Senior Project

Department of Economics

The Effects of Legalizing Cannabis on Youth Development

Sekou Larack Kaba

Spring 2023

Advisor: *Dr. Ali Enami*

Abstract

This study investigates the potential effects of legalizing cannabis on youth development by examining the associations between state-level policy changes and key indicators of academic performance, substance use, and cognitive functioning. Using data from the National Center for Educational Statistics (NCES), Substance Abuse and Mental Health Services Administration (SAMHSA), and Integrated Public Use Microdata Series (American Community Survey), we explore the impact of legalizing cannabis on youth marijuana use, SAT scores, cognitive difficulty, and dropout rates. Our findings suggest that legalizing cannabis, either medically or recreationally, has a negative impact on SAT math and writing/verbal scores and may lead to increased dropout rates among youth. However, there is no significant effect on marijuana use, and the effect on cognitive difficulty is limited. These findings emphasize the need for policymakers to consider the potential negative impact of legalizing cannabis on youth development and implement measures to mitigate any potential harm.

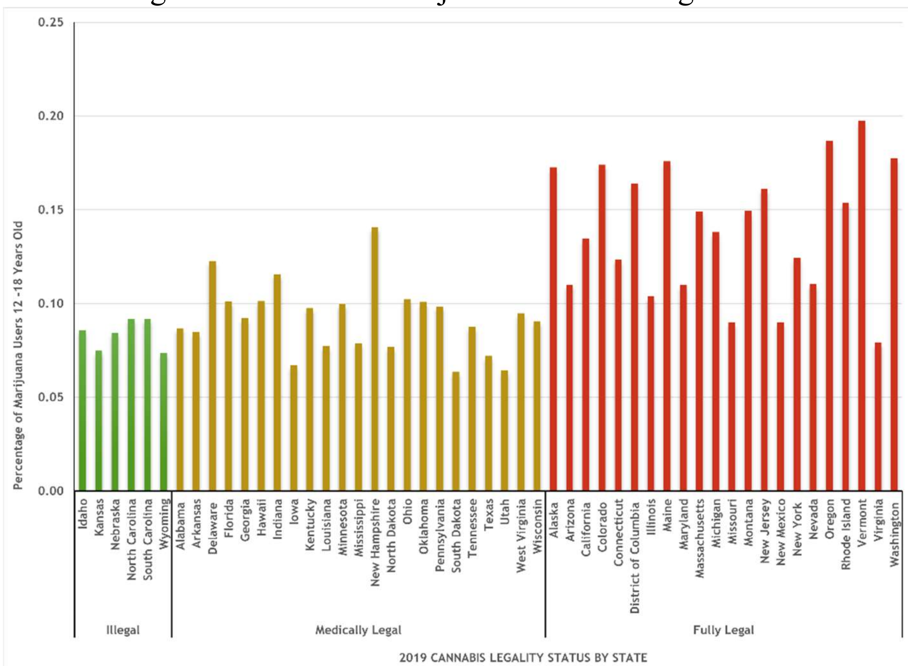
Table of Contents

I. Introduction

The use of cannabis among children and teenagers remains a serious concern, despite being illegal for those under the age of 21. In 2022, 30.7% of 12th graders report using marijuana in the past year and 6.3% reported using it daily. Additionally, 20.6% report vaping marijuana in the past year and 2.1% report doing so daily (NIDA, 2022). The increasing legalization of recreational cannabis in some states raises concerns about its potential impact on youth access, use, and development. Among youth receiving substance use disorder treatment, almost 50% of admissions among those aged 12 to 17 are attributed to marijuana, indicating the risk of addiction (NIDA, 2020). Therefore, it is crucial to address the potential impact of cannabis on youth and to prioritize education, prevention, and early intervention efforts. While advocates of recreational cannabis legalization argue for reducing criminalization, it is important to consider the accessibility of cannabis to young people through older peers or relatives, similar to cigarettes which are legal for adults but prohibited for minors.

The line graph in Figure 1 indicates that among states where marijuana is fully legal in 2019, some have a higher percentage of youth using marijuana, exceeding 15%, compared to states where marijuana is either illegal or only medically legal. On account of that, no state in the illegal or medical groups exceeds the 15% threshold. In 2019, the average rate of marijuana use among youth in prohibited states is 8%, in medical states, it is 9%, and in fully legal states it is 14%. It is important to consider these differences in marijuana use among states as they may have an impact on the outcome measures, such as academic achievement and cognitive functioning. The mean past month marijuana use is 6.70%, with a standard deviation of 4.13% from 2003-2021. Overall, the data found contains valuable information that allows for a thorough and comprehensive analysis of the effects of legalizing cannabis on youth development.

Figure 1. 2019 Percentage of Past Month Marijuana Users Among 12-18 Years Old by State



Source: Substance Abuse and Mental Health Services Administration (SAMHSA), *National Survey On Drug Use and Health.* Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

Currently, 39 states and the District of Columbia have legalized cannabis for medical use, and 21 states and the District of Columbia have legalized its recreational or adult use. Private research shows that cannabis has CBD properties that can alleviate schizophrenia symptoms in patients, which is more effective than conventional antipsychotic treatments (Davies et al., 2019). Additionally, in a placebo-controlled trial, the CBD properties of cannabis were found to reduce seizures in childhood epilepsy (Perucca, 2017). These medical discoveries have influenced the decision of many states to legalize cannabis medically or recreationally. However, the question now is what impact the legalization of cannabis will have on the development, access, and use of children growing up in an environment where they are continuously exposed to cannabis. This study provides a unique opportunity to assess the effects of cannabis legalization on youth development by analyzing recent data on youth cannabis users in light of previous research. The wide range of state legalization of cannabis allows for a comprehensive examination of the impact of legalization on youth.

II. Literature Review

The effects of cannabis on child development continue to be a significant concern, as early exposure to the drug can have lasting impacts on a child's growth. Studies have linked prenatal exposure to cannabis with various negative outcomes for infants and children. Research indicates that prenatal exposure to cannabis is associated with lower birth weight, cognitive development delays, and behavioral issues in infants and children. Baranger et al. (2022) found that prenatal cannabis exposure (PCE) after maternal awareness of pregnancy was linked to increased psychopathology during middle childhood. Fried (2004) discovered that infants exposed to cannabis during pregnancy had lower scores on measures of cognitive development and were more likely to experience delays in language development. The study also found that these effects were more prominent in children exposed to high levels of cannabis and that the effects were not as severe in children exposed to lower levels. Huizink et al. (2015) found that prenatal cannabis use was associated with an increased risk for behavioral problems, while Tarter et al. (2003) found that prenatal cannabis use was associated with an increased risk for mental health issues, such as depression and anxiety.

Furthermore, cannabis use during adolescence, when the brain is still developing, is currently linked to an increased risk of mental health problems such as depression, anxiety, and psychosis. Youth who use cannabis are also more likely to experience problems with memory, attention, and decision-making, which can negatively impact their school performance and overall life outcomes. Solowij et al. (2011) compare memory performance using the Rey Auditory Verbal Learning Test among cannabis and alcohol users of 181 adolescents.

Participants are matched for age, education, prior cognitive ability, and alcohol consumption. The study finds that cannabis users perform worse on memory tests compared to alcohol users and non-users, and the outcome is statistically significant. Harpin et al. (2017) find that after the state retail implementation of recreational marijuana in Colorado, adolescent marijuana uses behaviors and attitudes towards use remained unchanged over a one-year period. However, the perceived ease of accessing marijuana increased from 46% to 52%. Proximity to recreational marijuana stores is not found to be a significant factor in perceived ease of access. This is a secondary analysis of data from the Healthy Kids Colorado Survey collected from 40 schools, before and after recreational marijuana sales were allowed. The data is collected from a sample of 12,240 students in 2013 and 11,931 students in 2014.

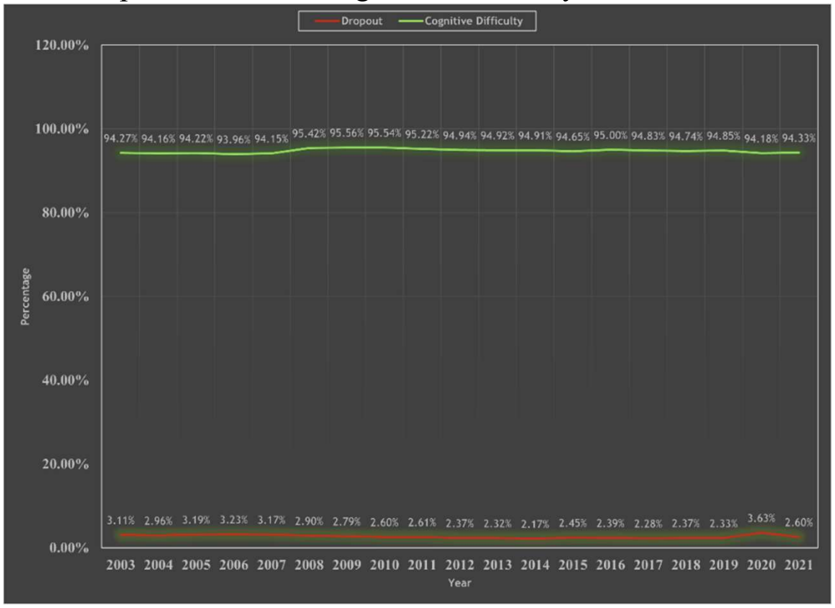
Moreover, some studies suggest that cannabis may have beneficial effects on youth development. Devinsky et al. (2017) find that children with a rare form of epilepsy called Dravet syndrome have fewer seizures when treated with a CBD-rich cannabis oil. On average, the change in seizure frequency amounts to a 39% decrease for the cannabidiol group patients, compared with a roughly 13% decrease among the placebo group. It is important to note that most studies on the effects of cannabis on child development and behavior focus on prenatal and postnatal heavy or chronic use of the drug, and those conducted on adolescent users focus on youth access and use. These studies were conducted during times when most states had laws banning the possession or use of cannabis in all forms. For example, the study conducted on the effects of legalizing cannabis on youth access in Colorado (Harpin et al., 2017), only had 8 states that had legalized cannabis recreationally when the study was conducted in 2017, compared to 21 states as of 2023. Solowij et al. (2011) study on the effect of cannabis on adolescents focuses on the age group of 18-24 years old, which gives room to analyze the effects of legalizing cannabis on younger age groups of 12-18 years old. It is important to note that these studies have limitations, such as small sample sizes, observational nature, and a lack of control for other factors that may influence child development and the use of cannabis, like the legalization status in different states. In terms, there are unanswered questions regarding the impact of legalizing cannabis on youth access, use, and development. The contribution to the literature on the effects of cannabis legalization on youth focuses on providing updated information on youth cannabis use and its impact on academic success and behavioral development, as well as informing clinical guidelines and public health policies to minimize potential harm.

III. Data Analysis

The data for this study contains 969 observations and valuable information regarding the relationship between legalizing cannabis medically or recreationally and its effects on youth development. Starting with the outcome variables, the average math and writing/verbal SAT scores by state from 2003-2021 data from the National Center for Education Statistics (NCES), serve as a measure of academic achievement. By examining changes in SAT scores in states that have legalized cannabis compared to those that have not, we can determine whether legalization has a positive or negative impact on students' academic performance. Similarly, another outcome variable in the dataset, high school dropout rates by state from 2003-2021, is created based on school attendance and educational attainment data from the Integrated Public Use Microdata Series (American Community Survey). The variable is created based on the condition that the teen is either not in school or has not completed high school to indicate the average percentage of high school dropouts by state, with a value of 0 indicating that the teen is in school, and a value of 1 indicating that the teen is not in school and did not complete high school. This variable provides information about the potential impact of legalization on students' engagement and motivation in education. Youth cognitive difficulty by state from 2003-2021 from the Integrated Public Use Microdata Series (American Community Survey) and youth marijuana use within the past month by state from 2003-2019 from the Substance Abuse and Mental Health Services Administration (SAMHSA) serve as measures of the effects of legalization on the cognitive functioning and substance use of youth. These variables are particularly important to include in the analysis as they provide direct evidence of the potential harm or benefit of legalizing cannabis on youth development.

The line graph in Figure 2 presents a clear comparison between the dropout rates and the prevalence of cognitive difficulty for 12-18 years old from 2003 to 2021. Based on the available data from 2003 to 2021, the average dropout rate among 12-18-year-olds is 2.71%, while the average prevalence of cognitive difficulty is 94.73%, which implies that 94.73% of 12-18-year-olds have no cognitive difficulty from 2003 to 2021. Figure 2 shows a correlation between the two variables, and the dropout rates have been stable and consolidates between 2.17% to 3.63% from 2003 to 2021 with the highest dropout rate in 2020, and the lowest in 2014. Moreover, the prevalence of cognitive difficulty remains relatively stable across the years, with a high average of 95.56% and a low average of 93.96%. This suggests that the majority of youth surveyed during this time period did not experience cognitive difficulties.

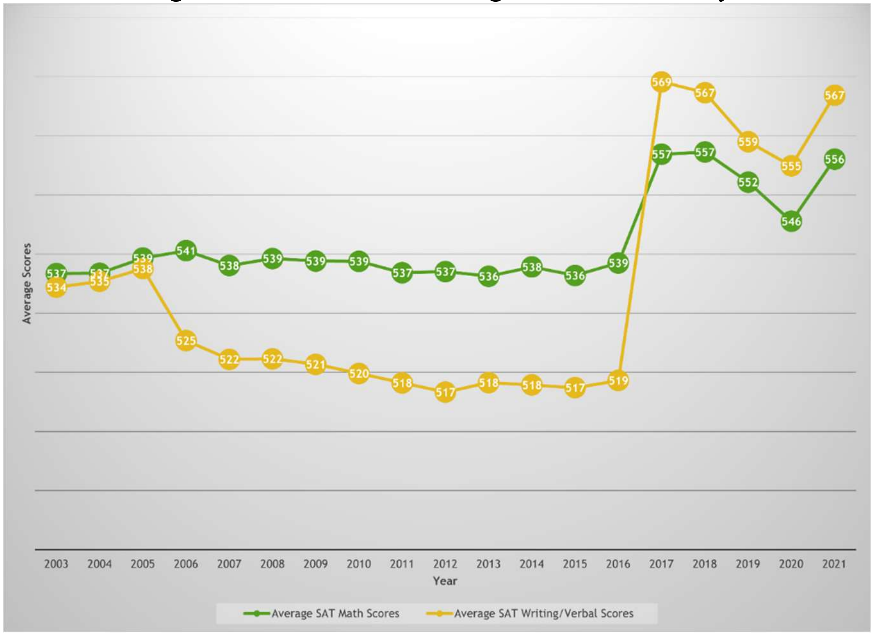
Figure 2. 2003-2021 Dropout Rates and Cognitive Difficulty for 12-18 Years Old by State



Source: Integrated Public Use Microdata Series (IPUMS: American Community Survey), *School Attendance, Cognitive Difficulty, and Educational Attainment (Various years), data available as of April 2023*

Additionally, the data shows that average SAT math scores remain relatively stable between 537 and 556 from 2003 to 2021 (Figure 3), with the highest score observed in 2017 and 2018. Average SAT writing/verbal scores also remain stable between 534 and 567 during the same period, with the highest score (Figure 3) observed in 2017. The mean SAT math score is 542.05, with a standard deviation of 43.60 from 2003-2021. The mean SAT writing/verbal score is 533.82, with a standard deviation of 45.13 from 2003-2021. Moreover, the stability of the average scores over time suggests that external factors, such as changes in educational policy or economic conditions, may not have a significant impact on student performance.

Figure 3. 2003-2021 Average SAT Math and Writing/Verbal Scores by State



Source: National Center of Educational Statistics (NCES), *Average SAT Math and Writing/Verbal Scores (Various years), data available as of April 2023*

The study employs control variables from the Integrated Public Use Microdata Series (American Community Survey) to help control for other factors that may influence the results. Family size, age, gender (male), race (white), employment status, usual hours worked per week, and family income from 2003-2021 are all relevant variables that may influence the outcome measures. By controlling these variables, the analysis's results can be ensured to be more robust and accurate. It is essential to note that the data is being cleaned and modified to address missing values and merging. This ensures the accuracy and reliability of the results obtained from the analysis. The outcome variables provide insight into the potential impact of legalization on education, substance use, and cognitive functioning, while the control variables in Figure 4 help to account for other factors that may influence the results.

Figure 4. Control Variables Statistics Summary

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
FAMSIZE	Family Size	969	4.13	0.27	2.76	5.16
AGE	Age	969	15.06	0.12	14.66	15.87
WHITE	Race	969	73.29	16.14	10.00	96.66
MALE	Gender	969	51.36	1.31	41.87	57.00
EMPLOYED	Employment Status	969	13.06	3.51	3.52	22.04
UHRSWORK	Work Hours Per Week	969	4.50	1.15	1.86	8.03
FTOTINC2	Family Income	969	89,819.94	22,981.35	49,224.01	192,350.72

Source: Integrated Public Use Microdata Series (IPUMS: American Community Survey), *Family Size, Age, Race, Sex, Employment Status, Work Hours Per Week, Family Income (Various years), data available as of April 2023*

In Figure 5, a balance of regressors analysis is being conducted to determine if there are significant differences in family size, age, race, gender, employment status, work hours per week, and family income between the control group and treatment group for treatmentF (states with recreational cannabis legalization) and treatmentM (states with medical cannabis legalization). The results of the analysis show that there is a statistically significant difference in the mean for family size, age, race (white), and gender (male) variables between the two groups in treatmentF (states with recreational cannabis legalization). However, only the difference in race (white) is economically significant due to the size of the differences. For the employment status and work hours per week variables, the results indicate a statistically significant difference between the two groups, but it is not economically significant due to the size of the differences. Finally, for the family income variable, the results show that there is a statistically significant difference between the two groups, and it is economically significant due to its size. For treatmentM (states with medical cannabis legalization), the difference in mean between the two groups is not statistically significant and not economically significant for family size, age, gender (male), employment status, and family income variables. However, for race (male) and work hours per week variables, the results indicate a statistically significant difference between the two groups. Still, due to the size of the differences in work hours per week, that variable is not economically significant compared to the race difference in Figure 5. Overall, these results demonstrate the importance of conducting a balance of regressors analysis to ensure that the treatment and control groups are comparable and to minimize the potential for confounding variables in treatmentF (states with recreational status) and treatmentM (states with medical status) variables. It is crucial to note that the accuracy and reliability of the results are ensured by cleaning and modifying the data to address missing values and merging.

Figure 5. Balance of Regressors for TreatmentF and Treatment M.

Variables	Control	TreatmentF	Difference (Control – TreatmentF)	TreatmentM	Difference (Control – TreatmentM)
Family Size	4.10	4.19	-0.09***	4.12	-0.02
Age	15.07	15.05	0.03***	15.07	0.01
White	75.92	71.46	4.47***	70.35	5.58***
Male	51.25	51.54	-0.30***	51.39	-0.15
Employed	13.37	12.42	0.95***	13.27	0.10
Work Hours Per Week	4.62	4.39	0.23***	4.42	0.20**
Family Income	86,876.70	97,022.80	-10,146.10***	86,702.90	173.80

Note: Robust standard errors are in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.
Source: Integrated Public Use Microdata Series (IPUMS: American Community Survey), *Family Size, Age, Race, Sex, Employment Status, Work Hours Per Week, Family Income.* Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

In Figures 6, 7, and 8, a parallel test is being conducted to evaluate the consistency level between treatmentF, treatmentM, and the control groups with respect to predictor variables before the post-treatment period, which is 2010 in this analysis. The analysis aims to determine the statistical significance of the association between the predictor variables and the outcome variables and compare the results of the data sets to assess their consistency and interchangeability. In Figure 6, the analysis involves testing the individual effects for each independent variable in treatmentF (states with recreational legalization status) and the interaction effects between treatmentF and year, treatmentF and y2, and treatmentF and y3. The results show that none of the interaction effects are statistically significant. In Figure 7, the analysis includes testing the individual effects for each independent variable in treatmentM (states with medical legalization status). The interaction effects between treatmentM and year, treatmentM and y2, and treatmentM and y3 are not statistically significant. In Figure 8, the visual parallel test for treatmentF (states with recreational legalization status), treatmentM (states with medical legalization status), and the control group (states that never legalized cannabis) show that before the post-treatment year in 2010, the predictor values for marijuana use among 12-18-year-olds within the past month for all three groups were moving in tandem. These results demonstrate the importance of assessing the consistency of the data sets and the relevance of conducting parallel tests to ensure the validity of the analysis.

Figure 6. Parallel Test for TreatmentF (states with recreational legalization status).

Regressors	Marijuana Use Within Past Month	Average SAT Math Scores	Average SAT Writing/Verbal Scores	Dropout Rate	Cognitive Difficulty
TreatmentF	-1.05	16.77	17.86	0.20	0.32
	(1.95)	(30.51)	(30.19)	(0.68)	(0.77)
Year	-0.31	18.89	35.32**	-0.24	-0.35
	(1.01)	(15.17)	(17.99)	(0.44)	(0.58)
Y2	-0.15	-4.26	-10.52**	0.10	0.02
	(0.29)	(4.48)	(5.29)	(0.12)	(0.17)
Y3	0.02	0.38	0.91**	-0.01	0.01
	(0.02)	(0.37)	(0.44)	(0.01)	(0.01)
TreatmentF*Year	2.37	8.42	-2.81	-0.01	-0.31
	(1.93)	(30.73)	(30.94)	(0.65)	(0.83)
TreatmentF*Year2	-0.67	-2.31	1.43	-0.01	0.10
	(0.53)	(8.67)	(8.85)	(0.17)	(0.24)
TreatmentF*Year3	0.05	0.17	-0.13	0.002	-0.01
	(0.04)	(0.71)	(0.74)	(0.01)	(0.02)
Intercept	-14.64	949.27	593.22	-14.64	93.27
	(34.57)	(388.07)	(414.07)	(19.89)	(28.16)
Control variables	Yes	Yes	Yes	Yes	Yes

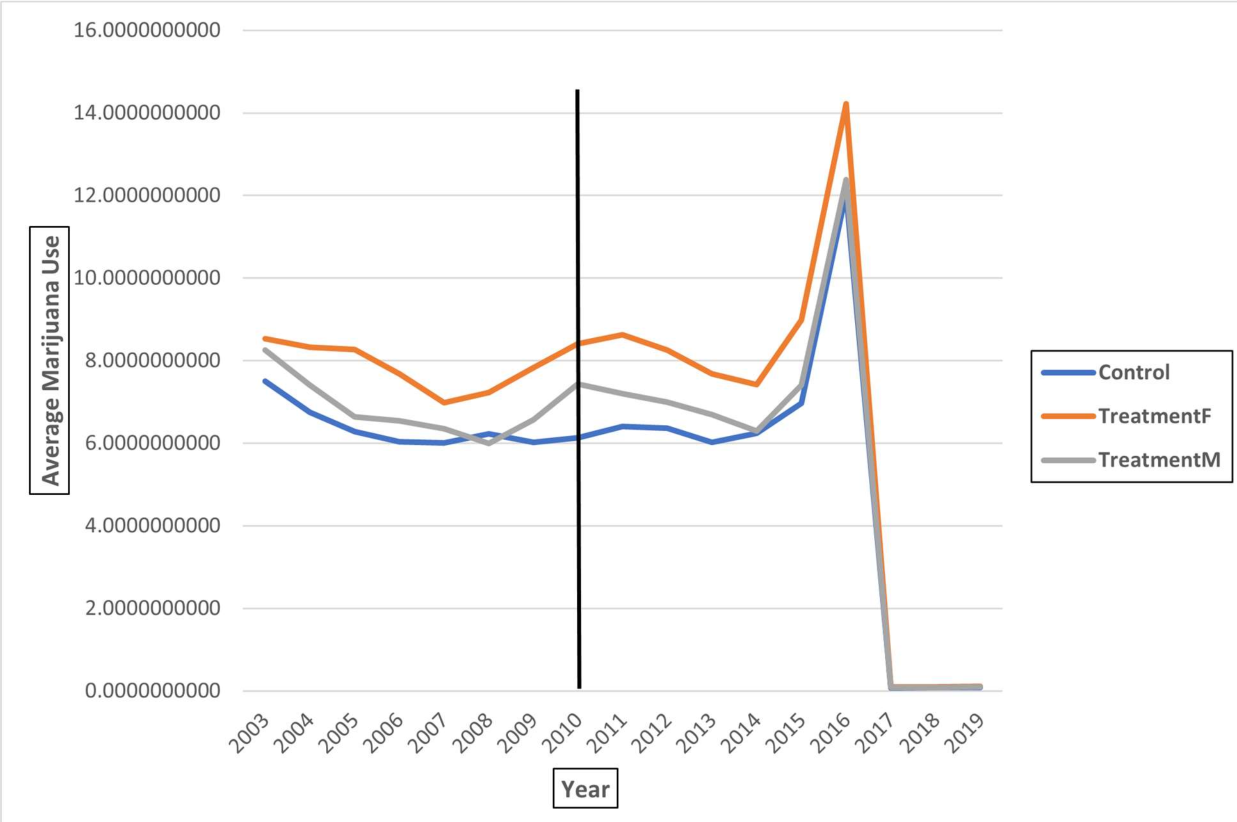
Note: Robust standard errors are in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.
Source: Integrated Public Use Microdata Series (IPUMS: American Community Survey), *Family Size, Age, Race, Sex, Employment Status, Work Hours Per Week, Family Income, School Attendance, Cognitive Difficulty, and Educational Attainment.* National Center of Educational Statistics (NCES), *Average SAT Math and Writing/Verbal Scores*. Substance Abuse and Mental Health Services Administration (SAMHSA), *National Survey On Drug Use and Health*. Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

Figure 7. Parallel Test for TreatmentM (states with medical legalization status).

Regressors	Marijuana Use Within Past Month	Average SAT Math Scores	Average SAT Writing/Verbal Scores	Dropout Rate	Cognitive Ability
TreatmentM	0.37	4.26	6.67	-0.70	-0.61
	(1.44)	(24.31)	(24.76)	(0.65)	(0.93)
Year	-0.50	21.67	33.63**	-0.40	-0.29
	(1.02)	(16.11)	(18.46)	(0.43)	(0.57)
Y2	-0.10	-4.89	-9.58*	0.14	0.01
	(0.29)	(4.78)	(5.46)	(0.12)	(0.17)
Y3	0.01	0.47	0.85**	-0.01	0.01
	(0.02)	(0.40)	(0.46)	(0.01)	(0.01)
TreatmentM*Year	0.25	-9.43	-12.81	0.76	0.31
	(1.46)	(25.50)	(26.09)	(0.63)	(0.99)
TreatmentM*Year2	-0.14	2.46	3.41	-0.21	-0.05
	(0.41)	(7.24)	(7.47)	(0.17)	(0.28)
TreatmentM*Year3	0.01	-0.17	-0.24	0.01	0.01
	(0.03)	(0.59)	(0.62)	(0.01)	(0.02)
Intercept	5.85	1030.67	699.39	-16.62	103.43
	(30.78)	(410.60)	(409.61)	(16.28)	(26.97)
Control variables	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors are in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.
Source: Integrated Public Use Microdata Series (IPUMS: American Community Survey), *Family Size, Age, Race, Sex, Employment Status, Work Hours Per Week, Family Income, School Attendance, Cognitive Difficulty, and Educational Attainment.* National Center of Educational Statistics (NCES), *Average SAT Math and Writing/Verbal Scores*. Substance Abuse and Mental Health Services Administration (SAMHSA), *National Survey On Drug Use and Health*. Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

Figure 8. Parallel Test



Source: Substance Abuse and Mental Health Services Administration (SAMHSA), *National Survey On Drug Use and Health.* Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

The current study analyzes the effects of legalizing cannabis on youth development using three databases merged by state. Prior to merging the databases, missing values were removed by deleting them from the dataset. The first database includes the outcome variables, which are SAT math and writing scores from the National Center for Education Statistics (NCES) from 2003-2021, and past month marijuana use from 2003-2019 from the Substance Abuse and Mental Health Services Administration (SAMHSA). The second database contains two outcome variables, cognitive difficulty, and dropout rates (created from school attendance and educational attainment data), and the averages of control variables, such as Family size, Age, Gender, Race, Family Income, Employment Status, and Hours worked per week from 2003-2021 from the Integrated Public Use Microdata Series (American Community Survey). The third database includes the legalization status of cannabis for states and the year it became legal medically and/or recreationally respectively for each state from the Cannabis Business Info. To analyze the effects of cannabis legality on youth development, the study uses a two-way fixed effects difference in differences model, which requires the following variables to measure the effect of different types of legalization: DIDm and DIDf. The did (difference in differences) variables are obtained based on the condition of states' cannabis legalization status and the year it became legalized. DIDm measures the difference in the effect of legalizing cannabis medically on the treatment group compared to the control group before and after the policy change, with a value of 1 indicating the treatment group, and a value of 0 indicating the control group. DIDf measures the difference in the effect of legalizing cannabis recreationally on the treatment group compared to the control group before and after the policy change, with a value of 1 indicating the treatment group, and a value of 0 indicating the control group. There are 969 observations in the dataset for DIDm and DIDf variables. The DIDm variable has two levels, 0 and 1, with 754 (77.81%) observations in the control group and 215 (22.19%) in the treatment group. The DIDf variable also has two levels, 0 and 1, with 332 (34.26%) observations in the control group and 637 (65.74%) in the treatment group. The average value of the DIDm variable is 0.22, indicating that there is a small effect of legalizing cannabis medically on the outcome variables. The average value of the DIDf variable is 0.65, indicating that there is a large effect of legalizing cannabis recreationally on the outcome variables.

IV. Conceptual Analysis

The legalization of cannabis remains a controversial topic in modern times. Advocates of legalization argue that it could lead to an increase in tax revenue and a decrease in crime, while critics suggest that it may contribute to substance abuse and have negative effects on youth development. The economic principle of supply and demand plays a crucial role in determining the price, quantity, and consumption of legalized cannabis in a market. In a perfectly competitive market, the increase in the supply of goods leads to a decrease in their price, which can ultimately result in an increase in consumption. When applied to the legalization of recreational cannabis, an increase in supply could lead to a decrease in price and an increase in quantity demanded, potentially among youth. Despite being legalized for adults, it is expected that youth will have easier access to it, as they do in the case of cigarettes and alcohol. However, the impact of cannabis legalization on youth development, access, and use continues to be a subject of debate and research. Some researchers argue that legalization may increase youth access to the drug, leading to higher levels of use and potential negative effects on their physical and mental health. Others argue that legalization could reduce the illicit market, thus limiting youth access to cannabis. Given these conflicting views, it is challenging to have clear a priori expectations regarding the impact of cannabis legalization on youth development, access, and use. The effects of legalization may vary depending on several factors such as the age of the youth, the type of legalization (medicinal vs. recreational), and the specific state or region where legalization occurs. A testable hypothesis that can be derived from this theory is that the legalization of cannabis for recreational use in a particular state will lead to an increase in cannabis use among youth aged 12-18. This hypothesis can be tested by analyzing the change in cannabis use among youth in that state before and after legalization, compared to states where cannabis remains illegal.

V. Empirical Methodology

The econometric model being used in this research is a two-way fixed difference-in-differences model with the following two equations: $Y = \beta_0 + \beta_1 DIDm_{st} + \beta_2 DIDf_{st} + Statefip_s +$

$Year_t$ to estimate the effects of legalization without accounting for the control variables but having fixed effects for state and year. And $Y = \beta_0 + \beta_1 DIDm_{st} + \beta_2 DIDf_{st} + Age_{st} + EmploymentStatus_{st} + FamilySize_{st} + FamilyIncome_{st} + White_{st} + Male_{st} + Hourworked_{st} + Statefip_s + Year_t$ to estimate the effects of legalization while accounting for the control variables and fixed effects for state and year. The dependent variables (Y) for the model include youth development, which is measured based on average SAT math and writing/verbal scores, marijuana used within the past month among 12-18 years old, high school dropout rates, and cognitive difficulty by state. The independent variables (DIDm and DIDf) are the difference in the effect of legalizing cannabis medically or recreationally on the treatment group compared to the control group. The control variables family size, age, race, sex, employment status, hours worked per week, and family income helps eliminate alternative explanations for the relationship between the independent variables and dependent variables. This model compares changes in youth development outcomes in states that have legalized cannabis with changes in states that have not legalized cannabis before and after legalization. The model includes fixed effects for states to control state-specific factors that may affect youth development, and year to control for unobserved time-invariant factors that vary across years but are constant within each year.

When conducting a two-way fixed difference in difference analysis with a fixed effect for state and year, there are several assumptions that must be met to ensure that the coefficients obtained from the regression represent causal effects rather than simply correlation. One of the key assumptions of this model is that the treatment effect (legalization of medical or recreational cannabis) is constant over time and across regions. This means that the effect of legalizing cannabis on youth development does not vary over time or between states. If this assumption is violated, it can lead to biased estimates of the treatment effect. Another important assumption is that there are no other confounding factors that influence the outcome variables besides the legal status of cannabis and other control variables included in the model. In other words, it needs to be assumed that the treatment and control groups are similar in terms of all other relevant factors that could affect youth development. This assumption is crucial to ensure that, after controlling for other variables included in the model, any differences in outcomes observed between the treatment and control groups are due to the legal status of cannabis and not other factors. Finally, we assume that there is no endogeneity in the model, meaning that the treatment variable is not correlated with the error term. Endogeneity can arise if there are unobserved factors that influence both the treatment variable and the outcome variable, leading to biased estimates of the treatment effect. Overall, meeting these assumptions is crucial to ensure that the coefficients obtained from the model represent the causal effects of legalizing cannabis on youth development.

VI. Results

The following presents the results of two-way fixed estimation models without accounting for control variables but with fixed effects for state and year. The results are presented in five different models in Figure 6, using the following equation: $Y = \beta_0 + \beta_1 DIDm_{st} + \beta_2 DIDf_{st} + Statefip_s + Year_t$. DIDm measures the difference in the effect of legalizing cannabis medically on the treatment group compared to the control group before and after the policy change, while DIDf measures the difference in the effect of legalizing cannabis recreationally on the treatment group compared to the control group before and after the policy change. In **Model 1**, which measures the effects of legalizing cannabis on the outcome variable of marijuana use within the past month, the coefficients of the difference in differences variables (DIDm = 0.09 and DIDf = 0.52) are not statistically significant. This suggests that legalizing cannabis for either medical or recreational use does not have a significant effect on marijuana use among youth. The small size of the coefficients indicates that the economic significance of these variables is also small, and the results are not in line with prior expectations. **Model 2** measures the effects of legalizing cannabis on the outcome variable of average SAT math scores. The coefficient for the difference in differences variable DIDm = -3.98 is not statistically significant, suggesting that legalizing cannabis for medical use does not have a significant effect on average SAT math scores. The coefficient for the difference in differences variable DIDf = -9.99* is statistically significant at the 10% level, suggesting that legalizing cannabis for recreational use has a small but significant negative effect on average SAT math scores. The negative sign of the coefficient suggests that legalizing cannabis has a negative impact on SAT math scores, and the size of the coefficient indicates that the economic significance of this variable is large. **Model 3** measures the effects of legalizing cannabis on the outcome variable of average SAT writing/verbal scores. The coefficient for the difference in differences variable DIDm = -6.19** is statistically significant at the 5% level, suggesting that legalizing cannabis for medical use has a significant negative effect on average SAT writing/verbal scores, and the size of the coefficient indicates that the economic significance of this variable is large. The coefficient for the difference in differences variable DIDf = -8.93 is not statistically significant, suggesting that legalizing cannabis for recreational use does not have a significant effect on average SAT writing/verbal scores. **Model 4** measures the effects of legalizing cannabis on the outcome variable dropout rates. The coefficient for the difference in differences variable DIDm = -0.08 is not statistically significant, suggesting that legalizing cannabis for medical use does not have a significant effect on dropout rates. The small size of the coefficient suggests that the economic significance of this variable is also small, despite the negative effect. The coefficient for the difference in differences variable DIDf = -0.34*** is statistically significant at the 1% level, suggesting that legalizing cannabis for recreational use does have a small but significant negative effect on dropout rates. The small size of the coefficient indicates that the economic significance of this variable is small. **Model 5** measures the effects of legalizing cannabis on the outcome variable cognitive difficulty. The coefficients of the difference in differences variables (DIDm = 0.07 and DIDf = -0.29) are not statistically significant, suggesting that legalizing cannabis for either medical or recreational use does not have a significant effect on cognitive difficulty. The small size of the coefficients suggests that the economic significance of these variables is also small.

Figure 9. Two-Way Fixed Effects Difference in Differences Results Without Control Variables

Regressors	Marijuana Use Within Past Month	Average SAT Math Scores	Average SAT Writing/Verbal Scores	Dropout Rates	Cognitive Difficulty
DIDm	0.09	-3.98	-6.19**	-0.08	0.07
	(0.27)	(2.89)	(2.94)	(0.08)	(0.11)
DIDf	0.52	-9.99*	-8.93	-0.34***	0.29
	(0.53)	(5.70)	(5.73)	(0.13)	(0.18)
Intercept	-1.33*	607.81***	617.45***	2.85***	94.70***
	(0.79)	(8.34)	(8.87)	(0.27)	(0.35)
Control Variables	No	No	No	No	No
State and Year Fixed	Yes	Yes	Yes	Yes	Yes
Number of Observations	867	969	969	969	969
Adjusted R-Square	0.8465	0.8233	0.7996	0.6058	0.6775

Overall Significance	63.41***	246.24***	150.49***	31.49***	46.77***

Note: Robust standard errors are in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.
***Source*:** Integrated Public Use Microdata Series (IPUMS: American Community Survey), *Family Size, Age, Race, Sex, Employment Status, Work Hours Per Week, Family Income, School Attendance, Cognitive Difficulty, and Educational Attainment.* National Center of Educational Statistics (NCES), *Average SAT Math and Writing/Verbal Scores.* Substance Abuse and Mental Health Services Administration (SAMHSA), *National Survey On Drug Use and Health.* Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

Additionally, the study presents the results of the two-way fixed estimation difference with the effects of control variables and fixed effects for state and year considered in five different models in Figure 7. The models follow the equation $Y = \beta_0 + \beta_1 DIDm_{st} + \beta_2 DIDf_{st} + Age_{st} + EmploymentStatus_{st} + FamilySize_{st} + FamilyIncome_{st} + White_{st} + Male_{st} + Hoursworked_{st} + Statefip_s + Year_t$. In **Model 6** (Marijuana Use), the coefficient for DIDm and DIDf is not statistically significant, indicating that legalizing cannabis medically or recreationally does not have a significant effect on the treatment group compared to the control group, and is not economically significant for substance use. However, the control variables like age and race (white) coefficients are statistically significant at the 5% and 1% levels respectively, and economically significant. employment status, family size, gender, and hours worked per week control variables are not statistically and economically significant. The coefficient for age and race is unexpected, indicating that age and race are associated with a higher probability of marijuana use, which could potentially be explained by other factors such as demographics and socioeconomic status. In **Model 7** (Average SAT Math), the coefficient for DIDm and DIDf is statistically significant at the 1% level, indicating that legalizing cannabis medically and recreationally does have a significant effect on the treatment group compared to the control group for the outcome variable average SAT math scores. The control variables coefficient in the model for age, employment status, family size, family income, race (white), and work hours per week are statistically significant at the 1% and 10% levels respectively, and some of the variables' coefficients are economically substantial due to their size, except for the family income and race (white) variables. **Model 8** measures the effects of legalization on the outcome variable average SAT writing/verbal scores. The coefficient for DIDm and DIDf is statistically significant at the 1% level, indicating that legalizing cannabis medically or recreationally does have a significant effect on the average SAT writing / verbal scores between the treatment group compared to the control group, and the coefficients are economically significant due to their sizes. The control variables coefficient in the model for age, employment status, and family income is statistically significant at the 1% level and economically significant for age and employment status but not for family income due to the coefficient size. All other control variables in the model are neither statistically substantial nor economically significant. **Model 9** measures the legalization effects on the outcome variable dropout rates. In the model, the DIDm coefficient has a negative effect on the dropout rate but this effect is not statistically and economically significant, indicating that legalizing cannabis medically does not have a significant effect on the dropout rates. The DIDf coefficient also has a negative effect on the dropout rates and this effect is statistically significant at the 10% level, but the coefficient is not economically significant due to its size, indicating that legalizing cannabis recreationally does have a significant effect on the dropout rates. The coefficient for the control variables age, family size, and family income are statistically significant at the 1% level in the model, but due to the family income coefficient size, that variable is not economically significant. The coefficient for employment status, gender (male), and work hours per week variables is statistically significant at the 10% level respectively, but due to the size of these coefficients, they're not economically significant. Finally, in **Model 10**, which measures the legalization effects on the outcome variable cognitive difficulty, the coefficient for DIDm is not statistically significant nor economically significant due to its size. However, the coefficient for DIDf is statistically significant at the 5% level, indicating that legalizing cannabis recreationally does have a negative and significant effect on cognitive difficulty, but has no economic significance due to the size of the coefficient. The coefficient for the control variable Family Size is statistically significant at the 1% level but not economically significant due to its size. All other control variables in the model are neither statistically substantial nor economically significant.

Figure 10. Two-Way Fixed Effects Difference in Differences Results with Control Variables

Regressors	Marijuana Use Within Past Month	Average SAT Math Scores	Average SAT Writing/Verbal Scores	Dropout Rates	Cognitive Difficulty
DIDm	0.09	-14.60***	-18.16***	0.16	-0.04
	(0.28)	(3.74)	(3.70)	(0.13)	(0.08)
DIDf	0.86	-9.70***	-11.93***	0.21*	-0.30**
	(0.56)	(3.60)	(3.68)	(0.12)	(0.13)
AGE	3.19**	-28.33**	-92.62***	2.01***	0.31
	(1.28)	(13.81)	(14.30)	(0.58)	(0.45)
Employment Status	-0.01	7.66***	7.03***	-0.04*	-0.04
	(0.09)	(0.79)	(0.81)	(0.02)	(0.03)
Family Size	1.75	20.43***	-8.46	2.71***	0.94***
	(1.43)	(5.55)	(6.00)	(0.21)	(0.34)
Family Income	-0.00**	-0.00***	-0.00***	0.00***	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
White	0.13***	0.21*	0.00	0.01	0.00
	(0.03)	(0.11)	(0.12)	(0.00)	(0.01)
Male	-0.05	0.58	-0.15	0.06*	0.03
	(0.06)	(1.06)	(1.03)	(0.04)	(0.02)
Hours worked per week	0.01	-6.89***	-3.23	-0.12*	0.11

	(0.25)	(1.95)	(2.00)	(0.07)	(0.09)
Intercept	-61.71***	-6,170.72***	-8,019.19***	121.60***	-8.16
	(23.58)	(669.44)	(683.85)	(18.42)	(7.45)
State and Year Fixed	Yes	Yes	Yes	Yes	Yes
Number of Observations	867	969	969	969	969
Adjusted R-Square	0.8528	0.3443	0.3476	0.2955	0.6144
Overall Significance	63.81***	54.43***	47.11***	33.32***	29.56***

Note: Robust standard errors are in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.
Source: Integrated Public Use Microdata Series (IPUMS: American Community Survey), *Family Size, Age, Race, Sex, Employment Status, Work Hours Per Week, Family Income, School Attendance, Cognitive Difficulty, and Educational Attainment.* National Center of Educational Statistics (NCES), *Average SAT Math and Writing/Verbal Scores.* Substance Abuse and Mental Health Services Administration (SAMHSA), *National Survey On Drug Use and Health.* Cannabis Business Info, *Where Marijuana is Legal in the United States (Various years), data available as of April 2023*

VII. Conclusion

In this study, we examine the effects of legalizing cannabis on youth development using a difference-in-difference approach with various outcome variables. Our findings indicate that legalizing cannabis for medical or recreational use does not have a significant impact on marijuana use among youth, as observed in models 1 and 6. However, we do observe a negative effect of legalizing cannabis, either medically or recreationally, on SAT math and writing/verbal scores, as indicated by models 2, 3, 7, and 8. The size of the coefficients suggests that these effects are economically significant. We also find that legalizing cannabis for recreational purposes has a statistically significant effect on dropout rates, as seen in models 4 and 9, with a p-value of 1% when control variables are not accounted for. When control variables are included in model 9, the coefficient for dropout rates becomes smaller and statistically significant at the 5% level. On the other hand, legalizing cannabis for recreational purposes does not have a significant impact on cognitive difficulty in model 5 when control variables are not included. However, once control variables are accounted for in model 10, we observe a statistically significant effect at the 5% level for the DIDf variable, although the effect size is not economically significant. Model 7 and 8, which incorporate control variables and fixed effects for state and year, appear to provide the most comprehensive explanation of the effects of legalizing cannabis for medical or recreational use on youth development. These models show that legalizing cannabis has a detrimental effect on the average SAT math and writing/verbal scores among youth, with statistically significant results at the 1% level. Additionally, we find that adding control variables improves the accuracy of some models, such as models 7 and 8 (SAT math and writing/verbal), but not all. Specifically, adding control variables does not change the results in model 6 (Marijuana Use). Overall, our findings suggest that legalizing cannabis for medical or recreational use negatively affects the academic performance and cognitive ability of youth.

To improve the model, future research could consider incorporating other control variables that could affect the association between legalizing cannabis and youth development, such as parenting styles, family values, or school quality. Additionally, future research could explore the long-term effects of legalizing cannabis on youth development. In summary, while legalizing cannabis for medical or recreational use does not have a substantial effect on marijuana use among youth, it does have a negative impact on SAT math and writing/verbal scores, cognitive difficulty, and dropout rate. Furthermore, the economic significance of the findings is considerable, and the connection between legalizing cannabis and youth development may be affected by other factors that require further investigation.

Based on the findings of the study, it is recommended that policymakers should carefully consider the potential negative impact of legalizing cannabis on youth development, specifically on their SAT math and writing/verbal scores. While legalizing cannabis may have potential benefits for adults, policymakers should take measures to prevent or mitigate any negative impact on youth, such as implementing strict age restrictions, limiting advertising and marketing to youth, and investing in education programs to raise awareness about the potential risks associated with cannabis use. Additionally, further research is needed to better understand the long-term effects of legalizing cannabis on youth development.

VIII. References

Anja C.Huizink, Eduard J.H.Mulder (2006). *Maternal smoking, drinking or cannabis use during pregnancy and neurobehavioral and cognitive functioning in human offspring*, https://doi.org/10.1016/j.neubiorev.2005.04.005

Cathy Davies, Sagnik Bhattacharyya (2019). *Cannabidiol as a potential treatment for psychosis,* 2019; 9: 2045125319881916, 10.1177/2045125319881916
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6843725/

David A. A. Baranger, PhD; Sarah E. Paul, MA; Sarah M. C. Colbert, BA; Nicole R. Karcher, PhD, Emma C. Johnson, PhD; Alexander S. Hatoum, PhD, Ryan Bogdan, PhD (2022). *Association of Mental Health Burden With Prenatal Cannabis Exposure From Childhood to Early Adolescence: Longitudinal Findings From the Adolescent Brain Cognitive Development (ABCD) Study*, 176(12), 1261-1265, 10.1001/jamapediatrics.2022.3191, Association of Mental Health Burden With Prenatal Cannabis Exposure From Childhood to Early Adolescence: Longitudinal Findings From the Adolescent Brain Cognitive Development (ABCD) Study | Adolescent Medicine | JAMA Pediatrics | JAMA Network

Emilio Perucca (2017). *Cannabinoids in the Treatment of Epilepsy: Hard Evidence at last?,* 7(2): 61–76, 10.14581/jer.17012, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5767492/

Nadia Solowij, Katy A. Jones, Megan E. Rozman, Sasha M. Davis, Joseph Ciarrochi, Patrick C. L. Heaven, Dan I. Lubman & Murat Yücel (2011). *Verbal learning and memory in adolescent cannabis users, alcohol users and non-users,* 131–144 , https://link.springer.com/article/10.1007/s00213-011-2203-x

National Institute on Drug Abuse (2020-2022). *Marijuana Facts for Teens,* https://nida.nih.gov/sites/default/files/mj_parents_facts_brochure.pdf

Orrin Devinsky, M.D., J. Helen Cross, Ph.D., F.R.C.P.C.H., Linda Laux, M.D., Eric Marsh, M.D., Ian Miller, M.D., Rima Nabbout, M.D., Ingrid E. Scheffer, M.B., B.S., Ph.D., Elizabeth A. Thiele, M.D., Ph.D., and Stephen Wright, M.D (2017). *Trial of Cannabidiol for Drug-Resistant Seizures in the Dravet Syndrome,* 10.1056/NEJMoa1611618, https://www.nejm.org/doi/full/10.1056/NEJMoa1611618?query=featured_home

Peter A. Fried Ph.D.; Amy J.Porath (2004). *Effects of prenatal cigarette and marijuana exposure on drug use among offspring,* https://doi.org/10.1016/j.ntt.2004.12.003

Ralph E. Tarter, Ph.D., Levent Kirisci, Ph.D., Ada Mezzich, Ph.D., Jack R. Cornelius, M.D., M.P.H., Kathleen Pajer, M.D., M.P.H., Michael Vanyukov, Ph.D., William Gardner, Ph.D., Timothy Blackson, Ph.D., and Duncan Clark, M.D., Ph.D (2003). *Neurobehavioral Disinhibition in Childhood Predicts Early Age at Onset of Substance Use Disorder*, https://doi.org/10.1176/appi.ajp.160.6.1078

Scott B. Harpin, Ashley Brooks-Russell, Ming Ma, Katherine A. James, Arnold H. Levinson (2017).*Adolescent Marijuana Use and Perceived Ease Of Access Before and After Recreational Marijuana Implementation In Colorado,* https://doi.org/10.1080/10826084.2017.1334069

IPUMS USA, Family size, Family income, Sex, Gender, Race, Hours, State, Employment Status, School attendance, Educational Attainment, Cognitive Difficulty (various years), data available as of April 2023. https://usa.ipums.org/usa/

NCES National Center for Education Statistics, Average SAT scores and percentage of graduates by state (Various Years), data available as of April 2023.
https://nces.ed.gov/programs/digest/d20/tables/dt20_226.60.asp

Substance Abuse and Mental Health Services Administration (SAMHSA), National Survey On Drug Use and Health (Various years), data available as of April 2023
https://www.samhsa.gov/data/nsduh/state-reports-NSDUH-2018

Cannabis Business Info, Where Marijuana is Legal in the United States (Various years), data available as of April 2023
https://mjbizdaily.com/map-of-us-marijuana-legalization-by-state/

IX. Sas Appendix.

```
/* Merging Datasets*/

Libname Honors '/home/u60656346/Honors';

/*Make sure variables have the same name and delete missing values*/

Data Honors.OutcomeVariables2 (drop=dropout);

Set Honors.OutcomeVariables1;

Statefip=Statefip;

if Statefip ='.' then delete;

Run; quit;

Data Honors.controlVariables4 (drop=_type_ _freq_);

Set Honors.controlVariables3;

statefip=statefip;

if statefip ='.' then delete;

Run; quit;

/*Proc Sort Data*/

Proc sort data=Honors.OutcomeVariables2;

By statefip year;

Run;

Proc sort data=Honors.controlVariables4;

By statefip Year;

Run;

/*Merge Data*/

Data Honors.DataMerge;
```

```
Merge Honors.OutcomeVariables2 Honors.controlVariables4;

By statefip;

Run;

/* Label the new variable */

Data Honors.dataMerge;

set Honors.datamerge;

label SATMath = "SAT Math Score";

label SATWV = "SAT Writing / Verbal Score";

label MarijuanaUse = "Past Month Marijuana Use";

label FAMSIZE = "Family Size";

label AGE = "Age";

label White = "Race";

label Male = "Gender";

label EMPLOYED= "Employment Status";

label UHRSWORK = "Work House Per Week";

label FTOTINC2 = "Family Income";

label CognitiveAbility = "Cognitive Difference";

Run;

/*Finding the Mean of the Outcome data*/

Proc means data=Honors.DataMerge;

VAR

SATMath

SATWV
```

```
Cognitiveability
MarijuanaUse
Dropout;
run;
/*Finding the Mean of the control data*/
Proc means data=Honors.DataMerge;
VAR
FAMSIZE
AGE
White
Male
Employed
UHRSWORK
FTOTINC2;
run;
/*Creating treatment variable for states who legalized recreational and medical cannabis*/
/* Merge the state data and legalization status */
data Honors.merged;
  merge Honors.legalization Honors.DataMerge Honors.Legalizationmfyears;
  by state;
run;
/* Assign treatment groups */
data Honors.Merged1;
```

```
  set Honors.merged;
  if legalization = 'fully' then group = 1;
  else if legalization = 'medical' then group = 2;
  else group = 0;
run;
/* Create separate treatment groups */
data Honors.Merged2;
  set Honors.Merged1;
  if group = 1 then treatment1 = 1;
  else treatment1 = 0;
run;
data Honors.Merged3;
  set Honors.Merged2;
  if group = 2 then treatment2 = 1;
  else treatment2 = 0;
run;
/* Create control group */
data Honors.Merged4;
  set Honors.Merged3;
  if group = 0 then control = 1;
  else control = 0;
run;
/*Finding the DID variable*/
```

```
/* Create DID1 variable based on fully legal states */

/*  Create the DID1 variable based on the year of legalization */

data Honors.Merged5;

  set Honors.Merged4;

  if (year >= myear) & (myear ne 0) & ((Year < fyear) or (fyear =0)) then DIDm = 1;

  else DIDm = 0;

run;

/*  Create the DID2 variable based on the year of legalization */

data Honors.Merged6;

  set Honors.Merged5;

  if (Year >= fyear) & (fyear ne 0) then DIDf = 1;

  else DIDf = 0;

run;

Proc Means Data= Honors.Merged6;

var Treatment1 Treatment2 DIDm DIDf;

run;

/*Balance of regressors*/

ods output ConfLimits=Honors.ParallelTest1 ttests=honors.Pvalue1;

Proc ttest data=Honors.Merged6;

var FAMSIZE AGE White Male Employed UHRSWORK FTOTINC2;

Class Treatment1;

Where Treatment2 ne 1;

Run;
```

```
ods output ConfLimits=Honors.ParallelTest2 ttests=honors.Pvalue2;

Proc ttest data=Honors.Merged6;

var FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2;

Class Treatment2;

Where Treatment1 ne 1;

Run;

Proc freq data = Honors.Merged6;

tables DIDm DIDf;

run;

Proc means data = honors.merged6;

var DIDm DIDf;

run;

/*Parallel Test for Treatment1*/

Data ParallelTrend1;

set honors.merged6;

where year<2010 and treatment2 ne 1 and statefip ne 6 and statefip ne 2 and statefip ne 8 and statefip ne
        53 and statefip ne 50 and statefip ne 44

and statefip ne 41 and statefip ne 35 and statefip ne 32 and statefip ne 30 and statefip ne 26

and statefip ne 23 and statefip ne 15;

Year = year - 2002;

y2= year*year;

y3=year*year*year;

y4=year*year*year*year;

y5=year*year*year*year*year;
```

```
run;

Proc SurveyReg data=ParallelTrend1 plots=none;

Model1: Model SATMath= Treatment1 Year Y2 Y3 Treatment1*Year Treatment1*Y2 Treatment1*Y3
        FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2 / adjrsq solution;

run;

Proc SurveyReg data=ParallelTrend1 plots=none;

Model2: Model SATWV= Treatment1 Year Y2 Y3 Treatment1*Year Treatment1*Y2 Treatment1*Y3
        FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2 / adjrsq solution;

run;

Proc SurveyReg data=ParallelTrend1 plots=none;

Model3: Model MarijuanaUse= Treatment1 Year Y2 Y3 Treatment1*Year Treatment1*Y2
        Treatment1*Y3 FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2 / adjrsq
        solution;

run;

Proc SurveyReg data=ParallelTrend1 plots=none;

Model4: Model Dropout= Treatment1 Year Y2 Y3 Treatment1*Year Treatment1*Y2 Treatment1*Y3
        FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2 / adjrsq solution;

run;

Proc SurveyReg data=ParallelTrend1 plots=none;

Model5: Model CognitiveAbility= Treatment1 Year Y2 Y3 Treatment1*Year Treatment1*Y2
        Treatment1*Y3 FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2/ adjrsq
        solution;

run;

/*Parallel Test for Treatment2*/

Data ParallelTrend2;

set honors.merged6;
```

```
where year<2010 and treatment1 ne 1 and statefip ne 6 and statefip ne 2 and statefip ne 8 and statefip ne
        53 and statefip ne 50 and statefip ne 44

and statefip ne 41 and statefip ne 35 and statefip ne 32 and statefip ne 30 and statefip ne 26

and statefip ne 23 and statefip ne 15;;

Year = year - 2002;

y2= year*year;

y3=year*year*year;

y4=year*year*year*year;

y5=year*year*year*year*year;

run;

Proc SurveyReg data=ParallelTrend2 plots=none;

Model6: Model SATMath= Treatment2 Year Y2 Y3 Treatment2*Year Treatment2*Y2 Treatment2*Y3
        FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2/ adjrsq solution;

run;

Proc SurveyReg data=ParallelTrend2 plots=none;

Model7: Model SATWV= Treatment2 Year Y2 Y3 Treatment2*Year Treatment2*Y2 Treatment2*Y3
        FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2/ adjrsq solution;

run;

Proc SurveyReg data=ParallelTrend2 plots=none;

Model8: Model MarijuanaUse= Treatment2 Year Y2 Y3 Treatment2*Year Treatment2*Y2
        Treatment2*Y3 FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2/ adjrsq
        solution;

run;

Proc SurveyReg data=ParallelTrend2 plots=none;

Model9: Model Dropout= Treatment2 Year Y2 Y3 Treatment2*Year Treatment2*Y2 Treatment2*Y3
        FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2/ adjrsq solution;
```

```
run;

Proc SurveyReg data=ParallelTrend2 plots=none;

Model10: Model CognitiveAbility= Treatment2 Year Y2 Y3 Treatment2*Year Treatment2*Y2
        Treatment2*Y3 FAMSIZE  AGE White Male Employed UHRSWORK FTOTINC2/ adjrsq
        solution;

run;

/* Pararell Trend Visual */

Data Honors.VisualPt /* (Drop anything that was treated before 2010)*/;

set honors.merged6;

where Year le 2019 and statefip ne 6 and statefip ne 2 and statefip ne 8 and statefip ne 53 and statefip ne
        50 and statefip ne 44

and statefip ne 41 and statefip ne 35 and statefip ne 32 and statefip ne 30 and statefip ne 26

and statefip ne 23 and statefip ne 15;

run;

proc sort data = Honors.VisualPt;

by group year;

run;

ods output summary=Honors.VisualPt2;

Proc Means data = Honors.VisualPt;

var marijuanause;

by group year Treatment1 Treatment2 Control;

run;

ods excel file="/home/u60656346/Honors/VisualPt.XLSX";

Proc print data = Honors.VisualPt2;

var Group Year MarijuanaUse_Mean;
```

```
run;

Ods excel close;

/*Multiple regression on the effects legalizing recreational cannabis on youth development no control
        variables*/

ods output ParameterEstimates=PEforModel1  DataSummary=ObsModel1  FitStatistics=AdjRsqModel1
        Effects=OverallSigModel1;

Proc SurveyReg data=Honors.Merged6 plots=none;

        class statefip year;

        Model1: Model MarijuanaUse=  DIDm DIDf Statefip Year/ adjrsq solution;

run;

ods output ParameterEstimates=PEforModel2  DataSummary=ObsModel2  FitStatistics=AdjRsqModel2
        Effects=OverallSigModel2;

Proc SurveyReg data=Honors.Merged6 plots=none;

        class statefip year;

        Model2: Model SatMath= DIDm DIDf Statefip Year / adjrsq solution;

run;

ods output ParameterEstimates=PEforModel3  DataSummary=ObsModel3  FitStatistics=AdjRsqModel3
        Effects=OverallSigModel3;

Proc SurveyReg data=Honors.Merged6 plots=none;

        class statefip year;

        Model3: Model SATWV= DIDm DIDf Statefip Year / adjrsq solution;

run;

ods output ParameterEstimates=PEforModel4  DataSummary=ObsModel4  FitStatistics=AdjRsqModel4
        Effects=OverallSigModel4;

Proc SurveyReg data=Honors.Merged6 plots=none;

        class statefip year;
```

```
	Model4: Model Dropout= DIDm DIDf Statefip Year/ adjrsq solution;

run;

ods output ParameterEstimates=PEforModel5  DataSummary=ObsModel5  FitStatistics=AdjRsqModel5
	Effects=OverallSigModel5;

Proc SurveyReg data=Honors.Merged6 plots=none;

	class statefip year;

	Model5: Model CognitiveAbility= DIDm DIDf Statefip Year / adjrsq solution;

run;

/*Multiple regressionnon the effects legal recreational cannabis on youth development with control
	variables*/

ods output ParameterEstimates=PEforModel6  DataSummary=ObsModel6  FitStatistics=AdjRsqModel6
	Effects=OverallSigModel6;

Proc SurveyReg data=Honors.Merged6 plots=none;

	class statefip year;

	Model6: Model MarijuanaUse= DIDm Didf FAMSIZE  AGE white male employed
	UHRSWORK FTOTINC2 YEAR Statefip / adjrsq solution;

run;

ods output ParameterEstimates=PEforModel7  DataSummary=ObsModel7  FitStatistics=AdjRsqModel7
	Effects=OverallSigModel7;

Proc SurveyReg data=Honors.Merged6 plots=none;

	Model7: Model SATMath= DIDm Didf FAMSIZE  AGE white male employed UHRSWORK
	FTOTINC2 YEAR Statefip / adjrsq solution;

run;

ods output ParameterEstimates=PEforModel8  DataSummary=ObsModel8  FitStatistics=AdjRsqModel8
	Effects=OverallSigModel8;

Proc SurveyReg data=Honors.Merged6 plots=none;
```

```
Model8: Model SATWV= DIDm Didf FAMSIZE  AGE white male employed UHRSWORK
FTOTINC2 YEAR Statefip / adjrsq solution;

run;

ods output ParameterEstimates=PEforModel9  DataSummary=ObsModel9  FitStatistics=AdjRsqModel9
Effects=OverallSigModel9;

Proc SurveyReg data=Honors.Merged6 plots=none;

Model9: Model CognitiveAbility= DIDm Didf FAMSIZE  AGE white male employed
UHRSWORK FTOTINC2 YEAR Statefip / adjrsq solution;

run;

ods output ParameterEstimates=PEforModel10  DataSummary=ObsModel10
FitStatistics=AdjRsqModel10  Effects=OverallSigModel10;

Proc SurveyReg data=Honors.Merged6 plots=none;

class statefip year;

Model10: Model Dropout= DIDm Didf FAMSIZE  AGE white male employed UHRSWORK
FTOTINC2 YEAR Statefip / adjrsq solution;

run;

******************************;

/* Build the results table */

******************************;

/* Step 1: clean-up the output of the regression analysis you have saved */

Data Honors.Results1;

length Model $30; /* Makes sure the variable Model has the right length and its values are not
truncated */

length Parameter $30; /* Makes sure the variable Parameter has the right length and its values
are not truncated */

set PEforModel1-PEforModel5 indsname=M; /*"indsname" creates an indicator variable (here I
call it "M") that tracks the name of databases use in the "set" statement */
```

```
	keep Model Parameter EditedResults;
	if	M="WORK.PEFORMODEL1" then Model="Model1";
	if	M="WORK.PEFORMODEL2" then Model="Model2";
	if	M="WORK.PEFORMODEL3" then Model="Model3";
	if	M="WORK.PEFORMODEL4" then Model="Model4";
	else if M="WORK.PEFORMODEL5" then Model="Model5";
	where Estimate ne 0;
	if Probt le 0.01 then Star="***";
		else if Probt le 0.05 then Star="**";
		else if Probt le 0.1 then Star="*";
	Results=Estimate;
	EditedResults=Cats(put(Results,comma16.2),Star);
	output;
	Results=stderr;
	EditedResults=Cats("(",put(Results,comma16.2),")");
	output;
run;
/* We sometimes need this sorting step when we have multiple regression models */
proc sort data=Honors.Results1 out=Honors.Results1_Sorted;
	by Model Parameter;
run;
/* Step 2: Create separate results columns (in the form of separate databases) corresponding to each
	model */
data Model1Results(rename=(EditedREsults=Model1))
```

```
        Model2Results(rename=(EditedREsults=Model2))
        Model3Results(rename=(EditedREsults=Model3))
        Model4Results(rename=(EditedREsults=Model4))
        Model5Results(rename=(EditedREsults=Model5));

        set Honors.Results1_Sorted;
        if Model="Model1" then output Model1Results;
        if Model="Model2" then output Model2Results;
        if Model="Model3" then output Model3Results;
        if Model="Model4" then output Model4Results;
        else if Model="Model5" then output Model5Results;
        drop Model;
run;
/* Step 3: Create the final results table that would include all models side-by-side*/
data Honors.Results1_Table_Wide;
        merge Model1Results Model2Results Model3Results Model4Results Model5Results;
        by Parameter;
        if mod(_n_,2)=1 then Regressors=Parameter;
        length Order 3;
        if Parameter="Intercept" then Order=1;
                else if Parameter="DIDm" then Order=2;
                else if substr(Parameter,1,6)="DIDf " then order=3;
                else Order=100;
```

```
run;

/* Order the variables in the results table */

proc sort data=Honors.Results1_Table_Wide out=Honors.Results1_WideSorted(drop=Order
        Parameter);

        by Order;

run;

/*Step 4: Create the rows for other statistics*/

/* The row for Number of Obs */

data Honors.NumofObs(keep=Label1 Model1 Model2 Model3 Model4 Model5);

        merge ObsModel1(rename=(nvalue1=NVMoel1)) ObsModel2(rename=(nvalue1=NVMoel2))

        ObsModel3(rename=(nvalue1=NVMoel3)) ObsModel4(rename=(nvalue1=NVMoel4))

        ObsModel5(rename=(nvalue1=NVMoel5));

        by Label1;

        where Label1="Number of Observations";

        Model1=put(NVMoel1,comma16.0);

        Model2=put(NVMoel2,comma16.0);

        Model3=put(NVMoel3,comma16.0);

        Model4=put(NVMoel4,comma16.0);

        Model5=put(NVMoel5,comma16.0);

run;

/* The row for Adj R-sq */

Data Honors.AdjRsq;

        merge AdjRsqModel1(rename=(cvalue1=Model1)) AdjRsqModel2(rename=(cvalue1=Model2))

        AdjRsqModel3(rename=(cvalue1=Model3)) AdjRsqModel4(rename=(cvalue1=Model4))
```

```
        AdjRsqModel5(rename=(cvalue1=Model5));
        by Label1;
        Where Label1="Adjusted R-Square";
        drop nvalue1;
run;
/* The row for Overall Significance */
data OSM1(rename=(EditedValue=Model1)) OSM2(rename=(EditedValue=Model2))
OSM3(rename=(EditedValue=Model3)) OSM4(rename=(EditedValue=Model4))
OSM5(rename=(EditedValue=Model5));
        set OverallSigModel1 OverallSigModel2 OverallSigModel3 OverallSigModel4
        OverallSigModel5 indsname=M;
        Where Effect="Model";
        Label1="Overall Significance";
        if ProbF le 0.01 then Star="***";
                else if ProbF le 0.05 then Star="**";
                else if ProbF le 0.1 then Star="*";
        EditedValue=Cats(Put(FValue,comma16.2),Star);
        if M="WORK.OVERALLSIGMODEL1" then output OSM1;
        if M="WORK.OVERALLSIGMODEL2" then output OSM2;
        if M="WORK.OVERALLSIGMODEL3" then output OSM3;
        if M="WORK.OVERALLSIGMODEL4" then output OSM4;
        else if M="WORK.OVERALLSIGMODEL5" then output OSM5;
        keep Label1 EditedValue;
```

```
run;

Data Honors.OverallSig;

	merge OSM1 OSM2 OSM3 OSM4 OSM5;

	by Label1;

run;

/* Combine all rows for other statistics */

Data Honors.OtherStat;

	set Honors.NumofObs Honors.AdjRsq honors.OverallSig;

	rename Label1=Regressors;

Run;

/* Step 5: Add other statistics to the results table */

Data Honors.Sorted_WithStat;

	set Honors.Results1_WideSorted Honors.OtherStat;

run;

/* Print the clean results table */

ods excel file="/home/u60656346/Honors/HonorsResults1.xlsx" options(Embedded_Titles="ON"
	Embedded_Footnotes="ON"); /*Use the path to your MySAS folder */

Title "Table: The Impact of Legalizing Cannabis on Youth Development";

footnote justify=left "Note: robust standard errors are in parentheses. *, **, and *** indicate

						10%, 5%, and 1% significance levels, respectively.";

proc print data=Honors.Sorted_WithStat noobs;

	var Regressors;

	var Model1-Model5 /style(header)={just=center} style(data)={just=center
	tagattr="type:String"};
```

```
format Regressors $VariableName.;

run;

ods excel close;

******************************;

/* Build the results table */

******************************;

/* Step 1: clean-up the output of the regression analysis you have saved */

Data Honors.Results2;

    length Model $30; /* Makes sure the variable Model has the right length and its values are not
    truncated */

    length Parameter $30; /* Makes sure the variable Parameter has the right length and its values
    are not truncated */

    set PEforModel6-PEforModel10 indsname=M; /*"indsname" creates an indicator variable (here I
    call it "M") that tracks the name of databases use in the "set" statement */

    keep Model Parameter EditedResults;

    if      M="WORK.PEFORMODEL6" then Model="Model6";

    if      M="WORK.PEFORMODEL7" then Model="Model7";

    if      M="WORK.PEFORMODEL8" then Model="Model8";

    if      M="WORK.PEFORMODEL9" then Model="Model9";

    else if M="WORK.PEFORMODEL10" then Model="Model10";

    where Estimate ne 0;

    if Probt le 0.01 then Star="***";

        else if Probt le 0.05 then Star="**";

        else if Probt le 0.1 then Star="*";

    Results=Estimate;
```

```
        EditedResults=Cats(put(Results,comma16.2),Star);

        output;

        Results=stderr;

        EditedResults=Cats("(",put(Results,comma16.2),")");

        output;

run;

/* We sometimes need this sorting step when we have multiple regression models */

proc sort data=Honors.Results2 out=Honors.Results2_Sorted;

        by Model Parameter;

run;

/* Step 2: Create separate results columns (in the form of separate databases) corresponding to each
model */

data Model6Results(rename=(EditedREsults=Model6))

        Model7Results(rename=(EditedREsults=Model7))

        Model8Results(rename=(EditedREsults=Model8))

        Model9Results(rename=(EditedREsults=Model9))

        Model10Results(rename=(EditedREsults=Model10));

        set Honors.Results2_Sorted;

        if Model="Model6" then output Model6Results;

        if Model="Model7" then output Model7Results;

        if Model="Model8" then output Model8Results;

        if Model="Model9" then output Model9Results;

        else if Model="Model10" then output Model10Results;

drop Model;
```

```
run;

/* Step 3: Create the final results table that would include all models side-by-side*/

data Honors.Results2_Table_Wide;

	merge Model6Results Model7Results Model8Results Model9Results Model10Results;

	by Parameter;

	if mod(_n_,2)=1 then Regressors=Parameter;

	length Order 3;

	if Parameter="Intercept" then Order=1;

		else if Parameter="DIDm" then Order=2;

		else if substr(Parameter,1,6)="DIDf " then order=3;

		else Order=100;

run;

/* Order the variables in the results table */

proc sort data=Honors.Results2_Table_Wide out=Honors.Results2_WideSorted(drop=Order
	Parameter);

	by Order;

run;

/*Step 4: Create the rows for other statistics*/

/* The row for Number of Obs */

data Honors.NumofObs2(keep=Label1 Model6 Model7 Model8 Model9 Model10);

	merge ObsModel6(rename=(nvalue1=NVMoel6)) ObsModel7(rename=(nvalue1=NVMoel7))

	ObsModel8(rename=(nvalue1=NVMoel8)) ObsModel9(rename=(nvalue1=NVMoel9))

	ObsModel10(rename=(nvalue1=NVMoel10));

	by Label1;
```

```
        where Label1="Number of Observations";
        Model6=put(NVMoel6,comma16.0);
        Model7=put(NVMoel7,comma16.0);
        Model8=put(NVMoel8,comma16.0);
        Model9=put(NVMoel9,comma16.0);
        Model10=put(NVMoel10,comma16.0);
run;

/* The row for Adj R-sq */
Data Honors.AdjRsq2;
        merge AdjRsqModel6(rename=(cvalue1=Model6)) AdjRsqModel7(rename=(cvalue1=Model7))
        AdjRsqModel8(rename=(cvalue1=Model8)) AdjRsqModel9(rename=(cvalue1=Model9))
        AdjRsqModel10(rename=(cvalue1=Model10));
        by Label1;
        Where Label1="Adjusted R-Square";
        drop nvalue1;
run;
/* The row for Overall Significance */
data OSM6(rename=(EditedValue=Model6)) OSM7(rename=(EditedValue=Model7))
OSM8(rename=(EditedValue=Model8)) OSM9(rename=(EditedValue=Model9))
OSM10(rename=(EditedValue=Model10));
        set OverallSigModel6 OverallSigModel7 OverallSigModel8 OverallSigModel9
        OverallSigModel10 indsname=M;
```

```
	Where Effect="Model";
	Label1="Overall Significance";
	if ProbF le 0.01 then Star="***";
		else if ProbF le 0.05 then Star="**";
		else if ProbF le 0.1 then Star="*";
EditedValue=Cats(Put(FValue,comma16.2),Star);
	if M="WORK.OVERALLSIGMODEL6" then output OSM6;
	if M="WORK.OVERALLSIGMODEL7" then output OSM7;
	if M="WORK.OVERALLSIGMODEL8" then output OSM8;
	if M="WORK.OVERALLSIGMODEL9" then output OSM9;
	else if M="WORK.OVERALLSIGMODEL10" then output OSM10;
	keep Label1 EditedValue;
run;
Data Honors.OverallSig2;
	merge OSM6 OSM7 OSM8 OSM9 OSM10;
	by Label1;
run;
/* Combine all rows for other statistics */
Data Honors.OtherStat2;
	set Honors.NumofObs2 Honors.AdjRsq2 honors.OverallSig2;
	rename Label1=Regressors;
Run;
/* Step 5: Add other statistics to the results table */
```

```
Data Honors.Sorted_WithStat2;

	set Honors.Results2_WideSorted Honors.OtherStat2;

run;

/* Print the clean results table */

ods excel file="/home/u60656346/Honors/HonorsResults2.xlsx" options(Embedded_Titles="ON"
	Embedded_Footnotes="ON"); /*Use the path to your MySAS folder */

Title "Table: The Impact of Legalizing Cannabis on Youth Development w/ Control Variables";

footnote justify=left "Note: robust standard errors are in parentheses. *, **, and *** indicate

	10%, 5%, and 1% significance levels, respectively.";

proc print data=Honors.Sorted_WithStat2 noobs;

	var Regressors;

	var Model6-Model10 /style(header)={just=center} style(data)={just=center
	tagattr="type:String"};

	format Regressors $VariableName.;

run;

ods excel close;
```