

# Silver Spoon Effect on Education: Analysis of The Effect of Household Wealth Toward Children's Educational Output

Hyecheol Jang<sup>\*1</sup>, Carleigh Heintz<sup>\*2</sup>, Chang Su<sup>\*3</sup>, Desmond Fung<sup>\*4</sup>, and Tsz Yau Iris Chow<sup>\*5</sup>

**Abstract**—This research shows the relationship between the wealthiness and willingness to invest toward the children's education and the educational output of those children. We utilize mean income, private school enrollment rate, demographics, and employees' highest diploma for an area as features and ACT score and high school graduation rate of people between 18 and 24 years old as output to fit our linear regression model.

Showing the statistical significance of mean income, we verified that wealthiness affects academic success. By combining the fact that people with higher degrees earns more money in future, the concerns about the silver spoon effect caused by education has been proved.

## I. INTRODUCTION

Nowadays, the majority of countries and societies are having a hard-time addressing long-resting social issues: polarization. Modern society rarely has an identifying system, but due to the extreme polarization, it is not difficult to see separated "classes", classified people's life-style by wealthiness, in the real world. With the instinct of mankind, everyone wants to be a part of upper-class society, and they, especially middle and lower class households believe education could be a ladder to go up to the upper class. According to Rugarber (2017, [1]), the higher-education degree holder gained about 1.5 times higher salary compared to those who only have a high-school diploma. Also, by showing two sets of data with different time-frames, in the years 1999 and 2015, Rugarber was able to address that the wage gap becomes wider over time (para. 1-3). Seeing the previous analysis of the data, it looks like the higher educational status makes a higher salary, which lets people move toward the upper class.

However, education opportunities are not equally provided for all people in society. With investment of additional resources on children's education, paying tuition for private school, hiring tutors, and even sending their kids to other countries, people expect to bring better opportunity to their kids; therefore, education could be one of the ways to transfer the wealth of parents' generation to their kids, making polarization even worse. By conducting this research, we want to verify the existence of the "silver spoon effect" - if parents are wealthy, their kids are more likely to be wealthy - caused by education. Specifically, we want to show that the areas

with higher economic status and with the strong willingness of investing more on their next generations' education have better educational output, expecting we can see a positive relationship between wealth and educational results (ACT Score, High School Graduation Rate, Statewide test score). **Our analysis shows that areas with higher economic status have better educational output in Wisconsin.**

By showing the relatedness of wealth and educational output, we, ultimately, want to verify whether education might be used as a tool for "upper classes" to indirectly shift their capitals and superior societal position to the next generation.

## II. DATA COLLECTION / ANALYSIS

Before going through how we collected and analyzed the datasets, note that all relevant dataset, codes, and documents are in our GitHub repository (URL: <https://github.com/hyecheol123/STAT333-Edugression>)

### A. Wisconsin Income by County

Before analyzing the relationship between the economic status and the educational output, we started to make a graph of annual wage per employee for each county in Wisconsin. This can be used gauge which citizens within the counties of the state are making more money compared to others, we are using this as a proxy variable for economic status.

The data has been retrieved from Bureau of Labor Statics (BLS)'s "Quarterly Census of Employment and Wages" data, and we used year 2017's census data (QCEW News Releases., n.d., [2]).

**Database search criteria** has been listed below.

- All Counties in a State, One Industry
- Counties in Wisconsin
- Year: 2017
- Quarter: Annual Averages
- Ownership: Total, All Ownerships
- Industry: 10: Total, All Industries
- Do not Include records with suppressed employment and wages

By selecting specific columns based on "QCEW Open Data Access: CSV Data Slices" (Data Slices by Industry., 2016, [3]), we are able to make two plots for "Total Annual Wage" and "Annual Wage per Employee" for each of the counties of the Wisconsin. The unit of income has been measured in USD(U.S. Dollar) and the unit of analysis is counties (FIPS code). Note that we understood the total/annual wage is synonym of total/annual income for each county, and from

<sup>\*</sup>Everyone is an undergraduate student of the Department of Statistics, University of Wisconsin - Madison (Madison, WI, 53706, USA), taking STAT 333 during FALL19, with professor Karl Rohe.

<sup>1</sup>Hyecheol Jang [hyecheol.jang@wisc.edu](mailto:hyecheol.jang@wisc.edu)

<sup>2</sup>Carleigh Heintz [cmheintz@wisc.edu](mailto:cmheintz@wisc.edu)

<sup>3</sup>Chang Su [csu29@wisc.edu](mailto:csu29@wisc.edu)

<sup>4</sup>Desmond Fung [dfung2@wisc.edu](mailto:dfung2@wisc.edu)

<sup>5</sup>Tsz Yau Iris Chow [tchow7@wisc.edu](mailto:tchow7@wisc.edu)

now on, this paper will use the term "income" as a synonym of wage, including the original meaning of the term income.

2017 Total Annual Wage of counties in the Wisconsin

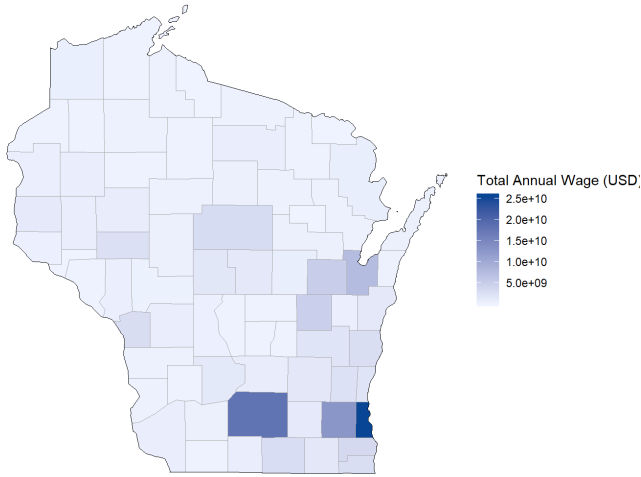


Fig. 1. Plot for Total Annual Wage(Income) of counties in the Wisconsin

Seeing the plot for Total Annual Wage(Income) of counties in Wisconsin (Fig. 1), it can be concluded that the Madison and Milwaukee areas can be considered the economic centers of the Wisconsin, but as this plot indicated cumulative total wage that **all** workers in each county made, it does not consider the populations of each county, which allows for counties with larger populations to have much greater total wage than counties with smaller populations.

2017 Annual Wage per Employee of counties in the Wisconsin

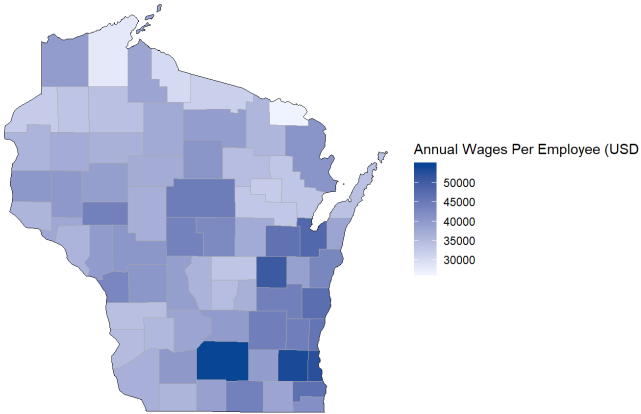


Fig. 2. Plot for Annual Wage(Income) per Employee of counties in the Wisconsin

Seeing the plot for Annual Wage(Income) per Employee of counties in Wisconsin (Fig. 2), it can still be concluded that the Madison and Milwaukee areas are considered as the economic center of the Wisconsin, but as this plot indicated average employee's income, it is more likely to represent household economic power compare to the previous plot.

Therefore, for further analysis, we decided to use average income to represent household economic status.

### B. Wisconsin Mean Income by zip code

To get more specific data, our team decided to use data set coded by zip code. To begin with, we analyzed the mean income for each zip code, which indicates household economic status.

The data has been retrieved from American Community Survey (ACS), and we specifically used year 2017's census data for Mean Income In The Past 12 Months (In 2017 Inflation-Adjusted Dollars) from 2013-2017 American Community Survey 5-Year Estimates (American FactFinder ..., 2010, [4]). From the cited database, we used **search criteria** below.

- Table Name: S1902
- Select 2017 data
- Added Geographic after Search: 5-Digit Zip Code Tabulation Area - 860
- State: Wisconsin
- All 5-Digit Zip Code Tabulation Areas

By selecting Mean Income - Estimate for each zip code, we were able to get the mean income for each zip code. The unit of income is USD, and the unit of analysis is each area using same zip code.

2017 Mean Income for each Zipcode in the Wisconsin

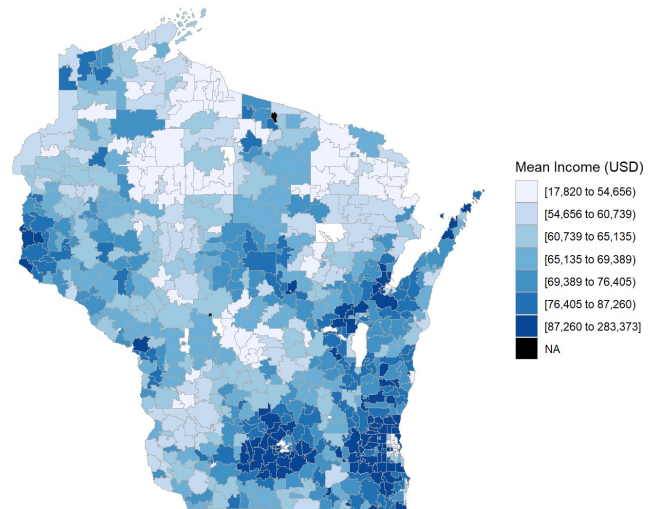


Fig. 3. Plot for 2017 Mean Income for each Zip Code in the Wisconsin

Seeing the plot of wage coded by zip code (Fig. 3), we can still conclude that Madison and Milwaukee area's mean incomes are still higher than other areas. However, we are able to intuitively see that there exist outliers in this data (See the legend), which is not represented in the plot clearly.

Looking through the data set, the area zip code 53047 is an outlier, where its mean income is reaching near 300 thousand, while others are less than 200 thousands. After some

research, we found that this zip code belongs to Lebanon, Wisconsin, of which population is only 1664. Concluding from this analysis, among those small number of citizens in Lebanon, on that zip code, very small number of people who earns a lot of money might live there. Additionally areas with smaller populations are less resistant to individual outliers within the population, which may skew the data.

#### C. Ratio of Children in Private School by Zip Code

To check the willingness of payment from parents towards their children's education, we decided to use private school enrollment ratio as a proxy. As there is no database that directly indicates the direct enrollment ratio for the students in private schools but we were able to find the number of female and male students enrolled in both private and public schools coded by zip code. Using this data, we calculated the ratio of children in private school for each zip code.

This data was retrieved from American Community Survey (ACS), and the condition is same for the other data set we retrieved from ACS (See Subsection B), except for the table name: **B14002** (American FactFinder ..., [4]).

From that data, we calculated the total number of all students living in the area coded by the designated zip code and the number of students attending private schools. Then, by dividing the latter one by the total number, we finally calculated the ratio which we were interested in.

#### D. Demographic Distribution by Zip Code

We used the demographic data of Wisconsin to see if race has an effect on people's willingness to pay for children's education and to see how racial makeup of an area plays a role in educational output. The data was retrieved from American Community Survey (ACS), and the condition is the same for the other data set we retrieved from the ACS, except for the table name: **B02001** (American FactFinder ..., 2010, [4]).

We used the proportion of Caucasians, African Americans, and Asians for each zip code. The data contains the number of the total population and the head counts for each demographic, we went through the data and calculated the ratio for each race represented.

#### E. Employee Education by Zip Code

In our models, we included education status as variables. We use this as an indicator to analyze the assumption that education status of parents (the employees) will imply how much money parents might invest in furthering their child's education. For instance, if a child's parents went to college, they will likely want their children to go to college as well.

The data we found for this included total number of bachelors, masters, PhD and professional degrees for each zip code in Wisconsin. This data was retrieved from American Community Survey (ACS), and the condition is same for the other data set we retrieved from ACS (See Subsection B), except for the table name: **B15003** (American FactFinder ..., [4]). We did some modification by dividing each count by the total, and found out the proportion for each level of education.

#### F. High School Graduation Rate by Zip Code

In addition to ACT score, we also used the graduation rate as an estimate for a school's academic output. The graduation rate may be an indirect measure of the quality of the education, and we expect that schools in areas with higher mean income will have higher graduation rate.

The data was retrieved from American Community Survey (ACS), and the condition is the same for the other data set we retrieved from ACS, except for the table name: **S1501**. Note that this data only contain graduation rate among the population ranging age of 18 to 24, which indicates recent graduates.

#### G. ACT Score by School's Zip Code

Since the ACT is issued statewide to public schools every year, it can be used as a good estimate for a school's academic output, as all public school students in Wisconsin are required to take it. This data from public schools is available every year from the Wisconsin Department of Education, and we used it as it was published by the Journal Sentinel for the year 2017 (Wisconsin ACT scores..., n.d., [5]).

In the data, the average ACT Score (between 0 and 36) for each school was reported. Since this data was coded by school name instead of zip code, unlike all of the other data sets we used, we manually looked up the zip code for each school to be able to use for the data analysis.

2017 Wisconsin Public School's ACT Score by zipcode

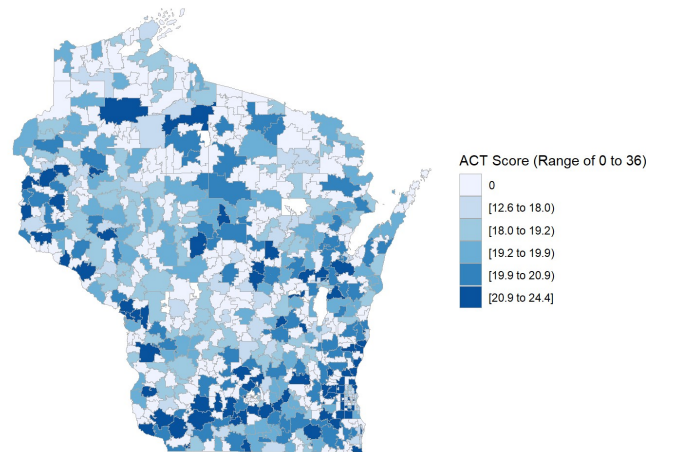


Fig. 4. Plot for 2017 ACT Average score for Wisconsin's Public Schools

Comparing this data's figure (Figure 4) to Figure 3, similar patterns can already be observed between areas in Wisconsin with higher incomes and areas which have higher average ACT Scores. This pattern is especially prevalent in inner city Milwaukee and its surrounding areas. Comparing both graphs, it can be seen that inner city Milwaukee, an area of generally significantly low income, also has significant low ACT scores. Contrasted with it's surrounding areas, of higher income, that have significantly high ACT scores.

### III. STATISTICAL ANALYSIS

To conduct our analysis, we decided to use two different multiple linear regression models for our two different output variables, ACT score and graduation rate.

#### A. Statistical Model for ACT Score

```
Call:
lm(formula = ACT ~ ., data = ACT)

Residuals:
    Min       1Q   Median       3Q      Max
-8.1275 -0.6315  0.1330  0.9446  6.2148

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.961e+00  1.049e+00   9.497 < 2e-16 ***
whiteRatio    5.621e+00  1.158e+00   4.856 1.61e-06 ***
balckRatio    8.660e-01  1.291e+00   0.671  0.5028
asianRatio    4.549e+00  4.406e+00   1.032  0.3024
Assoc.Ratio   6.612e+00  3.665e+00   1.804  0.0718 .
Bachelor.Ratio -2.039e+00  2.566e+00  -0.795  0.4273
Master.Ratio  4.620e+00  5.276e+00   0.876  0.3817
Prof.Degree.Ratio 1.036e+01  1.220e+01   0.849  0.3963
Doctor.Ratio  -4.580e+00  1.298e+01  -0.353  0.7243
Mean.income   4.952e-05  7.868e-06   6.293 6.81e-10 ***
Private.Ratio -2.798e-01  1.260e+00  -0.222  0.8243
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.765 on 501 degrees of freedom
(373 observations deleted due to missingness)
Multiple R-squared:  0.4346, Adjusted R-squared:  0.4233
F-statistic: 38.5 on 10 and 501 DF, p-value: < 2.2e-16
```

Fig. 5. Linear Regression Model for ACT  
Diagnostic Plots are located at the end of article

When making a linear model to analyze the relationship between our features and ACT score (See Fig. 5), we used a linear regression with ACT score as our dependent variable and demographic (White, Asian, Black ratio), parent's education level (Associate, Bachelor, Master, Prof.Degree, Doctor ratio of the zip code), mean income, and private school enrollment ratio as our independent variables.

The result shows that for the ACT test score, the significant variables were white ratio ( $p = 1.61e-16$ ) along with mean income of the zip code ( $p = 6.81e-10$ ).

We expected that mean income would have a significant effect on student's ACT Score, because wealthier students can afford to take the test several times, which has been known to increase a students' score. Besides that, wealthier students typically attend better-funded schools with a significant number of AP classes and are more likely to have access to tutors and standardized test preparation classes.

Interestingly, the willingness of putting additional money on the children's education, proxied by private school enrollment rate, does not play a significant role. Knowing that we only have the ACT score data for public schools and the willingness of investment toward education cannot only be illustrated by this simple factor, this result is not that surprising.

Note that, parent's education level is not significant in relation to student's ACT score, which is out of our expectation because we expected the higher the academic degree earned by parents, the higher the students' test scores.

Interestingly, the white ratio is the only significant variable among the demographic, we believe this is because areas of higher income and thus better schools, tend to be areas with a higher white ratio, such as the suburbs. Additionally, Milwaukee has one of the highest rates of segregation in the United States, which leads to low-income areas full of minority populations that have poor quality schools, and thus poor educational outputs. This factor could be influencing the model due to the large amount of Milwaukee Public Schools (MPS) that are represented in the data. When looking at the individual data points for MPS schools, many of their average ACT scores are significantly below the average. There are also significantly more schools in Milwaukee than any other area in Wisconsin.

```
Analysis of Variance Table

Model 1: ACT ~ ACT + Assoc.Ratio + Bachelor.Ratio + Master.Ratio + Prof.Degree.Ratio +
Doctor.Ratio + Mean.income + Private.Ratio
Model 2: ACT ~ whiteRatio + balckRatio + asianRatio + Assoc.Ratio + Bachelor.Ratio +
Master.Ratio + Prof.Degree.Ratio + Doctor.Ratio + Mean.income +
Private.Ratio
Res.Df    RSS Df Sum of Sq    F    Pr(>F)
1      504 1773.2
2      501 1560.0    3      213.22 22.825 7.272e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Fig. 6. ANOVA table for checking significance of demographic feature.

For ANOVA analysis on the model with demographic feature and one without it (See Fig.6), we found that demographic information is significant ( $p = 7.272e-14$ ) in relation to the linear model for students' ACT score, which shows that including demographic in our model influences students' ACT score.

To check how well our model represents the data, we plotted four diagnostic plots for our linear model for students' ACT score (See Fig.9). Our first plot, Residuals vs Fitted, shows that our residuals have a linear relationship. The residuals also appear to be mostly normally distributed as per the Q-Q plot with a few deviations present in the ends of the graph.

#### B. Statistical Model for Graduation Rate

When making a linear model to analyze the relationship between our features and the output (See Fig. 7), it was found that the significant variables were the ratio of white people ( $p = 0.000522$ ), the ratio of Asian people ( $p = 0.0231$ ), the ratio of associate's ( $p < 2e-16$ ), bachelor's ( $p < 2e-16$ ), and masters ( $p = 4.9e-5$ ) degrees, along with the mean income of the zip code ( $p = 0.00672$ ).

We expected that education levels of employees would be significant in relation to graduation rate, because we expected that parents (the employees) with higher levels of education would also want their children to pursue higher education and thus, have better educational outputs in high school. It is notable though that professional degrees and doctorates were

```

Call:
lm(formula = GraduationRate ~ ., data = Grad)

Residuals:
    Min       1Q   Median       3Q      Max
-24.9171  -1.2883   0.4397   1.7514  11.8066

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.211e+01  1.697e+00  42.479  < 2e-16 ***
whiteRatio    6.135e+00  1.761e+00   3.485  0.000522 ***
balckRatio   -1.413e+00  2.377e+00  -0.594  0.552502
asianRatio   -1.992e+01  8.750e+00  -2.276  0.023128 *
Assoc.Ratio   5.088e+01  4.147e+00  12.268  < 2e-16 ***
Bachelor.Ratio 2.845e+01  3.281e+00   8.670  < 2e-16 ***
Master.Ratio  2.692e+01  6.591e+00   4.085  4.9e-05 ***
Prof.Degree.Ratio -8.367e-01  1.610e+01  -0.052  0.958577
Doctor.Ratio   2.376e+01  1.706e+01   1.393  0.164000
Mean.income    2.901e-05  1.067e-05   2.718  0.006719 **
Private.Ratio  -7.446e-01  1.448e+00  -0.514  0.607202
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.436 on 740 degrees of freedom
(25 observations deleted due to missingness)
Multiple R-squared:  0.541, Adjusted R-squared:  0.5348
F-statistic: 87.21 on 10 and 740 DF, p-value: < 2.2e-16

```

Fig. 7. Linear Regression Model for Grad  
Diagnostic Plots are located at the end of article

not found to be significant, even though we still expected this trend to be relevant as education level of the parents increase.

Interestingly, race also play a significant role in this model, especially the white ratio. We suspect it is because better schools tend to be in areas with higher income and lower minority populations (suburbs), so schools with better educational outputs tend to have more white students. As it was mentioned earlier, there is a significant problem with segregation in Milwaukee. Because of the segregation, there are zip codes in low-income areas that have significant minority populations. This factor could be influencing the model due to the large amount of Milwaukee zip codes that are present in the data compared to other areas which generally only have one or two zip codes represent the area.

```

Analysis of Variance Table

Model 1: GraduationRate ~ GraduationRate + Assoc.Ratio + Bachelor.Ratio +
  Master.Ratio + Prof.Degree.Ratio + Doctor.Ratio + Mean.income +
  Private.Ratio
Model 2: GraduationRate ~ whiteRatio + balckRatio + asianRatio + Assoc.Ratio +
  Bachelor.Ratio + Master.Ratio + Prof.Degree.Ratio + Doctor.Ratio +
  Mean.income + Private.Ratio
  Res.Df    RGS Df Sum of Sq    F    Pr(>F)
1     743 9429.3
2     740 8738.7  3    690.63 19.494 3.581e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Fig. 8. ANOVA table for checking significance of demographic feature

When conducting ANOVA test on the demographic information (See Fig. 8), we created two models, one that featured white ratio, Asian ratio, and black ratio and one that did not. The ANOVA analysis was found to be significant ( $p=3.58e-12$ ), which shows that the demographic information is significant in relation to the linear model for graduation

rate.

The diagnostic plot for the data (See Fig. 10) appears relatively normal. The residuals vs. fitted plot shows an indication of a linear relationship in the data, which is what we want. The residuals also appear to be mostly normally distributed as per the Q-Q plot with a few deviations present in the ends of the graph.

#### IV. LIMITATIONS

The biggest limitation of our project is data limitation. The best way to have conducted this research would have been to get data for each individual person and conduct analysis on the individual level. Unfortunately, data such as this is just simply not available. Because of this, we had to use zip codes, but for each of these it is only averages within a specific area. Data such as this is not weighted to the area's populations, so areas that are less populated are less resilient to outliers and therefore could have their averages easily affected by extreme high or low values.

Furthermore, the ACT data we used was indexed by the zip code of the school. This limited the data we were able to use, as we could only use zip codes that had a corresponding public school (and ACT score) in our analysis. So, we were unable to use zip codes in which there was no school, even though there still may be students residing within that area. Secondly, this data is not able to account for the zip codes in which students actually live, only where they go to school. There is often shifting of where students go to school, or students who live in a different zip code than the school they attend. To deal with this issue, we tried to find zip codes of students attending each school, to better group the data, but we cannot group zip codes according to school districts because we were unable to find the data for this. Moreover, we also tried to find the data set indicating mean income, parents' highest diploma, and demographic for each school, but the public schools do not need to report this data. Researching more about this, the survey about income status of each student is restricted by Protection of Pupil Rights Amendment (PPRA), and so schools are unable to report this data (What is the..., n.d., [6]).

Additionally, it is almost impossible to find data for how much a parent invests in their child's education such as prep class for ACT and hiring private tutors. We could only use proxies such as mean income and private school enrollment, but that does not necessarily correlate to how much parents invest in a child's education. Furthermore, the mean income include childless family incomes and we are not able to remove that from our data.

#### V. CONCLUSIONS

Overall, the two statistical model that we fitted state that the mean income of each household significantly affects a child's academic output. Although, due to the existing limitation, we were not able to conclude the willingness to invest additional dollars towards children's education makes a significant difference on their academic success, we were

able to show that the income, the implied wealth of parents, makes difference on the educational output of their children.

With the discovered relationship between parental wealth and the educational output, we are able to conclude that education might be able to be utilized as a tool to transfer the wealth of the parents' generation to their kids, contributing toward the polarization. Sadly, the education causes the silver spoon effect, and this is the time for the government to think about providing fair opportunities on education in order to combat the long-lasting societal problem, polarization.

#### REFERENCES

- [1] Rugaber, C. S. (2017, January 12). Pay gap between college grads and everyone else at a record. Retrieved November 27, 2019, from <https://www.usatoday.com/story/money/2017/01/12/pay-gap-between-college-grads-and-everyone-else-record/96493348/>.
- [2] QCEW News Releases. (n.d.). Retrieved December 2, 2019, from <https://www.bls.gov/cew/>.
- [3] Data Slices by Industry. (2016, June 21). Retrieved December 2, 2019, from [https://data.bls.gov/cew/doc/access/csv\\_data\\_slices.htm#ANNUAL\\_LAYOUT](https://data.bls.gov/cew/doc/access/csv_data_slices.htm#ANNUAL_LAYOUT).
- [4] American FactFinder - Search. (2010, October 5). Retrieved December 2, 2019, from <https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t.???;::>
- [5] Wisconsin ACT scores 2017-18. (n.d.). Retrieved December 10, 2019, from <https://wisinfo.com/milwaukee/database/2018/wisconsin-act-scores-2017-18/#!/composite.desc.1/>.
- [6] What is the Protection of Pupil Rights Amendment (PPRA)? (n.d.). Retrieved December 10, 2019, from <https://studentprivacy.ed.gov/faq/what-protection-pupil-rights-amendment-ppra>.

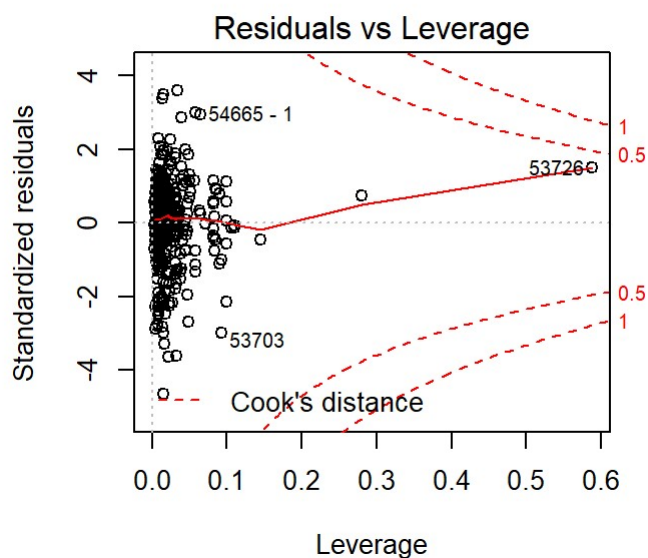
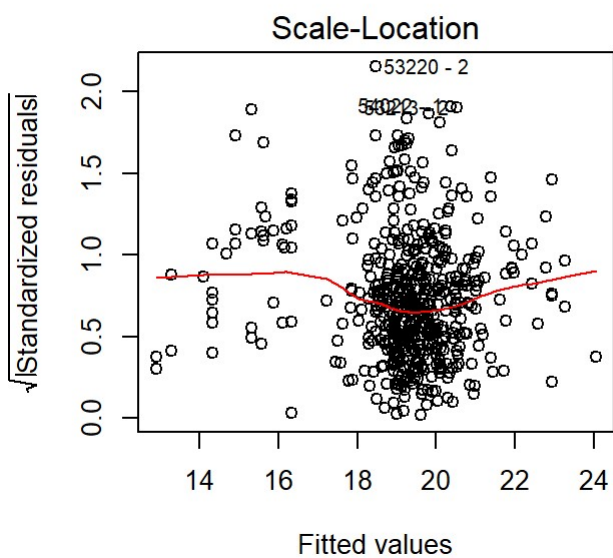
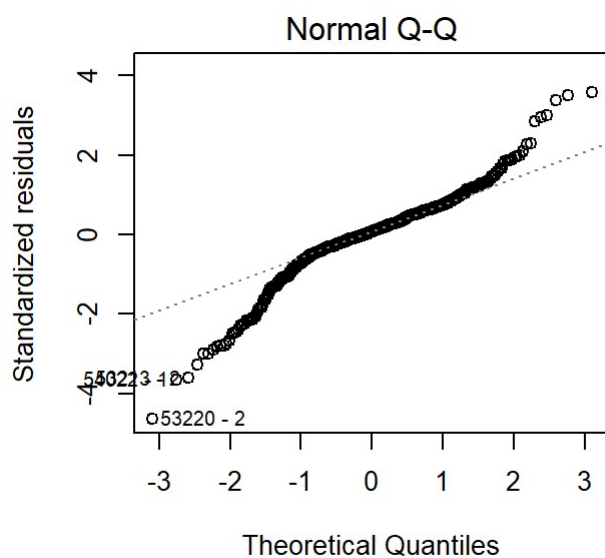
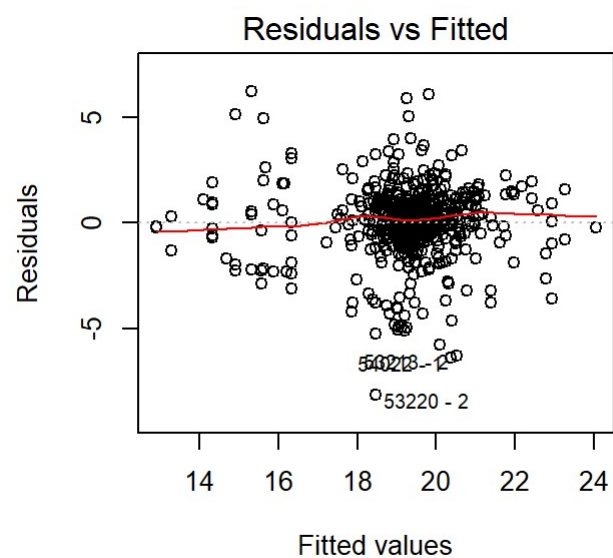


Fig. 9. Diagnostic Plot for ACT Model



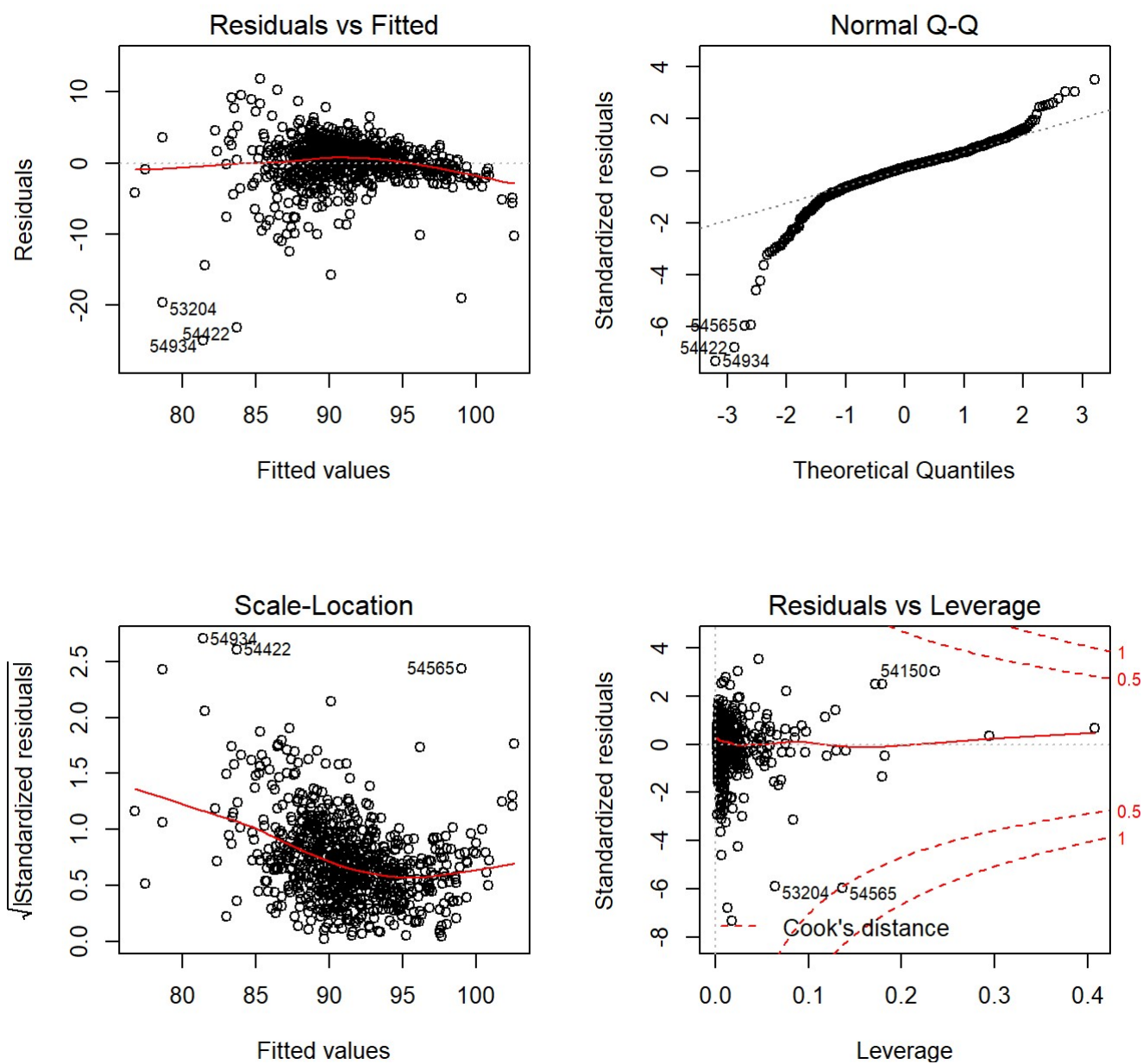


Fig. 10. Diagnostic Plot for Graduation Rate Model