The Effects of Marijuana Laws on High School Graduation Rates.

W. Hunter Giles

11/18/2022

Economics of the Public Sector

Fall 2022

ABSTRACT

This study compares how marijuana legalization affects educational attainment. A difference-and-difference model and inverse probability weighting model using BRFSS examines the relationship on high school graduation rates. The research is designed to increase knowledge in marijuana legalization options and to produce results with external validity.

INTRODUCTION

Past research has struggled to find a clear direction relationship between motivation and marijuana use. Researches have pointed out negative associations between early marijuana usage and educational attainment, but have failed to untangle the inherit reverse causality problem. It is unclear whether marijuana leads to low achievement, or low achievers chose to use marijuana. This paper attempts to solve the reverse causality problem with a quasi-experiment in which a counterfactual is used. Motivation, in this case, in proxied by high school graduation status, and marijuana usage is proxied by state marijuana legalization. As of 2022, twenty-two states (including District of Columbia) have legalized marijuana for recreational use, and 40 states legalized marijuana for medicinal use. Given individual level data in all states before and after legalization, a difference-and-difference model can be used to claim causal results.

The data used for this analysis is from the Behavioral Risk Factor Surveillance System annual survey, in which individuals from all U.S. states are randomly selected to participate. The population of interest is age 18-24 year olds between the survey years 2001 and 2020. This age group was chosen because 18 is the average high school graduation age, and the range to 24 includes people that got their GED, while still being similar enough to 18 year olds. From this data, the analysis found that marijuana legalization causes a decrease in high school graduation rates. This help true for both the difference-and-difference model, and the inverse probability weighting robustness check.

Background on Marijuana Liberalization in the U.S.

In 1970, the Comprehensive Drug Abuse Prevention and Control Act was passed. Title II of the act, the Controlled Substance Act (CSA), was the bases for the Nixon Administration’s “war on drugs” campaign. Under this title, drugs were categorized based on their potential for medical use and addictiveness. Schedule one labeling is reserved for drugs that have zero positive benefits on society, such as Marijuana, LSD, and heroin. In 1996, under California Proposition 215, California became the first state to legalize marijuana for medical use, contradicting federal law.

In 2012, Washington and Colorado became the first states to legalize marijuana for recreational use for residents over the age of 21 and to license vendors. Since marijuana legalization is becoming more common in condition as the political marijuana opposition cools. Eighteen states have legalized medical use, and nineteen more have legalized medical and recreational services. This new trend has occurred despite marijuana still being a schedule one drug at the federal level.

LITERATURE REVIEW

Whether marijuana legalization positively affects social well-being is still a question. Research has shown potential negative effects on physical and mental health. Van Ours and Williams (2012) find that Cannabis usage causes small reductions in mental health among men and women and decreases physical state in men. The same researchers found in 2009 that early exposure to cannabis can reduce educational attainment.

Measuring the effect of marijuana use on educational achievement is important because education indicates future wages and societal contributions. However, it is difficult to determine if marijuana appeals more to already lethargic people or if drug exposure decreases educational motivation (reverse causality). As the number of children using marijuana before 12th grade increases (42% of 12th graders), the demand for the true causal relationship grows (Johnston et al., 2006). Longitudinal studies can be used to help control for reverse causality. Beverly et al. (2019) used such data to find a positive relationship between the age at which marijuana was first tried and the last school grade completed. A similar result was found for the probability of employment.

In the past few decades, many studies have been conducted to quantify the relationship between drug use and crime. Because of the political climate surrounding marijuana legalization, it is one of the chief drugs of such studies. Unfortunately, the relationship sways between positive and negative depending on the study. For this reason, Meta-analysis studies are a helpful tool for studying the plethora of information on this topic. Bennet et al. (2008) found that marijuana users are 1.5 times more likely to commit a crime of any type than non-marijuana users. This is substantially smaller than heroin (3.8 times), crack (6.2 times), or amphetamines (1.86 times). Studies with meta-analysis have found a shocking trend that the relationship between drugs (i.e., marijuana) and crime steadily increased from 1980 to the 2000s. One plausible reason for this is the passing of the Comprehensive Drug Abuse Prevention and Control Act of 1970, which made marijuana a highly-controlled schedule 1 substance. The act’s adverse effects could have given illegal drug deals a new product to sell. Although, this relationship is speculative.

The issue with the meta-analysis is that the researchers failed to differentiate between correlation and causality. It was not until 1996 when California became the first state to legalize marijuana for recreational use, that a difference-and-difference method could be used. The quasi-experiments almost consistently show that the legalization of medical marijuana does not statistically increase crime and may reduce property and violent offenses (Chu & Townshend, 2018). Brinkman and Mok-Lamme (2019) found that an additional dispensary reduces crime by 17 offenses per 10,000 residents.

Beverly et al. (2019) attempted to show a relationship with two logistic regression models, but failed to identify a counterfactual. The present study can provide further evidence to distinguish a causal relationship between marijuana usage and educational achievement. In the future, when policy making are evaluating the effects of marijuana legalization, they can refer to this paper’s results.

DATA DESCRIPTION

The present study uses data from the Behavioral Risk Factor Surveillance System survey (BRFSS). This survey is conducted annually over all 50 U.S. states and three territories. The Center of Disease Control (CDC) governs over the survey because it is primarily used for health-related research. BRFSS started in 1984; however, not all states participated until 2001. They use stratified randomization to collect individual data via phone calls. They surveyed over 350,000 people about (1) demographics and current health conditions, (2) optional CDC models, and (3) any state-added questions.

This study uses a pooled cross-sectional sample from the BRFSS between 2001 and 2022 of individuals between ages 18 to 24[[1]](#footnote-1). The sample has 302,620 total observation across all years and groups at the individual level. When broken down by treatment and control, there are 92,236 observation in the pretreatment non-legalization group, 104,704 in the pretreatment legalization group, 84,155 in the posttreatment non-legalization group, and 21,525 in the posttreatment legalization group. Table 1 breaks down the average observed characteristics of each group. The high school graduation rate (hsgrad) is above 88% in each group. The non-legalization hsgrad increased from 88.51% to 93.50% (4.99 percentage point increase), and the legalization hsgrad from 92.21% to 93.89% (1.68 percentage point increase). Therefore, states without marijuana legalization increased by 3.31 percentage points more than states with marijuana legalization. Figure 1 displays hsgrad annually from 2001 to 2019.

The percent of Hispanics in legalization groups (10.85%) is approximately twice as non-legalization groups (19.92%), possibly due to many legalization states near the Mexico border. The percentage white is constant in each group, ~68% in the non-legalization states, and ~61% in the legalization states. The black population is higher in the non-legalization states (~11%), and drops from 8.08% to .044% between pretreatment and posttreatment in legalization states. The Asian population increases in both the treatment and control states from pretreatment to posttreatment. Hawaiian and Native American combined make up less than 3% of each group. The percent of married individuals falls from 21.45% to 12.38% in non-legalization states and from 13.18% to 9.19% in legalization states. Between pretreatment and posttreatment, the percentage of low income workers (less than $35,000) decrease in both legalization and non-legalization states, while higher earners (greater than $35,000) increased.

EMPERICAL STRATEGY

Two identification strategies and a naïve regression are used to analyze the effects of marijuana legalization on educational attainment. The first is a difference-and-difference model and the second is a naïve regression with inverse probability weighting (IPW). The main model is the DD and the IPW model is to serve as a robustness check. Both models use robust standard errors to account for heteroskedasticity.

*Difference and Difference*

The DD model is a quasi-experimental designed to measure the effects of a given treatment by assigning certain observation control group and others two the treatment group. The assignment of the observation is based on their self-selected enrollment into the treatment. This design can be advantageous when evaluating state-level policy effects. States that opt into a policy are used as the treatment group, and states that opt out of the policy are the control group. For this analysis states that have elected to legalize marijuana are assigned to the treatment group, and states that have not elected to legalize marijuana are assigned to the control group. Using *Equation 1*, the DD model is used to find, if any causal relationship exists between marijuana legalization and educational attainment.

**Equation 1**

The primary dependent variable, , is an indicator as to whether an individual (i) in time (t) is has graduated high school. is a state (s) level binary variable equal to one if the state has legalized marijuana and it is the treatment indicator for the model. It is important to note that is not time variant. is a binary variable that is equal to one if the treatment has occurred. The independent variable of interest, , is the interaction between the treatment (Marijuana) and if the observation falls after legalization (Post). is a vector of individual (i) demographics observed at time (t) such as age, race, or gender.

The model is interpreted by holding constant and splitting the data into four groups: pretreatment legalization state, posttreatment legalization state, pretreatment non-legalization state, and posttreatment non-legalization state. is the effect of being in a pretreatment legalization state. is the effect of being in the posttreatment stage regardless of group. The additional effect of being in a posttreatment legalization state is . is the average high school graduation rate of a pretreatment non-legalization. is a vector of coefficients for individual demographics. is the error term.

One key assumption for DD is that the trends in treatment and controls are parallel (*Figure 1*). In this case, the trends are not perfectly parallel, however, they do seem to increase at the same relative rate, and start to converge after 2013 (the selected post year for the control group). The model controls for certain demographics ( that could be disrupting the parallel trends. Unfortunately, it is difficult to visually show the trends while demographics are being controlled for, for this reason, we must rely on economic intuition.

Another possible limitation of the model is states stagger into legalization in difference years, which can create bias in the . To account for this a staggered DD approach is uses. Colorado and Washington were the first states to legalize recreational marijuana in 2012, since then 20 other states have legalized it. The stagger DD adjusts the years in which the new policy was started, so that the policy state date for all observations line up in year 2012.

*Inverse Probability Weighting and Naïve*

To serve as a robustness check for the DD, an IPW model is used. The data is filtered so just observation in the posttreatment period remain (). There tree steps in an IPW. First, a logistic regression is ran to find the probability () of being in the treatment group (*Equation 2*). This model is based off individual level demographics (), state fixed effects (), and year fixed effects (). Second, the probabilities are transformed into inverse probability ratio weights (*Equation 2*). And finally, a simple regression with IPWs is ran to calculated the estimated effect of marijuana legalization on high school graduation (*Equation 3*).

**Equation 2**

**Equation 3**

**Equation 4**

The IPW gives for emphasis to observation with irregular behavior. For example, if an observation has a high probability of being in a legalization state but is actually in a control state, then they receive a higher weight. This process corrects for any bias involving demographic, state, and year differences. In addition to the IPW model, a naïve regression without weights is ran for comparison (Equation 5).

**Equation 5**

*Expectations*

From previous research an economic intuition, I would expect marijuana legalization to have a negative effect on high school graduation. Even though the legal purchasing age is 21, legalizing marijuana increases drug exposure for the youth. It is speculative, but also plausible that youth access to marijuana will increase, therefore increasing the probability of early substance use.

RESULTS AND INTERPRETATIONS

In this analysis both a design-based (DD) and a model-based robustness check (IPW) were used to assess the effects of marijuana legalization on high school education.[[2]](#footnote-2) Additionally, a naïve regression was run for a result comparison. The regression share the null hypothesis that marijuana legalization does not affect high school graduation rates, insinuating a two-tailed t-test. All the coefficients and regression diagnostics are displayed in Table 2. The DD and IPW both controlled for race, marital status, house hold size, income, year fixed effects, and state fixed effects. The control variables coefficients are not displayed because they offer zero unbiased evidence to the present study. The regressions are designed to unbiased only the variable of interest (), and not their observed confounders.

*Difference and Difference*

From the DD results, we can reject the null hypothesis at the 95% confidence level (p < .01), meaning that marijuana legalization has a significant effect on high school graduation rates. Marijuana legalization decreases the high school graduation rate by 1.1 percentage points. Also, legalization state’s high school graduation rate is 2.5 percentage points higher than non-legalization states, and high school graduation rates increased on average posttreatment by 5.7% points. The model had an adjusted r-squared score of 0.073 meaning 7.3% of variation in the dependent variable is accounted for.

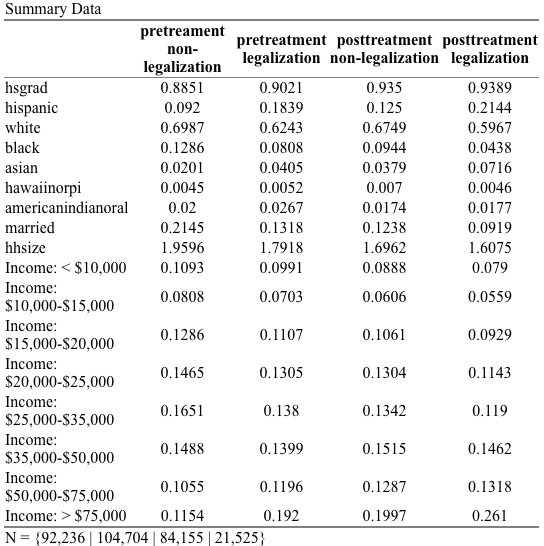
*Inverse Probability Weighting and Naïve*

The IPW produced similar results to the DD and the results are displayed in Table 2. From IPW, we can reject the null hypothesis, meaning that marijuana legalization significantly decreases high school graduation by 1.3 percentage points (p < .05). 1.4% of variation is accounted for from the IPW. Using the Naïve model, we cannot reject the null hypothesis, and 0% of the dependent variables variation is accounted for.

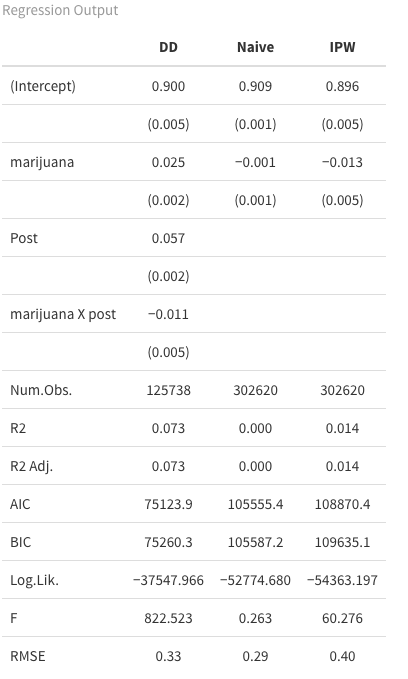
CONCLUSION

APPENDIX

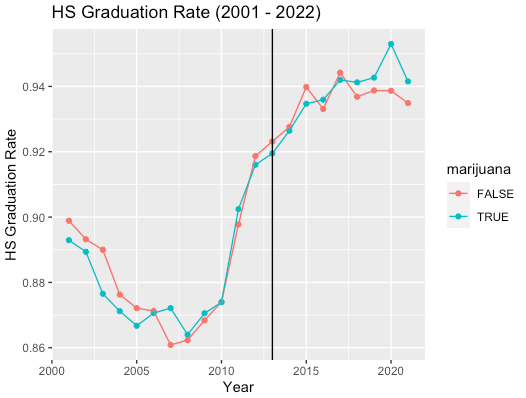
**Table 1**



**Table 2**



**Figure 1**



1. Data can be found at https://github.com/whgiles/GSU/tree/main/public\_sector/finalProject/Marijuana/data [↑](#footnote-ref-1)
2. Regression code can be found at https://github.com/whgiles/GSU/blob/main/public\_sector/finalProject/script.r [↑](#footnote-ref-2)