Online Appendix: Am I My Brother's Barkeeper? Sibling Spillovers in Alcohol Consumption at the Minimum Legal Drinking Age*

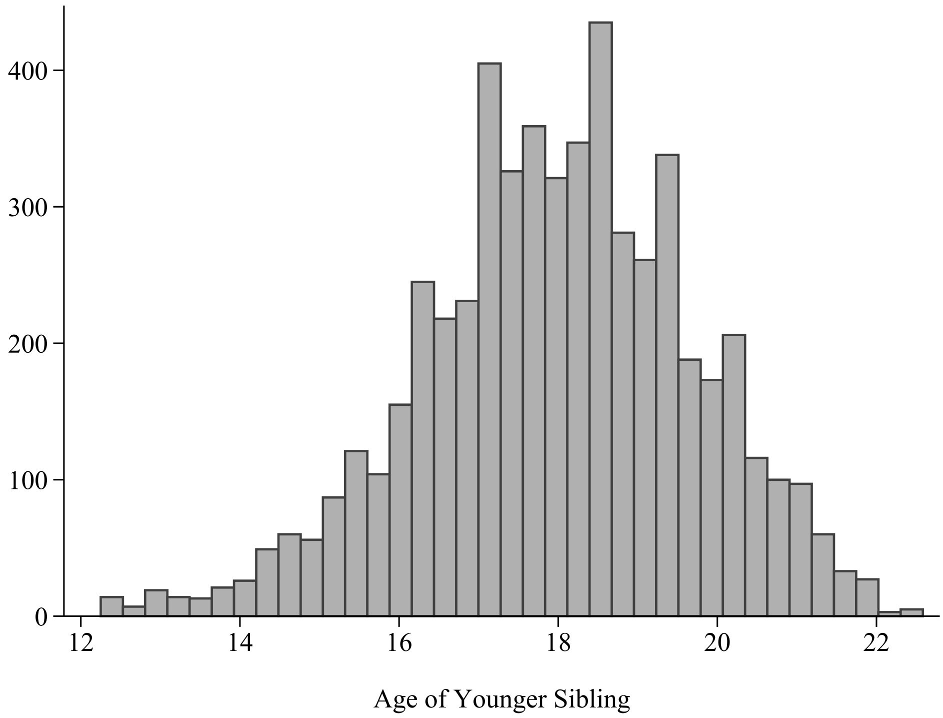
Eunju Lee
Louisiana State University

Geoffrey C. Schnorr
United States Military Academy at West Point

July 10, 2024

A Additional Figures and Tables

Figure A1: Age Distribution of Younger Sibling



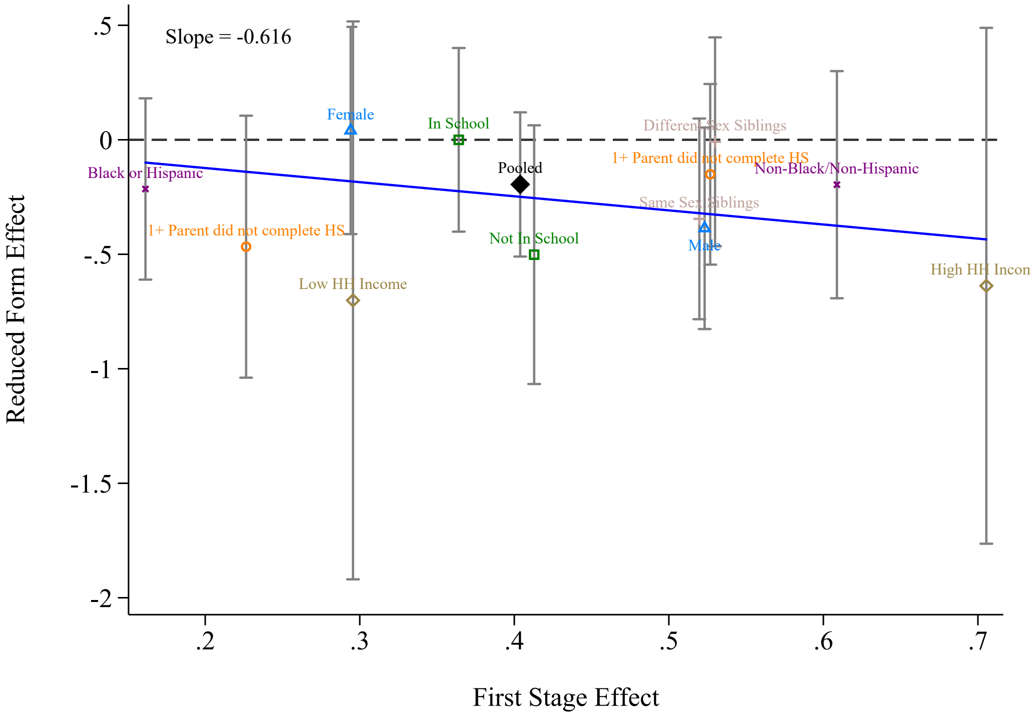
Note: Y-axis shows the count of younger siblings in our sample with each age value (age in months of younger sibling).

*Lee: eunjulee@lsu.edu. Schnorr (corresponding author): gcschnorr@gmail.com.

Figure A2: Density of Running Variable

Note: Y-axis shows the count of sibling-pairs with each running variable value (age in months of older sibling). Point estimates and standard errors are from a regression of the number of observations in each bin on the centered-at-21 running variable, the treatment cutoff, and their interaction (estimated with individual-level data).

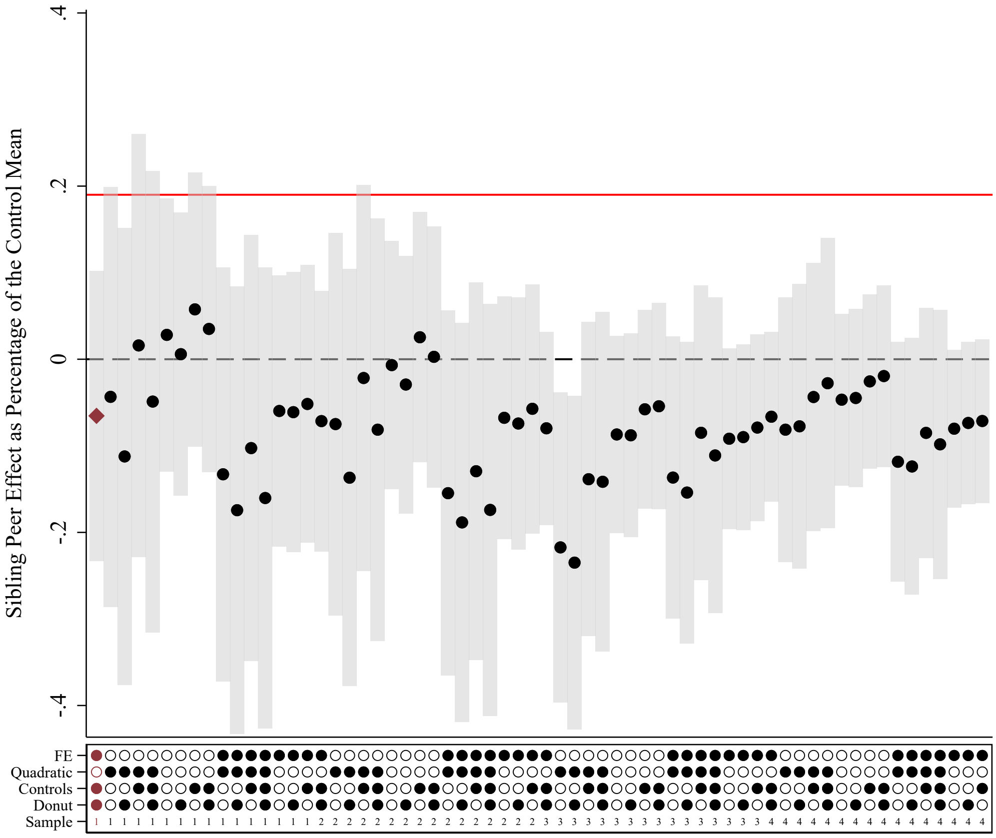
Figure A3: First Stage Effect vs Reduced Form Effect, Count of Binge Drinking Days, Without Any Covariates



Note: This figure plots the subgroup-level first stage effects (estimates from Table A4) against the reduced form effects for the same subgroups (estimates from Table A5). The "pooled" estimate in both plots is from the specification without any covariates in the full samples (Table 2 and Table 3). Whiskers mark 95% confidence intervals.

Figure A4: Robustness to Alternative Specifications

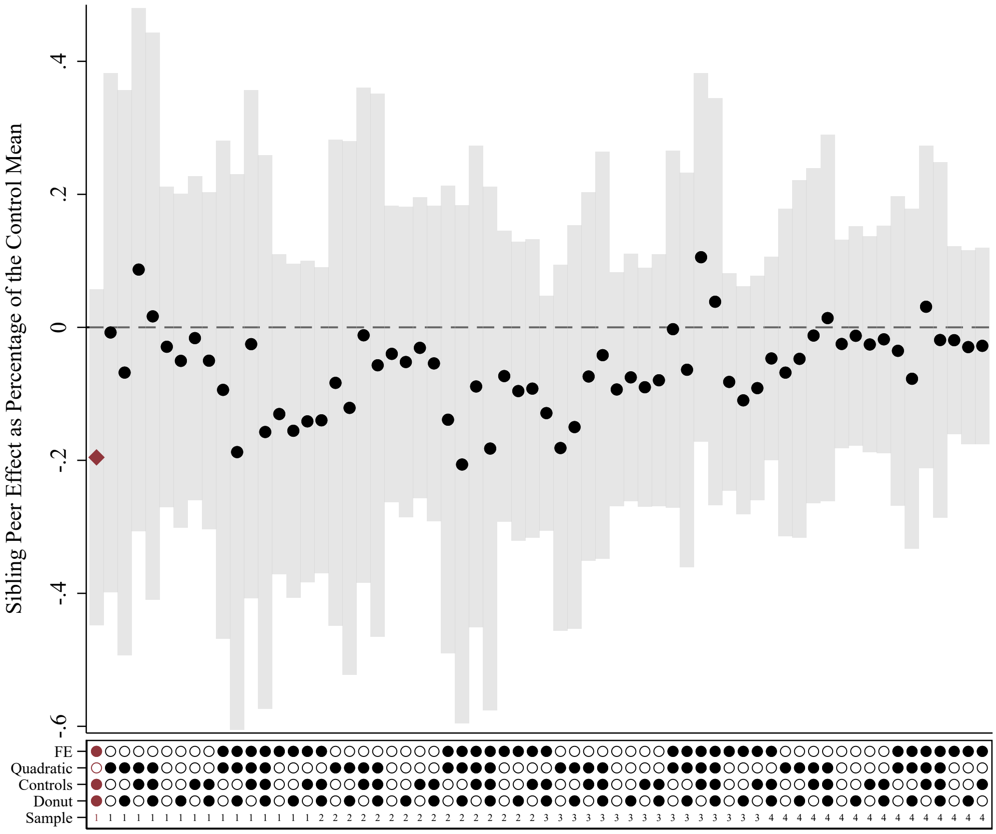
(Outcome = Prob. of Any Binge Drinking)



Note: Figure shows point estimates and confidence intervals from various specifications, all with an indicator for any binge days as the outcome, and all scaled by the pre-cutoff mean (constant in the corresponding regression on the centered-at-zero running variable, the treatment cutoff, and their interaction). The red line shows the point estimate from Eisenberg et al. (2014) for comparison. "Sample" refers to different subsets of sibling pairs, defined by (i) whether they live together, and (ii) whether they are the oldest two siblings in the household (or family). 1 = same household, oldest two siblings in the family. 2 = same household, the peer is closest older sibling but not necessarily the only older sibling. 3 = siblings do not necessarily live together, siblings are the oldest two siblings in the family. 4 = siblings do not necessarily live together, the peer is the closest older sibling but not necessarily the only older sibling.

Figure A5: Robustness to Alternative Specifications

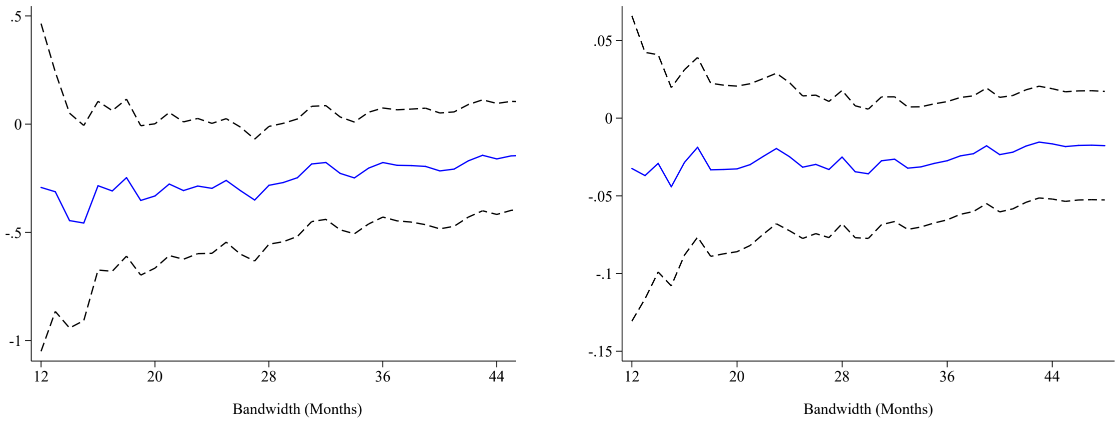
(Outcome = Count of Binge Drinking Days)



Note: Figure shows point estimates and confidence intervals from various specifications, all with the count of binge drinking days as the outcome, and all scaled by the pre-cutoff mean (constant in the corresponding regression on the centered-at-zero running variable, the treatment cutoff, and their interaction). "Sample" refers to different subsets of sibling pairs, defined by (i) whether they live together, and (ii) whether they are the oldest two siblings in the household (or family). 1 = same household, oldest two siblings in family. 2 = same household, peer is closest older sibling but not necessarily the only older sibling. 3 = siblings do not necessarily live together, siblings are the oldest two siblings in the family. 4 = siblings do not necessarily live together, peer is closest older sibling but not necessarily the only older sibling.

Figure A6: Robustness to Different Bandwidths

(a) Count of Binge Drinking Days (b) Prob. of Any Binge Drinking



<u>Note:</u> Figure shows point estimates and confidence intervals from the preferred specification in Table 3 (donut, linear, controls, FE) estimated with different bandwidths (in months).

Table A1: Correlations in Alcohol Consumption Between Siblings

	Younger Sibling Consumption			
	Count of Drinking Days		Any Drinking Days	
Older Sibling Consumption	0.166***	0.145***	0.208***	0.196***
	(0.023)	(0.017)	(0.021)	(0.017)
Mean	2.783	2.976	0.461	0.490
Observations	2,878	4,983	2,878	4,983
	Count of Binge Drinking Days		Any Binge Drinking Days	
Older Sibling Consumption	0.171***	0.135***	0.180***	0.153***
	(0.028)	(0.020)	(0.021)	(0.017)
Mean	1.378	1.450	0.273	0.291
Observations	2,877	4,976	2,877	4,976
Lag		X		X

Note: Each model has the younger sibling's consumption as the outcome and the older sibling's consumption as the covariate of interest. Model 1 uses the contemporaneous consumption of the older sibling by focusing on the sibling pairs whose interview dates are within 14 days. Model 2 uses the lagged (prior survey year) consumption. All models are estimated via OLS with cluster robust standard errors at the household level and are limited to households with exactly two NLSY97 respondent siblings, in which the older sibling is 23 years old or younger. All models include a vector of controls for both siblings similar to those described in Section 2. Means of each outcome (for younger siblings) are shown for each model. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A2: Smoothness of Covariates at Cutoff

	Predicted Outcomes			
	Drinking Days	Any Drinking Days	Binge Drinking Days	Any Binge Binge Days
Sibling 21+	-0.020	-0.005	-0.001	-0.003
	(0.065)	(0.007)	(0.039)	(0.006)
Constant	2.459***	0.445***	1.216***	0.260***
	(0.082)	(0.008)	(0.053)	(0.007)
N	5344	5344	5335	5335

Note: Outcomes are predicted values from a regression of the relevant consumption measure on the month and year of the survey, educational attainment and enrollment of the respondent, geography (urban/rural, census region), household size, an indicator for whether the respondent has children, and an indicator for whether the respondent worked in the past year. Results shown are from regressions of these predicted values on the age of the older sibling in months, an indicator for whether the older sibling is over 21, and their interaction. Standard errors are bootstrapped. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A3: Effect of Younger Sibling on Older Sibling

	Drinking Days	Any Drinking Days	Binge Drinking Days	Any Binge Drinking Days
	Two Youngest Siblings Living in the Same Household			
Sibling 21+	-0.262	-0.018	-0.006	0.046
	(0.410)	(0.031)	(0.247)	(0.032)
Control Mean	4.743	0.672	1.854	0.374
Observations	2,878	2,878	2,820	2,820
	Two Siblings Living in the Same Household			
Sibling 21+	-0.279	-0.002	-0.106	0.037
	(0.286)	(0.024)	(0.195)	(0.025)
Control Mean	4.342	0.615	1.643	0.338
Observations	4,694	4,694	4,616	4,616
	Two Youngest Siblings			
Sibling 21+	0.092	-0.003	-0.011	0.003
	(0.209)	(0.016)	(0.129)	(0.016)
Control Mean	4.691	0.694	1.726	0.383
Observations	8,970	8,970	8,778	8,778
	Two Siblings			
Sibling 21+	-0.066	-0.005	0.002	0.007
	(0.147)	(0.012)	(0.095)	(0.012)
Control Mean	4.461	0.652	1.603	0.358
Observations	16,319	16,319	16,027	16,027

Note: The first panel is estimated in a sample of NLSY97 respondents who have exactly one younger sibling in their household, and where that younger sibling is between the ages of 19 and 23. The second panel is estimated in a sample of NLSY97 respondents who are the second youngest siblings, where their younger sibling is between the ages of 19 and 23, and the younger sibling may not currently reside in the same household as the respondent. The third panel is estimated in a sample of NLSY97 respondents who have one or more younger siblings in their household where the closest younger sibling to the respondent is between the ages of 19 and 23. The last panel is estimated in a sample of NLSY respondents who have one or more younger siblings, where the closest younger sibling to the respondent is between the ages of 19 and 23, and the closest younger sibling may not currently reside in the same household as the respondent. Controls include the month and year of the survey, the age of the respondent, educational attainment and enrollment of respondents, geography (urban/rural, census region), household size, an indicator for whether the respondent has children, and an indicator for whether the respondent worked in the past year. All models include cluster robust standard errors at individual level. Sibling age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A4: Discontinuities in Older Sibling Alcohol Consumption in Subgroups, Binge Days, Without Any Covariates

	Parental Education		Household Income	
	< HS	≥ HS	< Median	≥ Median
Age 21+	0.226	0.527***	0.253*	0.846***
	(0.152)	(0.139)	(0.127)	(0.203)
Control Mean	1.120	1.992	1.513	1.910
Observations	5,535	9,873	10,893	6,479
	Older Sibling's School Enrollment		Older Sibling's Sex	
	Enrolled	Not Enrolled	Male	Female
Age 21+	0.364**	0.413**	0.523***	0.294**
	(0.136)	(0.126)	(0.151)	(0.098)
Control Mean	1.729	1.590	2.360	0.930
Observations	8,017	13,196	10,501	10,767
	Race		Sibling Sex Composition	
	Black or Hispanic	Non-Black & Non-Hispanic	Same	Different
Age 21+	0.161	0.609***	0.520*	0.530*
	(0.115)	(0.133)	(0.202)	(0.206)
Control Mean	1.055	2.136	1.626	1.757
Observations	9,789	11,479	4,478	4,618

Note: All models are estimated in the corresponding subgroup of a sample of NLSY97 respondents who are the oldest siblings currently residing in their household who are between the ages of 19 and 23. The median household income is $51,500. All models include cluster robust standard errors at individual level. Age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A5: Reduced Form in Subgroups, Binge Days, Without Any Covariates

	Parental Education		Household Income	
	< HS	≥ HS	< Median	≥ Median
Sibling 21+	-0.467	-0.151	-0.701	-0.638
	(0.291)	(0.201)	(0.621)	(0.574)
Control Mean	1.128	1.477	0.850	2.383
Observations	1,504	2,693	1,134	1,134
	Older Sibling's School Enrollment		Older Sibling's Sex	
	Enrolled	Not Enrolled	Male	Female
Sibling 21+	-0.000	-0.502+	-0.386+	0.041
	(0.204)	(0.288)	(0.224)	(0.231)
Control Mean	1.259	1.214	1.344	1.143
Observations	2,733	2,759	2,947	2,574
	Race		Sibling Sex Composition	
	Black or Hispanic	Non-Black & Non-Hispanic	Same	Different
Sibling 21+	-0.215	-0.196	-0.345	-0.009
	(0.202)	(0.253)	(0.223)	(0.232)
Control Mean	0.903	1.613	1.348	1.138
Observations	2,784	2,737	3,006	2,515

Note: All models are estimated in the corresponding subgroup of a sample of NLSY97 respondents who have exactly one older sibling in their household, and where that older sibling is between the ages of 19 and 23. The median household income is $51,500 in this sample. All models include cluster robust standard errors at individual level. Sibling age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A6: Discontinuities in Older Sibling Alcohol Consumption in Subgroups, Binge Days, Inverse Propensity Score Weighted

	Parental Education		Household Income	
	< HS	≥ HS	< Median	≥ Median
Age 21+	0.133	0.509***	0.087	1.055**
	(0.179)	(0.146)	(0.126)	(0.391)
Control Mean	1.162	1.759	1.303	1.818
Observations	5,241	9,333	10,359	6,104
	Older Sibling's School Enrollment		Older Sibling's Sex	
	Enrolled	Not Enrolled	Male	Female
Age 21+	0.420**	0.474***	0.587**	0.229*
	(0.141)	(0.135)	(0.178)	(0.109)
Control Mean	1.485	1.597	2.172	0.973
Observations	7,626	12,508	9,894	10,240
	Race		Sibling Sex Composition	
	Black or, Hispanic	Non-Black &, Non-Hispanic	Same	Different
Age 21+	0.175	0.622***	0.481*	0.536*
	(0.129)	(0.143)	(0.215)	(0.218)
Control Mean	0.962	1.962	1.715	1.652
Observations	9,268	10,866	4,257	4,381

Note: Controls include the month and year of the survey, the age of the respondent, educational attainment and enrollment of respondents, geography (urban/rural, census region), household size, an indicator for whether the respondent has children, and an indicator for whether the respondent worked in the past year. The same controls are used to estimate the propensity scores. Variables used to define the subgroups are omitted from the list of controls. All models are estimated in the corresponding subgroup of a sample of NLSY97 respondents who are the oldest siblings currently residing in their household who are between the ages of 19 and 23. The median household income is $51,500. Standard errors in all models are bootstrapped with 1000 replications. Re-sampling is at the younger-sibling level (a given younger sibling typically appears in multiple waves with the same older sibling). Age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A7: Reduced Form in Subgroups, Binge Days, Inverse Propensity Score Weighted

	Parental Education		Household Income	
	< HS	≥ HS	< Median	≥ Median
Sibling 21+	-0.042	-0.216	-0.528	-0.847
	(0.304)	(0.226)	(0.510)	(0.571)
Control Mean	0.931	1.666	0.788	2.117
Observations	1,396	2,566	1,061	1,062
	Older Sibling's School Enrollment		Older Sibling's Sex	
	Enrolled	Not Enrolled	Male	Female
Sibling 21+	-0.025	-0.537^{+}	-0.463*	0.034
	(0.239)	(0.324)	(0.210)	(0.218)
Control Mean	1.591	1.253	1.356	1.138
Observations	2,585	2,608	2,795	2,425
	Race		Sibling Sex Composition	
	Black or Hispanic	Non-Black & Non-Hispanic	Same	Different
Age 21+	-0.264	-0.385	-0.542*	0.095
	(0.206)	(0.321)	(0.240)	(0.220)
Control Mean	0.769	1.712	1.309	1.172
Observations	2,608	2,612	2,848	2,372

Note: Controls include the month and year of the survey, the age of the respondent, educational attainment and enrollment of respondents, geography (urban/rural, census region), household size, an indicator for whether the respondent has children, and an indicator for whether the respondent worked in the past year. The same controls are used to estimate the propensity scores. Variables used to define the subgroups are omitted from the list of controls. All models are estimated in the corresponding subgroup of a sample of NLSY97 respondents who have exactly one older sibling in their household, and where that older sibling is between the ages of 19 and 23. The median household income is $51,500 in this sample. Standard errors in all models are bootstrapped with 1000 replications. Re-sampling is at the younger-sibling level (a given younger sibling typically appears in multiple waves with the same older sibling). Sibling age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table A8: Switcher and Non-Switcher Characteristics

	Switchers mean (sd)	Non-switchers mean (sd)	p-value
Age difference between the siblings	2.81	3.01	0.00
	(1.36)	(1.65)	
Siblings have same gender	0.55	0.53	0.04
	(0.50)	(0.50)	
Younger Sibling's Characteristics:			
Female	0.47	0.48	0.44
	(0.50)	(0.50)	
Race: Black	0.27	0.24	0.04
	(0.44)	(0.43)	
Race: Hispanic	0.24	0.24	0.75
	(0.43)	(0.42)	
Urban	0.78	0.75	0.01
	(0.41)	(0.43)	
AFQT score	46,037.04	42,094.33	0.00
	(29,391.96)	(28,584.32)	
Household Income	$66,654.55	$52,071.82	0.00
	(57,999.39)	(48,741.99)	
N	3716	1859	
N Sibling Pairs	1180	1423	

Note: FE estimates of the effect of an older sibling turning 21 are identified only by groups who have variation in this treatment, i.e., younger siblings who show up in our sample at least once before their older sibling turns 21 and at least once after. This table compares these "switcher" sibling pairs to the corresponding "non-switcher" sibling pairs, in our main analysis sample (younger sibling is a NLSY-respondent, older sibling is between the ages of 19 and 23, siblings are the two oldest siblings in the household). The third column displays the p-value for the null hypothesis that the means in the switcher and non-switcher groups are equivalent.

Table A9: Switcher and Non-Switcher Characteristics in an Alternative Sample

	Switchers mean (sd)	Non-switchers mean (sd)	p-value
Age difference between the siblings	2.65	4.03	0.00
	(1.32)	(2.10)	
Siblings have same gender	0.52	0.53	0.68
	(0.50)	(0.50)	
Younger Sibling's Characteristics:			
Female	0.50	0.47	0.18
	(0.50)	(0.50)	
Race: Black	0.25	0.25	0.93
	(0.43)	(0.43)	
Race: Hispanic	0.20	0.28	0.00
	(0.40)	(0.45)	
Urban	0.76	0.76	0.76
	(0.43)	(0.43)	
AFQT score	45,620.70	40,877.99	0.00
	(28,991.13)	(26,712.87)	
Household Income	$53,699.29	$47,828.96	0.09
	(56,351.16)	(48,448.64)	
N	4240	345	
N Sibling Pairs	1788	278	

Note: FE estimates of the effect of an older sibling turning 21 are identified only by groups who have variation in this treatment, i.e., younger siblings who show up in our sample at least once before their older sibling turns 21 and at least once after. This table compares these "switcher" sibling pais to the corresponding "non-switcher" sibling pairs, in an alternative sample that does not require the siblings to live together (i.e., here the younger sibling is a NLSY-respondent, the older sibling is between the ages of 19 and 23, and the siblings are the two oldest siblings in the family). The third column displays the p-value for the null hypothesis that the means in the switcher and non-switcher groups are equivalent.

Table A10: Robustness to Mean-Squared Error Optimal Bandwidth from Calonico et al. (2020) and Bias-Corrected Robust Confidence Intervals from Calonico et al. (2014, 2018)

	Count of Binge Drinking Days			
	(1)	(2)	(3)	(4)
Sibling 21+	-0.252**	-0.242**	-0.120	-0.131
	[-0.502,-0.059]	[-0.588,-0.039]	[-0.522,0.283]	[-0.621, 0.586]
Bandwidth	19.069	24.000	22.828	24.000
Observations	4151	5024	3931	4124
	Any Binge Drinking Days			
Sibling 21+	-0.029*	-0.024+	-0.000	-0.001
	[-0.059,0.000]	[-0.076,-0.001]	[-0.062,0.074]	[-0.072,0.104]
Bandwidth	18.239	24.000	24.657	24.000
Observations	3939	5024	4278	4124
FE	X	X		
Quadratic				
Controls	X	X	X	X
Donut	X	X	X	X

Note: This table displays specifications using triangular kernels and include bias-corrected robust 95% confidence intervals described in Calonico et al. (2014, 2018). Models 1 and 3 use mean squared error (MSE) optimal bandwidths from Calonico et al. (2020), and the other models use our ad-hoc 24-month bandwidth. Controls include the month and year of the survey, the age of the respondent, educational attainment and enrollment of respondents, geography (urban/rural, census region), household size, an indicator for whether the respondent has children, and an indicator for whether the respondent worked in the past year. Standard errors are cluster robust at individual level. All models use Sibling age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

B Heterogeneity

B.1 Correlation between reduced form and first stage estimates across subgroups

Angrist et al. (2022) show that the exercise we carry out in Section 4.5 can be used to motivate an overidentification test of the null hypothesis that all across-subgroup variation in the treatment effect is driven by variation in the strength of the first stage. Results consistent with such a null hypothesis are also consistent with the exclusion restriction. To see this, consider a model in which the older sibling's legal drinking status $1\{age_{1h} \geq 21\}$ influences their alcohol consumption alc_{1h},[22] this first stage effect varies by some vector of sibling-pair characteristics S_h (dummies for various binary subgroups), and older sibling consumption then influences younger sibling consumption in the same way for all sibling-pair subgroups:

$$alc_{2h} = \beta_1 alc_{1h} + f(age_{1h}) + S_h' \beta_2 + \epsilon_{2h} \tag{B1}$$

$$alc_{1h} = \alpha_1 1\{age_{1h} \geq 21\} + (S_h' \cdot 1\{age_{1h} \geq 21\})\alpha_2 + f(age_{1h}) + S_h' \alpha_3 + \eta_{1h} \tag{B2}$$

Equation B2 is the first stage in a 2SLS specification that uses the cutoff dummy ($1\{age_{1h} \geq 21\}$) and its interactions with a vector of subgroup dummies S_h to instrument for older sibling alcohol consumption. Since the model allows the first stage effect to vary by these subgroups, we can define subgroup-specific first stage effects as:

$$\alpha(S_h) = \alpha_1 + S_h' \alpha_2 \tag{B3}$$

Since the 2SLS estimate for β_1 is equal to the ratio of the reduced-form effect over the first stage effect, it follows (using notation from our reduced form specification in Section 3.1 (γ_1)) that the subgroup-specific reduced-form effects are:

$$\gamma_1(S_h) = \alpha(S_h)\beta_1 \tag{B4}$$

[22] Note that while we exclude the time (t) subscript here for simplicity our subgroup level results will use the same preferred specification from Section 4.2 which includes individual level fixed effects, controls, and a linear donut RD specification.

This setup can then be used to implement an overidentification test as described by Angrist et al. (2022).

Instead of a formal test, we plot the subgroup level first stage estimates (e.g., $\widehat{E}[\widehat{\alpha}(S_h)|G_h = 1]$, where G_h is some subgroup dummy) against the subgroup-level reduced form estimates ($\widehat{E}[\widehat{\gamma}_1(S_h)|G_h = 1]$) to informally inspect the relationship between the two. We estimate these subgroup-level first stage and reduced form effects using the corresponding specifications described in Section 3.1 and Section 3.2 in subsamples defined by the value of G_h. We do not formally implement the overidentification test because our first stage and reduced form estimates come from separate samples.

B.2 Inverse propensity score weighting

As described in Section 4.5 we follow Carril et al. (2017) and Gerardino et al. (2017) to estimate propensity scores for each subgroup of interest, and use these to re-weight the samples used in our subgroup level regressions. Specifically, for each subgroup dummy G_h, we estimate $P(G_h = 1|X_h) \equiv P(X_h)$ as the fitted values from a probit model (with G_h as the outcome and X_h as the covariates) estimated via maximum likelihood. We then rerun our subgroup level first stage and reduced estimates with the following weights:

$$G_h \frac{p}{P(X_h)} + (1 - G_h) \frac{1 - p}{1 - P(X_h)} \tag{B5}$$

where p is the unconditional probability that $G_h = 1$. Standard errors are bootstrapped to account for the uncertainty in the estimated weights.

C Evidence against offsetting parental responses

As described in section 4.4 parental responses are one key pathway through which the older sibling's legal access to alcohol could directly (i.e., not via an endogenous peer effect) affect the younger sibling's consumption of alcohol. The NLSY97 includes two separate questions that can provide some suggestive evidence on the role that parental responses play in explaining the main results. First, all respondents are asked to score the degree to which they are monitored by their parents (including, if applicable, parents they live with and parents they do not live with) on a scale from 0-16, with higher scores indicating closer monitoring. Second, respondents are asked to classify their parents' parenting styles (again, including if applicable parents who do and do not live with the respondent) as either uninvolved, permissive, authoritarian, or authoritative.

In the first panel of Table C1 we present results for models similar to equation 1 with the parental monitoring score for the parents that the respondent lives with as the outcome. If the respondent lives with two parents, the outcome is their average score. In the second panel of Table C1, similar results are presented with an indicator for whether or not at least one parent was reported to be either authoritarian or authoritative.

The survey questions underlying these outcomes are asked only in survey rounds between 1997 and 2000. Luckily, a substantial portion of the sibling-pair-years that meet our sample restriction requirements are in this period. However, some are not, and sample sizes are therefore smaller than the previously described full sample results in Table 3. Models in Table C1 are the same as those in Table 3 to demonstrate the robustness of the results.

In each of the 5 models for each of the two outcomes, point estimates for a discontinuity in parental behavior at the cutoff are either negative (suggesting parents become less strict or engage in less monitoring), not statistically significant, or both. In the preferred model (1) neither point estimate is economically significant. We interpret this as suggestive evidence that parental responses are not driving the observed decrease in younger sibling alcohol consumption at the cutoff.

Table C1: Parenting style changes at the cutoff

	Degree of Parental Monitoring				
	(1)	(2)	(3)	(4)	(5)
Sibling 21+	-0.330	-0.726*	-0.418	-0.421	-0.302
	(0.299)	(0.353)	(0.478)	(0.292)	(0.277)
Control Mean	9.031	9.031	9.031	9.031	9.031
Observations	1,783	1,467	1,783	1,937	1,818
	Parents: Authoritarian/Authoritative				
Sibling 21+	0.002	-0.001	-0.004	0.007	0.001
	(0.023)	(0.023)	(0.037)	(0.024)	(0.022)
Control Mean	0.251	0.251	0.251	0.251	0.251
Observations	5,272	4,317	5,272	5,575	5,391
FE	X		X	X	X
Quadratic			X		
Controls	X	X	X		X
Donut	X	X	X	X	

Note: Controls include the month and year of the survey, the age of the respondent, educational attainment and enrollment of respondents, geography (urban/rural, census region), household size, an indicator for whether the respondent has children, and an indicator for whether the respondent worked in the past year. All models are estimated in a sample of NLSY97 respondents who are the second oldest siblings currently residing in their household, and where the oldest sibling is between the ages of 19 and 23. All models include cluster robust standard errors at individual level. Sibling age is centered at 21 years. +, *, **, and *** denote statistical significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

D Time use of siblings near the MLDA

As described in section 4.4, a concern in our setting is that younger siblings may not be sufficiently exposed to changes in older sibling alcohol consumption that occur at the MLDA. This might be the case if, for example, siblings in this age range do not spend meaningful amounts of time together.

While a small literature on the shared time use of siblings does exist (e.g., Dunifon et al., 2017; Wikle and Hoagland, 2020), there does not appear to be any work on the age ranges relevant to our analysis. We therefore address this concern with a simple descriptive analysis which demonstrates that siblings at ages near the MLDA spend substantial amounts of time together—even after excluding time with parents, and especially when focusing on same-gender sibling pairs.

The ATUS is a repeated cross-sectional survey, which regularly samples a subset of Current Population Survey (CPS) respondents and solicits a detailed time diary for one day's worth of activities (Hofferth et al., 2020). These time diaries detail the minute-level activities of respondents, including what activity was being performed, where, and with whom. Since the ATUS sample is drawn from the CPS, the two surveys can be linked so that the household structure of each ATUS respondent is observable. We use this information to select all respondents from the 2003-2019 waves of the ATUS who lived with *only one* older sibling who was 23 years old or younger. We then use the information on what activities were performed with whom to characterize the strength of these sibling relationships and provide suggestive evidence regarding the potential for younger siblings to be exposed to changes in older sibling alcohol consumption. Summary statistics for our analysis sample of 2,795 ATUS respondents are shown in Table D1.

We calculate the average time spent per day with an older sibling. For comparison, we also calculate the average time spent per day with friends.[23] Finally, we present these means in several different subsets of activities. First, we move from all activities reported in the time diary to activities performed while not at work or school—under the assumption that such "discretionary" time is more relevant for the behavior we are interested in. Second, we further restrict our focus to time spent not at work or school and without any parents present for the same reason. Finally, we break down this last category of discretionary and unsupervised time into the main activity groups defined by ATUS.[24] Results are shown in Figure D1.

The average respondent reports spending roughly 1 hour and 40 minutes with an older sibling during their diary day, just under one hour of which is discretionary and unsupervised

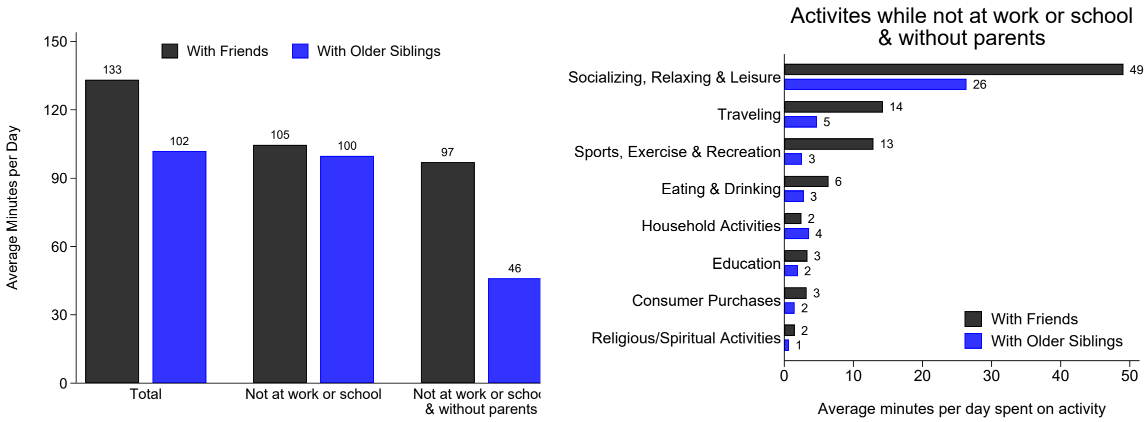
[23]We consider an activity to be performed with an older sibling (or similarly, with a friend) if the so-called "who" record in the ATUS data lists at least one older sibling (or friend).

[24]Note that it is not possible to directly observe alcohol consumption in the ATUS.

time, when alcohol consumption would be most likely to occur. Since a substantial portion of the average diary day in our sample is spent either at work, at school, or with parents (see Table D1), this is a meaningful amount of time.[25] Moreover, the activities that siblings engage in together during this discretionary and unsupervised time (right panel of the figure) are broadly similar to the activities engaged in with friends. While it is clear that respondents spend more time with friends (especially within discretionary and supervised time), we argue that this is strong evidence that the sibling pairs we study in our main analysis are a meaningful peer group for the behavior of interest.

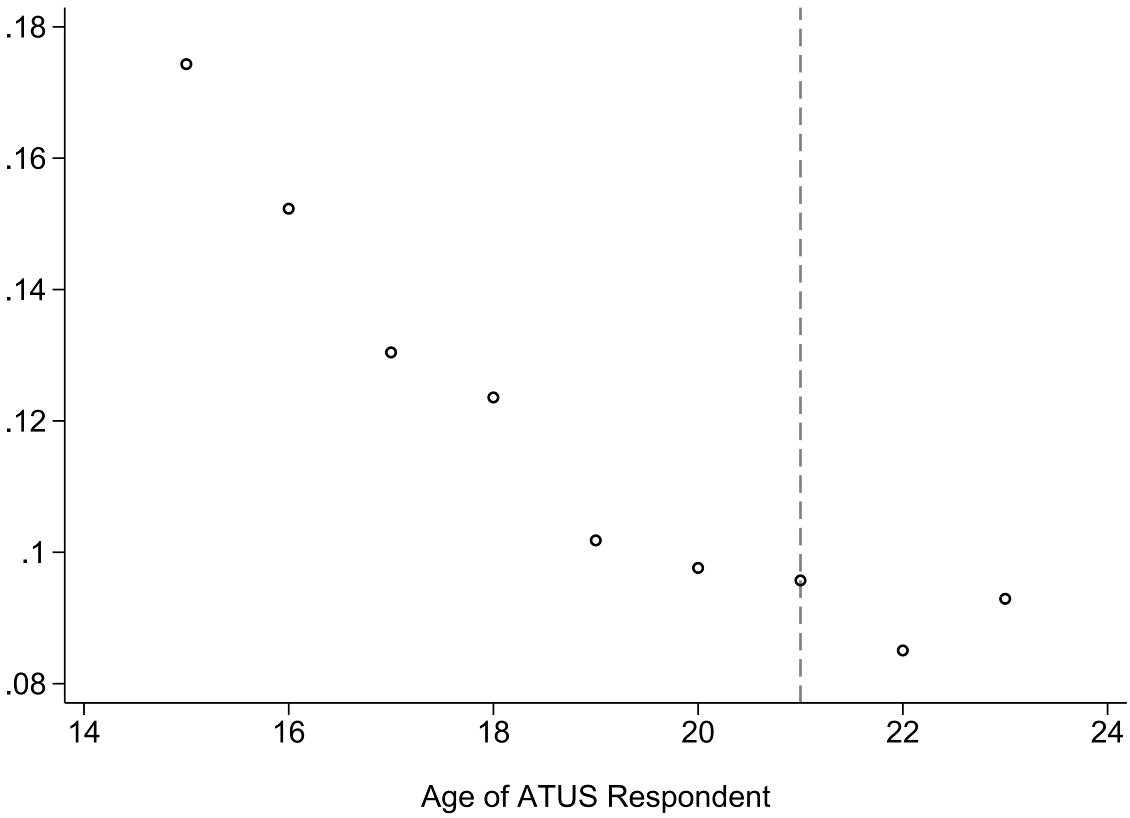
[25] This amount does not change substantially when the older sibling turns 21 (see Figure D2).

Figure D1: Time Spent with Older Siblings in the ATUS



Note: Descriptive statistics are from a sample of 2,795 respondents in the 2003-2019 waves of the ATUS who have one older sibling under the age of 23 in their household. Each panel shows the average (across respondent) time per day spent with either friends (black, left bars in each pair in the left panel and top bars in each pair in the right panel) or older siblings (blue, right bars in each pair in the left panel and bottom bars in each pair in the right panel) in various categories.

Figure D2: Mean Proportion of Day Spent with Any Younger Sibling



Note: The sample includes all 4,615 respondents in the 2003-2019 ATUS who were between the ages of 15 and 23 at the time of the survey and had at least one younger sibling between the ages of 14 and 22.

Table D1: ATUS Sample Summary Statistics

	Mean	SD	Min	Max
Respondent Characteristics:				
Age	17.15	(1.76)	15.0	22.0
Female	0.49	(0.50)	0.0	1.0
In high school	0.58	(0.49)	0.0	1.0
In college	0.18	(0.39)	0.0	1.0
Employed	0.38	(0.49)	0.0	1.0
Surveyed on a weekend or holiday	0.31	(0.46)	0.0	1.0
Age of closest older sibling	19.86	(1.89)	16.0	23.0
Minutes Spent (during ATUS-diary day):				
At home	333.69	(221.19)	0.0	1440.0
At school	168.94	(223.88)	0.0	1020.0
At work	67.23	(163.28)	0.0	890.0
Alone	223.38	(196.42)	0.0	1370.0
With an older sibling	103.71	(171.45)	0.0	1200.0
With a parent	147.03	(192.08)	0.0	1200.0
With a friend	135.97	(208.73)	0.0	1140.0
With anyone else	186.36	(227.81)	0.0	1210.0
N	2,795			

Note: Summary statistics from time diaries for all ATUS respondents in the 2003-2019 waves who lived with one older sibling 23 years old or younger. All summary statistics are calculated using survey weights. Data from Hofferth et al. (2020).

References

Angrist, J., D. Autor, and A. Pallais (2022). Marginal effects of merit aid for low-income students. The Quarterly Journal of Economics 137(2), 1039–1090. (Cited on pages 17 and 18.)

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. Journal of the American Statistical Association 113(522), 767–779. (Cited on page 16.)

Calonico, S., M. D. Cattaneo, and M. H. Farrell (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. The Econometrics Journal 23(2), 192–210. (Cited on page 16.)

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82(6), 2295–2326. (Cited on page 16.)

Carril, A., A. Cazor Katz, M. P. Gerardino, S. Litschig, and D. Pomeranz (2017). Rddsga: Stata module to conduct subgroup analysis for regression discontinuity designs. (Cited on page 18.)

Dunifon, R., P. Fomby, and K. Musick (2017). Siblings and children's time use in the united states. Demographic Research 37, 1611–1624. (Cited on page 21.)

Eisenberg, D., E. Golberstein, and J. L. Whitlock (2014). Peer effects on risky behaviors: New evidence from college roommate assignments. Journal of Health Economics 33, 126–138. (Cited on page 4.)

Gerardino, M. P., S. Litschig, and D. Pomeranz (2017). Can audits backfire? evidence from public procurement in chile. NBER Working Papers (23978). (Cited on page 18.)

Hofferth, S. L., S. M. Flood, M. Sobek, and D. Backman (2020). American time use survey data extract builder: version 2.8 [dataset]. College Park, MD: University of Maryland and Minneapolis, MN: IPUMS. https://doi.org/10.18128/D060.V2.8. (Cited on pages 21 and 25.)

Wikle, J. S. and A. Hoagland (2020). Adolescent interactions with family and emotions during interactions: Variation by family structure. Journal of Family Psychology 34(5), 544. (Cited on page 21.)