

# The Effect of the Minimum Legal Drinking Age on Alcohol Consumption and Crime

Tristan Mesaros

*University of California, Santa Cruz*

*Department of Economics*

October 15, 2021

## **Abstract**

This paper, with the use of statistical empirical methods, was able to show the effect that the minimum legal drinking age has on alcohol consumption and crime rates. This study was made in order to find causal effects of modifying the legal drinking age on different crime rates in order to find if reducing the minimum legal drinking age increases or decreases crime rates. A balance regression table, including different sample characteristic covariates, will be made to provide evidence that there are not significant differences between the two groups, people prior to turning 21 and people post turning 21. Regression discontinuities using the current minimum legal drinking age will be performed on different type of arrests to obtain a better understanding on the change in amount of arrests prior and post to turning 21. There are two main focuses are of the paper are; to determine how much the Minimum Legal Drinking Age (MLDA) reduces the proportion of the population that drinks, and how much the MLDA can reduce crime. Through the use of regression tables and regression discontinuity figures, we were able to find IV estimates for the effect of turning 21 on all arrests. It showed that there is an increase of about 891.304 arrests per 10,000 people that turn 21. We were also able to deduce that the MLDA increases the percent of days on which people drink by 8.66%. The results made it able to find that the reducing the MLDA increases crime and alcohol consumption further.

# 1 Introduction

The article *The Minimum Legal Drinking Age and Crime* is a statistics article by Christopher Carpenter and Carlos Dobkin that focuses on the effect that the modification of the minimum legal drinking age has on crime rates. The results of the paper show that "reducing the availability of alcohol to young adults has the potential to reduce crime." and that a reduction in the minimum legal drinking age has a strong possibility of increasing crime rates. Throughout this paper, I will use regression discontinuity graphs and regression tables to show a that there is causal effect between the minimum legal drinking age and crime.

The data used throughout this paper were gathered from two sources. The first dataset was gathered from the National Health Interview Sample Adult Files from 1997 to 2007 (NHIS.dta). The dataset consists of sample characteristics taken from the participants as well as if they drink, the percentage of days on which they drink and how many days they are from turning 21. This dataset shows the demographic of people around the age of 21 that drink. The second dataset was from the article *The Minimum Legal Drinking Age and Crime* by Carpenter and Dobkin. The dataset consists of variables based on the different types of arrests and their rates. With the combination of both datasets, I was able to find the difference in arrest rates for those prior to turning 21 and those post to turning 21.

In the early 2000's, states such as, Florida, Wisconsin, Vermont, and Missouri, were having policy debates on the proposal of reducing the MLDA. Reducing the MLDA has a chance to reduce underage drinking but may also increase crime rates and reduce the age at which young adults drive drunk. This paper helps understand the effect of reducing the MLDA has on crime rates.

In order to first determine whether bias can occur through sample differences, a balance table was generated using the data comparing the different age groups using many different covariates. The covariates are sample characteristics that

help determine if there are any demographic differences between the age groups. The sample characteristic covariates include race, gender, if they are attending school, if they are employed etc.

With the help of regression discontinuity charts, I was able to show that there is a clear effect of turning 21 on arrest directly related to alcohol and arrests indirectly related to alcohol. Regressions on different age groups and arrest types were made to detect a difference between effect of turning 21 on alcohol and non-alcohol related arrests. The arrests directly related to alcohol include, driving under the influence, liquor law violation, and drunk possible risk to self harm. The arrests indirectly related to alcohol include, robbery, disorderly conduct, simple assault, and aggravated assault.

Regression tables were made to find the point estimates in order to gain a better understanding of the effect that turning 21 has on the different types of crimes previously listed. These regression tables regress the different crimes on turning 21. Using these regression tables, we can gain a better understanding of the effect of turning 21 has on arrests directly related to alcohol in comparison to the arrests indirectly related to alcohol.

The questions being posed is done in order to find if the policies of reducing the minimum legal drinking age in a few different states would affect crime rates. This is done by looking into the change in crime as people turn 21 compared to people prior to turning 21. If acts of crime directly related to alcohol increase as people turn 21, then this would tell us that reducing the minimum legal drinking age to an age lower than 21 would result in an increase in crime rates. The results found in this paper were able to show that a reduction in the MLDA has the large possibility of increasing crime rates, more specifically crime rates related to alcohol. Unfortunately due to the IV assumptions, which I go deeper into later in the paper, the causal effect found cannot be completely trusted.

## 2 Data

The first data set I am using is from a sample drawn from the National Health Interview Sample Adult Files from 1997 to 2007. This data set includes sample characteristics about the participating individuals which can be found in the balance regression table (Table 1). This data set is used to determine two things, the first being the type of people that report drinking and how much, the second being to show that there are little to no differences in the group of individuals prior and post to turning 21. This shows the elimination of selection bias due to difference in sample groups.

Table 1: **Balance Regression Table**

	(1) uninsured	(2) hs_diploma	(3) hispanic	(4) black	(5) white	(6) employed	(7) married	(8) male	(9) going_school
Over21	-0.0178 (0.0134)	0.0106 (0.0113)	-0.00952 (0.0124)	-0.0144 (0.0104)	0.0160 (0.0144)	0.00729 (0.0140)	-0.0301** (0.0101)	0.0154 (0.0144)	0.00730 (0.0110)
Age-21	0.0262** (0.00836)	0.0229** (0.00757)	0.0000670 (0.00791)	0.00864 (0.00659)	-0.00739 (0.00917)	0.0590*** (0.00911)	0.0511*** (0.00557)	-0.0232* (0.00918)	-0.0581*** (0.00766)
(Age-21)*Over21	-0.0171 (0.0116)	-0.0200* (0.00999)	-0.000891 (0.0108)	-0.00383 (0.00900)	0.00563 (0.0126)	-0.0250* (0.0122)	0.00186 (0.00868)	0.0278* (0.0126)	0.0180 (0.00964)
Constant	0.318*** (0.00962)	0.821*** (0.00836)	0.241*** (0.00897)	0.157*** (0.00755)	0.554*** (0.0104)	0.642*** (0.0102)	0.152*** (0.00699)	0.428*** (0.0104)	0.166*** (0.00827)
Observations	18801	18801	18801	18801	18801	18801	18801	18801	18801

Robust standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

However there are still different forms of bias that can arise. The first being measurement error, which would relate to the way that the ages are calculated. In this case I believe that the way the ages are calculated is not very prone to bias because the calculation of days until turning 21 makes the ages very specific. The second would be recall bias, in this case people might forget exactly how many drinks they had on a certain evening due to not being sober enough to count. Recall bias is very possible in this case which cannot be really accounted for. The third bias is desirability bias, which relates to the fact that people might under report their drinking because it promotes taboo behavior and to make themselves desirable.

The second data set I am using to run the regressions is from the statistical article, *The Minimum Legal Drinking Age and Crime* by Christopher Carpenter and Carlos Dobkin from the Massachusetts Institute of Technology Press. The article uses a census on arrests from California to gather the data they need to perform regressions. More specifically, the crime data came from California's Monthly Arrest and Citation Register (MACR) for arrests from 1979 to 2006. This data is taken from a sample that contains arrest rates per 100,000 by age and is broken down by cause. The different causes of arrests are used as the sample characteristic covariates. These covariates can be found in the regression discontinuity figures, (Figures 4-10). This data set is used for finding the difference between the effect of turning 21 on arrests directly related to alcohol and arrests indirectly related to alcohol.

### 3 Empirical Methods

The regression discontinuity is the main form of regressions that will be made throughout this paper. A regression discontinuity is a quasi-experimental design that uses pre-test and post-test statistics to find the effect of an experiment before and after the treatment was given. When it comes to finding the effect of the MLDA on alcohol consumption and crime rates, the regression discontinuity will be able to show the change in alcohol consumption and crime rates right at the cutoff. In this case, the cutoff takes place once the individual turns 21. Using regression discontinuity will be able to precisely determine the magnitude of the cutoff for each arrest.

To begin making the regression discontinuity figures, I had to find the right binwidth and bandwidth to use. In this case, the binwidth is in regard to the amount of days we are analyzing in the discontinuity. The bandwidth is in regard to the age range we are looking at when analyzing the discontinuity. As seen in Figure 1, the figure shows four graphs all similar except with different bin sizes for the binwidth. I decided to chose the bin size of 30 days. Although

the other bin sizes, 10, 20, and 40 days could work, I found that the graph with a bin size of 30 days seems to be the cleanest and easiest to read the discontinuity. Figure 2 shows the same graphs with the same binwidth of 30 days, except the bandwidths vary. I decided to chose the bandwidth range of 19 to 23 days. There are two reasons as to why I did not chose any other bandwidth range. The first being that the other graphs regression discontinuity line seem to be too cluttered. The second being that the discontinuity is not well centered in the other bandwidth ranges. Having a bandwidth range that is too large will start to show estimates that are not specific to the cutoff because a 30 year old will behave quite differently than a 21 year old. It is very important to find the right fit of binwidth and bandwidth before making regression discontinuities. A wrong bin and bandwidth can cause the discontinuities to be unclear and wrong in terms of their estimates.

Figure 1: Alcohol Drinks Consumed on Age, Binwidth

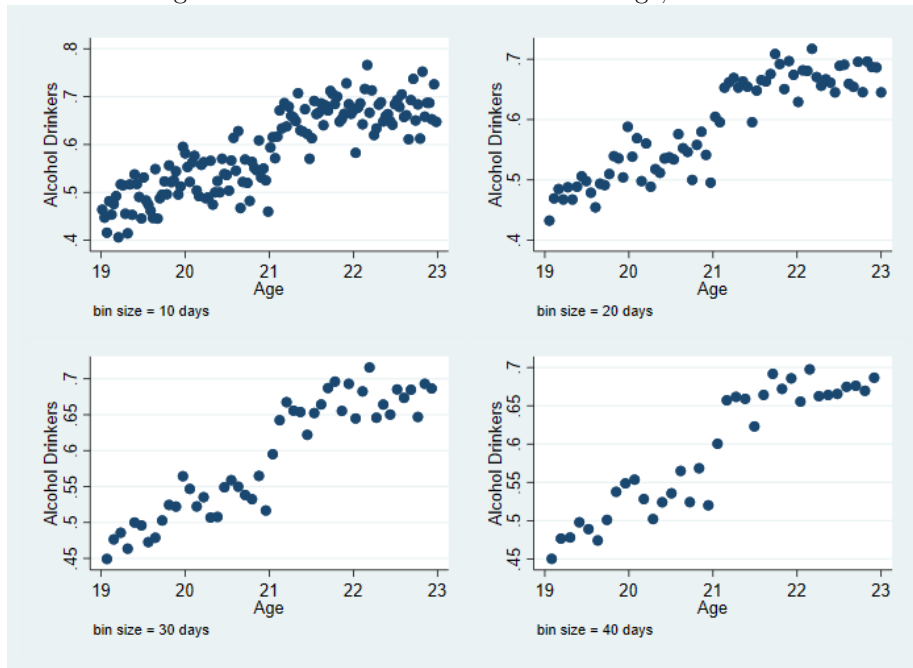


Figure 2: Alcohol Drinks Consumed on Age, Bandwidth

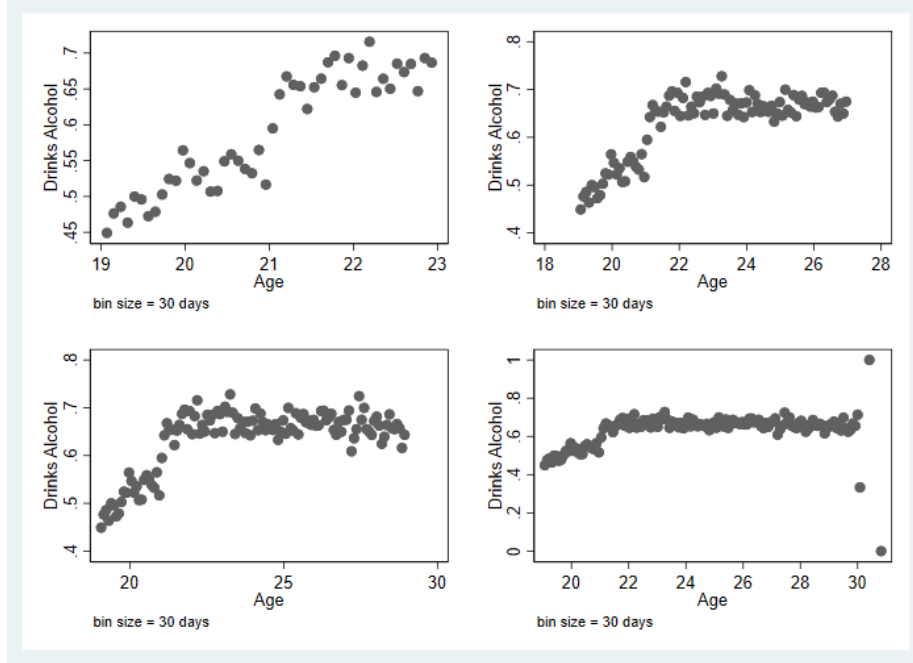


Figure 3, shows the age profile along with the percentage of days on which they report drinking. Superimposed onto the graph are three polynomial regression lines. A linear, a quadratic, and a cubic regression line. I believe the line of best fit is the linear regression line. I chose the linear regression line because it seems to be the cleanest which makes it easier to read the discontinuity, although the quadratic and cubic regression lines seem to be more specific.

Heteroscedasticity can be explained by a systematic change in the spread of the residuals over the range of measured values. This means that there is unequal scatter in the regression figures which causes a higher error term. Heteroscedasticity is an issue when running regressions because it tends to make the error term inflated and overestimated. In order to eliminate the chance of an exaggerated error term due to heteroscedasticity, we made all the standard errors in Table 1, 2, and 3 robust. Using robust standard errors basically compresses the estimator of variance in order to obtain unbiased standard errors.

Figure 3: Age Profile of Drinking



## 4 Results

### 4.1 Regression Table Results

In order to get rid of possible bias due to difference in the age groups, those prior to turning 21, and those post to turning 21. This bias can be eliminated by showing how small the differences in sample characteristics are between the age groups. The sample characteristics include high school diploma (*HS\_diploma*), hispanic (*hispanic*), white (*white*), black (*black*), uninsured (*uninsured*), employment status (*employed*), marital status (*married*), working\_lw, if they are attending school (*going\_school*), and male (*male*). The two key variables that will be used to regress using the sample characteristics are, how many days they are away from turning 21 (*days\_21*), and if they report drinking alcohol (*drinks\_alcohol*). A balance regression table was made (Table 1) which shows the decimal estimates of the sample characteristic covariates and how similar they are for the different age groups. Looking at the first covariate, *uninsured*,



we can see that there is a 1.8% decrease in the amount of people insured as they turn 21 compared to those prior to turning 21. This difference is not significant when it comes to such a large sample size. The second covariate I will mention is *employed*, which shows that there is a 0.7% increase in the amount of people that are employed once they are over 21 compared to those under 21. This is expected because most people are still in college before turning 21 and get a job once they are out. This difference is not significant enough to matter. The third and final covariate I will mention is *married*. This is the only covariate that has a significant difference because it is statistically significant at the 99% significance level. The 3% increase in the number of people who get married after turning 21 is expected because the older someone gets, the more likely they are to get married. After previously analyzing the estimates in Table 1, I was able to come to the conclusion that there are no significant differences in the sample characteristic covariates between the group over 21 and the group under 21, besides their marital status.

We must take into account the one issue about performing a regression balance table on this many covariates. That is the issue of multiple inference. Multiple inference is when multiple significance tests are performed at once. The issue with this is that there is a chance that at least one covariate results in a statistically significant difference between the two groups. When performing multiple inference, the probability that one or more covariates are significant is a 40.1%. The large number of significance test being performed means we must hold our individual tests to higher standard in order to limit the issue of multiple inference. Due to multiple inference, some significance will occur in the balance table.

Table 2, Effect of Turning 21 on Types of Arrests, shows the change in the different type of arrests as a person turns 21. The key here is to observe the differences in estimates for the arrests directly related to alcohol compared to the arrests indirectly related to alcohol. To begin, the regression table shows that there was a decrease of 71.33 liquor law violation arrests post to turning

21. There is also an increase of about 31.74 arrests related to drunk self harm post to turning 21. Lastly we can see an increase of about 51.47 driving under the influence arrests post to turning 21. We cannot determine if that is a large amount without comparing it to the arrest that are indirectly related to alcohol. When looking at the estimates for non-alcohol related arrests, we can see the estimates are not nearly as large as arrests related to alcohol. It was estimated that there is an increase of about 2.047 arrests that come from robbery. There is an increase of about 5.551 arrests for simple assault and an increase of 3.777 arrests for aggravated assault. Lastly there is an increase of about 2.169 arrests related to disorderly conduct or vagrancy.

Table 2: **Effect of Turning 21 on Types of Arrests**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	all	dui	liquor_laws	drunk_risk	robbery	combined_oth	ot_assault	aggravated_assault
Over21	77.38*** (5.851)	51.47*** (1.595)	-71.33*** (0.528)	31.74*** (1.986)	2.047*** (0.279)	2.169*** (0.255)	5.551*** (0.526)	3.777*** (0.503)
Age-21	-24.71*** (1.984)	30.01*** (0.523)	-18.11*** (0.458)	4.380*** (0.468)	-5.368*** (0.177)	-2.524*** (0.142)	-1.017*** (0.261)	1.567*** (0.289)
(Age-21)*Over21	-48.99*** (4.573)	-26.54*** (1.242)	14.65*** (0.476)	-11.96*** (1.527)	1.858*** (0.243)	0.924*** (0.215)	-1.998*** (0.435)	-1.965*** (0.435)
Constant	1559.0*** (2.112)	195.9*** (0.589)	86.34*** (0.503)	104.4*** (0.515)	23.72*** (0.192)	14.78*** (0.161)	54.94*** (0.287)	65.31*** (0.317)
Observations	1460	1460	1460	1460	1460	1460	1460	1460

Robust standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

As for the arrests not related to alcohol from Table 2, it is clear that the estimates are much smaller. Arrests related to committing robbery see only an increase of about 2 arrests which is not very significant. Arrests related to simple and aggravated assault saw about a 5.5 increase and a 3.8 increase

respectively. As for the last non-alcohol related arrest, disorderly conduct or vagrancy, there is roughly a 2.2 increase in arrests. All the increases of the arrests that are unrelated to alcohol are significantly smaller than the increase in the arrests that are directly related to alcohol. This proves that, not only is there is a large increase in arrests related to alcohol as people turn 21, but there is a significantly larger increase in alcohol related arrests compared to those that are not related to alcohol.

Table 3: **Estimates of Change in Drinking**

	(1)	(2)	(3)	(4)	(5)	(6)
	People Who Drink Alcohol	People Who Drink Alcohol	People Who Drink Alcohol	People Who Drink Alcohol	People Who Drink Alcohol	People Who Drink Alcohol
Over21	0.0866*** (0.0142)	0.0908*** (0.0215)	0.0815** (0.0291)	0.0866*** (0.0143)	0.0901*** (0.0217)	0.0794** (0.0295)
Age-21	0.0439*** (0.00922)	-0.0236 (0.0369)	-0.0509 (0.0926)	0.0439*** (0.00922)	-0.0236 (0.0369)	-0.0509 (0.0926)
(Age-21)*Over21	-0.0243* (0.0124)	0.0969 (0.0497)	0.207 (0.125)	-0.0243 (0.0124)	0.0982* (0.0500)	0.215 (0.126)
(Age-21) <sup>2</sup>		-0.0340 (0.0180)	-0.0681 (0.107)		-0.0340 (0.0180)	-0.0681 (0.107)
(Age-21) <sup>2</sup> * Over21		0.00727 (0.0241)	-0.0618 (0.145)		0.00673 (0.0242)	-0.0695 (0.146)
(Age-21) <sup>3</sup>			-0.0114 (0.0354)			-0.0114 (0.0354)
(Age-21) <sup>3</sup> * Over21			0.0457 (0.0476)			0.0479 (0.0479)
Birthday				0.00192 (0.0825)	0.0206 (0.0833)	0.0360 (0.0845)
Constant	0.559*** (0.0105)	0.536*** (0.0158)	0.532*** (0.0213)	0.559*** (0.0105)	0.536*** (0.0158)	0.532*** (0.0213)
Observations	18801	18801	18801	18801	18801	18801

Robust standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 3, Estimates of Change in Drinking, is a table that shows the estimates of the increase in the proportion of people who drink, along with different order polynomials. The table shows that at the cutoff which is the *Over21* variable there is an 8.66 percentage point increase in the amount of people who drink after

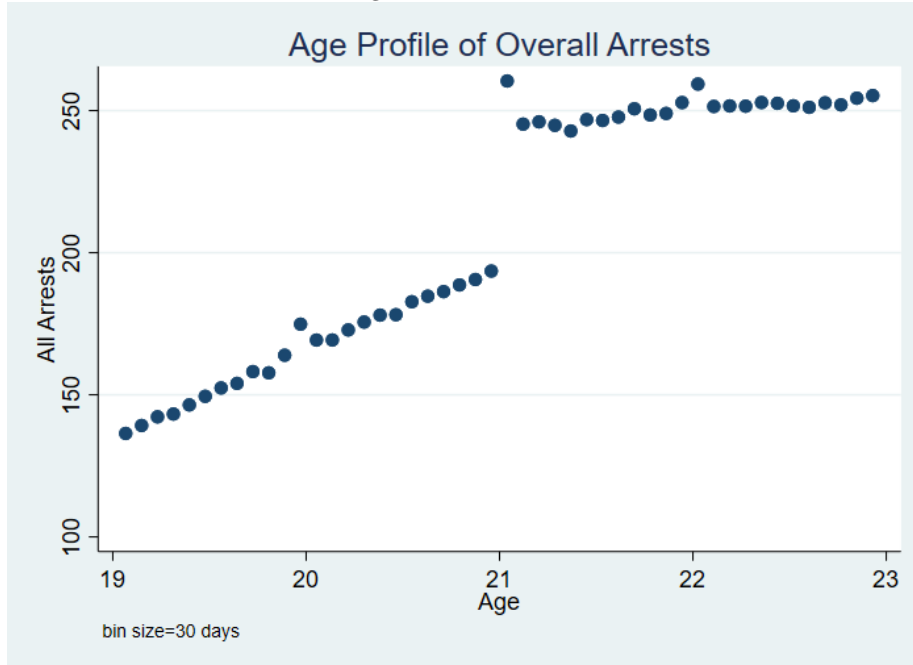
they turn 21. As we add more flexibility with the different order polynomials, we decided the best fit is the linear fit. As we added the variable *Birthday* to the regression, the only estimates that stay the same are the estimates for the linear fit, which stands at 8.66 percentage points increase, in the amount of people who drink alcohol, for both.

## 4.2 Regression Discontinuity Figures Results

The regression discontinuity figures (Figures 5-10), show the discontinuity in the different types of arrests as people turn 21. The different arrests include; all arrests (*all*), drunk possible risk to self (*drunk\_risk*), driving under the influence (*dui*), violation of the liquor laws (*liquor\_laws*), disorderly conduct or vagrancy (*combined\_oth*), robbery (*robbery*), simple assault (*ot\_assault*), and aggravated assault (*aggravated\_assault*). Looking over the regression discontinuity figures, I found a difference in the amount of arrests that occur as people turn 21. The discontinuities for the regression figures on arrests that are directly related to alcohol are clearly much larger than the discontinuities on the regression figures on arrests indirectly related to alcohol just by analyzing it visually. As previously discussed, Table 2 shows the exact estimates of the discontinuity which explains the larger increase of arrests related to alcohol as opposed to those unrelated to alcohol.

Figure 3 shows a clear regression discontinuity that starts at about 55% to about 65% of people who drink alcohol. This shows that there is roughly a 10% increase in the amount of people who drink after turning 21. This helps understand that people drink more as they turn 21 because it is now legal.

Figure 4: All Arrests



We will start off by looking at the regression discontinuity graph on alcohol related arrests. These include Figure 5, Figure 6, and Figure 7. To start with Figure 5, the black dotted regression line signifies amount of arrests related to driving under the influence. We can see a clear discontinuity of about roughly 50 arrests. This shows that as a group of people turn 21, there are about 51 more arrests. This might be due to the fact that since they can drinking legally now, they believe there is less consequences from being arrested for driving under the influence. In Figure 6, The black dotted regression line shows that there is a decrease of roughly 70 arrests that relate to liquor law violation because turning 21 allows people to buy alcohol legally. Now if we look at Figure 7, we can see a clear regression discontinuity on the drunk risk to self harm. The discontinuity shows an increase of roughly 30 arrests related to self harm when drinking. All these estimates coincide with the regression table on the estimates of drinking on types of arrests (Table 2).

Figure 5: Driving Under the Influence

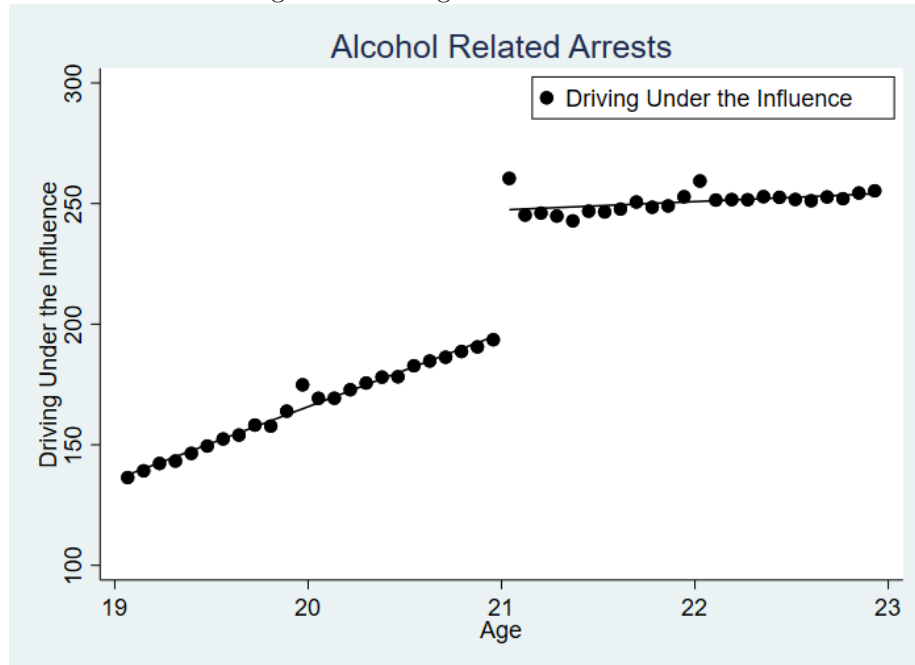


Figure 6: Violation of Liquor Law

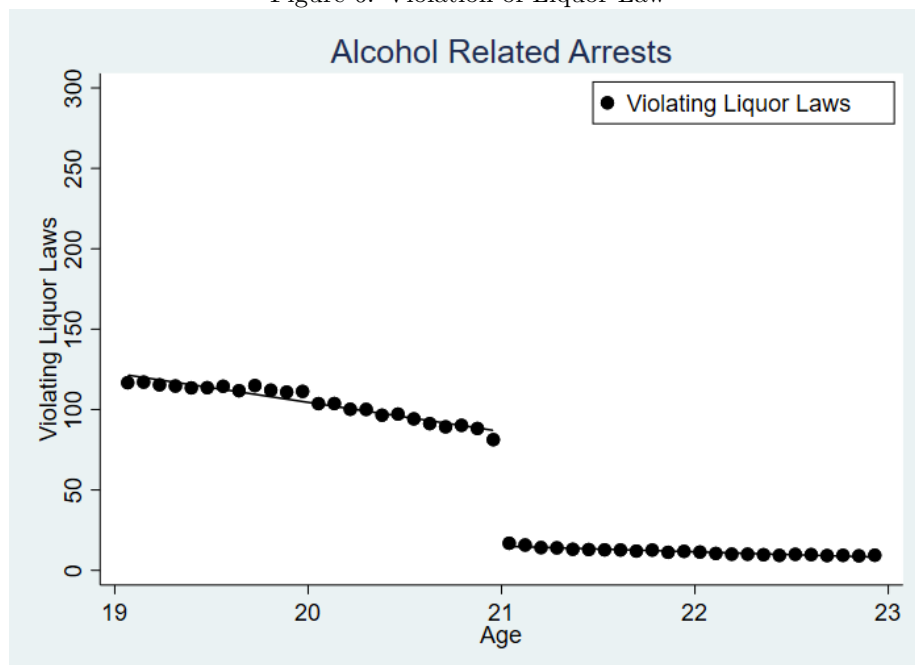
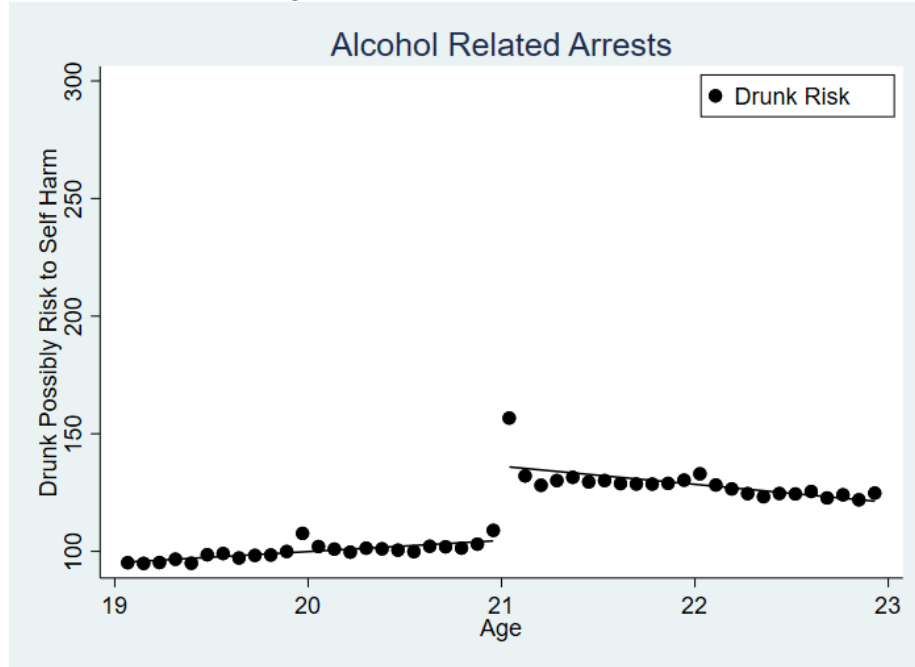


Figure 7: Drunk Risk to Self Harm



Now that we have analyzed the regression discontinuities of arrests directly related to alcohol, we must analyze the regression discontinuities that are indirectly related to alcohol in order to show the difference in magnitude. Figure 8 shows the regression discontinuity on two types of assault, simple and aggravated assault. When looking at the red regression discontinuity, we can see a small discontinuity for aggravated assault. The discontinuity is about 3.8 which signifies an increase of about 3.8 arrests post turning 21 which is not very large. When looking at the black regression line, which is simple assault arrests, we can see a small discontinuity which signifies an increase of about 5.6 arrests post turning 21. Another type of arrest that we will analyze is disorderly conduct or vagrancy. Figure 9 shows the regression discontinuity that was performed on arrests unrelated to alcohol, more specifically disorderly conduct or vagrancy. The discontinuity in Figure 9 is very small, it is about 2.2 arrests. This signifies that there is an increase of about 2.2 arrest related to disorderly conduct as people turn 21. This is too small of an estimate to be significant. The last arrest

that is not related to alcohol is robbery. Figure 10 is a regression discontinuity that shows the change in arrests related to robbery post turning 21. We can see little to no discontinuity in this figure, to be specific, is about an increase of about 2 arrests related to robbery. However, if we look at the regression line, the amount of arrests related to robbery overall decrease after turning 21. This can most likely be explained by the fact that as people get older, they have more money which shows no need to commit robbery.

Figure 8: Assault

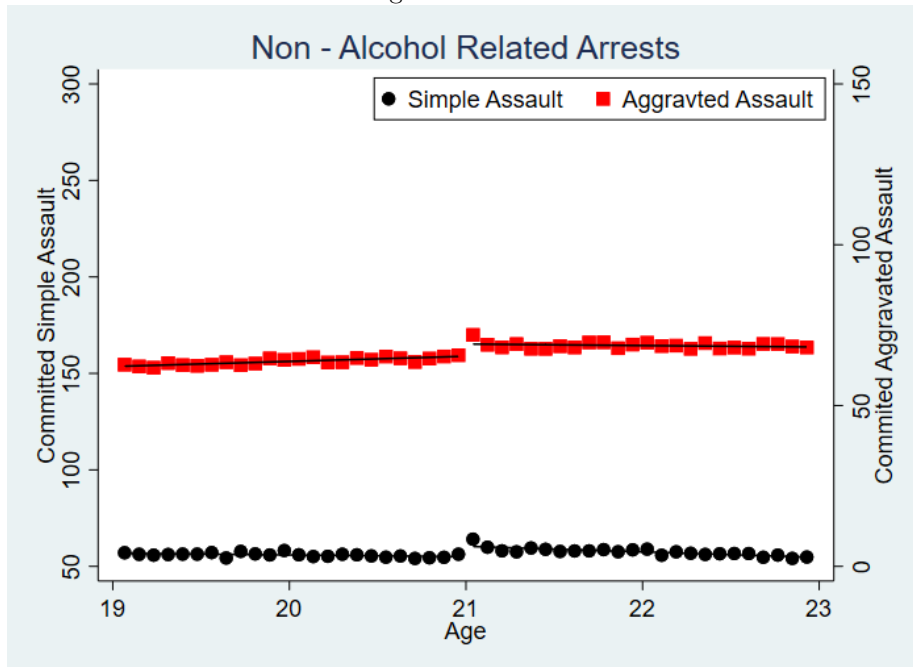




Figure 9: Disorderly Conduct

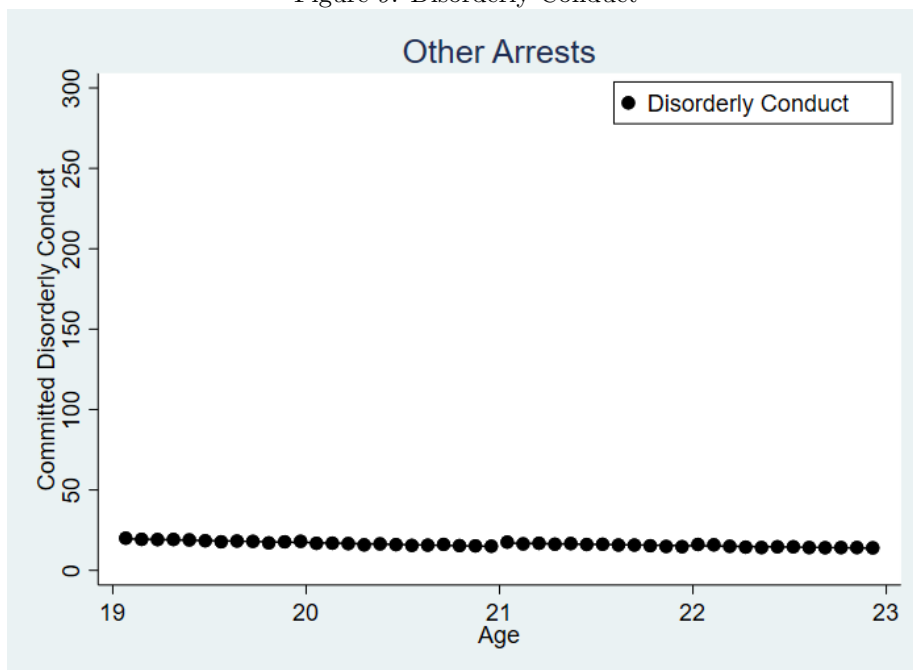
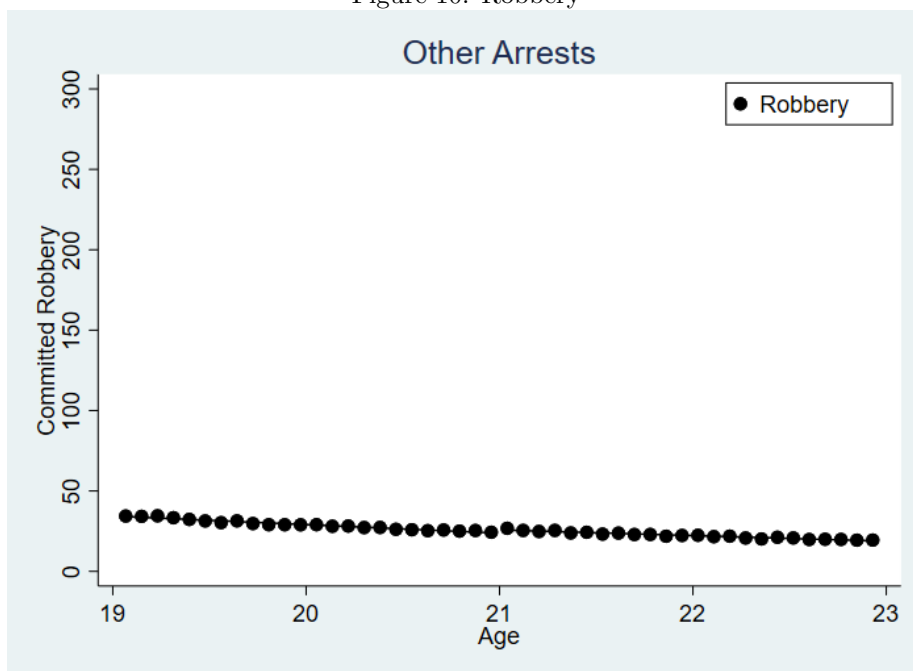


Figure 10: Robbery



## 5 Conclusion

A regression table including the instrumental variable estimates of turning 21 on the different types of arrests. The instrumental variable regression is made to find the effect of turning 21 has on each of the types of arrests while controlling for confounding variables and measurement error. As seen in Table 4, IV Estimates of Turning 21 on Types of Arrests, we can see that the three largest changes in arrests were the arrests directly related to alcohol. Driving under the influence saw an increase of about 595.6 arrests for every 10,000 people. Violation of liquor laws saw a decrease of about 828.32 arrests for every 10,000 people. As for the arrests unrelated to alcohol, the changes in arrests are much smaller. Robbery saw an increase of about 23.7 arrests per 10,000 people. Disorderly conduct saw an increase of about 25.1 arrests for every 10,000 people. Simple and aggravated assault saw an increase of roughly 64 arrests and 43.6 arrests respectively for every 10,000 people. The much larger differences in arrests for alcohol related arrests compared to non-alcohol related arrests show that turning 21 has a much larger effect on alcohol related arrests. Very importantly, all arrests saw the highest increase of about 891.3 arrests per 10,000 people. This is good evidence that the MLDA has a strong effect on crime rates of all arrests.

Table 4: **IV Estimates of Turning 21 on Types of Arrests**

	All Arrests	DUI	Liquor Laws	Robbery	Disorderly Conduct	Simple Assault	Aggravated Assault
IV Coefficient	891.304	595.596	-828.32	23.739	25.102	63.992	43.557
IV Standard Error	154.381	98.631	135.901	5.107	4.971	11.801	9.1
IV T-Statistic	5.773	6.039	-6.095	4.649	5.05	5.432	5.423

To conclude, the results that were found made it able to deduce a few conclusions. To begin, as a whole, the minimum legal drinking age has a clear effect on crime rates and alcohol consumption. With the analysis of Table 3, we were

able to find evidence that the MLDA, which in this case is age 21, increases alcohol consumption. Table 2, Effect of Turning 21 on Types of Arrests, was able to show that the MLDA increases crime rates. Using the IV estimates, It was found that there is a large increase in arrests related to alcohol post to turning 21. Since turning 21 increases the arrests related to alcohol, we are able to say that decreasing the MLDA would most likely increase crime rates even further. However for the IV estimates to hold, the three IV assumptions must be met. The first being that the reduced form and the first stage estimates are consistent, this means that there is good randomization assignment. The first assumption was previously proved using Table 1. The second assumption is that the first stage estimate cannot be zero, which in this case none of the first stage estimates are zero. The third and final is that the instrument Z, which in this case is the minimum legal drinking age, has no direct impact on any other variables that could affect our Y outcome, which is alcohol consumption and crime rates. In our case, the instrument does in fact have a direct effect on the outcome. The instrument Z in this case is the MLDA and the outcome is alcohol consumption and crime rates. It is quite clear that the MLDA has a direct effect on alcohol consumption and crime rates because of the estimates. This means that our IV estimates are not consistent. Therefore it is quite possible that reducing the MLDA would increase alcohol consumption and crime rates, but due to the third IV assumption not being met, we cannot directly state the effect to be entirely true.

## 6 References

Carpenter, Christopher, and Carlos Dobkin. “THE MINIMUM LEGAL DRINKING AGE AND CRIME.” *The Review of Economics and Statistics*, vol. 97, no. 2, The MIT Press, 2015, pp. 521–24, <http://www.jstor.org/stable/43556190>.