Isaiah Sarria

Professor Dewey

Econometrics

6 December 2023

Project Report

The Causal Effect of Medical Marijuana Laws on Alcohol Consumption per State

# Abstract

This work finds the relationship between medical marijuana liberalization (MML) and its effect on alcohol consumption per state. The approach used was researching into many articles that spoke on either the predictors of a state becoming a treatment/MML state or predictors of alcohol consumption per state. Then, using available data for predictors of treatment for the year 1990 and available data for predictors of alcohol consumption per state for the year 2020, regressions were used in combination with inverse probability weighting. Inverse probability weighting is a means of matching that estimates the effect treatment had on the states that were treated and states if they would’ve been treated. This analysis found that treatment effect to have a positive correlation between the two variables of interest (predictor variables), which could be interpreted as that the two substances may be complementary as other research has concluded, but not all.

# Introduction

It is important for policymakers to have a rough estimate of the effects of their decisions so that they may justify them reasonably and empirically. In recent years, the liberalization of medical marijuana is becoming more diffused among the United States. There are many reasons why a state may choose to implement or reject such policies. Some are based upon evidence and facts, while other reasons may stem from bias that is related to other things that are as surface level as personal bias and social stigmas. Literature research on the topic reveals many different outcomes and variables that predict those outcomes. In this paper, the focused outcome is the effect of a state adopting medical marijuana laws/legalization (MML) has on the state’s alcohol consumption and related variables. The related variables are factors that impact why a state might have or not have MML, which will be the treatment in this analysis, and other variables that impact alcohol consumption.

# 

Figure : Direct Acrylic Graph

After the review of many articles on the topics of predictors of MML treatment and alcohol consumption, we can see some predictors and their potential causal effects in the DAG above in Figure 1. These will be the predictors present in models during analysis and their implications and origins will be further discussed. Research has already been done on this topic and have concluded different results such as that alcohol and marijuana may be substitutes, have no relationship, or are complements. The findings in this paper will show that the two substances may in fact be complements for each other.

# Predictors: Selection, Discussion, and Collection

The observed predictors of alcohol consumption are unemployment, marital status, poverty, and median income. In that list of predictors of alcohol consumption, some of those variables are confounding as they are also influencing the probability of a state to be treated as shown visually in Figure 1, and they include unemployment and income. The only non-confounding variable used in this paper to predict the likelihood of MML treatment is AIDS per capita in the state.

## Unemployment

Unemployment was found to be either a promoter or a reducer of alcohol consumption, and this makes sense as not having a job may make someone reduce their expenses, but it is possible that an unemployment spell may become so intense to a point that a job seeker might escape to alcohol among other factors (Popvici). Unemployment is one of those confounding variables as it is positively related to the likelihood of adopting the MML treatment, by 1.2% of adoption likelihood for every 1% increase in unemployment (Bradford). This may indicate that the stigma of making users lazy is not prevalent in these MML states, or it at least supersedes a notion that legalizing medical marijuana brings in more benefits to the state in the form of a new revenue stream which will require more employees to operate. Like all the other variables used in this analysis, data for the unemployment rate was collected for the years of 1990 and 2020 were sourced from the U.S Bureau of Labor Statistics who use the Current Population to estimate unemployment bases on the percentage of the labor force in that state who are actively looking for employment (DePersio; “Annual Unemployment Rates by State”; “Unemployment Rates by State”). For reference, the Current Population “is a monthly survey of households conducted by the Bureau of Census and co-sponsered by the Bureau of Labor Statistics”. It consists of sampling 60,000 household via telephone and in-person interviews (“Current Population Survey”).

## Marital Status

Marital status is another predictor of alcohol consumption, and when I begun to search if this was a relevant predictor, I thought it would be more related as a good estimator of someone’s higher age or their likelihood to be less health conscious which would lead to increased alcohol consumption. While that could be somewhat true, marital status was more of an important estimator of the outcome of alcohol consumption due to the joint incomes of two people, especially those of higher joint incomes (Minardi). Data for marriage was presented by the Centers for Disease Control and Prevention who conducts a series of data collection programs, but the relevant once here was the National Survey of Family Growth which consisted of personal interviews of men and women between the ages of 15-44 (“NCHS – Surveys and Data Collection Systems”; “Marriage Rates by State: 1990, 1995, and 1999-2021”; Marriage Rates by State: 2019-2021). That selection is unfortunate though, as I think there could be a lot of individuals who are older who may be single or divorced while at the same time would have the income to sustain any binge drinking they may succumb to due to unfortunate divorces or loneliness.

## Poverty

The effect of poverty on alcohol consumption behaves very similarly to unemployment as the longer the duration of poverty or involuntary unemployment in a span of 13 years serves as significant predictors of heavy drinking (Mossakowski). In hindsight, this was probably to be expected, but it was certainly surprising that despite how related to the other similar confounders of unemployment and income, poverty “does not have a statistically significant effect on the likelihood of MML adoption” (Bradford). The data for percent of persons in poverty in 1990 was originated from a source that cited the Current Population survey from the 1989 census of the Current Population just as the 2020 data of poverty rates for this analysis (“Mapping Poverty”; “Digest of Education).

## Income

Income determines the access to resources and the exposure of financial problems and deprivations, and it also relates to the frequency of drinking alcohol. Those who are not susceptible to financial issues may enjoy more frequent light drinking whereas those in poverty for long durations will respond to high stress via heavy drinking, but findings are inconsistent (Cerda, “The Relationship between Neighborhood; Mossakowski). Unfortunately, it’s gotten to a point that three confounding, including income, are variables that behave the same way. As for income’s effect of MML treatment, “each 10% increase in income leads to a 2.6% increase in adoption” (Bradford). That correlation is possibly explained by the Baby Boomer generation becoming very aged, a desire for tax revenue, and an entrepreneurial desire to capitalize on medical marijuana (Bradford). The income data for 1990, again is a source that cited the Current Population survey, but the 2020 data comes from an author who did income calculations based on “microdata from the United States Census Bureau’s Annual ASEC survey” (“Digest of Education Statistics, 2010”; Pk). The ASEC, Annual Social and Economic Supplement, is an extension of the Current Population survey of more than 75,000 households covering detailed questions of social and economic characteristics of each household member (Bureau).

## AIDS

More of an accidental find, it turns out that AIDS cases per 100,00 residents could be a predictor of state MML because marijuana helps alleviate symptoms of the disease which leads to pressure on legislators (Bradford). Initially, I misinterpreted the article because it wasn’t until later in the article that it claimed the rates of AIDS was not a significant predictor, though in some logistical regression models in this analysis, it has a low p-value and I was just interested in its effect, so I kept it in. The data for 2020 AIDS per state was only found as the HIV per 100,000 and it comes from data reported to the National HIV Surveillance System that is reported to the CDC (“HIV Diagnoses Rates by State U.S. 2020”).

## Unobserved Effect: Population

Population as a variable for either the outcome or treatment had no supporting literature. I include population because I think it is a good indication of access to the number of bars, social gatherings, events, or other situations where alcohol use is encouraged. Furthermore, supporting the legalization of medical marijuana is more of a liberal ideology. It is difficult to find the percentage of liberals per state in 1990, but I think that a higher population attracts more liberal thinking, so I will be using it as a predictor of treatment as well. I emphasize that the possible confounding relationships state population has one the outcome and the propensity for treatment are unobserved. The 2020 population data is sourced from census data and the 1990 population data was found in homework data, but another source was found that has similar numbers, but they don’t state where they got it from (“United States Census, 2020”; “Area, Population & Density by US State: 1990”).

# Analysis and Results

## Setup

The first step taken for the analysis was determining which state would be classified as a treatment or a control state. If the state never had liberalized medical marijuana, their *treat* variable will be assigned to 0. If a state has been an MML state for less than five years then the state would be ignored, but if they were an MML state for more than five years then their *treat* variable will be assigned to 1. The reasoning for this is because an MML state who too recently became an MML state might not be a good basis for comparison purposes between other states who were an MML state for a longer time or for control state comparisons. To further clarify why the control state comparisons would be an issue for new MML states is because data for the year of 1990 is serving as our baseline for state predictors that estimate the likelihood of a state becoming an MML state, and 1990 is a good year to choose because no state had liberalized medical marijuana until 1996, providing an ample time gap of six years of the baseline year and the year the first state became an MML state. Every variable has a version for 2020 and 1990, but the 1990 variables were only used to estimate the likelihood of MML, and the 2020 variables were used to feed into the outcome model of alcohol consumption to measure the impact of MMLs. Finally, the last two steps that was taken was to transform the alcohol consumption per state in per capita terms and then take the natural log of all the variables for the analysis.

## Exploratory Data Analysis

|  |  |  |  |
| --- | --- | --- | --- |
|  | treat | | |
|  | 0 | 1 | Total |
| N | 20 (46.5%) | 23 (53.5%) | 43 (100.0%) |
| alcper2020 | 31.947 (40.966) | 53.802 (71.290) | 43.637 (59.523) |
| aidspercap1990 | 9.285 (10.159) | 16.300 (23.917) | 13.037 (18.944) |
| medinc1990 | 44,166.550 (5,471.166) | 54,837.739 (8,946.219) | 49,874.395 (9,190.699) |
| medinc2020 | 88,385.279 (10,148.209) | 104,614.547 (15,011.232) | 97,066.050 (15,222.011) |
| povertyper2020 | 11.820 (2.762) | 10.287 (2.612) | 11.000 (2.761) |
| unemprate1990 | 5.150 (1.153) | 5.517 (1.016) | 5.347 (1.085) |
| unemprate2020 | 6.455 (1.273) | 8.213 (1.977) | 7.395 (1.889) |
| marrate2020 | 5.660 (0.880) | 5.878 (3.650) | 5.777 (2.710) |
| pop1990 | 5.046 (4.668) | 4.440 (6.321) | 4.722 (5.557) |
| pop2020 | 6.897 (6.871) | 5.812 (8.130) | 6.317 (7.502) |

Figure 2: Summary Statistics

This table shows the averages and standard deviations of all the predictor data by states that are MML states, are not MML states, and then total number of states what were not dropped. Also, we can see that eight of the states (this data does include the District of Columbia as a state) were dropped due to being an MML state for a short duration at the point of 2020.

## Treatment Model Evaluation

I tested multiple combinations of different variables for both the treatment and the outcome models. I found that for both models, including all the variables I have available to me were resulting in the highest R-Squared, meaning that their models were explained more from the variables compared to other models that were created.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | treat | treat | treat | treat |
| lnpop1990 | 0.032 | 0.233 | 0.205 | 0.327 |
|  | (0.036) | (0.143) | (0.125) | (0.171) |
|  | 0.00 | 0.02 | 0.01 | 0.03 |
| lnaidspercap1990 | 5.509 | 6.889 | 7.143 |  |
|  | (6.871) | (4.667) | (4.714) |  |
|  | 0.17 | 0.00 | 0.00 |  |
| lnmedinc1990 | 2.71e+10 |  |  | 1.05e+06 |
|  | (1.84e+11) |  |  | (4.14e+06) |
|  | 0.00 |  |  | 0.00 |
| lnunemprate1990 | 1.84e+06 |  | 12.808 |  |
|  | (9.49e+06) |  | (24.816) |  |
|  | 0.01 |  | 0.19 |  |
| Intercept | 0.000 | 0.112 | 0.002 | 0.000 |
|  | (0.000) | (0.117) | (0.006) | (0.000) |
|  | 0.00 | 0.04 | 0.06 | 0.00 |
| Pseudo R-squared | 0.62 | 0.22 | 0.26 | 0.41 |

Figure 3: Treatment Models

The treatment logistical regression model (logistical regression being chosen due to a binary dependent variable) that included all the variables has a Pseudo R-Squared of .62. That Pseudo R-Squared value is an acceptable score for the effectiveness of the model, but with substantial room for improvement. Furthermore, in the first treatment model, the only significant predictors were a state’s population, income, and unemployment.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Model 1 Propensity | | Model 2 Propensity | | Model 3 Propensity | | Model 4 Propensity | |
| Treat | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| 0 | 17 | 3 | 15 | 5 | 15 | 5 | 17 | 3 |
| 1 | 3 | 20 | 5 | 18 | 4 | 19 | 5 | 18 |

Figure 4: Confusion Matrices for Models 1-4

Furthermore, Figure 4 shows the confusion matrices of the four treatment models when attempting to predict the probability of treatment by assigning a Propensity of treatment of 0 or 1 for each state based on predicted propensities. For example, the first model assigns assigned 20 (17 + 3) states a 0 for Propensity of treatment, meaning that it predicts 20 states for treatment. However, it was only correct for 17 of those states. Then, it predicted 23 states to have treatment, and again, it was wrong for 3 states because 20 of those states were treated (with MML). Essentially, it is wrong in some accounts because some of the propensities towards treatment of different treated states in different models are like those propensities towards treatment of states without treatment. There are states that have the probability to be treated but are still not treated. This suggests some overlap that can be used for matching techniques and can also be represented by Figures 5, despite the small amount of overlap as it may not be enough for a great matching application.



Figure 5: Histogram of Propensity Score Overlap

## Outcome Model Evaluation

The outcome regression model for alcohol consumption with all the variables has the highest R-Squared of .375. In this case, the treatment model seems to be a more decent model compared to the best outcome model we have available here, but the R-Squared is still not pretty at a .375, suggesting better variables would be beneficial. It’s doing something, but not enough to be an acceptable enough model to give any significance to and would have benefitted from other variables mentioned. Furthermore, in the first outcome model, the only significant predictor was a state’s population.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | lnalcper2020 | lnalcper2020 | lnalcper2020 | lnalcper2020 |
| lnmarrate2020 | -0.635 | -0.512 |  | -0.459 |
|  | (0.551) | (0.523) |  | (0.512) |
|  | 0.26 | 0.33 |  | 0.38 |
| lnpop2020 | -0.736 | -0.723 | -0.659 | -0.686 |
|  | (0.167) | (0.165) | (0.154) | (0.152) |
|  | 0.00 | 0.00 | 0.00 | 0.00 |
| lnunemprate2020 | 0.817 | 0.404 | 0.636 |  |
|  | (0.756) | (0.652) | (0.742) |  |
|  | 0.29 | 0.54 | 0.40 |  |
| lnpovertyper2020 | -0.475 |  | -0.533 |  |
|  | (0.754) |  | (0.756) |  |
|  | 0.53 |  | 0.48 |  |
| lnmedinc2020 | -1.447 |  | -1.155 |  |
|  | (1.270) |  | (1.250) |  |
|  | 0.26 |  | 0.36 |  |
| Intercept | 21.322 | 4.190 | 17.291 | 4.846 |
|  | (15.133) | (1.441) | (14.787) | (0.970) |
|  | 0.17 | 0.01 | 0.25 | 0.00 |
|  |  |  |  |  |
| R-Squared | .3725 | .3505 | .3500 | .3441 |

Figure 6: Outcome Models

## Treatment Effects

To find the effect of treatment on alcohol consumption, inverse probability weighting was used combined with regression. Regressions for the selected treatment model were conducted on states that were treated and then separately with the states that were not treated to make some predictions on the potential treatment effects. The regression on the treatment states was used to predict the potential treated outcomes for the states in the control group, and then the regression on the control states was used to predict potential treated outcomes for the control group. Resulting in actual and unobserved counterfactual outcomes to estimate the following treatment effects on the untreated (ATU), treated (ATT), and on average (ATE).

In Figure 8, the ATU, ATT, and the ATE are shown respectively in the “Mean” column. Now, these are log values, so the values would more specifically be 228%, 157%, and 188%. So, if the untreated were treated, alcohol consumption is estimated to increase by 228%. The treated states had their alcohol consumption increase by 157% after liberalization. Generally due to liberalizations being introduced to the United States, it is estimated that alcohol consumption has increased by 188%.

|  |  |  |  |
| --- | --- | --- | --- |
| Summary of teffect | | | |
| treat | Mean | Std. Dev | Frequency |
| 0 | 1.189265 | 0.41028 | 20 |
| 1 | 0.944834 | 0.917141 | 23 |
| Total | 1.058523 | 0.729362 | 43 |

Figure 7: ATU, ATT, ATE

# Conclusion

The positive values from the treatment effect table demonstrate a positive relationship of alcohol consumption and medical marijuana liberalization. This could suggest that both substances are complements of one another. If policymakers are considering implementing medical marijuana liberalization only for the sake of providing an alternative to alcohol, then liberalization is not the solution. Liberalization could possibly be more ideal for economic reasons if state legislators were interested in that outcome of treatment, so another analysis like this one would be useful. Though, the regression models in the analysis would’ve greatly benefited from other predictors that were associated with accessible data which would decrease p-values and increase the R-squared to produce a more reliable analysis. Some predictors that I wanted to include but couldn’t find data for or the article mentioning them was too restricted for me to feel satisfied with using them was smoking rates by state, male populations by state, number of liberalized states near a state, and number of liberals per state.

# References

“Annual Unemployment Rates by State.” *Annual Unemployment Rates by State | Iowa Community Indicators Program*, [www.icip.iastate.edu/tables/employment/unemployment-states. Accessed 7 Dec. 2023](http://www.icip.iastate.edu/tables/employment/unemployment-states.%20Accessed%207%20Dec.%202023).

“Area, Population & Density by US State: 1990”, demographia.com/db-landstatepop.htm. Accessed 9 Dec. 2023.

Bradford, Ashley C., and W. David Bradford. “Factors driving the diffusion of medical marijuana legalisation in the United States.” *Drugs: Education, Prevention and Policy*, vol. 24, no. 1, 2016, <https://doi.org/10.3109/09687637.2016.1158239>.

Bureau, US Census. “Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS).” *Census.Gov*, 8 Oct. 2021, www.census.gov/programs-surveys/saipe/guidance/model-input-data/cpsasec.html.

Cerda, Magdalena, et al. “Lifetime Income Patterns and Alcohol Consumption: Investigating The ...” *PubMed Central*, National Center for Biotechnology Information, www.ncbi.nlm.nih.gov/pmc/articles/PMC3185179/. Accessed 7 Dec. 2023.

Cerdá, Magdalena, et al. “The Relationship between Neighborhood Poverty and Alcohol Use: Estimation by Marginal Structural Models.” *Epidemiology (Cambridge, Mass.)*, U.S. National Library of Medicine, July 2010, www.ncbi.nlm.nih.gov/pmc/articles/PMC3897210/.

“Current Population Survey.” *U.S. Bureau of Labor Statistics*, U.S. Bureau of Labor Statistics, 24 Jan. 2023, www.bls.gov/cex/cecomparison/cps\_profile.htm#:~:text=The%20Current%20Population%20Survey%20(CPS)%20is%20a%20monthly%20sample%20survey,telephone%20and%20in%2Dperson%20interviews.

DePersio, Greg. “How Is the U.S. Monthly Unemployment Rate Calculated?” *Investopedia*, Investopedia, www.investopedia.com/ask/answers/063015/how-does-us-bureau-labor-statistics-calculate-unemployment-rate-published-monthly.asp#:~:text=The%20U.S.%20Bureau%20of%20Labor,underestimates%20the%20true%20unemployment%20rate. Accessed 8 Dec. 2023.

“Digest of Education Statistics, 2007.” *National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education*, nces.ed.gov/programs/digest/d07/tables/dt07\_020.asp. Accessed 7 Dec. 2023.

“Digest of Education Statistics, 2010.” *National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education*, nces.ed.gov/programs/digest/d10/tables/dt10\_025.asp. Accessed 7 Dec. 2023.

“Digest of Education Statistics, 2010.” *National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education*, nces.ed.gov/programs/digest/d10/tables/dt10\_025.asp. Accessed 7 Dec. 2023.

“HIV Diagnoses Rates by State U.S. 2020.” *Statista*, 26 Apr. 2023, www.statista.com/statistics/257734/us-states-with-highest-aids-diagnosis-rates/.

“HIV Surveillance.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 17 Oct. 2023, www.cdc.gov/hiv/library/reports/hiv-surveillance.html.

Macha, Vandana, et al. “Association of Recreational Cannabis Legalization With Alcohol Use Among Adults in the US, 2010 to 2019.” *JAMA Network*, jamanetwork.com/journals/jama-health-forum/fullarticle/2798860?utm\_source=For\_The\_Media&utm\_medium=referral&utm\_campaign=ftm\_links&utm\_term=111822. Accessed 7 Dec. 2023.

“Mapping Poverty: America’s Poverty Rates over Time.” *Self*, Self Financial, 23 June 2023, www.self.inc/info/poverty-rates-in-each-state/.

“Marriage Rates by State: 1990, 1995, and 1999-2021”, www.cdc.gov/nchs/data/dvs/marriage-divorce/state-marriage-rates-90-95-99-21.pdf. Accessed 7 Dec. 2023.

“Marriage Rates by State: 2019-2021.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 10 Feb. 2023, www.cdc.gov/nchs/pressroom/sosmap/marriage\_by\_state/marriage\_rates.htm.

Minardi, Benjamin. *Predicting Alcohol Use Behaviors in the United States: A Complex Survey Analysis*, 12 May 2022, d-scholarship.pitt.edu/42910/.

Mossakowski, Krysia. “Is the Duration of Poverty and Unemployment a Risk Factor for Heavy Drinking?” *Social Science & Medicine*, Pergamon, 21 June 2008, www.sciencedirect.com/science/article/abs/pii/S027795360800258X.

“NCHS - Surveys and Data Collection Systems.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 3 Feb. 2023, www.cdc.gov/nchs/surveys.htm.

Pk. “2020 Average Income by State plus Median, Top 1%, and All Income Percentiles.” *DQYDJ*, 11 Dec. 2021, dqydj.com/2020-average-income-by-state-median-top-percentiles/.

Popovici, Ioana, and Michael T French. “Does Unemployment Lead to Greater Alcohol Consumption?” *Industrial Relations*, U.S. National Library of Medicine, Apr. 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3609661/#:~:text=Khan%20et%20al%20(2002)%20analyze,term%20unemployment%20increases%20alcohol%20use.

Slater, Megan. “Surveillance Reports.” *National Institute on Alcohol Abuse and Alcoholism*, U.S. Department of Health and Human Services, www.niaaa.nih.gov/sites/default/files/pcyr1970-2021.txt. Accessed 7 Dec. 2023.

“Unemployment Rates for States.” *U.S. Bureau of Labor Statistics*, U.S. Bureau of Labor Statistics, www.bls.gov/lau/lastrk20.htm. Accessed 7 Dec. 2023.

“United States Census, 2020.” *Ballotpedia*, ballotpedia.org/United\_States\_census,\_2020. Accessed 9 Dec. 2023.

# Appendix A: Stata Code – Do File

cd "E:\Work 2.0\School Work\College Work\Econometrics" // Stating working directory

log using project\_draft.log, replace // Start a log file for recording

clear // Clear existing data in memory

import delimited econ\_project\_data.csv // Import data from a CSV file

\* Generate treatment variable

gen treat=0 if fymml==. // Set initial value to 0 for missing values

replace treat=1 if fymml<2015 // Set to 1 if liberalized by 2014

drop if treat==. // Exclude observations with missing values or less than 5 years under an MML

\* Develop treatment status model

// Generating variables for the treatment status model

gen alcper1990 = alcbevcon1990/(pop1990 \* 1000000) // vmt per capita, 1990

gen alcper2020 = alcbevcon2020/(pop2020 \* 1000000) // fatalities per million residents, 1990

gen lnalcper1990 = ln(alcper1990)

gen lnalcper2020 = ln(alcper2020)

gen lnaidspercap1990 = ln(aidspercap1990)

gen lnaidspercap2020 = ln(aidspercap2020)

gen lnmedinc2020 = ln(medinc2020)

gen lnmedinc1990 = ln(medinc1990)

gen lnpovertyper2020 = ln(povertyper2020)

gen lnpovertyper1990 = ln(povertyper1990)

gen lnunemprate1990 = ln(unemprate1990)

gen lnunemprate2020 = ln(unemprate2020)

gen lnmarrate2020 = ln(marrate2020)

gen lnmarrate1990 = ln(marrate1990)

gen lnpop2020 = ln(pop2020)

gen lnpop1990 = ln(pop1990)

// Display summary statistics of key variables

summarize treat alcper2020 aidspercap1990 medinc1990 medinc2020 povertyper2020 unemprate1990 unemprate2020 marrate2020 pop1990 pop2020

// Display data in a table

dtable treat alcper2020 aidspercap1990 medinc1990 medinc2020 povertyper2020 unemprate1990 unemprate2020 marrate2020 pop1990 pop2020

// Display data in a table, by treatment status, and export to a Word document

dtable alcper2020 aidspercap1990 medinc1990 medinc2020 povertyper2020 unemprate1990 unemprate2020 marrate2020 pop1990 pop2020, by(treat) export(sftable2.docx, replace)

// Fit logistic regression models to estimate propensity scores

logistic treat lnpop1990 lnaidspercap1990 lnmedinc1990 lnunemprate1990, robust

estimates store logistic1 // stores this model in memory for later

predict ps1 // propensity score, also this regression model has the best R-squared and was selected

logistic treat lnpop1990 lnaidspercap1990 , robust

estimates store logistic2

predict ps2

logistic treat lnpop1990 lnaidspercap1990 lnunemprate1990, robust

estimates store logistic3

predict ps3

logistic treat lnpop1990 lnmedinc1990 , robust

estimates store logistic4

predict ps4

// Generate binary variables based on propensity scores

gen p1=0 if ps1<0.5 // prediction of control based on model 1

replace p1=1 if ps1>=0.5

gen p2=0 if ps2<0.5

replace p2=1 if ps2>=0.5

gen p3=0 if ps3<0.5

replace p3=1 if ps3>=0.5

gen p4=0 if ps4<0.5

replace p4=1 if ps4>=0.5

\* Evaluate accuracy

tabulate treat p1

tabulate treat p2

tabulate treat p3

tabulate treat p4

// Export a table with model statistics to a Word document

etable , estimates(logistic\*) cstat(\_r\_b) cstat(\_r\_se)cstat(\_r\_p) ///

mstat(r2\_p) export(logisticmodels1.docx, replace)

// Display histogram of propensity scores

twoway (hist ps1 if treat==1, fraction fcolor(gs8) lcolor(none) ///

start(0) width(0.2) ) ///

(hist ps1 if treat==0, fraction fcolor(none) lcolor(black) ///

start(0) width(0.2) ) , xtitle("Propensity Score") scheme(s1mono) ///

legend(label(1 "1") label(2 "0")) xtitle("Propensity Score") ///

title("Propensity Score Overlap") saving(hw4pshist, replace)

\* Covariate overlap scatter plot

twoway (scatter lnaidspercap1990 lnpop1990 if treat==1, ms(th) ) ///

(scatter lnaidspercap1990 lnpop1990 if treat==0, ms(oh) ) , ///

ytitle("AIDS Per Capita" "(1990, log)") ///

xtitle("Population (1990, log)") ///

title("Overlap in Two Dimensions") ///

legend(label(1 "Treatment") label(2 "Control")) ///

scheme(s1mono) saving(pr\_draftscatter, replace)

\* Implement ipwra estimator step by step

gen ipw=1/ps1 if treat==1

replace ipw=1/(1-ps1) if treat==0

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 ///

if treat==0 [aw=ipw] // [aw = ipw] is the weighting for balancing

predict pycontrol

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 ///

if treat==1 [aw=ipw]

predict pytreat

\*Calculate ATU, ATT, ATE

gen teffect=pytreat-pycontrol

// Display summary statistics of treatment effects

tabulate treat, summarize(teffect)

\* ipwra estimate using teffects

teffects ipwra (lnalcper2020 lnmarrate2020 lnpovertyper2020 lnmedinc2020 lnunemprate2020) ///

(treat lnaidspercap2020 lnpop1990 lnmedinc2020 lnunemprate2020, logit)

\* using propensity score matching estimator using teffects

teffects psmatch (lnalcper2020) (treat lnaidspercap1990 lnpop1990, logit) , nn(1)

\* Lets make some outcome model regressions

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

estimate store regression1

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020

estimate store regression2

regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

estimate store regression3

regress lnalcper2020 lnmarrate2020 lnpop2020

estimate store regression4

// Display results in a table and export to a Word document

etable, estimates(regression\*) cstat(\_r\_b) cstat(\_r\_se) cstat(\_r\_p) mstat(r2\_p) ///

export(regcmodels1.docx, replace)

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 1

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 0

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 if treat == 1

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 if treat == 0

regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020

regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 1

regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 0

regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

regress lnalcper2020 lnmarrate2020 lnpop2020 if treat == 1

regress lnalcper2020 lnmarrate2020 lnpop2020 if treat == 0

regress lnalcper2020 lnmarrate2020 lnpop2020

// Close the log file

log close

# Appendix B – Stata Log – Raw Outputs

--------------------------------------------------------------------------------------------------------------------

name: <unnamed>

log: C:\Users\Isaiah\OneDrive - Florida Polytechnic University\Documents\College\_Work\Econometrics\project\_d

> raft.log

log type: text

opened on: 10 Dec 2023, 12:22:33

.

. clear

.

. import delimited econ\_project\_data.csv

(encoding automatically selected: ISO-8859-1)

(20 vars, 51 obs)

. \*a) Generate treatment variable

. gen treat=0 if fymml==. // control never liberalized

(31 missing values generated)

. replace treat=1 if fymml<2015 // treatment liberalized by 2014

(23 real changes made)

. drop if treat==. // excluding those observed less than 5 years under an MML

(8 observations deleted)

. \*b) Develop treatment sttus model

. \*generating variables

.

. gen alcper1990 = alcbevcon1990/(pop1990 \* 1000000) // vmt per capita, 1990

. gen alcper2020 = alcbevcon2020/(pop2020 \* 1000000) // fatalities per million residents, 1990

. gen lnalcper1990 = ln(alcper1990)

. gen lnalcper2020 = ln(alcper2020)

. gen lnaidspercap1990 = ln(aidspercap1990)

. gen lnaidspercap2020 = ln(aidspercap2020)

. gen lnmedinc2020 = ln(medinc2020)

. gen lnmedinc1990 = ln(medinc1990)

. gen lnpovertyper2020 = ln(povertyper2020)

. gen lnpovertyper1990 = ln(povertyper1990)

. gen lnunemprate1990 = ln(unemprate1990)

. gen lnunemprate2020 = ln(unemprate2020)

. gen lnmarrate2020 = ln(marrate2020)

. gen lnmarrate1990 = ln(marrate1990)

. gen lnpop2020 = ln(pop2020)

. gen lnpop1990 = ln(pop1990)

.

. summarize treat alcper2020 aidspercap1990 medinc1990 medinc2020 povertyper2020 unemprate1990 unemprate2020 marrate

> 2020 pop1990 pop2020

Variable | Obs Mean Std. dev. Min Max

-------------+---------------------------------------------------------

treat | 43 .5348837 .5046846 0 1

alcper2020 | 43 43.62785 59.42528 1.314248 303.965

aidsper~1990 | 43 13.03721 18.94369 .6 121.1

medinc1990 | 43 49874.4 9190.699 33631 69682

medinc2020 | 43 97066.05 15222.01 65648.61 134385.3

-------------+---------------------------------------------------------

poverty~2020 | 43 11 2.760694 6.2 17.5

unempra~1990 | 43 5.346512 1.0846 2.3 7.6

unempra~2020 | 43 7.395349 1.888934 4.2 13.5

marrate2020 | 43 5.776744 2.709668 3.2 21

pop1990 | 43 4.722848 5.557007 .45369 29.95951

-------------+---------------------------------------------------------

pop2020 | 43 6.316497 7.502153 .577605 39.50165

.

. dtable treat alcper2020 aidspercap1990 medinc1990 medinc2020 povertyper2020 unemprate1990 unemprate2020 marrate202

> 0 pop1990 pop2020

--------------------------------------

Summary

--------------------------------------

N 43

treat 0.535 (0.505)

alcper2020 43.628 (59.425)

aidspercap1990 13.037 (18.944)

medinc1990 49,874.395 (9,190.699)

medinc2020 97,066.050 (15,222.011)

povertyper2020 11.000 (2.761)

unemprate1990 5.347 (1.085)

unemprate2020 7.395 (1.889)

marrate2020 5.777 (2.710)

pop1990 4.723 (5.557)

pop2020 6.316 (7.502)

--------------------------------------

.

. dtable alcper2020 aidspercap1990 medinc1990 medinc2020 povertyper2020 unemprate1990 unemprate2020 marrate2020 pop1

> 990 pop2020, by(treat) export(sftable2.docx, replace)

---------------------------------------------------------------------------------------

treat

0 1 Total

---------------------------------------------------------------------------------------

N 20 (46.5%) 23 (53.5%) 43 (100.0%)

alcper2020 31.976 (41.048) 53.760 (71.104) 43.628 (59.425)

aidspercap1990 9.285 (10.159) 16.300 (23.917) 13.037 (18.944)

medinc1990 44,166.550 (5,471.166) 54,837.739 (8,946.219) 49,874.395 (9,190.699)

medinc2020 88,385.279 (10,148.209) 104,614.547 (15,011.232) 97,066.050 (15,222.011)

povertyper2020 11.820 (2.762) 10.287 (2.612) 11.000 (2.761)

unemprate1990 5.150 (1.153) 5.517 (1.016) 5.347 (1.085)

unemprate2020 6.455 (1.273) 8.213 (1.977) 7.395 (1.889)

marrate2020 5.660 (0.880) 5.878 (3.650) 5.777 (2.710)

pop1990 5.047 (4.668) 4.441 (6.321) 4.723 (5.557)

pop2020 6.896 (6.871) 5.813 (8.130) 6.316 (7.502)

---------------------------------------------------------------------------------------

(collection DTable exported to file sftable2.docx)

.

.

. logistic treat lnpop1990 lnaidspercap1990 lnmedinc1990 lnunemprate1990, robust

Logistic regression Number of obs = 43

Wald chi2(4) = 19.18

Prob > chi2 = 0.0007

Log pseudolikelihood = -11.195505 Pseudo R2 = 0.6231

----------------------------------------------------------------------------------

| Robust

treat | Odds ratio std. err. z P>|z| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop1990 | .0314279 .0351826 -3.09 0.002 .0035028 .2819747

lnaidspercap1990 | 5.533847 6.899397 1.37 0.170 .4805963 63.71973

lnmedinc1990 | 2.96e+10 2.01e+11 3.55 0.000 49673.69 1.77e+16

lnunemprate1990 | 1944746 1.01e+07 2.80 0.005 77.49155 4.88e+10

\_cons | 1.0e-123 8.0e-122 -3.55 0.000 1.4e-191 7.27e-56

----------------------------------------------------------------------------------

Note: \_cons estimates baseline odds.

Note: 0 failures and 1 success completely determined.

. estimates store logistic1

. predict ps1 // propensity score

(option pr assumed; Pr(treat))

. logistic treat lnpop1990 lnaidspercap1990 , robust

Logistic regression Number of obs = 43

Wald chi2(2) = 8.30

Prob > chi2 = 0.0158

Log pseudolikelihood = -23.05754 Pseudo R2 = 0.2237

----------------------------------------------------------------------------------

| Robust

treat | Odds ratio std. err. z P>|z| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop1990 | .2328759 .1429934 -2.37 0.018 .0698971 .775872

lnaidspercap1990 | 6.883271 4.663615 2.85 0.004 1.824237 25.97219

\_cons | .1119562 .1171 -2.09 0.036 .0144125 .8696773

----------------------------------------------------------------------------------

Note: \_cons estimates baseline odds.

. estimates store logistic2

. predict ps2

(option pr assumed; Pr(treat))

. logistic treat lnpop1990 lnaidspercap1990 lnunemprate1990, robust

Logistic regression Number of obs = 43

Wald chi2(3) = 9.50

Prob > chi2 = 0.0233

Log pseudolikelihood = -22.081698 Pseudo R2 = 0.2565

----------------------------------------------------------------------------------

| Robust

treat | Odds ratio std. err. z P>|z| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop1990 | .2053671 .1247501 -2.61 0.009 .0624409 .6754489

lnaidspercap1990 | 7.137591 4.710798 2.98 0.003 1.957763 26.02215

lnunemprate1990 | 12.80074 24.79313 1.32 0.188 .2874645 570.0145

\_cons | .0017431 .0058862 -1.88 0.060 2.33e-06 1.305111

----------------------------------------------------------------------------------

Note: \_cons estimates baseline odds.

. estimates store logistic3

. predict ps3

(option pr assumed; Pr(treat))

.

. logistic treat lnpop1990 lnmedinc1990 , robust

Logistic regression Number of obs = 43

Wald chi2(2) = 12.99

Prob > chi2 = 0.0015

Log pseudolikelihood = -17.656075 Pseudo R2 = 0.4055

------------------------------------------------------------------------------

| Robust

treat | Odds ratio std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

lnpop1990 | .3264402 .1699655 -2.15 0.032 .117656 .9057186

lnmedinc1990 | 1058404 4192044 3.50 0.000 450.0687 2.49e+09

\_cons | 4.70e-65 1.99e-63 -3.49 0.000 3.7e-101 5.97e-29

------------------------------------------------------------------------------

Note: \_cons estimates baseline odds.

. estimates store logistic4

. predict ps4

(option pr assumed; Pr(treat))

. gen p1=0 if ps1<0.5 // prediction of control based on model 1

(23 missing values generated)

. replace p1=1 if ps1>=0.5

(23 real changes made)

. gen p2=0 if ps2<0.5

(23 missing values generated)

. replace p2=1 if ps2>=0.5

(23 real changes made)

. gen p3=0 if ps3<0.5

(24 missing values generated)

. replace p3=1 if ps3>=0.5

(24 real changes made)

. gen p4=0 if ps4<0.5

(21 missing values generated)

. replace p4=1 if ps4>=0.5

(21 real changes made)

. \*evaluate accuracy

. tabulate treat p1

| p1

treat | 0 1 | Total

-----------+----------------------+----------

0 | 17 3 | 20

1 | 3 20 | 23

-----------+----------------------+----------

Total | 20 23 | 43

. tabulate treat p2

| p2

treat | 0 1 | Total

-----------+----------------------+----------

0 | 15 5 | 20

1 | 5 18 | 23

-----------+----------------------+----------

Total | 20 23 | 43

. tabulate treat p3

| p3

treat | 0 1 | Total

-----------+----------------------+----------

0 | 15 5 | 20

1 | 4 19 | 23

-----------+----------------------+----------

Total | 19 24 | 43

. tabulate treat p4

| p4

treat | 0 1 | Total

-----------+----------------------+----------

0 | 17 3 | 20

1 | 5 18 | 23

-----------+----------------------+----------

Total | 22 21 | 43

. etable , estimates(logistic\*) cstat(\_r\_b) cstat(\_r\_se)cstat(\_r\_p) ///

> mstat(r2\_p) export(logisticmodels.docx, replace)

-------------------------------------------------------

treat treat treat treat

-------------------------------------------------------

lnpop1990 0.031 0.233 0.205 0.326

(0.035) (0.143) (0.125) (0.170)

0.00 0.02 0.01 0.03

lnaidspercap1990 5.534 6.883 7.138

(6.899) (4.664) (4.711)

0.17 0.00 0.00

lnmedinc1990 2.96e+10 1.06e+06

(2.01e+11) (4.19e+06)

0.00 0.00

lnunemprate1990 1.94e+06 12.801

(1.01e+07) (24.793)

0.01 0.19

Intercept 0.000 0.112 0.002 0.000

(0.000) (0.117) (0.006) (0.000)

0.00 0.04 0.06 0.00

Pseudo R-squared 0.62 0.22 0.26 0.41

-------------------------------------------------------

(collection ETable exported to file logisticmodels.docx)

.

.

. twoway (hist ps1 if treat==1, fraction fcolor(gs8) lcolor(none) ///

> start(0) width(0.2) ) ///

> (hist ps1 if treat==0, fraction fcolor(none) lcolor(black) ///

> start(0) width(0.2) ) , xtitle("Propensity Score") scheme(s1mono) ///

> legend(label(1 "1") label(2 "0")) xtitle("Propensity Score") ///

> title("Propensity Score Overlap") saving(hw4pshist, replace)

file hw4pshist.gph saved

. \*d) Covariate overlap scatter plot

. twoway (scatter lnaidspercap1990 lnpop1990 if treat==1, ms(th) ) ///

> (scatter lnaidspercap1990 lnpop1990 if treat==0, ms(oh) ) , ///

> ytitle("AIDS Per Capita" "(1990, log)") ///

> xtitle("Population (1990, log)") ///

> title("Overlap in Two Dimensions") ///

> legend(label(1 "Treatment") label(2 "Control")) ///

> scheme(s1mono) saving(pr\_draftscatter, replace)

file pr\_draftscatter.gph saved

. \*f) Implement ipwra estimator step by step

. gen ipw=1/ps1 if treat==1

(20 missing values generated)

.

. replace ipw=1/(1-ps1) if treat==0

(20 real changes made)

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 ///

> if treat==0 [aw=ipw]

(sum of wgt is 29.61731898784637)

Source | SS df MS Number of obs = 20

-------------+---------------------------------- F(5, 14) = 1.22

Model | 6.32046982 5 1.26409396 Prob > F = 0.3489

Residual | 14.4622964 14 1.03302117 R-squared = 0.3041

-------------+---------------------------------- Adj R-squared = 0.0556

Total | 20.7827662 19 1.0938298 Root MSE = 1.0164

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnmarrate2020 | -.1058837 1.972294 -0.05 0.958 -4.336033 4.124266

lnpop2020 | -.5984192 .3967816 -1.51 0.154 -1.449431 .2525926

lnunemprate2020 | -.5485316 1.70591 -0.32 0.753 -4.207345 3.110282

lnpovertyper2020 | 1.764303 2.013618 0.88 0.396 -2.554478 6.083083

lnmedinc2020 | 2.064948 3.722435 0.55 0.588 -5.918881 10.04878

\_cons | -22.85958 45.39259 -0.50 0.622 -120.217 74.49784

----------------------------------------------------------------------------------

. predict pycontrol

(option xb assumed; fitted values)

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 ///

> if treat==1 [aw=ipw]

(sum of wgt is 33.34508800506592)

Source | SS df MS Number of obs = 23

-------------+---------------------------------- F(5, 17) = 2.37

Model | 7.96608081 5 1.59321616 Prob > F = 0.0833

Residual | 11.4217444 17 .671867317 R-squared = 0.4109

-------------+---------------------------------- Adj R-squared = 0.2376

Total | 19.3878252 22 .881264781 Root MSE = .81968

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnmarrate2020 | -.6975877 .5506901 -1.27 0.222 -1.859442 .4642668

lnpop2020 | -.5914577 .219874 -2.69 0.015 -1.055351 -.127564

lnunemprate2020 | .7481337 .9680531 0.77 0.450 -1.29428 2.790547

lnpovertyper2020 | -.4117494 .9782077 -0.42 0.679 -2.475587 1.652088

lnmedinc2020 | -2.36491 1.455797 -1.62 0.123 -5.436374 .7065536

\_cons | 31.9431 17.27806 1.85 0.082 -4.510408 68.39661

----------------------------------------------------------------------------------

. predict pytreat

(option xb assumed; fitted values)

.

. gen teffect=pytreat-pycontrol

. tabulate treat, summarize(teffect)

| Summary of teffect

treat | Mean Std. dev. Freq.

------------+------------------------------------

0 | .44564011 .36826646 20

1 | .3519005 .70120987 23

------------+------------------------------------

Total | .39550032 .56669639 43

. \*e) ipwra estimate using teffects

. teffects ipwra (lnalcper2020 lnmarrate2020 lnpovertyper2020 lnmedinc2020 lnunemprate2020) ///

> (treat lnaidspercap2020 lnpop1990 lnmedinc2020 lnunemprate2020, logit)

Iteration 0: EE criterion = 2.487e-20

Iteration 1: EE criterion = 2.418e-29

Treatment-effects estimation Number of obs = 43

Estimator : IPW regression adjustment

Outcome model : linear

Treatment model: logit

----------------------------------------------------------------------------------

| Robust

lnalcper2020 | Coefficient std. err. z P>|z| [95% conf. interval]

-----------------+----------------------------------------------------------------

ATE |

treat |

(1 vs 0) | .922247 .2272609 4.06 0.000 .4768238 1.36767

-----------------+----------------------------------------------------------------

POmean |

treat |

0 | 2.699702 .1781649 15.15 0.000 2.350506 3.048899

----------------------------------------------------------------------------------

.

. \*g) using propensity score matching estimator using teffects

. teffects psmatch (lnalcper2020) (treat lnaidspercap1990 lnpop1990, logit) , nn(1)

Treatment-effects estimation Number of obs = 43

Estimator : propensity-score matching Matches: requested = 1

Outcome model : matching min = 1

Treatment model: logit max = 1

------------------------------------------------------------------------------

| AI robust

lnalcper2020 | Coefficient std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

ATE |

treat |

(1 vs 0) | .588405 .3040504 1.94 0.053 -.0075229 1.184333

------------------------------------------------------------------------------

.

.

.

. \* Lets make some outcome model regressions

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(5, 37) = 4.39

Model | 20.8108949 5 4.16217898 Prob > F = 0.0031

Residual | 35.0467005 37 .947208123 R-squared = 0.3726

-------------+---------------------------------- Adj R-squared = 0.2878

Total | 55.8575955 42 1.32994275 Root MSE = .97325

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnmarrate2020 | -.6351816 .5510678 -1.15 0.256 -1.751751 .4813878

lnpop2020 | -.7357015 .1669134 -4.41 0.000 -1.0739 -.3975029

lnunemprate2020 | .8175466 .7557256 1.08 0.286 -.7136988 2.348792

lnpovertyper2020 | -.4750968 .7540668 -0.63 0.533 -2.002981 1.052788

lnmedinc2020 | -1.447726 1.269845 -1.14 0.262 -4.020676 1.125223

\_cons | 21.33262 15.13255 1.41 0.167 -9.328839 51.99408

----------------------------------------------------------------------------------

.

. estimate store regression1

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(3, 39) = 7.02

Model | 19.5794889 3 6.52649629 Prob > F = 0.0007

Residual | 36.2781066 39 .930207862 R-squared = 0.3505

-------------+---------------------------------- Adj R-squared = 0.3006

Total | 55.8575955 42 1.32994275 Root MSE = .96447

---------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

----------------+----------------------------------------------------------------

lnmarrate2020 | -.5118581 .522897 -0.98 0.334 -1.569517 .5458008

lnpop2020 | -.7232129 .164586 -4.39 0.000 -1.05612 -.3903064

lnunemprate2020 | .4038029 .6520358 0.62 0.539 -.9150639 1.72267

\_cons | 4.190027 1.44134 2.91 0.006 1.274641 7.105414

---------------------------------------------------------------------------------

.

. estimate store regression2

.

. regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(4, 38) = 5.12

Model | 19.5524586 4 4.88811466 Prob > F = 0.0021

Residual | 36.3051368 38 .955398338 R-squared = 0.3500

-------------+---------------------------------- Adj R-squared = 0.2816

Total | 55.8575955 42 1.32994275 Root MSE = .97744

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop2020 | -.6590497 .1537541 -4.29 0.000 -.9703086 -.3477908

lnunemprate2020 | .6364146 .7423956 0.86 0.397 -.8664866 2.139316

lnpovertyper2020 | -.533901 .7555846 -0.71 0.484 -2.063502 .9957001

lnmedinc2020 | -1.156171 1.249765 -0.93 0.361 -3.686189 1.373846

\_cons | 17.30539 14.78719 1.17 0.249 -12.62972 47.24049

----------------------------------------------------------------------------------

.

. estimate store regression3

.

. regress lnalcper2020 lnmarrate2020 lnpop2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(2, 40) = 10.49

Model | 19.2227289 2 9.61136444 Prob > F = 0.0002

Residual | 36.6348666 40 .915871665 R-squared = 0.3441

-------------+---------------------------------- Adj R-squared = 0.3113

Total | 55.8575955 42 1.32994275 Root MSE = .95701

-------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

--------------+----------------------------------------------------------------

lnmarrate2020 | -.4589204 .5118721 -0.90 0.375 -1.493452 .5756117

lnpop2020 | -.6856371 .1518102 -4.52 0.000 -.9924569 -.3788173

\_cons | 4.845967 .9699955 5.00 0.000 2.885533 6.806401

-------------------------------------------------------------------------------

.

. estimate store regression4

.

. etable, estimates(regression\*) cstat(\_r\_b) cstat(\_r\_se) cstat(\_r\_p) mstat(r2\_p) ///

> export(regcmodels.docx, replace)

--------------------------------------------------------------------

lnalcper2020 lnalcper2020 lnalcper2020 lnalcper2020

--------------------------------------------------------------------

lnmarrate2020 -0.635 -0.512 -0.459

(0.551) (0.523) (0.512)

0.26 0.33 0.38

lnpop2020 -0.736 -0.723 -0.659 -0.686

(0.167) (0.165) (0.154) (0.152)

0.00 0.00 0.00 0.00

lnunemprate2020 0.818 0.404 0.636

(0.756) (0.652) (0.742)

0.29 0.54 0.40

lnpovertyper2020 -0.475 -0.534

(0.754) (0.756)

0.53 0.48

lnmedinc2020 -1.448 -1.156

(1.270) (1.250)

0.26 0.36

Intercept 21.333 4.190 17.305 4.846

(15.133) (1.441) (14.787) (0.970)

0.17 0.01 0.25 0.00

--------------------------------------------------------------------

(collection ETable exported to file regcmodels.docx)

.

.

.

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 1

Source | SS df MS Number of obs = 23

-------------+---------------------------------- F(5, 17) = 2.44

Model | 9.74257116 5 1.94851423 Prob > F = 0.0769

Residual | 13.579906 17 .798818001 R-squared = 0.4177

-------------+---------------------------------- Adj R-squared = 0.2465

Total | 23.3224772 22 1.0601126 Root MSE = .89377

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnmarrate2020 | -.6261934 .5559008 -1.13 0.276 -1.799041 .5466547

lnpop2020 | -.5699706 .2220118 -2.57 0.020 -1.038375 -.1015667

lnunemprate2020 | .3629292 1.017763 0.36 0.726 -1.784363 2.510222

lnpovertyper2020 | -.6375577 .9277169 -0.69 0.501 -2.594869 1.319754

lnmedinc2020 | -2.36955 1.395529 -1.70 0.108 -5.313859 .5747598

\_cons | 33.20855 16.51191 2.01 0.060 -1.628544 68.04563

----------------------------------------------------------------------------------

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 0

Source | SS df MS Number of obs = 20

-------------+---------------------------------- F(5, 14) = 1.91

Model | 11.7204016 5 2.34408032 Prob > F = 0.1556

Residual | 17.1452706 14 1.22466218 R-squared = 0.4060

-------------+---------------------------------- Adj R-squared = 0.1939

Total | 28.8656722 19 1.5192459 Root MSE = 1.1066

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnmarrate2020 | .6693506 2.025726 0.33 0.746 -3.675399 5.0141

lnpop2020 | -.9277712 .4958734 -1.87 0.082 -1.991314 .1357715

lnunemprate2020 | .0781177 2.027603 0.04 0.970 -4.270658 4.426893

lnpovertyper2020 | 1.480345 1.879748 0.79 0.444 -2.551313 5.512004

lnmedinc2020 | 1.865099 3.761988 0.50 0.628 -6.203564 9.933761

\_cons | -21.87844 45.97561 -0.48 0.642 -120.4863 76.72944

----------------------------------------------------------------------------------

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(5, 37) = 4.39

Model | 20.8108949 5 4.16217898 Prob > F = 0.0031

Residual | 35.0467005 37 .947208123 R-squared = 0.3726

-------------+---------------------------------- Adj R-squared = 0.2878

Total | 55.8575955 42 1.32994275 Root MSE = .97325

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnmarrate2020 | -.6351816 .5510678 -1.15 0.256 -1.751751 .4813878

lnpop2020 | -.7357015 .1669134 -4.41 0.000 -1.0739 -.3975029

lnunemprate2020 | .8175466 .7557256 1.08 0.286 -.7136988 2.348792

lnpovertyper2020 | -.4750968 .7540668 -0.63 0.533 -2.002981 1.052788

lnmedinc2020 | -1.447726 1.269845 -1.14 0.262 -4.020676 1.125223

\_cons | 21.33262 15.13255 1.41 0.167 -9.328839 51.99408

----------------------------------------------------------------------------------

.

.

.

.

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 if treat == 1

Source | SS df MS Number of obs = 23

-------------+---------------------------------- F(3, 19) = 2.94

Model | 7.39342281 3 2.46447427 Prob > F = 0.0595

Residual | 15.9290544 19 .838371283 R-squared = 0.3170

-------------+---------------------------------- Adj R-squared = 0.2092

Total | 23.3224772 22 1.0601126 Root MSE = .91563

---------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

----------------+----------------------------------------------------------------

lnmarrate2020 | -.4675988 .5615606 -0.83 0.415 -1.642959 .7077611

lnpop2020 | -.5567951 .2154595 -2.58 0.018 -1.007757 -.1058331

lnunemprate2020 | -.1232189 .9374468 -0.13 0.897 -2.085318 1.83888

\_cons | 5.107807 1.847882 2.76 0.012 1.240147 8.975468

---------------------------------------------------------------------------------

.

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020 if treat == 0

Source | SS df MS Number of obs = 20

-------------+---------------------------------- F(3, 16) = 3.26

Model | 10.9466943 3 3.6488981 Prob > F = 0.0492

Residual | 17.9189779 16 1.11993612 R-squared = 0.3792

-------------+---------------------------------- Adj R-squared = 0.2628

Total | 28.8656722 19 1.5192459 Root MSE = 1.0583

---------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

----------------+----------------------------------------------------------------

lnmarrate2020 | 1.114571 1.810524 0.62 0.547 -2.723568 4.95271

lnpop2020 | -.787444 .4096555 -1.92 0.073 -1.655875 .0809869

lnunemprate2020 | .5086173 1.818748 0.28 0.783 -3.346956 4.364191

\_cons | 1.191528 3.831331 0.31 0.760 -6.930532 9.313587

---------------------------------------------------------------------------------

.

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 lnunemprate2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(3, 39) = 7.02

Model | 19.5794889 3 6.52649629 Prob > F = 0.0007

Residual | 36.2781066 39 .930207862 R-squared = 0.3505

-------------+---------------------------------- Adj R-squared = 0.3006

Total | 55.8575955 42 1.32994275 Root MSE = .96447

---------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

----------------+----------------------------------------------------------------

lnmarrate2020 | -.5118581 .522897 -0.98 0.334 -1.569517 .5458008

lnpop2020 | -.7232129 .164586 -4.39 0.000 -1.05612 -.3903064

lnunemprate2020 | .4038029 .6520358 0.62 0.539 -.9150639 1.72267

\_cons | 4.190027 1.44134 2.91 0.006 1.274641 7.105414

---------------------------------------------------------------------------------

.

.

.

.

.

.

. regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 1

Source | SS df MS Number of obs = 23

-------------+---------------------------------- F(4, 18) = 2.69

Model | 8.72896271 4 2.18224068 Prob > F = 0.0642

Residual | 14.5935145 18 .810750804 R-squared = 0.3743

-------------+---------------------------------- Adj R-squared = 0.2352

Total | 23.3224772 22 1.0601126 Root MSE = .90042

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop2020 | -.4544098 .1983529 -2.29 0.034 -.8711337 -.0376858

lnunemprate2020 | -.012192 .968897 -0.01 0.990 -2.047769 2.023385

lnpovertyper2020 | -.5687046 .9325895 -0.61 0.550 -2.528002 1.390593

lnmedinc2020 | -2.110418 1.386681 -1.52 0.145 -5.023726 .8028906

\_cons | 29.65719 16.32875 1.82 0.086 -4.648234 63.96262

----------------------------------------------------------------------------------

.

. regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020 if treat == 0

Source | SS df MS Number of obs = 20

-------------+---------------------------------- F(4, 15) = 2.51

Model | 11.5866921 4 2.89667302 Prob > F = 0.0855

Residual | 17.2789801 15 1.15193201 R-squared = 0.4014

-------------+---------------------------------- Adj R-squared = 0.2418

Total | 28.8656722 19 1.5192459 Root MSE = 1.0733

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop2020 | -.9888579 .4462501 -2.22 0.043 -1.940017 -.0376985

lnunemprate2020 | .1466376 1.956162 0.07 0.941 -4.022824 4.316099

lnpovertyper2020 | 1.631702 1.768118 0.92 0.371 -2.136953 5.400357

lnmedinc2020 | 1.862175 3.64856 0.51 0.617 -5.914547 9.638897

\_cons | -21.09452 44.53011 -0.47 0.643 -116.0082 73.81917

----------------------------------------------------------------------------------

.

. regress lnalcper2020 lnpop2020 lnunemprate2020 lnpovertyper2020 lnmedinc2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(4, 38) = 5.12

Model | 19.5524586 4 4.88811466 Prob > F = 0.0021

Residual | 36.3051368 38 .955398338 R-squared = 0.3500

-------------+---------------------------------- Adj R-squared = 0.2816

Total | 55.8575955 42 1.32994275 Root MSE = .97744

----------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

-----------------+----------------------------------------------------------------

lnpop2020 | -.6590497 .1537541 -4.29 0.000 -.9703086 -.3477908

lnunemprate2020 | .6364146 .7423956 0.86 0.397 -.8664866 2.139316

lnpovertyper2020 | -.533901 .7555846 -0.71 0.484 -2.063502 .9957001

lnmedinc2020 | -1.156171 1.249765 -0.93 0.361 -3.686189 1.373846

\_cons | 17.30539 14.78719 1.17 0.249 -12.62972 47.24049

----------------------------------------------------------------------------------

.

.

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 if treat == 1

Source | SS df MS Number of obs = 23

-------------+---------------------------------- F(2, 20) = 4.63

Model | 7.37893849 2 3.68946925 Prob > F = 0.0223

Residual | 15.9435387 20 .797176934 R-squared = 0.3164

-------------+---------------------------------- Adj R-squared = 0.2480

Total | 23.3224772 22 1.0601126 Root MSE = .89285

-------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

--------------+----------------------------------------------------------------

lnmarrate2020 | -.4910063 .5193263 -0.95 0.356 -1.574302 .5922894

lnpop2020 | -.5693837 .1882016 -3.03 0.007 -.9619653 -.1768021

\_cons | 4.905324 .995181 4.93 0.000 2.829413 6.981235

-------------------------------------------------------------------------------

.

. regress lnalcper2020 lnmarrate2020 lnpop2020 if treat == 0

Source | SS df MS Number of obs = 20

-------------+---------------------------------- F(2, 17) = 5.13

Model | 10.8591092 2 5.42955459 Prob > F = 0.0181

Residual | 18.006563 17 1.05920959 R-squared = 0.3762

-------------+---------------------------------- Adj R-squared = 0.3028

Total | 28.8656722 19 1.5192459 Root MSE = 1.0292

-------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

--------------+----------------------------------------------------------------

lnmarrate2020 | 1.244835 1.701481 0.73 0.474 -2.344977 4.834648

lnpop2020 | -.7032205 .2700587 -2.60 0.019 -1.272995 -.1334464

\_cons | 1.776684 3.121257 0.57 0.577 -4.808593 8.361961

-------------------------------------------------------------------------------

.

. regress lnalcper2020 lnmarrate2020 lnpop2020

Source | SS df MS Number of obs = 43

-------------+---------------------------------- F(2, 40) = 10.49

Model | 19.2227289 2 9.61136444 Prob > F = 0.0002

Residual | 36.6348666 40 .915871665 R-squared = 0.3441

-------------+---------------------------------- Adj R-squared = 0.3113

Total | 55.8575955 42 1.32994275 Root MSE = .95701

-------------------------------------------------------------------------------

lnalcper2020 | Coefficient Std. err. t P>|t| [95% conf. interval]

--------------+----------------------------------------------------------------

lnmarrate2020 | -.4589204 .5118721 -0.90 0.375 -1.493452 .5756117

lnpop2020 | -.6856371 .1518102 -4.52 0.000 -.9924569 -.3788173

\_cons | 4.845967 .9699955 5.00 0.000 2.885533 6.806401

-------------------------------------------------------------------------------

. log close

name: <unnamed>

log: C:\Users\Isaiah\OneDrive - Florida Polytechnic University\Documents\College\_Work\Econometrics\project\_d

> raft.log

log type: text

closed on: 10 Dec 2023, 12:22:48

--------------------------------------------------------------------------------------------------------------------