UROP Final Report

The consumption of medical care has been constantly increasing as income per capita rises while technology and the means of production in health industries improves. Medical care expenditures as a percent of GDP have risen from 5% in 1960 to almost 18% in 2018. One purpose of my research was to assess ways to reduce the percentage of GDP spent on medical care expenditures. By reducing the largest single sector of GDP, the savings can be spent on improving the lives of lower socioeconomic status individuals both domestically and abroad. I ran two classic linear regression models to determine if early education attendance would affect total Medicaid expenditure and the total number of inmates per state. A Priori, it was expected that as states increased the percentage of children attending a pre-K program, the total cost of Medicaid expenditures per capita and total number of inmates incarcerated would decrease.

My results did not support my hypotheses; no relationship was found between pre-K attendance and either outcome. The first regression that looked at how pre-K attendance affected total Medicaid spending looked as follows:

$$Y_i = -258.34 + 99.73x_1 - 1.84x_2 - 14.63x_3$$

(41.15)*** (15.58)*** (9.12) (12.10) Adjusted R Square = .9919

Y_i = Total Medicaid Spending (Billions of dollars)

x_1 =Health consumption expenditures per capita (millions of dollars)

x_2 =Average state Pre-K attendance (percentage)

x_3 =CHIP spending (millions of dollars)

While at first it may seem promising there is one variable statistically significant, total Medicaid spending and Health consumption expenditures per capita overlap in measurement. The overlapping in measurement is the reason the adjusted R-squared value is so high. This R-squared variable measures the amount of variance the regression function captures. My second

regression function that looked at how pre-K attendance affected the total number of inmates incarcerated looked as follows:

Y_i = 11.10** - .097x_1 - .133x_2 - 18.66x_3 R-squared adjusted = .981
(4.16) (.058) (.074) (45.50)
Y_i = Total number of inmates incarcerated
x_1 = Average state Pre-K attendance (percentage)
x_2 = Average CHIP spending per state ($millions)
x_3 = State population that is African American (percentage)

My second regression function did not exhibit any statistically significant variables except the intercept, which means the variables cannot be used to accurately predict the number of inmates incarcerated in any of the states. In both regression functions, I found it difficult to capture available data on independent variables that impact the dependent variable. There are many variables that make up the total Medicaid spending per capita that cannot be accounted for accurately with just a few variables included in a regression function. This same thought can be applied toward the lack of statistically significant variables in my second regression as well. My research sponsor and I talked multiple times about why there were difficulties detecting a relationship between the independent and dependent variables in this research project. When individuals enter pre-K education programs they are one to three years old. From such a young age all the way to thirty-one years old, individuals participating in the most successful pre-K programs were measured on indicators such as income, educational attainment, crimes committed, and unemployment history. And, while those researchers found promising results associated with pre-K attendance, at the population level, there are a multitude of factors that can impact income, educational attainment, crimes committed, and unemployment history over a

28-31 year span that makes it difficult to narrow down to a few variables included in my regression function.

During my time spent working on the UROP I found it challenging to find direction completing a project that has very little academic interest and therefor little corresponding literature. The correlation between early education and Medicaid spending is not a hot-button issue that is garnering a lot of academic support. I sought to find a relationship that has never been proved before and came up short. Regardless of the result, I learned quite a bit about a topic that is not covered in a conventional classroom setting. This roundabout way of failing in my ultimate goal of determining a causal relationship helped me learn how difficult it is to carry out research on my own. I reflected upon this with my research sponsor multiple times. We came to the conclusion that this failure to produce the desired result will help me learn from my mistakes so I can go on to create a better encompassing method to finding relationships amongst variables. It's too bad every student isn't required to complete a UROP, this was a great benefit to my academic growth. This project taught me patience, research skills, honed my econometric modelling, and increased my ability to read and summarize academic literature.