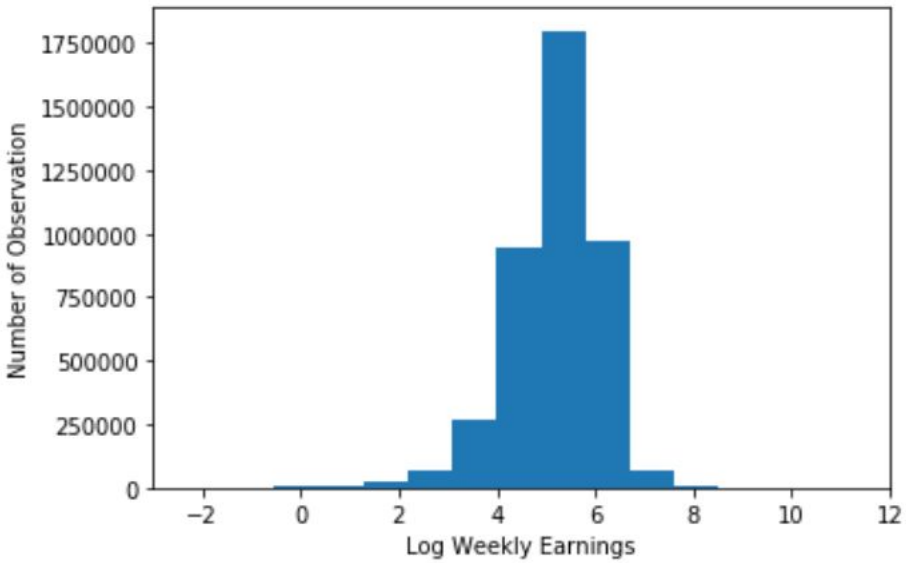

Figure 2 : Distributions of Years of Schooling

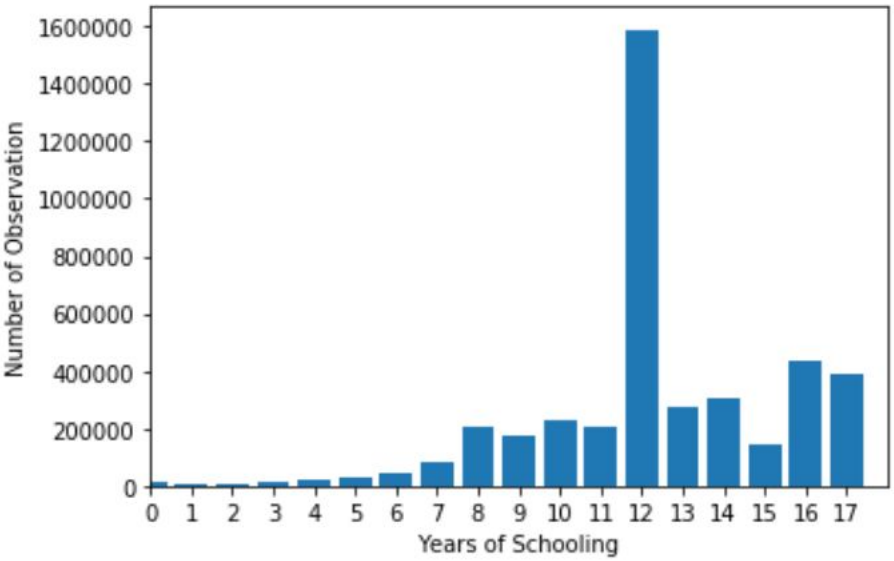

Figure 3 : Distributions of Mandated Years of Schooling