**Research Hypotheses:**

1. The higher the number of middle school suspensions a student receives the more likely he/she is to not graduate high school, after controlling for age and gender differences.

**Data:**

Data are from the NLSY97. The survey began in 1997 with a nationally representative sample of 8,984 men and women born during the years 1980 through 1984 and living in the United States. Participants were aged from 12 to 16 years old as of December 31st, 1996. From the Bureau of Labor Statistics website, “interviews were conducted annually from 1997 to 2001 and biennially since then. The ongoing cohort has been surveyed 18 times as of date … with oversamples of Hispanics and non-Hispanic blacks.” (NLS) For the purpose of the analyses, a subject is said to have graduated high school if when surveyed, the highest grade completed is either the twelfth grade or above. In the sample used for analyses, 2,003 (22.34%) participants did not graduate high school while 6,963 (77.66%) students did. Unlike other large-scale national-level delinquency studies, the NLSY97 selects random samples at the household level instead of at the school level.

**Measures:**

The analysis utilized four socio-demographic variables which were included in all further analyses. *Female* is a dummy variable where females were coded one (48.81%) and males were coded zero (51.19%). *Age* is an interval ratio variable that measures the respondent’s age at the 1997 interview with a mean of 13.99 and a standard deviation of 1.40 years old. *Black* is a dummy variable where black participants were labelled one (25.99%) and non-black participants were labelled zero (88.06%). *Immigrant* is a dichotomous variable measuring whether respondents were immigrants to the United States where immigrants were coded one (11.94%) and non-immigrants were coded zero (88.06%). *High-school graduate* is a dichotomous variable measuring whether students had graduated high school where graduates were coded one (77.66%) and non-graduates were coded zero (22.34%).

*Trauma*

There are several measures of trauma, all asked in 2002, for the last five years of experience when respondents were between the ages of 18 to 22; all trauma variables were dichotomous and coded one for a “Yes” response and coded zero for a “No” response. For this analysis, *Trauma* consisted of four observed variables: homeless, victim of violent crime, household member hospitalized, and household member in jail. *Homeless,* asked in 2002, measures whether the respondent had ever been homeless or lived in a homeless shelter for at least two days in the past five years. *Victim of violent crime*, asked in 2002, measures whether the respondent has ever been a victim of a violent crime in the past 5 years. *Household member hospitalized*, asked in 2002, is a dichotomous variable measuring whether the respondent had a family member admitted to the hospital within the past five years. Finally, household member in jail, asked in 2002, is a dichotomous variable measuring whether the respondent had a family member incarcerated within the past five years.

*Delinquency*

Delinquency is measured using a respondent’s engagement in the following three behaviors in the year 1997 or before: ever in a gang, ever smoked, and ever arrested for an illegal/delinquent offense. All delinquency variables were measured as dichotomous variables and coded one for “Yes” responses and zero for “No” responses. Respondents were between the ages of 12 and 16 when these questions were asked.

**Results**

The tables below are descriptive statistics of the variables the analysis used separated by gender and racial groups. The chi-square test of association was used for the categorical variables and the one-sample t-test was used for quantitative variables. The singular hypothesis was examined using the Structural Equation Model (SEM) and the Random Forest (RF) ensemble.

Table 1: High School Graduation Rates by Race

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Non-graduate** | **Graduate** |  |
| **Black** | 667 (28.66%) | 1660 (71.34%) | 167.9 / <0.001 |
| **Hispanic** | 532 (28.01%) | 1367 (71.99%) |
| **Mixed** | 14 (16.87%) | 69 (83.13%) |
| **White** | 790 (16.96%) | 3867 (83.04%) |

Table 2: High-school Graduation Rates by Immigration Status

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Non-graduate** | **Graduate** |  |
| **Non-immigrant** | 1737 (22.00%) | 6157 (78.00%) | 4.13 / 0.042 |
| **Immigrant** | 266 (24.81%) | 806 (75.19%) |

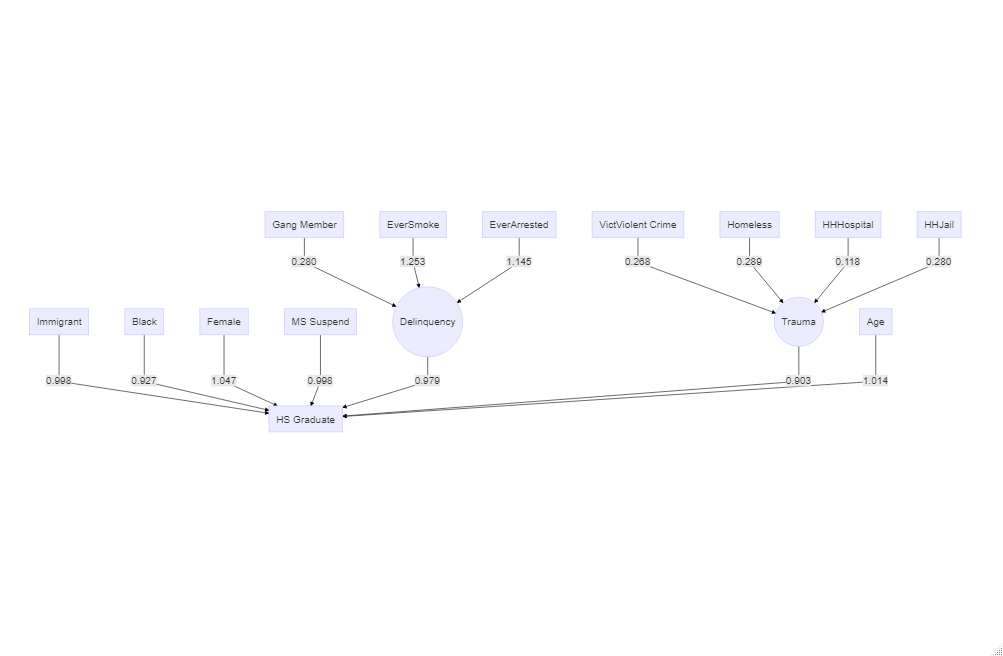
Table 3: High-school Graduation Rates by Gender

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Non-graduate** | **Graduate** |  |
| **Male** | 1187 (25.87%) | 3402 (74.13%) | 66.96 / <0.001 |
| **Female** | 816 (18.64%) | 3561 (81.36%) |

Table 4: Age by Gender

|  |  |  |  |
| --- | --- | --- | --- |
|  | Age | | ***t/pvalue*** |
|  | **Mean** | **SD** | -0.807/0.420 |
| **Male** | 13.98 | 1.39 |
| **Female** | 14.00 | 1.40 |

*Structural Equation Model*

The Structural Equation Model (SEM) is often used to assess unobservable “latent” constructs via observed variables in the data. The are two major latent variables present in the analysis: *Delinquency* and *Trauma*, both of which were discussed earlier. A single regression model was incorporated in the SEM, regressing high school graduation against the latent and observed variables discussed earlier. The SEM results indicate that all the variables mentioned thus far have some sort of a statistically significant effect, whether positive or negative, on high school graduation. The following figure represents the SEM built from this analysis with the numbers representing the odds ratio relating to either the latent variable or the response variable.All the variables present in the analysis, latent or observed, were statistically significant at the 5% level of significance and above. Having experienced a traumatic event led to a decrease in the odds of graduating high school. Partaking in a delinquent activity was associated with a lower odds of graduating high school.

*The Role of Demographics on High School Graduation*

The odds of graduating high school are 1.05 (95% CI: 1.03, 1.08) times lower for immigrants than for his/her non-immigrant counterparts, 1.04 (95% CI: times higher for females compared to males, 1.014 (95% CI: times higher for every one-year increase in age, 1.014 (95% CI: times lower for every year increase in age beyond 14, and 1.09 (95% CI: times lower for blacks compared to other races.

*The Role of Trauma on High School Graduation*

For students who experienced at least one trauma during his/her childhood, the odds of graduating high school were 5% lower (95% CI: ) than for students who did not experience any childhood trauma.

*The Role of Delinquency on High School Graduation*

For students who participated in at least one delinquent activity during his/her childhood, the odds of graduating high school were 11% lower (95% CI: ) than for students who did not participate in delinquent activities.

**Limitations**

There are a few limitations to consider when looking at this analysis. First, the analysis did not use weights for the SEM, which may limit the generalization of the findings. Secondly, the research hypothesis is rather narrow in its scope and other hypotheses should be considered before taking any action regarding the results. Finally, additional predictor variables should be used to increase the model’s performance and predictive power, like adverse childhood experiences (ACEs).

**Conclusion**

**Resources:**

Bureau of Labor Statistics

Tedor, Miyuki et. al. *Explaining the school-to-prison pipeline via race and gender: The impact of trauma, school suspension, school dropout, and delinquency.* 2020