Jon Murphy
EDU-210 SP 24
Final Paper Citations and Appendix

Reference List

Bloch, M., & Cox, A., et al. (2015). *Mapping Segregation*. New York Times. https://www.nytimes.com/interactive/2015/07/08/us/census-race-map.html

Chicago Public Schools (2020). *Chicago Public Schools - School Profile Information SY1819*. City of Chicago. https://data.cityofchicago.org/Education/Chicago-Public-Schools-School-Profile-Information-/kh4r-387c/about_data

Dennis, A. D. (2022). "The Intellectual Emancipation of the Negro": Madeline Morgan and the Mandatory Black History Curriculum in Chicago during World War II. *History of Education Quarterly*, 62(2), 136–160. https://doi.org/10.1017/heq.2022.2

Justice, B. (2023). Schooling as a White Good. *History of Education Quarterly*, 63(2), 154–178. https://doi.org/10.1017/heq.2023.7

Kochhar, R. & Cilluffo, A. (2018). *Key findings on the rise in income inequality within America's racial and ethnic groups*. Pew Research Center. https://pewrsr.ch/2m61p9D

Lukes, Dylan, and Christopher Cleveland. (2021*). The Lingering Legacy of Redlining on School Funding, Diversity, and Performance*. (EdWorkingPaper: 21-363). Annenberg Institute at Brown University: https://doi.org/10.26300/qeer-8c25

Nelson, R. K., Winling, L, et al. (2023). *Mapping Inequality: Redlining in New Deal America*. Digital Scholarship Lab. https://dsl.richmond.edu/panorama/redlining.

Pounder, C. C. H., Herbes-Sommers, Christine., Strain, T. Heather., & Smith, Llewellyn. (2003). *Race: The Power of an Illusion*. California Newsreel.

Sampson, R. (2024). *Great American City: Chicago and the Enduring Neighborhood Effect*. Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226834016

Winling, L. C., & Michney, T. M. (2021). The Roots of Redlining: Academic, Governmental, and Professional Networks in the Making of the New Deal Lending Regime. *The Journal of American History,* 108(1), 42–69. https://doi.org/10.1093/jahist/jaab066

Campus Resources Used:

Consultation with Prof. Wells (Department of Statistics), 4/16/24

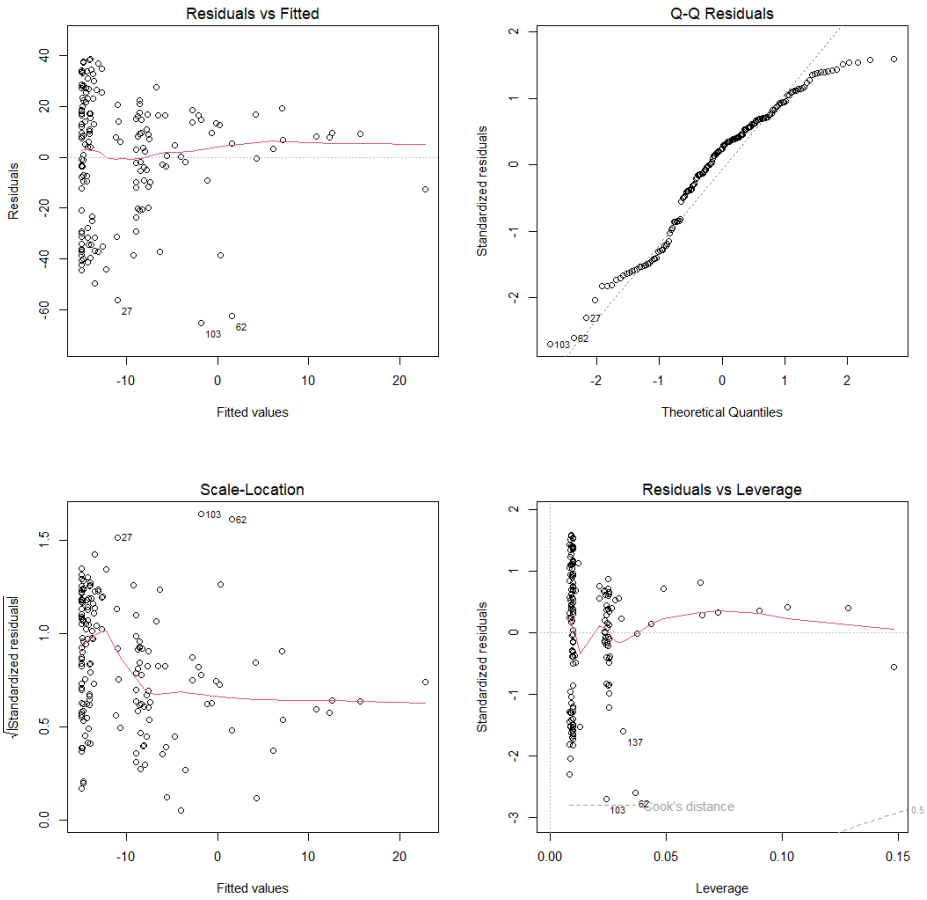
Jon Murphy
EDU-210 SP 24
Final Paper Citations and Appendix

Appendix A: Model Conditions and Validation

The linear model used in this analysis has the following conditions:

1. The error terms are independent.
2. The error terms are normally distributed with mean zero.
3. The error terms have equal variance across the range of the explanatory.

The conditions are mostly satisfied using the below plots:



1. Based on the collection method of the data, where every row is a different school, and all schools are measured, this condition appears to be satisfied.
2. While the Q-Q Residuals plot on the top right is not perfectly linear, it appears appropriate to continue with this analysis, while keeping in mind that the conditions may not be perfectly satisfied. Basically, this condition may not be fully met, although it is likely met enough to continue with the analysis.

3. The Residuals vs Fitted graph on the top left shows that the bulk of the data is on the low end of the explanatory. The graph, however, does not show evidence of a strong wedge shape, as the variance is pretty similar, except for at the very high end of the fitted values.

Overall, the conditions may not be perfectly met for a linear model, as the data is not perfect. This does not mean, however, that the model has no value, as long as it is undertaken with some reservations. As is commonly stated, all models are wrong, but some are useful, and this is definitely a useful model.

With these reservations in mind, below is the full model output for the fitted linear model.

```
Call:
lm(formula = College_Enrollment_Rate_Mean_Distance ~ Percent_Not_White +
    Attendance_Boundaries, data = cps_hs)

Residuals:
    Min      1Q  Median      3Q     Max
-65.414 -19.920   5.956  17.151  38.716

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                54.8755    20.7348   2.647  0.00893 **
Percent_Not_White          -0.6985     0.2161  -3.232  0.00149 **
Attendance_BoundariesTRUE   6.0627     4.3495   1.394  0.16527
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 24.51 on 162 degrees of freedom
  (13 observations deleted due to missingness)
Multiple R-squared:  0.07284,  Adjusted R-squared:  0.06139
F-statistic: 6.364 on 2 and 162 DF,  p-value: 0.002185
```

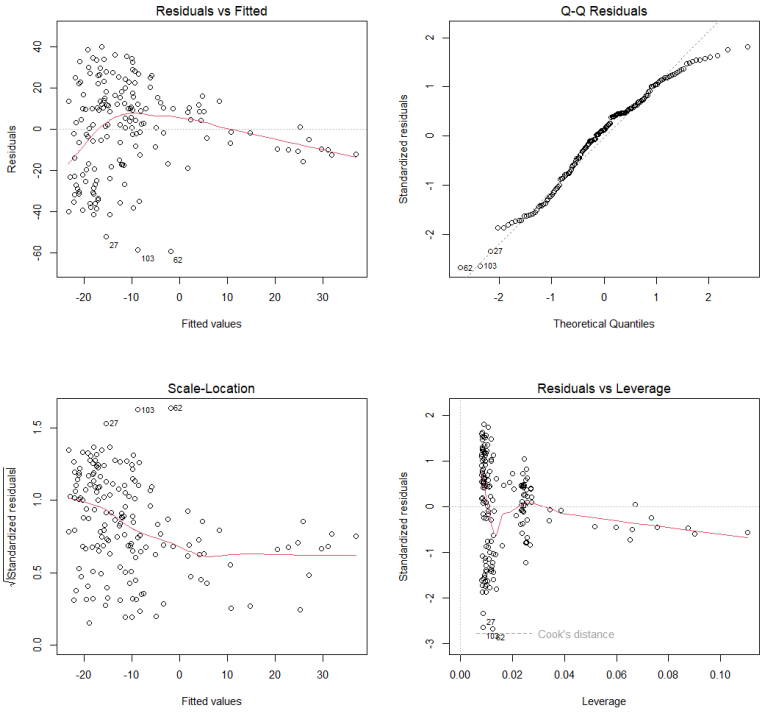
The overall model is significant, and the coefficient for Percent_Not_White is significant. The coefficient for Attendance Boundaries is not, so no inference was drawn on that point, it was simply used as a control variable. Additionally, inference was undertaken with the reservations of the model condition, and later the model is debunked.

For the Pearson’s Correlation test, the output is below.

```
        Pearson's product-moment correlation

data:  cps_hs$Percent_Not_White and cps_hs$Percent_Low_Income
t = 19.373, df = 174, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7732099 0.8683390
sample estimates:
      cor
0.8265901
```

For the linear model using income instead of race, the same conditions were tested using the below plots:



1. As before, based on the collection method of the data, where every row is a different school, and all schools are measured, this condition appears to be satisfied.
2. Based on the Q-Q residuals plot, this model fits the data even better, and appears to be mostly linear with some departure at the extremes of the quantiles. Basically, this condition is likely met, with some reservations.
3. Based on the residuals vs fitted plot, there may be more a wedge shape, although with the bulk of the data on the low end, it is difficult to say. Basically, this condition is likely met.

Overall, the conditions are mostly met, with a few reservations. With that in mind, the output of this model is below:

```
Call:
lm(formula = College_Enrollment_Rate_Mean_Distance ~ Attendance_Boundaries +
    Percent_Low_Income, data = cps_hs)

Residuals:
    Min      1Q  Median      3Q     Max
-59.230 -16.810   2.853  15.446  40.105

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                61.7232    10.9506   5.637 7.53e-08 ***
Attendance_BoundariesTRUE   8.0759     3.9583   2.040   0.0429 *
Percent_Low_Income         -0.8544     0.1252  -6.827 1.65e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 22.28 on 162 degrees of freedom
  (13 observations deleted due to missingness)
Multiple R-squared:  0.2336,   Adjusted R-squared:  0.2241
F-statistic: 24.68 on 2 and 162 DF,  p-value: 4.394e-10
```