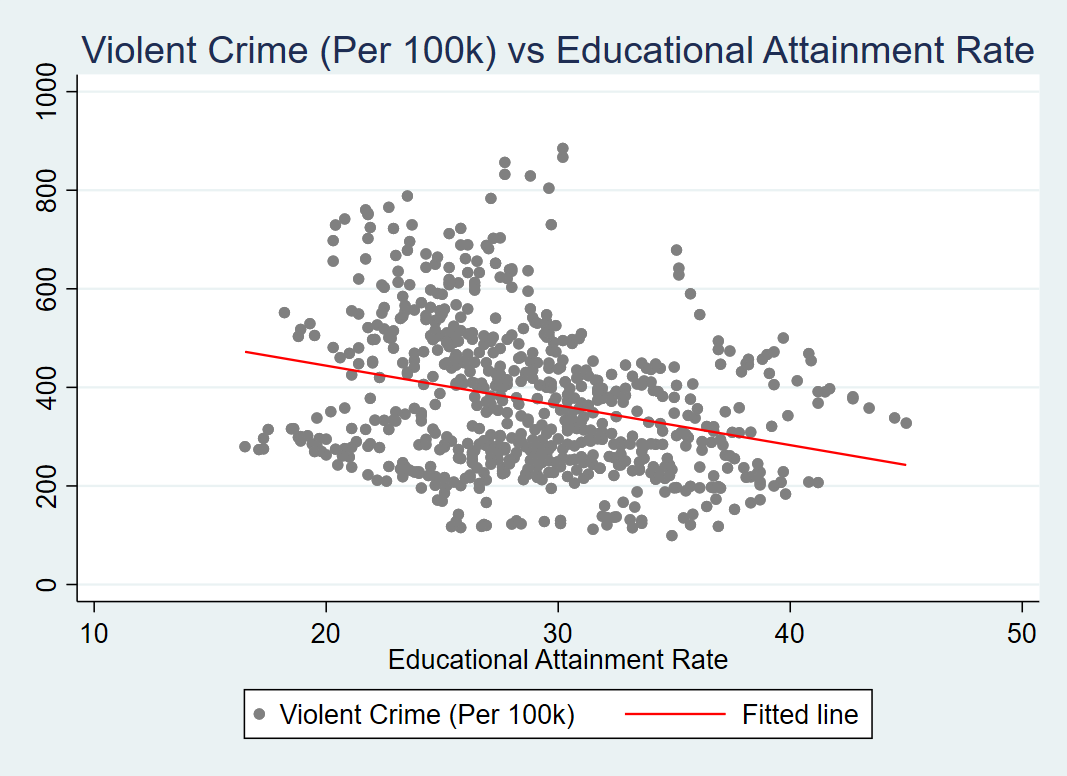
Eugene Baldonado

Professor Lang

ECON108 Final Regression Project: Educational Attainment Rate on Violent Crime Rate

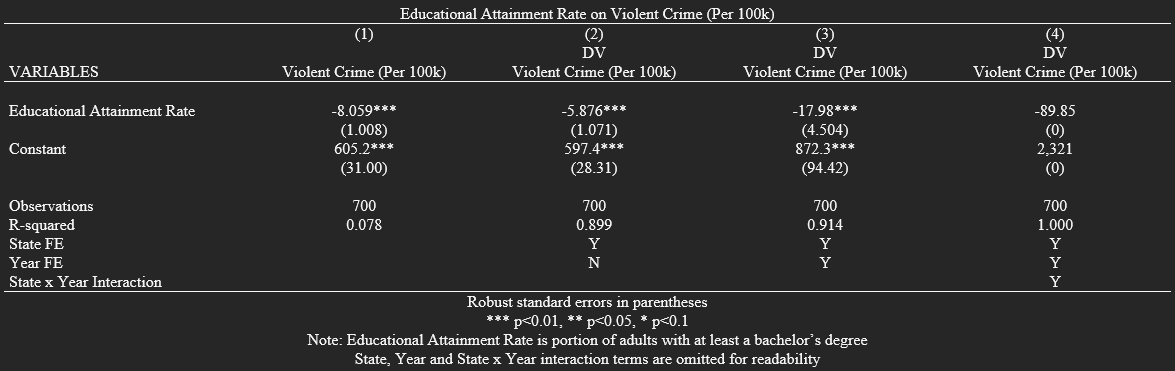
The question I posed for this regression project was: Is a state’s educational attainment rate (adults obtaining bachelor’s degree) related to its violent crime rate? Assuming that more educated citizens tend to follow the law; I assumed a higher educational attainment rate would provide a state with more educated law-abiding citizens. Under these assumptions I made the hypothesis that States with higher educational attainment rates would expect lower violent crime rates and similarly States with a lower educational attainment rate would expect a higher violent crime rate. This would mean that a State's violent crime rate would have an inverse relationship to the educational attainment rate.

The data I collected was from the Federal Reserve Economic Data (FRED) datasets and from the FBIs open government data. I collected all of the reported violent crimes from each state from 2009 - 2019, their population, and their violent crime per 100k persons from FBI.gov. Additionally, I collected the educational attainment rates (bachelor's degree or higher for each state) for each year from 2009 - 2019 from the FRED. I then cleaned both datasets and merged them based on the State name for my final dataset.



Scatter plot of all State’s violent crime (per 100k) against their educational attainment rate from 2009 - 2019

I planned on controlling for certain fixed effects, so I knew from the start that I was going to do a fixed effects regression model. Each state could have naturally higher educational attainment rates or violent crime rates, so to control for these state based differences (assuming that each state’s violent crime rate doesn’t drastically change from year to year) I could control for state fixed effects by creating dummy variables for each state; additionally, I also wanted to control for varying time fixed effects since certain years may have varying violent crime rates or educational attainment rates across the whole U.S. After controlling for both state and time fixed effects I expected a higher correlation between educational attainment rate and violent crime rate for a given state and time period. Below are the results from regressing violent crime rates on educational attainment rate without fixed effects, state fixed effects only, state and time fixed effects, and state and time fixed effects with interaction terms.



Although statistically significant, the first regression revealed a very low R^2 of 0.078, showing that the variation of the violent crime rate was very loosely explained by the educational attainment rate. As we included fixed effects on States and time (years) not only did our R-squared increase (from 0.078 to 0.914), our 𝛃1 coefficient also showed that a State’s violent crime rate had an increasingly negative (inverse) relationship with it’s educational attainment rate (from -8.059 to -17.98). This interpretation of 𝛃1 is in whole percentages (17% instead of .17) and shows that on average a single percent increase in a state’s educational attainment rate decreased the violent crime (per 100k) for that particular state by 17.98, controlling for fixed effects. It can also be noted that 𝛃0 represents the state of Wyoming in the year 2019 (the last State/Year), this is to prevent multicollinearity issues when regressing on each of the state and year dummy variables. I was interested in adding an interaction term for state and year fixed effects and I realized this produced the regression equation that would accurately predict every point. This makes sense since we are accounting for the fixed effect of the state, year as well as the interaction term between them in the regression model, so given our dataset the combination of the dummy variables/interaction terms would yield the exact point for a given state and year. What I found interesting was that this made 𝛃1 even more negatively correlated, which states that a single percentage increase would decrease a given State and year’s violent crime by 89.85 (per 100k population) when controlling for fixed effects, on average. Despite these findings clearly showing an inverse relationship between a state’s educational attainment rate and violent crime rate, there can still be omitted variable bias that can make this model less accurate in predicting violent crime rate.

After questioning the relationship between a state’s educational attainment rate and its violent crime rate, I found statistical significance in their inverse relationship. However, the issues of omitted variable bias’ can make it difficult to state if it is exactly educational attainment causing certain effects on a State’s violent crime rate. While looking at other possible factors (that have relationships with educational attainment rate of a state), I found that possible omitted variables that could be included in the regression model could be the State’s poverty level and/or Unemployment rate. Each of these variables could have been included in the regression to make a model that accurately predicts a State’s violent crime rate. However, the purpose of these regressions wasn’t to make a predicting regression model (as with our last regression we fit the model to the specific data perfectly). After controlling for state and year fixed effects as well as their interaction term, 𝛃1 reveals that under these controls State’s with lower educational attainment rates had higher violent crime rates, with similar effects with higher educational attainment rates yielding lower violent crime rates on average.