

Online Appendices

Appendix A. Studies related to cyberbullying victimization and mental distress

| Study | Data | Theory | Method | Related findings |
|---------------------------------------|--|------------------------------|--|--|
| Beckman et al. (2012) | 3820 13–16 years old students in Sweden | No theory | ANOVA; Multinomial regression analysis; Linear regression analysis | Regardless of the type of bullying involvement, a positive association between bullying and psychosomatic problems was established. The hypothesis that cyberbullying poses a stronger effect on mental health than traditional bullying was not supported. |
| Beran & Li (2007) | 432 students from grades 7–9 in Canadian schools | Social rank theory | Logistic regression | Cyberbullying victims at school experienced difficulties such as low marks, poor concentration, and absenteeism. |
| Bottino et al. (2015) | A systematic review of two databases: PubMed and Virtual Health Library (BVS) | No theory | Literature review | Cyberbullying perpetrators and victims had more emotional and psychosomatic problems, and social difficulties and felt unsafe and not being cared for in school. Cyberbullying victimization was related to moderate to severe depressive symptoms, substance use, ideation, and suicide attempts. |
| Cassiani-Miranda et al. (2021) | 1462 students between 13 and 17 years in Colombia | No theory | Logistic regression | Cyberbullying victimization is associated with post-traumatic stress disorder risk, cigarette smoking, female gender, family dysfunction, and poor-fair general health condition |
| Cénat et al. (2016) | Teenagers who participated in two waves of the Quebec Youths' Romantic Relationships Survey (N = 8194 and 6780 at Wave I and II, respectively) | No theory | Bivariate analyses | Cyberbullying and bullying victimization mediated the relationship between child sexual abuse and mental health. |
| Kowalski et al. (2014) | A list of related research and reports on both the assessment methods and prevalence rates of cyberbullying across varying samples | The general aggression model | A meta-analytic review | The strongest correlates with cyberbullying victimization were stress and suicidal ideation. Several methodological and sample characteristics moderate these relationships. |

| | | | | |
|--------------------------------|--|--|--|---|
| Nixon (2014) | Existing literature on the effects of cyberbullying on adolescent health across multiple studies worldwide | No theory | Literature review | Adolescent cyberbullying victims reported increased depressive affect, anxiety, loneliness, suicidal behavior, and somatic symptoms |
| Parris et al. (2022) | 169 high school students in the U.S. | Response styles theory | Structural equation modeling | Social media rumination mediated the relationship between bullying perceptions at school and distress, but mediation was not supported in the relationship between cyberbullying perceptions and distress |
| Pettalia et al. (2013) | 260 students in the 12 schools of Northeastern Ontario, Canada | No theory | Logistic regression | Cyberbullying victims reported higher scores than non-victims on the cognitive empathy scale |
| Schneider et al. (2012) | 20406 12–19 grade students in MetroWest Massachusetts | No theory | Cross-tabulations & binomial logistic regression | Cyberbullying victims reported elevated levels of distress. |
| Sourander et al. (2010) | 2215 13 to 16 years old Finnish adolescents | No theory | Logistic regression | Cyberbullying victimization is associated with psychiatric and psychosomatic problems. |
| Wang et al. (2011) | 7313 students in grades 6–10 in the U.S. | No theory | Regression analysis | Cyberbullying victims reported a higher level of depression than bullying victims |
| Worsley et al. (2018) | 476 adolescents aged 13–19 years | Adult attachment theory; stress-buffering model; Beck's cognitive theory | Regression analyses | Cyberbullying victims experienced higher levels of depression and anxiety and endorsed more self-statements indicative of anxious attachment than nonvictims |

Appendix B. The data collection process and tests that ensure the reliability of the data in YRBS

The YRBS paid great attention to the data collection process to ensure the reliability of the data (Brener *et al.* 2013). Readers are referred to Brener *et al.* (2013) for a detailed treatment of the methodology of YRBS. We highlight and discuss two aspects of the data collection method that are relevant to potential concerns regarding our study. Trained YRBS data collectors traveled to participating schools in the U.S. to administer the YRBS questionnaires to students. In all surveys, students completed the self-administered questionnaire during one class period and recorded their responses directly in a computer-scannable booklet or on a computer-scannable answer sheet (Brener *et al.* 2013). These measures guaranteed that students who were more familiar with internet devices or more active on the internet had the same chance of being interviewed as those who were not, dissipating the concern that the data collection process might cause potential selection bias because students who were familiar with internet devices might have been more likely to answer the questionnaire and at the same time more likely to be cyberbullied.

In addition, measures were taken to ensure the privacy of students and the anonymity of answers: first, students' desks were spread throughout the classroom to minimize the chance that students could see each other's responses; second, students were encouraged to cover their responses as they complete the questionnaire; third, in the national survey and certain state surveys, when students completed the questionnaire, they were asked to seal their questionnaire booklet or answer sheet in an envelope before placing it in a box (Brener *et al.* 2013). These measures helped guarantee that students did not have the incentive to lie in their answers.

Furthermore, we chose a marker variable in YRBS, the number of times the respondents ate carrots in the last 7 days, which we used to further test whether common-method bias exists (Lindell & Whitney 2001). A marker variable is one that is prone to social desirability bias (i.e., people are likely to respond highly in the affirmative) but unrelated to the theoretical model being tested. The idea here is that if common method bias existed, all (or most) constructs would be highly correlated, including the marker variable (Pavlou *et al.* 2007). The simplest way to test this is to run the marker variable in the correlation matrix of all variables. We did so against all the major constructs (the eight measures of substance use and cyberbullying victimization) of our model and found trivial to low correlations (The coefficients of correlation between the marker variable and the major constructs of interest were all below 0.1). This provides further evidence that common-method bias is not a legitimate threat to this study.

Appendix C. Data Sources for State-level Policies and Economic Conditions

- 1) State median income. Source: U.S. Census Bureau, Historical Income Tables: Households, “Table H-8. Median Household Income by State.” (<http://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html>). Most recent date of access: Jan 19, 2021.
- 2) Unemployment rate. Source: Annual Unemployment Rates by State. U.S. Department of Labor, Bureau of Labor Statistics. Local Area Unemployment Statistics (LAUS). (<https://www.bls.gov/lau/#tables>) and Iowa Community Indicators Program, Iowa State University. (<https://www.icip.iastate.edu/tables/employment/unemployment-states>). Most recent date of access: Jan 19, 2021.
- 3) Expenditures per pupil for public elementary and secondary education. Source: U.S. Department of Education National Center for Education Statistics (<http://nces.ed.gov/ccd/elsi/>). Data from 2017-2019 are not available and are replaced with data from 2016. Most recent date of access: Jan 19, 2021.
- 4) Pupil-teacher ratio. Source: U.S. Department of Education National Center for Education Statistics (<http://nces.ed.gov/ccd/elsi/>) Most recent date of access: Jan 19, 2021.
- 5) Cigarette tax. Source: The Tax Burden on Tobacco, 1970-2018. Orzechowski and Walker. Tax Burden on Tobacco. Tax burden data was obtained from the annual compendium on tobacco revenue and industry statistics, The Tax Burden on Tobacco. Data are reported on an annual basis; Data include federal and state-level information regarding taxes applied to the price of a pack of cigarettes. Updated August 13, 2020. (<https://chronicdata.cdc.gov/Policy/The-Tax-Burden-on-Tobacco-1970-2018/7nwe-3aj9/data>) Note: Since the data of 2019 is not available in the data source, the data of 2018 is used instead of the data of 2019.
- 6) Beer tax. Source: Alcohol Policy Information System (https://alcoholpolicy.niaaa.nih.gov/taxes_beer.html). All tax rates are the rates on the first day of each year. Numbers are rounded up to the cent to achieve consistency of the two datasets. Based on the drinking pattern of youth, using beer taxes to approximate the tax burden for youth alcohol consumption is more accurate than the liquor tax or wine tax. Since the data of 2019 in Utah is not available in the data source, the data of 2018 is used instead of the data of 2019 in Utah.
- 7) Anti-Bully Laws. Source: Rees, D. I., Sabia, J. J., & Kumpas, G. (2020). *Anti-bullying laws and suicidal behaviors among teenagers* (No. w26777). National Bureau of Economic Research.
- 8) Anti-Cyberbully Laws. Source: Nikolaou, D. (2017). Does cyberbullying impact youth suicidal behaviors? *Journal of health economics*, 56, 30-46.
- 9) Medical and recreational marijuana legalization year. <https://www.procon.org/> ProCon.org. “Legal Medical Marijuana States and DC” (<https://medicalmarijuana.procon.org/legal-medical-marijuana-states-and-dc/>). “Legal Recreational Marijuana States and DC” (<https://marijuana.procon.org/legal-recreational-marijuana-states-and-dc/>) Most recent date of access: Jan 19, 2021.
- 10) State high school student enrollment by year, state, grade, gender, and race. Source: U.S. Department of Education National Center for Education Statistics (<http://nces.ed.gov/ccd/elsi/>) Most recent date of access: Jan 19, 2021.

Appendix D. Summary statistics

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
|--|-------------|----------|--------------|----------|------------------|----------|
| | Full sample | | Cyberbullied | | Not cyberbullied | |
| | N | mean | N | mean | N | mean |
| Substance use | | | | | | |
| Did binge drink in last 30 days | 252,848 | 0.16 | 37,311 | 0.24 | 215,537 | 0.15 |
| Did drink in last 30 days | 549,345 | 0.32 | 80,190 | 0.45 | 469,155 | 0.3 |
| Did smoke in last 30 days | 574,813 | 0.12 | 84,787 | 0.20 | 490,026 | 0.11 |
| Did use marijuana in last 30 days | 580,279 | 0.21 | 86,326 | 0.28 | 493,953 | 0.2 |
| Number of days binge drinking in last 30 days | 252,848 | 0.59 | 37,311 | 1.03 | 215,537 | 0.52 |
| Number of days drinking in last 30 days | 549,345 | 1.50 | 80,190 | 2.50 | 469,155 | 1.35 |
| Number of days smoking in last 30 days | 574,813 | 1.59 | 84,787 | 2.62 | 490,026 | 1.43 |
| Times used marijuana in last 30 days | 580,279 | 2.90 | 86,326 | 3.79 | 493,953 | 2.76 |
| Proxies for mental distress and risk attitude | | | | | | |
| Sad/hopeless in last 12 months | 588,823 | 0.30 | 88,210 | 0.59 | 500,613 | 0.26 |
| Never or rarely wear seat belt | 400,892 | 0.07 | 61,232 | 0.09 | 339,660 | 0.07 |
| Cyberbullying and bullying victimization | | | | | | |
| Bullied in the last 12 months | 595,498 | 0.18 | 89,545 | 0.62 | 505,953 | 0.11 |
| Cyberbullied in the last 12 months | 595,498 | 0.14 | 89,545 | 1.00 | 505,953 | 0.00 |
| Individual-level controls | | | | | | |
| Under-weight | 595,498 | 0.23 | 89,545 | 0.25 | 505,953 | 0.23 |
| Normal weight | 595,498 | 0.48 | 89,545 | 0.41 | 505,953 | 0.50 |
| Over-weight | 595,498 | 0.24 | 89,545 | 0.26 | 505,953 | 0.24 |
| Obese | 595,498 | 0.05 | 89,545 | 0.07 | 505,953 | 0.04 |
| Female | 595,498 | 0.49 | 89,545 | 0.66 | 505,953 | 0.46 |
| Hispanic | 595,498 | 0.25 | 89,545 | 0.19 | 505,953 | 0.26 |
| Black | 595,498 | 0.16 | 89,545 | 0.11 | 505,953 | 0.17 |
| White | 595,498 | 0.50 | 89,545 | 0.59 | 505,953 | 0.48 |
| Age | 595,498 | 16.03 | 89,545 | 15.93 | 505,953 | 16.04 |
| Grade | 595,498 | 10.44 | 89,545 | 10.36 | 505,953 | 10.45 |
| State-level controls | | | | | | |
| Anti-bully laws | 595,498 | 0.98 | 89,545 | 0.98 | 505,953 | 0.98 |
| Anti-cyberbully laws | 595,498 | 0.96 | 89,545 | 0.96 | 505,953 | 0.96 |
| State median income (\$) | 595,498 | 58045.20 | 89,545 | 57894.24 | 505,953 | 58070.12 |
| Per pupil spending (\$) | 595,498 | 13070.36 | 89,545 | 13094.97 | 505,953 | 13066.30 |
| Pupil-teacher ratio | 595,498 | 16.78 | 89,545 | 16.74 | 505,953 | 16.78 |
| Unemployment rate | 595,498 | 6.09 | 89,545 | 6.09 | 505,953 | 6.08 |
| Cigarette tax (\$) | 595,498 | 1.60 | 89,545 | 1.61 | 505,953 | 1.60 |
| Beer tax (\$) | 595,498 | 0.30 | 89,545 | 0.29 | 505,953 | 0.30 |
| Medical marijuana legalization | 595,498 | 0.40 | 89,545 | 0.40 | 505,953 | 0.40 |
| Recreational marijuana legalization | 595,498 | 0.10 | 89,545 | 0.10 | 505,953 | 0.10 |

Notes: Individual-level data are from 2011-2019 state and national YRBS. State-level controls are from various sources as illustrated in Online Appendix C. Marijuana use is measured in “occasions” instead of “days” in YRBS. Students reported their risky behaviors in YRBS categorically. We took the midpoint of each category range as an approximation of the actual binge drinking/drinking/smoking days and marijuana consumption occasions. Samples were weighted using gender-by-race-by-grade-by-year-by-state high school student population. Weights were generated using state high school student enrollment numbers by year, state, grade, gender, and race from the U.S. Department of Education National Center for Education Statistics (<http://nces.ed.gov/ccd/elsi/>). Most recent date of access: Jan 9, 2024.

From the summary statistics of the full sample, the proportions of students who binge drank, drank, smoked, and consumed marijuana in the last 30 days are 16%, 32%, 12%, and 21%, respectively. During our sample period, high-school students engaged in an average of 0.6 days of binge drinking, 1.5 days of drinking, 1.6 days of smoking, and 2.9 occasions of marijuana use in the last 30 days. Regarding the likelihood of being cyberbullied, we see that 14% of students have been cyberbullied. Students who were cyberbullied show a higher level of substance use than students who were not cyberbullied. We see that students who were cyberbullied were more likely to experience sadness or hopelessness than those who were not cyberbullied. 59 percent of students who were cyberbullied experienced sadness or hopelessness.

Appendix E. OLS regression results of cyberbullying victimization on individual controls

| | (1) | (2) |
|---|-------------------------------|----------------------|
| | Cyberbullied in the past year | |
| Gender | | |
| Female | 0.099*** (0.005) | 0.099*** (0.005) |
| Age (“12 or younger” is omitted and served as baseline) | | |
| 13 | -0.389*** (0.059) | -0.390*** (0.059) |
| 14 | -0.399*** (0.041) | -0.398*** (0.041) |
| 15 | -0.393*** (0.040) | -0.393*** (0.040) |
| 16 | -0.388*** (0.040) | -0.387*** (0.039) |
| 17 | -0.380*** (0.041) | -0.379*** (0.040) |
| 18 or older | -0.377*** (0.042) | -0.376*** (0.042) |
| Grade (9th grade is omitted and served as baseline) | | |
| 10 | -0.016*** (0.005) | -0.016*** (0.005) |
| 11 | -0.036*** (0.006) | -0.037*** (0.006) |
| 12 | -0.047*** (0.008) | -0.048*** (0.008) |
| Race/Ethnicity (“All other races” is omitted and served as baseline) | | |
| Hispanic | -0.039*** (0.003) | -0.039*** (0.003) |
| Black | -0.050*** (0.004) | -0.050*** (0.004) |
| White | 0.016*** (0.004) | 0.016*** (0.004) |
| Weight (normal weight is omitted and served as baseline) | | |
| under-weight | 0.043*** (0.004) | 0.044*** (0.004) |
| over-weight | 0.028*** (0.004) | 0.028*** (0.004) |
| obese | 0.096*** (0.013) | 0.096*** (0.013) |
| State-level controls | No | Yes |
| State dummies | Yes | Yes |
| Year dummies | Yes | Yes |
| Observations | 595,498 | 595,498 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Estimates were obtained using weighted OLS with standard errors (in parentheses) clustered by state. The weighting method is provided in the empirical method section. State-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, beer tax, cigarette tax, dummies for antibullying laws and anti-cyberbullying laws, medical marijuana legalization, and recreational marijuana legalization status.

Appendix F. Effects of cyberbullying victimization on substance use, OLS

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|--|--|--|---------------------------------|---|------------------------------------|--|--|--|
| | Did binge drink in the last 30 days | Number of days binge drinking in last 30 days | Did drink in last 30 days | Number of days drinking in last 30 days | Did smoke in last 30 days | Number of days smoking in last 30 days | Did use marijuana in last 30 days | Times used marijuana in last 30 days |
| Cyberbullied in the last 12 months | 0.093*** (0.007) | 0.523*** (0.055) | 0.142*** (0.006) | 1.140*** (0.062) | 0.081*** (0.004) | 1.101*** (0.084) | 0.099*** (0.005) | 1.437*** (0.099) |
| Observations | 252,848 | 252,848 | 549,345 | 549,345 | 574,813 | 574,813 | 580,279 | 580,279 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. These estimates were obtained using weighted OLS with standard errors clustered by state. Regression sample weights were generated using state high school student enrollment numbers by year, state, grade, gender, and race from the U.S. Department of Education National Center for Education Statistics (<http://nces.ed.gov/ccd/elsi/>). Covariates include year- and state-fixed effects along with individual-level and state-level controls.

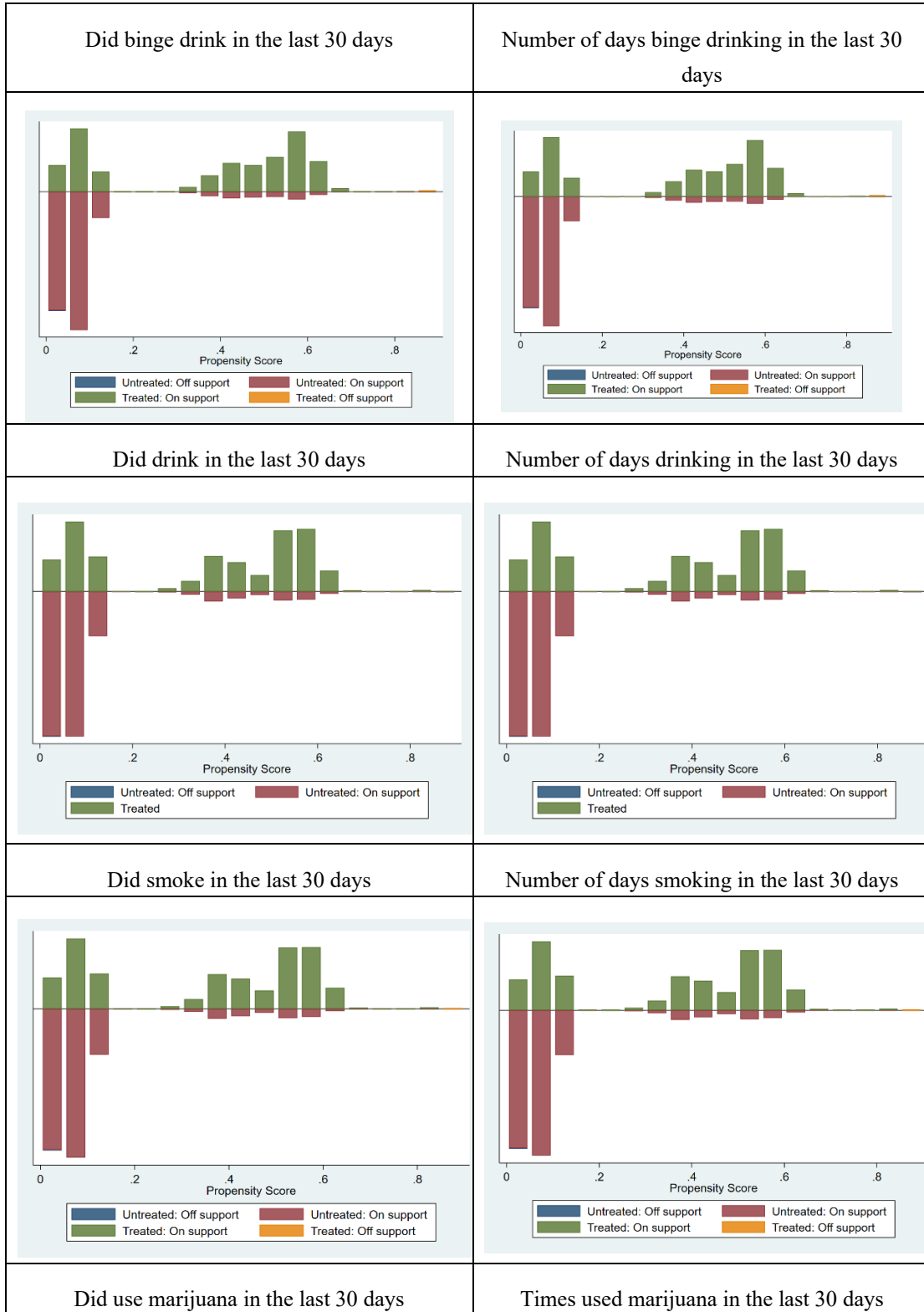
Appendix G.

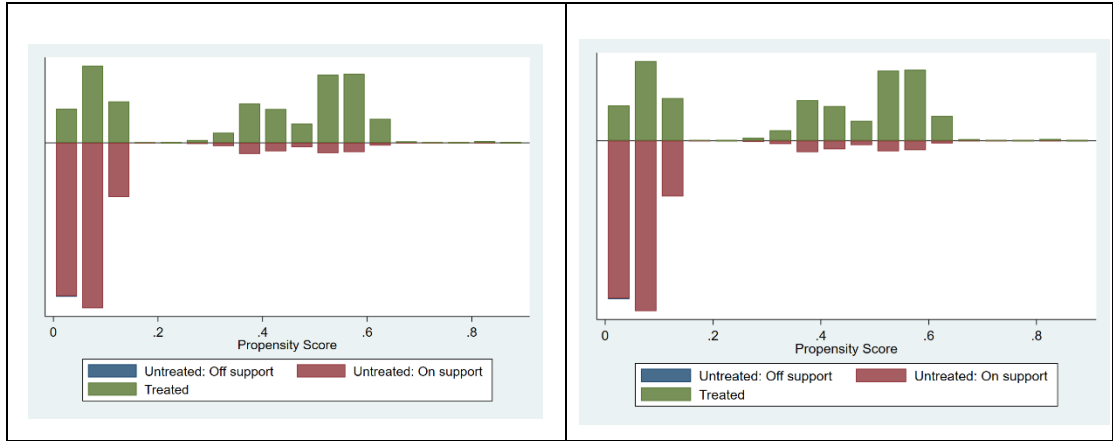
Table G1: Test for balancing of covariates

| | Cyberbullying victimization | | |
|--|-----------------------------|--|--|
| | Pseudo R-Squared | Mean of absolute standardized bias (%) | Median of absolute standardized bias (%) |
| Did binge drinking/ binge drinking days | | | |
| Before matching (unmatched) | 0.249 | 4.8 | 1.7 |
| After Matching (matched) | 0.001 | 0.7 | 0.6 |
| Did drinking/ drinking days | | | |
| Before matching (unmatched) | 0.232 | 4.7 | 1.7 |
| After Matching (matched) | 0.001 | 0.6 | 0.5 |
| Did smoke/smoking days | | | |
| Before matching (unmatched) | 0.231 | 4.7 | 1.8 |
| After Matching (matched) | 0.001 | 0.6 | 0.4 |
| Did use marijuana/marijuana use times | | | |
| Before matching (unmatched) | 0.229 | 4.7 | 1.7 |
| After Matching (matched) | 0.001 | 0.6 | 0.5 |

Notes: Statistics were calculated using Stata routine “psmatch2” developed by Leuven and Sianesi (2018). The mean and median of absolute standardized bias are reduced drastically, and the Pseudo R-squared numbers are reduced to almost zero after matching. This indicates that the matching process removed almost all the explanatory power of the covariates in the matched sample.

Figure G1. Propensity score histogram by treatment status





Appendix H. Robustness checks

Our PSM results are robust to various sets of covariates and matching methods. We performed various robustness checks including 1) using selection on observables to assess the bias from unobservables; 2) testing for reverse causality, and 3) testing for the threats of potential sample selection bias, measurement error, and procedural bias, to show that it is unlikely that the above mentioned issues drive the positive associations. These analyses suggest that the positive associations between cyberbullying and substance use are indicative of a causal relationship. Study 3 adopted a scenario-based experiment and further bolstered this notion.

a. Control for student's electronic device use time

Table H1 presents the PSM estimates with students' electronic device use time ("hours playing video or computer games or using a computer for something that is not school work") serving as an additional matching variable. The results are consistent with our main results shown in **Table 2**. There are only slight differences in point estimates and standard errors and many estimates are nearly identical. These results support the hypothesis that adolescent cyberbullying victimization is associated with higher levels of substance use.

Table H1. Effects of cyberbullying victimization on substance use, propensity score matching, adding electronic device use time in propensity score

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|------------------------------------|-------------------------------------|--|---------------------------|--|---------------------------|---|-----------------------------------|--------------------------------------|
| | Did binge drink in the last 30 days | Number of days of binge drinking in last 30 days | Did drink in last 30 days | Number of days of drinking in last 30 days | Did smoke in last 30 days | Number of days of smoking in last 30 days | Did use marijuana in last 30 days | Times used marijuana in last 30 days |
| Cyberbullying victimization | | | | | | | | |
| ATT | 0.087*** | 0.528*** | 0.144*** | 1.197*** | 0.082*** | 1.195*** | 0.108*** | 1.591*** |
| S.E. | (0.005) | (0.027) | (0.004) | (0.035) | (0.003) | (0.045) | (0.003) | (0.062) |
| T-statistics | 19.07 | 19.24 | 37.06 | 34.60 | 31.00 | 26.60 | 34.89 | 25.46 |
| Observations on support | 247,673 | 247,673 | 536,666 | 536,666 | 561,146 | 561,146 | 566,594 | 566,594 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Estimates were obtained using PSM using nearest-neighbor matching (one-to-one matching with replacement). Variables included in propensity score are state and year-fixed effects, individual-level and state-level variables. Individual-level variables include dummies for age, grade, gender, race, underweight, overweight, obese, and electronic device use time. The state-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, beer tax, cigarette tax, dummies for state anti-bully laws and anti-cyberbully laws, medical marijuana legalization, and recreational marijuana legalization status. Additionally, when calculating the propensity of cyberbullying victimization, the variable bullying victimization is included.

Table H2 presents the corresponding balancing test of covariates before and after matching in models presented in **Table H1**. The results show that we can produce balanced samples after the matching process for all three model specifications: the mean and median of absolute standardized bias shown in these tables are also reduced drastically, and the pseudo-R-squares are reduced to almost zero after matching.

Table H2: Test for balancing of covariates, propensity score matching, adding electronic device use time in propensity score

| | Cyberbullying victimization | | |
|--|-----------------------------|--|--|
| | Pseudo R-Squared | Mean of absolute standardized bias (%) | Median of absolute standardized bias (%) |
| Did binge drinking/ binge drinking days | | | |
| Before matching (unmatched) | 0.25 | 4.9 | 1.6 |
| After Matching (matched) | 0.002 | 1.1 | 0.9 |
| Did drinking/ drinking days | | | |
| Before matching (unmatched) | 0.234 | 4.8 | 1.7 |
| After Matching (matched) | 0.002 | 1 | 0.8 |

| | | | |
|--|-------|-----|-----|
| Did smoke/smoking days | | | |
| Before matching (unmatched) | 0.233 | 4.8 | 1.8 |
| After Matching (matched) | 0.002 | 1 | 0.8 |
| Did use marijuana/marijuana use times | | | |
| Before matching (unmatched) | 0.231 | 4.8 | 1.7 |
| After Matching (matched) | 0.002 | 1 | 0.8 |

Notes: Statistics were calculated using Stata routine “psmatch2” developed by Leuven and Sianesi (2018).

b. Control for students' risk attitude ("never/rarely wear a seatbelt")

As individuals' risk attitudes may influence their subsequent risky behaviors such as substance use (Wärneryd 1996), **Table H3** presents the PSM estimates with a proxy for students' risk attitude, "never/rarely wear a seatbelt," along with other variables used in our main specification. While the question about seat belt wearing is not asked in some state YRBS surveys and significant numbers of observations are lost across all substance use measures, the estimated effects of cyberbullying victimization are largely similar to our main results shown in **Table 2**.

Table H3: Effects of cyberbullying victimization on substance use, propensity score matching, adding proxy for risk attitude in propensity score

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|------------------------------------|-------------------------------------|---|---------------------------|---|---------------------------|--|-----------------------------------|--------------------------------------|
| | Did binge drink in the last 30 days | Number of days binge drinking in last 30 days | Did drink in last 30 days | Number of days drinking in last 30 days | Did smoke in last 30 days | Number of days smoking in last 30 days | Did use marijuana in last 30 days | Times used marijuana in last 30 days |
| Cyberbullying victimization | | | | | | | | |
| ATT | 0.094*** | 0.557*** | 0.143*** | 1.250*** | 0.084*** | 1.279*** | 0.106*** | 1.606*** |
| S.E. | (0.005) | (0.033) | (0.006) | (0.043) | (0.004) | (0.078) | (0.005) | (0.091) |
| T-statistics | 18.04 | 16.89 | 22.71 | 29.07 | 19.04 | 16.49 | 22.17 | 17.56 |
| Observations on support | 149,224 | 149,224 | 366,770 | 366,770 | 383,970 | 383,970 | 389,951 | 389,951 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Estimates were obtained using PSM using nearest-neighbor matching (one-to-one matching with replacement). Variables included in propensity score are state and year-fixed effects, individual-level and state-level variables. Individual-level variables include dummies for age, grade, gender, race, underweight, overweight, obese, and the proxy for risk attitude "never or rarely wear a seatbelt". The state-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, beer tax, cigarette tax, dummies for state anti-bully laws and anti-cyberbully laws, medical marijuana legalization, and recreational marijuana legalization status. Additionally, when calculating the propensity of cyberbullying victimization, the variable bullying victimization is included.

Table H4 presents the corresponding balancing test of covariates before and after matching in models presented in **Table H3**. The results show that we can produce balanced samples after the matching process for all three model specifications: the mean and median of absolute standardized bias shown in these tables are also reduced drastically, and the pseudo-R-squares are reduced to almost zero after matching.

Table H4: Test for balancing of covariates, adding proxy for risk attitude in propensity score

| | Pseudo R-Squared | Cyberbullying victimization | |
|--|------------------|--|--|
| | | Mean of absolute standardized bias (%) | Median of absolute standardized bias (%) |
| Did binge drinking/ binge drinking days | | | |
| Before matching (unmatched) | 0.243 | 5.4 | 2.2 |
| After Matching (matched) | 0.003 | 1 | 0.8 |
| Did drinking/ drinking days | | | |
| Before matching (unmatched) | 0.229 | 5 | 1.9 |
| After Matching (matched) | 0.002 | 0.8 | 0.6 |
| Did smoke/smoking days | | | |
| Before matching (unmatched) | 0.229 | 5.1 | 2 |
| After Matching (matched) | 0.002 | 0.9 | 0.7 |
| Did use marijuana/marijuana use times | | | |
| Before matching (unmatched) | 0.227 | 5 | 2 |
| After Matching (matched) | 0.002 | 0.8 | 0.6 |

Notes: Statistics were calculated using Stata routine "psmatch2" developed by Leuven and Sianesi (2018).

c. Control for a proxy of students' home environment

As factors such as home environment and parental behaviors may affect how adolescents react to cyberbullying, we want to further control for family's background. Unfortunately, there is only one variable in state YRBS (but not national YRBS) that we can use to proxy for the family's economic background, i.e., "During the past 30 days, how often did you go hungry because there was not enough food in your home?" We note that this proxy for a family's economic background only captures a fraction of variation in adolescents' home environments, but it is the closest thing we have to a measure of household economic conditions. Thus, as a further robustness check, we controlled for this variable in the PSM estimation. The results shown in **Tables H5 and H6** are consistent with our main results. We did not include this variable in the baseline analyses because using it results in a large loss in observations (given that it is not included in the national version of the survey).

Table H5: Effects of cyberbullying victimization on substance use, propensity score matching, controlling for the home environment

| | (1) Did binge drink in the last 30 days | (2) Number of days binge drinking in last 30 days | (3) Did drink in last 30 days | (4) Number of days drinking in last 30 days | (5) Did smoke in last 30 days | (6) Number of days smoking in last 30 days | (7) Did use marijuana in last 30 days | (8) Times used marijuana in last 30 days |
|--|---|---|--|--|---|---|---|---|
| cyberbullying victimization | | | | | | | | |
| ATT | 0.088*** | 0.517*** | 0.128*** | 1.036*** | 0.075*** | 0.998*** | 0.100*** | 1.285*** |
| S.E. | 0.009 | 0.062 | 0.009 | 0.074 | 0.006 | 0.107 | 0.007 | 0.140 |
| T-statistics | 9.67 | 8.32 | 14.70 | 13.92 | 12.67 | 9.28 | 14.31 | 9.17 |
| Observations on support | 40,210 | 40,210 | 79,461 | 79,461 | 86,314 | 86,314 | 83,993 | 83,993 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. The data are from 2011-2019 state YRBS. Estimates were obtained using PSM using nearest-neighbor matching (one-to-one matching with replacement). Variables included in the propensity score are state and year-fixed effects, and individual-level and state-level variables. Individual-level variables include dummies for age, grade, gender, race, underweight, overweight, obese, and how often the individual went hungry because there was not enough food in the home (never, rarely, sometimes, most of the time, or always). The state-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, beer tax, cigarette tax, dummies for state anti-bully laws and anti-cyberbully laws, medical marijuana legalization, and recreational marijuana legalization status. Additionally, when calculating the propensity of cyberbullying victimization, the variable bullying victimization is included.

Table H6: Test for balancing of covariates, propensity score matching, controlling for the home environment

| | Cyberbullying victimization | | |
|--|-----------------------------|---|---|
| | Pseudo R-Squared | Mean of absolute standardized bias (%) | Median of absolute standardized bias (%) |
| Did binge drinking/ binge drinking days | | | |
| Before matching (unmatched) | 0.243 | 9 | 4.5 |
| After Matching (matched) | 0.003 | 1.7 | 1.7 |
| Did drinking/ drinking days | | | |
| Before matching (unmatched) | 0.236 | 7.7 | 3.4 |
| After Matching (matched) | 0.003 | 1.4 | 0.9 |
| Did smoke/smoking days | | | |
| Before matching (unmatched) | 0.234 | 7.5 | 3 |
| After Matching (matched) | 0.003 | 1.4 | 1.1 |
| Did use marijuana/marijuana use times | | | |
| Before matching (unmatched) | 0.233 | 7.6 | 3.4 |
| After Matching (matched) | 0.003 | 1.4 | 1 |

Notes: Statistics were calculated using Stata routine "psmatch2" developed by Leuven and Sianesi (2018)

d. Using selection on observables to assess the bias from unobservables

In this section, we return to the baseline linear regressions shown in **Appendix F**. We adopt the method proposed by Oster (2019) and directly gauge how much stronger the selection on unobserved factors relative to observed factors must be to explain away the estimated effects from these OLS regressions. This sensitivity test is useful to further assess how likely our results are biased by the unaccounted unobserved factors.

Borrowing the insight from Altonji *et al.* (2005), Oster (2019) considers three regressions, one without control variables ($Y = \beta X$, the short regression), one with a full set of observed covariates ($Y = \beta X + \Psi\omega^\circ = \beta X + W_1$, the intermediate regression), and a hypothetical one with all possible observable and unobservable covariates ($Y = \beta X + W_1 + W_2$, the hypothetical regression). The estimates of the coefficient on the treatment X from the short and intermediate regressions are biased due to omitted variables. The key insight is that changes in the estimates of β and the R-squares from the short regression to the intermediate regression are useful to measure both how observed controls affect the results and how the unobserved controls would affect the results through selection (if they were observed and controlled). Oster (2019) proved that (in Proposition 3, Oster (2019)) by assuming a value for the R-square of the hypothetical regression including all possible observable and unobservable covariates (denoted as R_{max}^2), a ratio δ (Oster’s δ) of the effect of unobserved controls to the effect of observed controls on the treatment ($\delta \frac{Cov(W_1, X)}{Var(W_1)} = \frac{Cov(W_2, X)}{Var(W_2)}$) represents the (relative) selection on unobservables that would be necessary to reduce the estimated treatment effect, β , to zero. The ratio depends on several factors including 1) the choice of R_{max}^2 ; 2) coefficients on the treatment and R-squares from the short and the intermediate regressions; 3) variances of the outcome and the treatment; and 4) variance of the treatment that cannot be explained by the observed controls.

Following the recommendation of Oster (2019), we assume R_{max}^2 equals 1.3 times the R^2 of the regression with all observed covariates. The 1.3 multiple is chosen so that 90% of randomized results from top economics journals could survive and is then widely adopted across other academic disciplines in practice (see, for example, Campos-Mercade *et al.* 2021; Ebenstein *et al.* 2017; Ferraro & Simorangkir 2020; Mazrekaj *et al.* 2020; Tsai *et al.* 2020).

We calculate the ratios for cyberbullying for all eight measures of risky behaviors shown in **Appendix F**. The results are reported in **Table H7**. Oster’s δ s for cyberbullying range from 5.0 to 32.3, in absolute values. Therefore, to attribute the entire estimated effects of cyberbullying on risky behaviors to omitted variable bias, the selection on unobservables would have to be at least five times larger than the selection on observed variables, and on average, about ten times larger. Given our included controls encompass a rich set of observed individual characteristics and state-level policies that are shown in previous literature to predict cyberbullying and students’ risky behaviors, and we also included state and year-fixed effects in the OLS estimation, these results suggest that the effects of cyberbullying are extremely unlikely to be accounted for by unobservable factors, further bolstering the reliability of our results regarding **H1**. In addition, we also calculated Oster’s ratios using several different maximum R-squared values ($R_{max}^2 = 1.5 \times R^2$ and $R_{max}^2 = 2 \times R^2$). The new results also point to the same story and suggest that our results are unlikely to be driven by unobserved factors. These results are presented in **Table H8**.

Table H7. Oster (2019)’s δ , test of coefficient robustness to selection on unobservables based on Oster (2019)

| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|-------------------------------------|---|---------------------------|---|---------------------------|--|-----------------------------------|--------------------------------------|
| Did binge drink in the last 30 days | Number of days binge drinking in last 30 days | Did drink in last 30 days | Number of days drinking in last 30 days | Did smoke in last 30 days | Number of days smoking in last 30 days | Did use marijuana in last 30 days | Times used marijuana in last 30 days |
| 6.4 | 7.1 | 5.0 | 5.7 | 5.8 | 6.5 | 13.7 | -32.3 |

Notes: Each cell of the table reports Oster (2019)’s δ , which represents the relative importance of unobservables compared to observables necessary to explain the entire estimated effects of cyberbullying in **Table 2**. $R_{max}^2 = 1.3 \times R^2$.

Table H8: Oster (2019)'s δ , test of coefficient robustness to selection on unobservables based on Oster (2019), assuming different R_{max}^2 's

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|--|-------------------------------------|---|---------------------------|---|---------------------------|--|-----------------------------------|--------------------------------------|
| | Did binge drink in the last 30 days | Number of days binge drinking in last 30 days | Did drink in last 30 days | Number of days drinking in last 30 days | Did smoke in last 30 days | Number of days smoking in last 30 days | Did use marijuana in last 30 days | Times used marijuana in last 30 days |
| Panel 1: $R_{max}^2 = 1.5 \times R^2$ | 4.3 | 4.8 | 3.5 | 4 | 3.8 | 4.2 | 9.7 | -21.5 |
| Panel 2: $R_{max}^2 = 2 \times R^2$ | 2.3 | 2.6 | 1.9 | 2.3 | 2.1 | 2.2 | 5.7 | -11.7 |

Notes: Each cell of the table reports Oster (2019)'s δ , which represents the relative importance of unobservables compared to observables necessary to explain the entire estimated effects of cyberbullying in **Table 2**.

e. Testing for reverse causality

As another robustness check, we want to test whether reverse causality is threatening our conclusions; that is, if higher risky behaviors lead to higher chances of being cyberbullied and whether our main results are only a reflection of these positive associations. This classical endogeneity issue will also arise if we simply treat the eight measures of substance use under consideration as the independent variables and cyberbullying victimization as dependent variables and perform OLS regressions. These simple OLS regressions might suffer from biasedness caused by omitted variables and/or the simultaneous equation issue. One way to test for reverse causality is to use instrumental variables that affect risky behaviors (substance use) to identify the effects of such behaviors on cyberbullying victimization (i.e., the “reverse” causal effect). Valid IVs should affect risky behaviors but not be correlated with unobserved factors that might affect both risky behaviors and cyberbullying victimization. To this end, we tested some potential candidate IVs for risky behaviors (drinking, smoking, and marijuana use) that are used in the literature, including state-level beer tax (Dee 1999; Xuan *et al.* 2013), cigarette tax (Carpenter & Cook 2008; DeCicca *et al.* 2008; Hansen *et al.* 2017), and marijuana legalization status (Anderson *et al.* 2015; Chu 2014; Pacula *et al.* 2015; Wen *et al.* 2015).

Following the standard criteria for weak IV's (Staiger & Stock 1997; Stock & Yogo 2002), we find that the candidate IV's listed above for drinking and marijuana use suffer from obvious weak IV issues in our sample. We could only find valid IVs for smoking behaviors (we will discuss the robustness of these IVs later). Although not ideal, this enables us to test whether one substance use behavior, i.e., smoking, would cause changes in cyberbullying victimization.

The IV estimates of the effects of smoking on cyberbullying victimization are shown in **Tables H9 and H10**. The excluded instruments for smoking status and smoking days are state-level cigarette tax, medical marijuana legalization status, and recreational marijuana legalization status (using these three IVs together achieves the highest Kleibergen-Paap Wald F-statistics for the first stage) and we performed two-stage least-square (2SLS) IV estimations with standard errors clustered at the state level.¹ The first stage regressions (**Table H9**) suggest

¹ Theoretically, medical marijuana legalization status and recreational marijuana legalization status should have a direct effect on marijuana use and an indirect effect on smoking, either because marijuana and cigarettes are substitutes or complements. Sharp readers may question why, in practice, these marijuana-related laws are weak IVs for marijuana use in our sample? We note that even though the most obvious effects of these marijuana laws should be on marijuana use, these candidate instruments are not strong enough in predicting marijuana use in the first stage. Such results are consistent with the well-established evidence from the literature on marijuana laws and juvenile marijuana use suggesting that, while medical and recreational marijuana increase the use of marijuana by adults, there is little evidence showing that these laws increase marijuana use for teens (Anderson *et al.* 2015; Choo *et al.* 2014; Harper

that higher cigarette taxes decrease the likelihood of smoking and smoking days, while medical marijuana legalization and recreational marijuana legalization increase students' smoking status and smoking days (suggesting that marijuana and tobacco are complements). The associated Kleibergen-Paap Wald F statistics range from 7 to 15, indicating the IVs are fairly strong (especially when state-level controls are excluded). Interestingly, from the second stage regression (**Table H10**), we find that all estimates for the effects of smoking status and smoking days on cyberbullying victimization are negative. Many estimates are statistically insignificant, and those estimates that are significant (only at the 10% level) do not consistently show whether smoking affects cyberbullying more. Furthermore, the point estimates on cyberbullying are similar in magnitude. These results indicate that the causal effects of smoking on cyberbullying are likely to be small or negative. To summarize, we do not find evidence that smoking or more smoking days lead to higher chances of being cyberbullied. These results support the notion that the reverse causality issue is not a serious threat to our conclusion.

Admittedly, our test using the IV method comes with two caveats: first, one could argue that our IVs are not strong enough when state-level controls are included because although the associated Kleibergen-Paap Wald F statistics are large (7.0 and 7.9), they do not exceed the traditional rule of thumb value of 10 (Staiger & Stock 1997); second, we could not perform the same 2SLS-IV test for drinking and marijuana behavior due to a lack of strong IV's. Nevertheless, we do not find evidence suggesting that the reverse causality issue is driving our results. We note that with available IVs for drinking and marijuana use, the question of whether drinking and marijuana use could affect cyberbullying remains an important one to answer.

Table H9: First-stage estimates of the effects of smoking on cyberbullying victimization, 2SLS

| | Dependent variable | | | |
|--|---------------------------|--|---------------------------|--|
| | Did smoke in last 30 days | Number of days smoking in last 30 days | Did smoke in last 30 days | Number of days smoking in last 30 days |
| Excluded Instrumental Variables | | | | |
| Cigarette tax | -0.031*** (0.010) | -0.616*** (0.143) | -0.028** (0.013) | -0.564*** (0.205) |
| Medical marijuana legalization | 0.018** (0.007) | 0.321** (0.144) | 0.021*** (0.006) | 0.363** (0.147) |
| Recreational marijuana legalization | 0.031* (0.018) | 0.615** (0.258) | 0.043* (0.022) | 0.804** (0.351) |
| Additional control variables | | | | |
| Individual-level controls | | | | |
| Female | -0.028*** (0.005) | -0.433*** (0.063) | -0.029*** (0.005) | -0.433*** (0.063) |
| Age 13 | -0.338*** (0.055) | -7.897*** (1.349) | -0.340*** (0.055) | -7.936*** (1.349) |
| Age 14 | -0.348*** (0.053) | -8.213*** (1.400) | -0.349*** (0.053) | -8.225*** (1.401) |
| Age 15 | -0.325*** (0.051) | -7.888*** (1.386) | -0.326*** (0.051) | -7.893*** (1.387) |
| Age 16 | -0.291*** (0.052) | -7.275*** (1.384) | -0.291*** (0.052) | -7.273*** (1.387) |
| Age 17 | -0.266*** (0.053) | -6.831*** (1.399) | -0.265*** (0.053) | -6.821*** (1.402) |

et al. 2012; Hao & Cowan 2020; Lynne-Landsman et al. 2013), probably because the relative risks and legal cost of selling marijuana to underage teens increases (Anderson & Rees 2014) when compared with legally selling to adults after marijuana legalization.

| | | | | |
|----------------------------------|-----------|-----------|-----------|-----------|
| Age 18 or older | -0.208*** | -5.878*** | -0.207*** | -5.861*** |
| | (0.054) | (1.391) | (0.054) | (1.395) |
| Grade 10 | -0.008 | -0.076 | -0.008 | -0.083 |
| | (0.006) | (0.067) | (0.006) | (0.067) |
| Grade 11 | -0.011 | -0.109 | -0.012* | -0.125 |
| | (0.007) | (0.106) | (0.007) | (0.105) |
| Grade 12 | -0.009 | -0.104 | -0.010 | -0.127 |
| | (0.008) | (0.126) | (0.008) | (0.124) |
| Hispanic | -0.012*** | -0.302*** | -0.013*** | -0.305*** |
| | (0.004) | (0.093) | (0.004) | (0.090) |
| Black | -0.059*** | -0.902*** | -0.059*** | -0.901*** |
| | (0.005) | (0.070) | (0.004) | (0.069) |
| White | 0.019*** | 0.291*** | 0.019*** | 0.289*** |
| | (0.003) | (0.067) | (0.003) | (0.067) |
| Under weight | 0.006 | -0.011 | 0.005 | -0.017 |
| | (0.004) | (0.093) | (0.004) | (0.091) |
| Overweight | 0.018*** | 0.245*** | 0.018*** | 0.246*** |
| | (0.003) | (0.039) | (0.003) | (0.039) |
| Obese | 0.069*** | 1.055*** | 0.069*** | 1.056*** |
| | (0.012) | (0.193) | (0.012) | (0.193) |
| State-level controls | | | | |
| Anti-cyberbully law | | | -0.024*** | -0.359* |
| | | | (0.009) | (0.186) |
| Median income | | | 0.000 | 0.000 |
| | | | (0.000) | (0.000) |
| Per pupil spending | | | -0.000 | -0.000 |
| | | | (0.000) | (0.000) |
| Pupil teacher ratio | | | 0.006 | 0.067 |
| | | | (0.006) | (0.116) |
| Unemployment | | | 0.007 | 0.126 |
| | | | (0.005) | (0.108) |
| State fixed effect | Yes | Yes | Yes | Yes |
| Year fixed effect | Yes | Yes | Yes | Yes |
| Observations | 574,813 | 574,813 | 574,813 | 574,813 |
| Cragg-Donald Wald F statistic | 140.088 | 168.619 | 115.667 | 140.941 |
| Kleibergen-Paap Wald F statistic | 9.877 | 15.345 | 6.956 | 7.878 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parenthesis and are clustered by state. The weighting method is provided in the empirical method section. Covariates include year and state-fixed effects, along with individual-level and state-level controls. Individual-level controls include dummies for age, grade, gender, race, underweight, overweight, and obese. State-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, dummies for anti-bullying laws, and anti-cyberbullying laws. Excluded instruments are cigarette tax, medical marijuana legalization, and recreational marijuana legalization status.

Table H10: Second-stage estimates of the effects of smoking on cyberbullying victimization, 2SLS

| | Dependent variable | | | |
|---|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| | cyberbullying victimization | cyberbullying victimization | cyberbullying victimization | cyberbullying victimization |
| Key explanatory variable | | | | |
| Did smoke in last 30 days | -0.160 (0.135) | | -0.218* (0.116) | |
| Number of days smoking in last 30 days | | -0.008 (0.007) | | -0.011* (0.006) |
| Additional control variables | | | | |
| Individual-level controls | | | | |
| Female | 0.096*** (0.006) | 0.097*** (0.005) | 0.094*** (0.006) | 0.096*** (0.006) |
| Age 13 | -0.451*** (0.066) | -0.461*** (0.073) | -0.471*** (0.083) | -0.483*** (0.086) |
| Age 14 | -0.464*** (0.060) | -0.475*** (0.070) | -0.484*** (0.070) | -0.497*** (0.076) |
| Age 15 | -0.455*** (0.055) | -0.467*** (0.066) | -0.474*** (0.066) | -0.488*** (0.072) |
| Age 16 | -0.445*** (0.055) | -0.458*** (0.065) | -0.462*** (0.064) | -0.477*** (0.070) |
| Age 17 | -0.433*** (0.053) | -0.446*** (0.063) | -0.448*** (0.063) | -0.464*** (0.069) |
| Age 18 or older | -0.420*** (0.048) | -0.434*** (0.057) | -0.431*** (0.057) | -0.450*** (0.064) |
| Grade 10 | -0.019*** (0.005) | -0.018*** (0.005) | -0.020*** (0.005) | -0.019*** (0.005) |
| Grade 11 | -0.038*** (0.006) | -0.037*** (0.006) | -0.039*** (0.007) | -0.038*** (0.006) |
| Grade 12 | -0.050*** (0.008) | -0.050*** (0.009) | -0.051*** (0.008) | -0.051*** (0.008) |
| Hispanic | -0.044*** (0.004) | -0.044*** (0.004) | -0.045*** (0.005) | -0.045*** (0.005) |
| Black | -0.062*** (0.008) | -0.060*** (0.007) | -0.066*** (0.008) | -0.062*** (0.006) |
| White | 0.020*** (0.004) | 0.019*** (0.004) | 0.021*** (0.004) | 0.019*** (0.004) |
| Under weight | 0.040*** (0.004) | 0.039*** (0.004) | 0.040*** (0.005) | 0.039*** (0.005) |
| Ove weight | 0.031*** (0.005) | 0.030*** (0.004) | 0.032*** (0.005) | 0.031*** (0.004) |
| Obese | 0.104*** (0.022) | 0.102*** (0.020) | 0.108*** (0.018) | 0.105*** (0.016) |
| State-level controls | | | | |
| Anti-cyberbully law | | | -0.012 (0.011) | -0.010 (0.011) |
| Median income | | | 0.000 (0.000) | 0.000 (0.000) |
| Per pupil spending | | | -0.000 | -0.000 |

| | | | | |
|---------------------|---------|---------|---------|---------|
| | | | (0.000) | (0.000) |
| Pupil teacher ratio | | | 0.003 | 0.002 |
| | | | (0.005) | (0.005) |
| Unemployment | | | 0.005 | 0.005 |
| | | | (0.004) | (0.004) |
| State fixed effect | Yes | Yes | Yes | Yes |
| Year fixed effect | Yes | Yes | Yes | Yes |
| Observations | 574,813 | 574,813 | 574,813 | 574,813 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parenthesis and are clustered by state. The weighting method is provided in the empirical method section. Covariates include year and state fixed effects, along with individual-level and state-level controls. Individual-level controls include dummies for age, grade, gender, race, underweight, overweight, and obese. State-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, dummies for anti-bullying laws and anti-cyberbullying laws. Excluded instruments are cigarette tax, medical marijuana legalization, and recreational marijuana legalization status.

f. Dealing with potential sample selection bias, measurement error, and procedural bias

Several remaining concerns regarding the robustness of our analyses are addressed in this section.² Specifically, we deal with three concerns regarding the YRBS survey that may cause bias in our estimations, namely: a. (Sample selection bias) The completion of the survey may depend on the experiences of cyberbullying; b. (Measurement error) The risky behaviors students actually did may be different from what they claimed in the YRBS survey; and c. (Procedural bias) If the questions regarding the cyberbullying experience occur right before the substance use questions in the YRBS, the student may suffer from pressure to fill in more severe risky behaviors.

Regarding the first concern (sample selection bias), we note that YRBS is a long questionnaire focusing on risky behaviors including driving habits, weapon carrying, fighting, dating violence, suicide, tobacco use, alcohol use, marijuana use, and sexual behaviors. A typical state or national YRBS questionnaire contains over 80 questions (for the national YRBS, it is nearly 100), and only one question asked about the experience of being a victim of cyberbullying. Since it is not a questionnaire that focuses on the experiences of cyberbullying, and respondents fill out the questionnaire anonymously, it is unlikely that questions about cyberbullying affect the completion rate of those who have been cyberbullied compared to those who have been not. We did two formal tests to examine this point. First, we examine if the characteristics of those who answer the cyberbullying questions are similar to the full sample. Second, we examine whether there is a higher non-response rate for cyberbullying questions compared to other questions on mental health and risky behaviors (if cyberbullying experiences incite more severe mental distress and force students to quit the survey, we may see lower response rates for those questions).

To perform these two analyses, we only use the national YRBS as some state YRBS surveys do not ask questions on cyberbullying but the national YRBS always include these questions from 2011 to 2019. This will not matter for our main analyses, but for the socioeconomic characteristics comparison and the response rate calculation, it would require us to know whether cyberbullying questions were included in every YRBS questionnaire of every state for every year if we also wanted to use the state YRBS sample. **Table H11** shows the results of our first analysis. The socioeconomic backgrounds and risky behaviors of those who do and do not answer the cyberbullying questions are remarkably similar, indicating that the sample of individuals who answer those questions do not appear to be different from others. **Table H12** shows the results of our second test. The response rates for questions on cyberbullying (97.06%) are very high. In fact, these response rates are among the highest for the 11 questions (on cyberbullying experience, risky behaviors, and mental health) under consideration. We found no evidence that cyberbullying questions have higher non-response rates.

² We greatly appreciate one reviewer for raising these questions.

Table H11: Comparison of characteristics of those who answer the cyberbullying questions versus the full sample

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|-------|------|-------------|-------|------|
| | Students who answered the questions regarding cyberbullying | | | Full sample | | |
| | N | mean | SD | N | mean | SD |
| Socioeconomic background | | | | | | |
| Under-weight | 65,030 | 0.33 | 0.47 | 67,065 | 0.34 | 0.47 |
| Normal weight | 65,030 | 0.42 | 0.49 | 67,065 | 0.41 | 0.49 |
| Over-weight | 65,030 | 0.21 | 0.41 | 67,065 | 0.21 | 0.40 |
| Obese | 65,030 | 0.04 | 0.20 | 67,065 | 0.04 | 0.20 |
| Female | 65,030 | 0.51 | 0.50 | 67,065 | 0.51 | 0.50 |
| Hispanic | 65,030 | 0.11 | 0.31 | 67,065 | 0.11 | 0.32 |
| Black | 65,030 | 0.17 | 0.37 | 67,065 | 0.17 | 0.38 |
| White | 65,030 | 0.38 | 0.48 | 67,065 | 0.37 | 0.48 |
| Age | 65,030 | 16.05 | 1.23 | 67,065 | 16.06 | 1.24 |
| Grade 9 | 65,030 | 0.26 | 0.44 | 67,065 | 0.26 | 0.44 |
| Grade 10 | 65,030 | 0.25 | 0.43 | 67,065 | 0.25 | 0.43 |
| Grade 11 | 65,030 | 0.25 | 0.43 | 67,065 | 0.25 | 0.43 |
| Grade 12 | 65,030 | 0.24 | 0.43 | 67,065 | 0.24 | 0.43 |
| Risky behaviors | | | | | | |
| Did use marijuana in the last 30 days | 64,017 | 0.22 | 0.42 | 65,884 | 0.22 | 0.42 |
| Did binge drink in the last 30 days | 62,480 | 0.17 | 0.38 | 64,329 | 0.18 | 0.38 |
| Did drink in the last 30 days | 59,277 | 0.33 | 0.47 | 60,956 | 0.33 | 0.47 |
| Did smoke in the last 30 days | 62,070 | 0.14 | 0.35 | 63,850 | 0.14 | 0.35 |
| Number of days smoking in last 30 days | 62,070 | 1.92 | 6.21 | 63,850 | 1.92 | 6.21 |
| Number of days drinking in last 30 days | 59,277 | 1.57 | 3.96 | 60,956 | 1.58 | 3.99 |
| Number of days binge drinking in last 30 days | 62,480 | 0.67 | 2.49 | 64,329 | 0.67 | 2.51 |
| Times used marijuana in last 30 days | 64,017 | 3.28 | 9.47 | 65,884 | 3.29 | 9.47 |

Note: The full sample refers to all students in 2011-2019 national YRBS that have no missing variables regarding the socioeconomic background including gender, age, grade, race, and weight. Same criteria apply when defining students who answered the questions regarding cyberbullying.

Table H12: Response rate of questions on mental health and risky behaviors

| | |
|---|--------|
| Cyberbullied in the last 12 months | 0.9706 |
| Sad/hopeless in last 12 months | 0.9876 |
| Did use marijuana in the last 30 days | 0.9747 |
| Did binge drink in the last 30 days | 0.9398 |
| Did drink in the last 30 days | 0.9024 |
| Did smoke in the last 30 days | 0.9491 |
| Number of days smoking in last 30 days | 0.9491 |
| Number of days drinking in last 30 days | 0.9024 |
| Number of days binge drinking in last 30 days | 0.9398 |
| Times used marijuana in last 30 days | 0.9747 |

Note: Sample is from 2011-2019 national YRBS. All observation is included in order to calculate the response rate of each question.

Regarding the second concern (measurement error), we make several remarks. First, if the measurement error is not systematic and the measurement error is in our outcome variables, our estimates will still be unbiased (but with a larger standard error). Second, a more serious concern would be systematic measurement error (i.e., systematic over-reporting or under-

reporting). Regarding this, as we mentioned earlier in the data section, specific measures are taken to ensure the privacy of students and the anonymity of answers in YRBS, and these measures guarantee that students do not have the incentive to lie (over-reporting or under-reporting) in their answers. Nevertheless, there is evidence of under-reporting of marijuana use in other datasets (Mensch & Kandel 1988; Bessa et al. 2010), potentially because marijuana use is arguably accompanied by harsher social criticism than drinking or smoking. We note that under-reporting of risky behaviors might result in a downward bias in our estimates. If this were the case, our estimates would provide lower bounds of the causal effects of cyberbullying on marijuana use. However, this possibility does not negate the causal interpretation.

Regarding the third concern (procedural bias), it is true that questions about smoking behaviors, drinking behaviors, and marijuana use are asked after the questions about cyberbullying in YRBS. Although we cannot directly test whether the effect on risky behaviors in our analyses is indeed caused by procedural bias, we can make use of questions that are related to substance use in YRBS that are asked *before* the questions on cyberbullying (which are thus immune from such procedural bias). We can then test whether these behaviors that are related to substance use follow the same pattern as those that come after the questions on cyberbullying. Luckily, we have one such question that comes before the questions on cyberbullying and bullying in YRBS: “During the past 30 days, how many times did you drive a car or other vehicle when you had been drinking alcohol?” We make use of this question and construct two new dependent variables: “Did drive after drinking during the last 30 days” (0 or 1) and “Times driving after drinking during the last 30 days.” We then estimate the effects of cyberbullying on these two variables using the same PSM method and settings as in our main analyses. These results are shown in **Tables H13 and H14**. We find that being cyberbullied also causes increases in the likelihood of driving after drinking and the number of times driving after drinking. These results suggest that our main results are unlikely to be driven by procedural bias.

Table H13: Effects of cyberbullying victimization on driving after drinking during the last 30 days, propensity score matching

| | (1) Did drive after drinking during the last 30 days | (2) Times driving after drinking during the last 30 days |
|------------------------------------|---|---|
| Cyberbullying victimization | | |
| ATT | 0.034*** | 0.121*** |
| S.E. | 0.003 | 0.008 |
| T-statistics | 11.21 | 15.65 |
| Observations on support | 446,579 | 446,516 |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Estimates were obtained using PSM using nearest-neighbor matching (one-to-one matching with replacement). Variables included in the propensity score are state and year-fixed effects, and individual-level and state-level variables. Individual-level variables include dummies for age, grade, gender, race, underweight, overweight, and obese. The state-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, beer tax, cigarette tax, dummies for state anti-bully laws and anti-cyberbully laws, medical marijuana legalization, and recreational marijuana legalization status. Additionally, when calculating the propensity of cyberbullying victimization, the variable bullying victimization is included.

Table H14: Test for balancing of covariates, propensity score matching, driving after drinking during the last 30 days

| | Pseudo R-Squared | Mean of absolute standardized bias (%) | Median of absolute standardized bias (%) |
|---|------------------|--|--|
| Did drive after drinking during the last 30 days/ Times driving after drinking during the last 30 days | | | |
| Before matching (unmatched) | 0.237 | 4.8 | 1.7 |
| After Matching (matched) | 0.001 | 0.6 | 0.5 |

Notes: Statistics were calculated using Stata routine “psmatch2” developed by Leuven and Sianesi (2018).

g. Using alternative matching methods

On top of the main PSM matching method (nearest-neighbor one-to-one matching), we additionally used 5 more matching methods to assess whether our results are sensitive to different matching methods (Caliendo and Kopeinig, 2008):

- 1) nearest-neighbor 1-to-1 matching with a caliper of 0.01;
- 2) nearest-neighbor 1-to-1 matching with a caliper of 0.001;
- 3) nearest-neighbor 1-to-5 matching;
- 4) nearest-neighbor 1-to-5 matching with a caliper of 0.01;
- 5) nearest-neighbor 1-to-5 matching with a caliper of 0.001.

As shown in **Table H15**, these results are remarkably similar to our main results.





Table H15. Effects of cyberbullying victimization and bullying victimization on substance use, different matching methods

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|--|-------------------------------------|---|---------------------------|---|---------------------------|--|-----------------------------------|--------------------------------------|
| | Did binge drink in the last 30 days | Number of days binge drinking in last 30 days | Did drink in last 30 days | Number of days drinking in last 30 days | Did smoke in last 30 days | Number of days smoking in last 30 days | Did use marijuana in last 30 days | Times used marijuana in last 30 days |
| Main specification: | | | | | | | | |
| nearest-neighbor 1-to-1 matching | | | | | | | | |
| ATT | 0.092*** | 0.610*** | 0.149*** | 1.330*** | 0.084*** | 1.254*** | 0.114*** | 1.739*** |
| S.E. | 0.008 | 0.043 | 0.007 | 0.050 | 0.004 | 0.071 | 0.005 | 0.092 |
| T-statistics | 11.00 | 14.12 | 21.33 | 26.53 | 19.77 | 17.69 | 21.15 | 18.98 |
| Observations on support | 252,569 | 252,569 | 549,314 | 549,314 | 574,720 | 574,720 | 580,227 | 580,227 |
| Specification 2: | | | | | | | | |
| nearest-neighbor 1-to-1 matching with a caliper of 0.01 | | | | | | | | |
| ATT | 0.092*** | 0.610*** | 0.149*** | 1.329*** | 0.084*** | 1.254*** | 0.114*** | 1.739*** |
| S.E. | 0.008 | 0.043 | 0.007 | 0.050 | 0.004 | 0.071 | 0.005 | 0.092 |
| T-statistics | 11.00 | 14.12 | 21.33 | 26.52 | 19.77 | 17.69 | 21.15 | 18.98 |
| Observations on support | 252,568 | 252,568 | 549,313 | 549,313 | 574,720 | 574,720 | 580,221 | 580,221 |
| Specification 3: | | | | | | | | |
| nearest-neighbor 1- | | | | | | | | |

| to-1 matching with a caliper of 0.001 | | | | | | | | | |
|--|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|--|
| ATT | 0.092*** | 0.594*** | 0.149*** | 1.319*** | 0.084*** | 1.251*** | 0.114*** | 1.730*** | |
| S.E. | 0.008 | 0.043 | 0.007 | 0.050 | 0.004 | 0.071 | 0.005 | 0.091 | |
| T-statistics | 10.95 | 13.78 | 21.37 | 26.45 | 19.80 | 17.73 | 21.37 | 19.05 | |
| Observations on support | 252,376 | 252,376 | 549,087 | 549,087 | 574,541 | 574,541 | 580,053 | 580,053 | |
| Specification 4: nearest-neighbor 1-to-5 matching | | | | | | | | | |
| ATT | 0.087*** | 0.547*** | 0.142*** | 1.231*** | 0.084*** | 1.195*** | 0.109*** | 1.679*** | |
| S.E. | 0.004 | 0.027 | 0.004 | 0.029 | 0.002 | 0.039 | 0.003 | 0.051 | |
| T-statistics | 20.44 | 19.95 | 39.89 | 42.50 | 36.29 | 31.02 | 39.28 | 32.86 | |
| Observations on support | 252,569 | 252,569 | 549,314 | 549,314 | 574,720 | 574,720 | 580,227 | 580,227 | |
| Specification 5: nearest-neighbor 1-to-5 matching with a caliper of 0.01 | | | | | | | | | |
| ATT | 0.087*** | 0.549*** | 0.142*** | 1.229*** | 0.084*** | 1.193*** | 0.109*** | 1.679*** | |
| S.E. | 0.004 | 0.027 | 0.004 | 0.029 | 0.002 | 0.039 | 0.003 | 0.051 | |
| T-statistics | 20.44 | 20.10 | 39.88 | 42.44 | 36.22 | 30.87 | 39.26 | 32.85 | |
| Observations on support | 252,568 | 252,568 | 549,313 | 549,313 | 574,720 | 574,720 | 580,221 | 580,221 | |
| Specification 6: nearest-neighbor 1-to-5 matching with a caliper of 0.001 | | | | | | | | | |
| ATT | 0.087*** | 0.538*** | 0.142*** | 1.215*** | 0.083*** | 1.180*** | 0.110*** | 1.680*** | |
| S.E. | 0.004 | 0.025 | 0.004 | 0.029 | 0.002 | 0.038 | 0.003 | 0.051 | |
| T-statistics | 20.81 | 21.60 | 40.12 | 41.85 | 35.54 | 30.68 | 39.82 | 33.12 | |
| Observations on support | 252,376 | 252,376 | 549,087 | 549,087 | 574,541 | 574,541 | 580,053 | 580,053 | |

Notes: *** p<0.01, ** p<0.05, * p<0.1. Estimates were obtained using PSM using different matching methods. Variables included in propensity score are state and year fixed effects, individual-level and state-level variables. Individual-level variables include dummies for age, grade, gender, race, underweight, overweight, obese. The state-level controls include state median income, unemployment rate, per-pupil education spending, pupil-teacher ratio, beer tax, cigarette tax, dummies for state anti-bully laws and anti-cyberbully laws, medical marijuana legalization, and recreational marijuana legalization status. Additionally, when calculating the propensity of cyberbullying victimization, the variable bullying victimization is included.

Appendix I. The scenarios of the post-hoc study
Table I1³. Experimental design (observed by an audience and temporally separated)

| Scenario design 1 | | Temporally separated | |
|-------------------------|-----------------------------------|---|---|
| | | High | Low |
| Observed by an audience | Large audience (High publicity) | John is a high school student and member of swim team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram with his 300 followers. The teammate further commented on the post " <u><i>So gay!</i></u> " The comment was <u>posted 30 days ago.</u> | John is a high school student and member of swim team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram with his 300 followers. The teammate further commented on the post " <u><i>So gay!</i></u> " The comment was <u>posted an hour ago.</u> |
| | Photo used in above two vignettes |  |  |
| | Small audience (Low publicity) | John is a high school student and member of swim team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram with his 2 followers. The teammate further commented on the post " <u><i>So gay!</i></u> " The comment was <u>posted 30 days ago.</u> | John is a high school student and member of swim team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram with his 2 followers. The teammate further commented on the post " <u><i>So gay!</i></u> " The comment was <u>posted an hour ago.</u> |
| | Photo used in above two vignettes |  |  |
| Manipulation check | Cyberbullying | Do you think John was cyberbullied? (Yes/No) | |
| | Observed by an audience | How many others might observe the comment and the photo? (2/100/200/300) | |
| | Temporally separated | When was the comment posted? (an hour ago/ a day ago/ 3 days ago/5 days ago) | |

³ All pictures in the experiments were created by the AI tool Stable Diffusion and the software Photoshop.

Table I2. Experimental design (joined in by an audience and temporally separated)




| Scenario design 2 | | Temporally separated | |
|-----------------------|-----------------------------------|---|---|
| | | High | Low |
| Joined by an audience | Large audience (High publicity) | John is a high school student and member of wrestling team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram. The teammate further commented on the post " <u>So gay</u> ". Following that, <u>300</u> other students <u>commented</u> on the photo. The teammate's comment was <u>posted 30 days ago</u> . | John is a high school student and member of wrestling team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram. The teammate further commented on the post " <u>So gay</u> ". Following that, <u>300</u> other students <u>commented</u> on the photo. The teammate's comment was <u>posted an hour ago</u> . |
| | Photo used in above two vignettes |  |  |
| | Small audience (Low publicity) | John is a high school student and member of wrestling team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram. The teammate further commented on the post " <u>So gay</u> ". Following that, <u>2</u> other students <u>commented</u> on the photo. The teammate's comment was <u>posted 30 days ago</u> . | John is a high school student and member of wrestling team. One of his teammates took a photo of him after practice (as shown below) and shared it on Instagram. The teammate further commented on the post " <u>So gay</u> ". Following that, <u>2</u> other students <u>commented</u> on the photo. The teammate's comment was <u>posted an hour ago</u> . |
| | Photo used in above two vignettes |  |  |
| Manipulation check | Cyberbullying | Do you think John was cyberbullied? (Yes/No) | |
| | Joined in by an audience | How many others commented on the photo? (2/100/200/300) | |
| | Temporally separated | When was the teammate's comment posted? (an hour ago/ a day ago/ 3 days ago/5 days ago) | |

Table I3. Experimental design (observed by an audience and spatially separated)

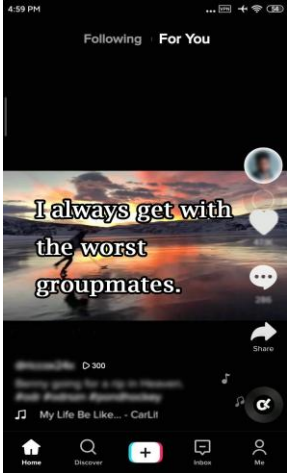
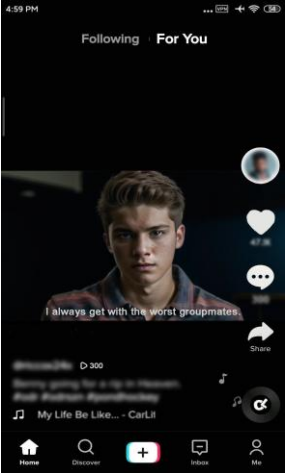

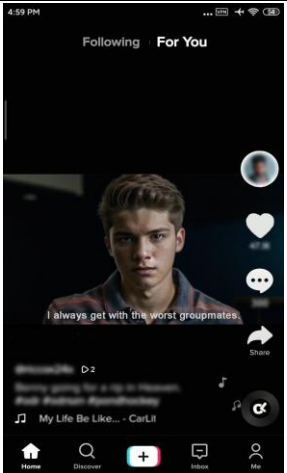
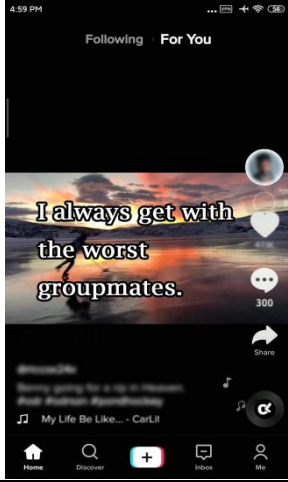
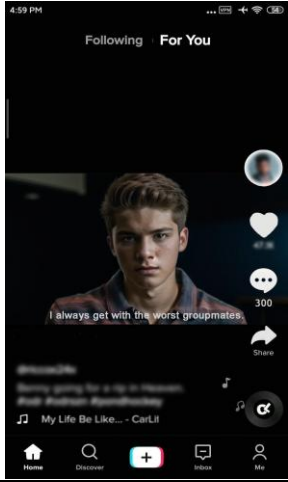
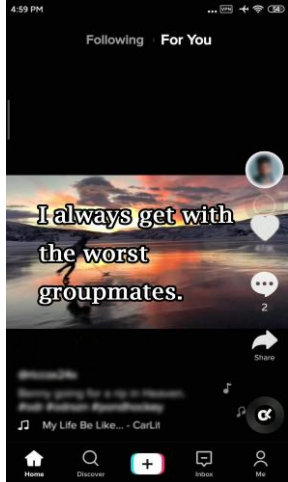
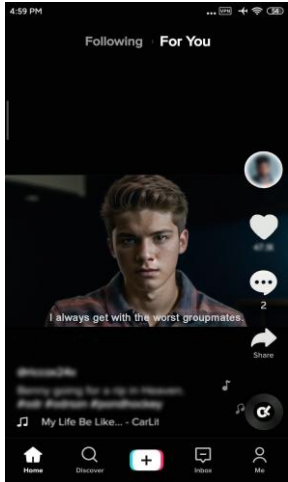
| Scenario design 3 | | Spatially separated | |
|-------------------------|----------------------------------|---|--|
| | | High | Low |
| Observed by an audience | Large audience (High publicity) | <p>Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) reposted by one of her groupmates that reads: "<u><i>I always get with the worst groupmates.</i></u>" Emma noticed that the TikTok was <u>viewed</u> by 300 others.</p> | <p>Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) posted by one of her groupmates who says: "<u><i>I always get with the worst groupmates.</i></u>" Emma noticed that the comment was <u>viewed</u> by 300 others.</p> |
| | Photo used in the above vignette |  |  |
| | Small audience (Low publicity) | <p>Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) reposted by one of her groupmates that reads: "<u><i>I always get with the worst groupmates.</i></u>" Emma noticed that the TikTok was <u>viewed</u> by 2 others.</p> | <p>Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) posted by one of her groupmates who says: "<u><i>I always get with the worst groupmates.</i></u>" Emma noticed that the TikTok was <u>viewed</u> by 2 others.</p> |
| | Photo used in the above vignette |  |  |
| Manipulation check | Cyberbullying | Do you think Emma was cyberbullied? (Yes/No) | |
| | Observed by an audience | How many others might observe the comment? (2/100/200/300) | |
| | Spatially separated | Did Emma see the face of the teammate when reading the post? (Yes/No) | |

Table I4. Experimental design (joined in by an audience and spatially separated)

| Scenario design 4 | | Spatially separated | |
|--------------------------|----------------------------------|---|---|
| | | High | Low |
| Joined in by an audience | Large audience (High publicity) | Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) reposted by one of her groupmates that reads: " <u><i>I always get with the worst groupmates.</i></u> " Emma noticed that the TikTok was commented by 300 others. | Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) posted by one of her groupmates who said: " <u><i>I always get with the worst groupmates.</i></u> " Emma noticed that the TikTok was commented by 300 others. |
| | Photo used in the above vignette |  |  |
| | Small audience (Low publicity) | Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) reposted by one of her groupmates that reads: " <u><i>I always get with the worst groupmates.</i></u> " Emma noticed that the TikTok was commented by 2 others. | Emma is a high school student. Emma worked with four other classmates on a project for their math class that lasted the whole semester. She was the only team member who missed an important meeting. The team received a low score for their project. Emma saw a TikTok (shown in the picture below) posted by one of her groupmates who said: " <u><i>I always get with the worst groupmates.</i></u> " Emma noticed that the TikTok was commented by 2 others. |
| | Photo used in the above vignette |  |  |
| Manipulation check | Cyberbullying | Do you think Emma was cyberbullied? (Yes/No) | |
| | Joined in by an audience | How many others commented on the TikTok? (2/100/200/300) | |
| | Spatially separated | Did Emma see the face of the teammate when reading the post? (Yes/No) | |

Dependent Variables

Prompt, scaling, and items for mental distress

Prompt: In the following week, John/Emma experienced some of the negative feelings below. If you were John/Emma, what is the chance that you would have the same feeling?

Scaling: 1=very unlikely; 2=unlikely; 3=neutral; 4=likely; 5=very likely

Anxiety

A1: Felt tense or 'wound up'

A2: Got a sort of frightened feeling as if something awful was about to happen

A3: Worrying thoughts went through my mind

A4: Couldn't sit at ease and relax

A5: Got a sort of frightened feeling like 'butterflies' in the stomach

A6: Felt restless and had to be on the move

A7: Got sudden feelings of panic

Depression

D1: Didn't enjoy things I used to enjoy

D2: Couldn't laugh and see the funny side of things

D3: Couldn't feel cheerful

D4: Felt as if slowed down

D5: Lost interest in personal appearance

D6: Couldn't look forward with enjoyment to things

D7: Couldn't enjoy a good book or radio or TV program

Prompt, scaling, and items for substance use

- If you were John/Emma, which of the following method you are likely to use to cope with the incident? (Choices are distributed to different participants in random order)
- Doing nothing
- Using substance (i.e., alcohol, cigarette, marijuana)
- Participating in physical activities
- Seeking help from others
- Retaliating against the perpetrator
- Avoiding cyber world
- Other (please specify)

Notes: (1) Following Zigmond & Snaith (1983), levels of anxiety/depression for each respondent were measured by adding the seven scores of all the corresponding items. As shown above, scaling for each item is 1 to 5. Hence, the range for levels of anxiety/depression is 7 to 35, in which 7 indicates the lowest level of anxiety/depression and 35 indicates the highest level of anxiety/depression. (2) Respondents can choose multiple answers in the substance use question. When a response includes the option "Using substance (i.e., alcohol, cigarette, marijuana)", the response will be coded as "1" (use substances). Otherwise, the response will be coded as "0" (not use substances).

References

- Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, *113*(1), 151–184.
- Anderson, D. M., Hansen, B., & Rees, D. I. (2015). Medical marijuana laws and teen marijuana use. *American Law and Economics Review*, *17*(2), 495–528.
- Anderson, D. M., & Rees, D. I. (2014). The legalization of recreational marijuana: How likely is the worst-case scenario? *Journal of Policy Analysis and Management*, *33*(1), 221–232.
- Beckman, L., Hagquist, C., & Hellström, L. (2012). Does the association with psychosomatic health problems differ between cyberbullying and traditional bullying? *Emotional and Behavioural Difficulties*, *17*(3–4), 421–434.
- Beran, T., & Li, Q. (2007). The relationship between cyberbullying and school bullying. *Journal of Student Wellbeing*, *1*(2), 15–33.
- Bessa, M. A., Mitsuhiro, S. S., Chalem, E., Barros, M. M., Guinsburg, R., & Laranjeira, R. (2010). Underreporting of use of cocaine and marijuana during the third trimester of gestation among pregnant adolescents. *Addictive Behaviors*, *35*(3), 266–269.
- Bottino, S. M., Bottino, C. M., Regina, C. G., Correia, A. V., & Ribeiro, W. S. (2015). Cyberbullying and adolescent mental health: Systematic review. *Cadernos de Saúde Pública*, *31*(3), 463–475.
- Brener, N. D., Kann, L., Shanklin, S., Kinchen, S., Eaton, D. K., Hawkins, J., & Flint, K. H. (2013). Methodology of the youth risk behavior surveillance system—2013. *Morbidity and Mortality Weekly Report: Recommendations and Reports*, *62*(1), 1–20.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.
- Campos-Mercade, P., Meier, A. N., Schneider, F. H., & Wengström, E. (2021). Prosociality predicts health behaviors during the COVID-19 pandemic. *Journal of Public Economics*, *195*, 104367.
- Carpenter, C., & Cook, P. J. (2008). Cigarette taxes and youth smoking: New evidence from national, state, and local Youth Risk Behavior Surveys. *Journal of Health Economics*, *27*(2), 287–299.
- Cassiani-Miranda, C. A., Campo-Arias, A., & Caballero-Domínguez, C. C. (2022). Factors associated with cyberbullying victimisation among Colombian high-school adolescents. *Journal of Child & Adolescent Trauma*, *15*(1), 27–36.
- Cénat, J. M., Blais, M., Hébert, M., Lavoie, F., & Guerrier, M. (2015). Correlates of bullying in Quebec high school students: The vulnerability of sexual-minority youth. *Journal of Affective Disorders*, *183*, 315–321.
- Choo, E. K., Benz, M., Zaller, N., Warren, O., Rising, K. L., & McConnell, K. J. (2014). The impact of state medical marijuana legislation on adolescent marijuana use. *Journal of Adolescent Health*, *55*(2), 160–166.
- Chu, Y. W. L. (2014). The effects of medical marijuana laws on illegal marijuana use. *Journal of Health Economics*, *38*, 43–61.
- DeCicca, P., Kenkel, D., & Mathios, A. (2008). Cigarette taxes and the transition from youth to adult smoking: Smoking initiation, cessation, and participation. *Journal of Health Economics*, *27*(4), 904–917.

- Dee, T. S. (1999). State alcohol policies, teen drinking and traffic fatalities. *Journal of Public Economics*, 72(2), 289–315.
- Ebenstein, A., Fan, M., Greenstone, M., He, G., & Zhou, M. (2017). New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy. *Proceedings of the National Academy of Sciences*, 114(39), 10384–10389.
- Ferraro, P. J., & Simorangkir, R. (2020). Conditional cash transfers to alleviate poverty also reduced deforestation in Indonesia. *Science Advances*, 6(24), eaaz1298.
- Hansen, B., Sabia, J. J., & Rees, D. I. (2017). Have cigarette taxes lost their bite? New estimates of the relationship between cigarette taxes and youth smoking. *American Journal of Health Economics*, 3(1), 60–75.
- Hao, Z., & Cowan, B. W. (2020). The cross-border spillover effects of recreational marijuana legalization. *Economic Inquiry*, 58(2), 642–666.
- Harper, S., Strumpf, E. C., & Kaufman, J. S. (2012). Do medical marijuana laws increase marijuana use? Replication study and extension. *Annals of Epidemiology*, 22(3), 207–212.
- Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, 140(4), 1073–1137.
- Leuven, E., & Sianesi, B. (2018). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Retrieved from <https://econpapers.repec.org/software/bocbocode/S432001.htm>
- Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, 86(1), 114–121.
- Lynne-Landsman, S. D., Livingston, M. D., & Wagenaar, A. C. (2013). Effects of state medical marijuana laws on adolescent marijuana use. *American Journal of Public Health*, 103(8), 1500–1506.
- Mazrekaj, D., De Witte, K., & Cabus, S. (2020). School outcomes of children raised by same-sex parents: Evidence from administrative panel data. *American Sociological Review*, 85(5), 830–856.
- Mensch, B. S., & Denise, B. K. (1988). Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Public Opinion Quarterly*, 52(1), 100–124.
- Nikolaou, D. (2017). Does cyberbullying impact youth suicidal behaviors? *Journal of Health Economics*, 56, 30–46.
- Nixon, C. L. (2014). Current perspectives: The impact of cyberbullying on adolescent health. *Adolescent Health, Medicine and Therapeutics*, 5, 143–158.
- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2), 187–204.
- Pacula, R. L., Powell, D., Heaton, P., & Sevigny, E. L. (2015). Assessing the effects of medical marijuana laws on marijuana use: The devil is in the details. *Journal of Policy Analysis and Management*, 34(1), 7–31.
- Parris, L., Lannin, D. G., Hynes, K., & Yazedjian, A. (2022). Exploring social media rumination: Associations with bullying, cyberbullying, and distress. *Journal of Interpersonal Violence*, 37(5–6), NP3041–NP3061.

- Pavlou, P., Liang, H., & Xue, Y. (2007). Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective. *MIS Quarterly*, 31(1), 105–136.
- Pettalia, J. L., Levin, E., & Dickinson, J. (2013). Cyberbullying: Eliciting harm without consequence. *Computers in Human Behavior*, 29(6), 2758–2765.
- Rees, D. I., Sabia, J. J., & Kumpas, G. (2020). Anti-bullying laws and suicidal behaviors among teenagers (No. w26777). National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w26777>
- Schneider, S. K., O'Donnell, L., Stueve, A., & Coulter, R. W. S. (2012). Cyberbullying, school bullying, and psychological distress: A regional census of high school students. *American Journal of Public Health*, 102(1), 171–177.
- Sourander, A., Brunstein, K. A., Ikonen, M., Lindroos, J., Luntamo, T., Koskelainen, M., Ristkari, T., & Helenius, H. (2010). Psychosocial risk factors associated with cyberbullying among adolescents: A population-based study. *Archives of General Psychiatry*, 67(7), 720–728.
- Staiger, D., & Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557–586.
- Stock, J. H., & Yogo, M. (2002). Testing for weak instruments in linear IV regression. National Bureau of Economic Research. Retrieved from <https://www.nber.org/papers/t0284.pdf>
- Tsai, L. L., Morse, B. S., & Blair, R. A. (2020). Building credibility and cooperation in low-trust settings: Persuasion and source accountability in Liberia during the 2014–2015 Ebola crisis. *Comparative Political Studies*, 53(10–11), 1582–1618.
- Wang, J., Nansel, T. R., & Iannotti, R. J. (2011). Cyber and traditional bullying: Differential association with depression. *Journal of Adolescent Health*, 48(4), 415–417.
- Wärneryd, E. (1996). Risk attitudes and risky behavior. *Journal of Economic Psychology*, 17(6), 749–770.
- Wen, H., Hockenberry, J. M., & Cummings, J. R. (2015). The effect of medical marijuana laws on adolescent and adult use of marijuana, alcohol, and other substances. *Journal of Health Economics*, 42, 64–80.
- Worsley, J. D., McIntyre, J. C., & Corcoran, R. (2018). Cyberbullying victimisation and mental distress: Testing the moderating role of attachment security, social support, and coping styles. *Emotional & Behavioural Difficulties*, 24(3), 1–16.
- Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370.