

# Lecture 9 - Regression Discontinuity Examples

Samuel DeWitt

# Regression Discontinuity Examples

- 1) Yoruk and Yoruk (2011) - Minimum Legal Drinking Age (MLDA)
- 2) Crost and Rees (2013) - New Estimates
- 3) Yoruk and Yoruk (2013) - Revision

## Why is 21 the legal drinking age in the US?

- ▶ Most states voluntarily set age at 21 until the 1970s
- ▶ Some states lowered the MLDA during the 1970s to as low as 18
- ▶ Drunk driving and traffic fatalities became prominent issues in early 1980s
- ▶ Reagan signs National Minimum Drinking Age Act in 1984

# What is the Effect of MLDA?

- ▶ We *think*...:
  - Higher MLDA reduces drunk driving and traffic fatalities
  - Higher MLDA reduces teen drinking (but does not preclude it)
  - Higher MLDA reduces teen smoking (tobacco and marijuana)
- ▶ We *know* that...
  - These impacts were largely driven by early-adoption states of 21 as MLDA

# Complications for Causal Effects

- ▶ How we try to estimate causal effects of laws
  - We assume individual-level behavior is not a cause of law, but a consequence
  - This can result in **exogeneity** in the relationship between law and behavior
  - We can then capitalize on cross-state differences in the type and timing MLDA laws
- ▶ Is MLDA exogenous?
  - Probably not - Law is not unrelated to teen drinking behavior

# Does MLDA Have a True Causal Effect on Smoking?

Yoruk and Yoruk (2011)

- ▶ Use age as a running variable to estimate within-person causal effect
  - I.e., what is the effect of being able to consume alcohol legally on smoking behaviors?
- ▶ Measure of age
  - Obtained exact birth and interview dates (public data gives you month at best)
  - Calculated the numbers of days pre/post turning 21 as of the date of the latest NLSY97 interview

## Does MLDA Have a True Causal Effect on Smoking? (cont)

- ▶ Age bandwidths (multiple)
  - 480 days
  - 240 days
  - 120 days
  - 60 days
  - 30 days
- ▶ Estimation
  - Local linear regressions above/below bandwidth
    - ▶ This just means the regressions are estimated right below and above the cutoffs according to the chosen bandwidths.

# Yoruk and Yoruk (2011) - Covariate Balance

**Table 1**

Test of the smoothness of the control variables around the 21st birthday.

	Outcome											
	HS Grad.	GED	Some college	College	Graduate	Student	ln(income)	Black	Hispanic	Female	Employed	Married
<i>T</i>	-0.028 (0.017)	-0.011 (0.008)	-0.003 (0.006)	0.002 (0.009)	-0.002 (0.002)	0.015 (0.018)	0.055 (0.076)	-0.009 (0.008)	-0.001 (0.008)	0.014 (0.015)	0.019 (0.016)	0.000 (0.012)
Constant	0.602 (0.012)***	0.078 (0.006)***	0.034 (0.004)***	0.067 (0.007)***	0.004 (0.001)***	0.4000 (0.013)***	10.349 (0.054)	0.153 (0.006)***	0.133 (0.006)***	0.485 (0.011)***	0.678 (0.011)***	0.121 (0.008)***
No. of obs.	27,698	27,698	27,698	27,698	27,698	2769	19,939	29,527	29,527	29,527	29,526	26,877
<i>R</i> <sup>2</sup>	0.0003	0.0003	0.0060	0.0016	0.0004	0.0026	0.0005	0.0001	0.0001	0.0001	0.0081	0.0009

Notes: *T* is a binary treatment variable which is equal to one if the respondent is at least 21 years old. All regressions include age and age-squared and their interactions with the treatment variable as discussed in the text. Sample weights are used in all regressions. The sign \*\*\* indicates the statistical significance at the 1 percent significance level, respectively. Robust standard errors clustered at the individual level are reported in parenthesis.



## Yoruk and Yoruk (2011) - Covariate Balance

Just a slide for listing what I say about Table 1

- ▶ Ignore significance of constant coefficient - this is mainly indicating a consistent estimate of where the line crosses the y-axis
- ▶ Most importantly - no covariates appear to be imbalanced around the 21st birthday cutoff
- ▶ This is largely expected - birthdate is not something people will try to manipulate for the purposes of a survey
- ▶ However, this does not imply that youth do not try to manipulate their apparent age to get access to alcohol!

# Yoruk and Yoruk (2011) - Effect of MLDA on Alcohol Consumption

**Table 2**

The effect of the MLDA on alcohol consumption.

	Outcome											
	Alcohol		Binge		DaysAlcohol		DaysBinge		NofDrinks		AvgDrinks	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>T</i>	0.058	0.049	0.057	0.052	1.500	1.684	0.355	0.509	0.088	−0.065	0.192	0.255
	(0.013)***	(0.015)***	(0.021)***	(0.023)**	(0.270)***	(0.292)***	(0.214)*	(0.228)**	(0.166)	(0.187)	(0.101)*	(0.108)**
Constant	0.853	0.783	0.592	0.558	5.655	5.209	3.018	3.092	4.349	4.185	1.212	1.150
	(0.011)***	(0.018)***	(0.016)***	(0.026)***	(0.190)***	(0.312)***	(0.152)***	(0.250)***	(0.120)***	(0.201)***	(0.072)***	(0.115)***
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of obs.	20,477	20,477	17,469	17,469	20,477	20,477	17,469	17,469	17,255	17,255	17,210	17,220
$R^2$	0.007	0.028	0.001	0.065	0.001	0.059	0.001	0.065	0.004	0.087	0.001	0.053
<i>Local linear</i>												
<i>T</i>	0.076		0.073		1.358		0.409		0.043		0.223	
	(0.017)***		(0.027)***		(0.352)***		(0.277)		(0.288)		(0.138)*	

## Yoruk and Yoruk (2011) - Effect of MLDA on Alcohol Consumption

Another slide for listing what I say about Table 2

- ▶ Underlying question being tested is if the MLDA has an effect on alcohol use above/below the cutoffs with bandwidth of 240 days.
- ▶ Results are about what we would expect:
  - Use of alcohol (binary) increases
  - Binge drinking (binary) also increases
  - Days using alcohol increases along with days of binge drinking
  - No effect on number of total drinks consumed, though

# Yoruk and Yoruk (2011) - Effect of MLDA on Smoking Habits

**Table 3**

The effect of the MLDA on smoking and marijuana use.

	Outcome											
	Smoking		Marijuana		DaysSmoking		DaysMarijuana		NofCigarettes		AvgCigarettes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>T</i>	0.009 (0.019)	0.002 (0.020)	0.066 (0.031)**	0.073 (0.034)**	0.409 (0.648)	0.347 (0.672)	-1.053 (0.738)	-0.645 (0.786)	0.095 (0.483)	0.034 (0.511)	0.149 (0.498)	0.053 (0.524)
Constant	0.857 (0.014)***	0.759 (0.023)***	0.722 (0.023)***	0.631 (0.039)***	19.031 (0.468)***	13.531 (0.814)***	9.576 (0.545)***	6.880 (1.008)***	9.432 (0.353)***	7.445 (0.636)***	8.631 (0.363)***	6.412 (0.659)***
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of obs.	12,700	12,700	6975	6975	12,700	12,700	6975	6975	10,748	10,748	10,672	10,672
<i>R</i> <sup>2</sup>	0.001	0.034	0.002	0.023	0.003	0.126	0.001	0.001	0.001	0.112	0.001	0.120
<i>Local linear</i>												
<i>T</i>	0.027 (0.023)		0.048 (0.040)		0.565 (0.839)		-1.409 (0.991)		-0.289 (0.647)		-0.241 (0.681)	

## Yoruk and Yoruk (2011) - Effect of MLDA on Smoking Habits

Another slide for listing what I say about Table 3

- ▶ Question being tested is whether legal access to alcohol also affects smoking (tobacco/marijuana) habits
- ▶ Results are a little strange
  - No effect on probability of smoking tobacco or amount, for that matter
  - About a 7% increase in the probability of using marijuana

Seems to contradict anecdotal evidence re: amplification of tobacco use and the alternative interpretation of marijuana being the *gateway* drug.

## Crost and Rees (2013) - Why It Is Important to Read Data Codebooks Carefully

The evidence from Yoruk and Yoruk (2011) was somewhat surprising, but Crost and Rees (2013) provide a good explanation for why the original results were somewhat counterintuitive.

- ▶ Some markers for why it looked like something was wrong:
  - Reported marijuana use in the past 30 days in Yoruk and Yoruk was **very** high (76%!!!)
  - Sample drops precipitously in marijuana discontinuity model (about 20,000 observations lost)

From this, Crost and Rees (2013) surmise that Yoruk and Yoruk (2011) ignored an important lead-in question. In fact, they accidentally restricted their sample only to those that used marijuana since the last survey!

# Crost and Rees (2013) - Corrected Findings

**Table 1**

Comparison of conditional and unconditional results.

	Dependent variable: used marijuana in past 30 days			
	(1) Conditional on use in past year	(2)	(3)	(4) Unconditional
Age>21 ( $\tau_{RD}$ )	0.069* (0.040)	0.075* (0.040)	0.011 (0.019)	0.017 (0.019)
Age	0.003 (0.062)	-0.004 (0.062)	-0.021 (0.030)	-0.022 (0.029)
Age <sup>2</sup>	-0.004 (0.029)	-0.007 (0.029)	-0.004 (0.014)	-0.005 (0.014)
Age $\times$ Age>21	-0.120 (0.090)	-0.112 (0.089)	-0.016 (0.041)	-0.018 (0.040)
Age <sup>2</sup> $\times$ Age>21	0.051 (0.043)	0.060 (0.043)	0.015 (0.019)	0.020 (0.018)
Constant	0.755*** (0.028)	0.835*** (0.034)	0.190*** (0.014)	0.254*** (0.018)
Mean	0.757 (0.006)	0.757 (0.006)	0.191 (0.003)	0.191 (0.003)
Controls	No	Yes	No	Yes
Observations	6551	6551	28,089	28,089

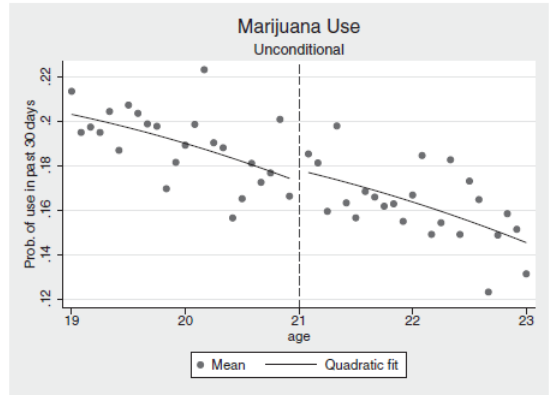
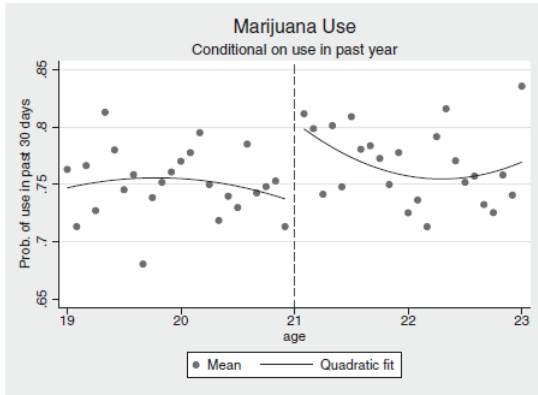
## Crost and Rees (2013) - Corrected Findings

Another slide for listing what I say about Table 1 from Crost and Rees (2013)

- ▶ Conditional estimates are only for youth who **have** used marijuana since the last interview
- ▶ Unconditional estimates are for all youth regardless of prior marijuana use
- ▶ What changes?
  - The 6-7% increase in marijuana use in the past 30 days disappears!



## Crost and Rees (2013) - Corrected Findings (Figure)



## Crost and Rees (2013) - Corrected Findings (Figure)

So, are alcohol and marijuana use interdependent?

- ▶ Kind of...but not in the way Yoruk and Yoruk thought of it
  - MLDA does not have an impact on the **probability** of using marijuana at all
  - MLDA does **look like** it may have an impact on the recency of using marijuana **if** you have used marijuana before

## Yoruk and Yoruk (2013) - Revised Estimates

Yoruk and Yoruk (2013) is a response to Crost and Rees (2013) and re-reports revised findings for all of the hypotheses tested in their original study.

- ▶ As an aside, I chose this series of papers because:
  - I figured you might be interested in MLDA laws given your average ages
  - It's a good example of a collegial back and forth where incorrect scientific evidence is corrected (this is an example of science working properly!)
  - A reminder not to take scientific evidence at face value but also not to dismiss it entirely.
    - ▶ It's important to be skeptical of **surprising** findings - sometimes there's a good reason for this

# Yoruk and Yoruk (2013) - Revised Estimates

**Table 1**

The effect of the MLDA on alcohol consumption, smoking, and marijuana use.

	Y&EY (restricted sample)			Full sample		
	No. of obs.	Mean	$\delta$	No. of obs.	Mean	$\delta$
Alcohol	20477	0.869	0.049 (0.015)***	29111	0.652	0.103 (0.016)***
DaysAlcohol	20477	6.091	1.684 (0.292)***	29111	4.566	1.739 (0.240)***
NofDrinks	17255	4.601	-0.065 (0.187)	28885	2.967	0.484 (0.144)***
Binge	17469	0.610	0.052 (0.023)**	28971	0.397	0.094 (0.018)***
DaysBinge	17469	3.121	0.509 (0.228)**	28971	2.029	0.642 (0.161)***
AvgDrinks	17210	1.305	0.255 (0.108)**	28839	0.841	0.295 (0.075)***
Smoking	12700	0.867	0.002 (0.020)	29215	0.406	-0.006 (0.016)
DaysSmoking	12700	19.246	0.347 (0.672)	29215	9.020	0.061 (0.424)
NofCigarettes	10748	9.706	0.034 (0.511)	29222	3.860	0.017 (0.251)
AvgCigarettes	10672	8.935	0.053 (0.524)	29146	3.542	0.033 (0.247)
Marijuana	6975	0.759	0.073 (0.034)**	29510	0.193	0.005 (0.014)
DaysMarijuana	6975	9.516	-0.645 (0.786)	29510	2.422	-0.347 (0.237)

## Yoruk and Yoruk (2013) - Revised Estimates

- ▶ So, a brief conclusion:
  - MLDA has a clear impact on the probability and frequency of alcohol use (as anticipated)
  - But, it does not have an impact on the probability and frequency of smoking or marijuana use in the unconditional samples
    - ▶ Its effect on marijuana use may be conditional though

The End

