Referee Comments

1 Response to comments from Referee #1

1.1 Suggestion for Improvement

> In the section about spillover analysis using Campus Safety and Security (CSS) data, the author has not specified if they used arrests for liquor law violations, disciplinary actions for liquor law violations, or some (weighted) average of both. Under Clery Act 1990, both the number of arrests and the number of disciplinary actions in residence halls (and elsewhere) are collected and reported by colleges and universities. However, the author has failed to specify which of the two they have used in this section of the paper. Not only should the author clearly specify which of these two variables is being used for the analysis in this section, but I would suggest that the author run separate analyses using each of the two variables (the number of arrests and disciplinary actions due to liquor law violations in residence halls) so that the reader can get some additional insights.

This is an excellent suggestion as the delineation between disciplinary actions and arrests is important for interpretation, clarity, and replication purposes. In the reviewed draft, the Campus Safety and Security (CSS) analysis uses disciplinary actions for liquor law violations and excludes arrests. As suggested, I have extended the analysis to include both disciplinary actions and arrests for liquor law violations in Appendix B and updated the final two paragraphs of Section 5.2 with additional commentary. Both of these extensions are further explained in the following paragraphs.

In Table 1 (Appendix B: Table B1 in the paper), I delineate between disciplinary actions and arrests in the CSS data for alcohol offenses by adding Columns 4 and 5 which splits all reports and residence hall incidents respectively. Columns 4 and 5 show that there is little evidence of arrests changing at universities, regardless of the location. In particular, for each additional moratorium day in a calendar year, alcohol arrests per-25000 enrolled students do not exhibit any statistically significant decrease and the point estimates are relatively small when compared to disciplinary actions.

Given the lack of significance and small point estimates in the arrest columns, I update the final two paragraphs of Section 5.2 to address the delineation between arrests and disciplinary actions and save the additional analysis of arrests for Appendix B. However, I am happy to move the analysis to the main paper upon editor's request. The following is an excerpt from the updated Section 5.2 in the main paper:

> As the second set of analysis, I analyze the CSS data to examine if students substitute partying at fraternity houses to different on-campus locations during moratoriums. The CSS data contains all disciplinary actions and arrests corresponding to liquor law violations in addition to reports of sexual assaults that occur in a calendar-year...

The extension continues in the final paragraph where I more directly assert the findings:

> Using the CSS data, there is evidence that moratoriums move drinking from fraternity houses to residence halls. Residence halls show a 0.270 *increase* in yearly disciplinary actions of alcohol offenses for each additional moratorium day in a calendar-year. Interestingly, this is accompanied by a 0.033 *decrease* in yearly residence hall sexual assault reports.

Note that I also update the percentages in the sentence "on average, the Daily Crime Logs contain approximately 30% and 50% of the yearly alcohol offenses and sexual assaults re-

ported in the CSS data respectively" to account for the addition of arrests in the CSS data analysis.

Finally, the text of Appendix B more closely describes the CSS data used and the corresponding results.

2 Response to comments from Referee #2

2.1 Strength of Results

2.2 Suggestion #1: Weighting

> To weight by total enrollment (or perhaps undergraduate enrollment if you feel that is more appropriate). Larger schools should have less residual variation, so weighting by size should reduce standard errors. The variance in school size should be large enough for this to make a difference.

Thank you for the thoughtful suggestion on how to tighten/strengthen the main results of the paper. In Table 2 (Table BLANK in the main paper) I have included a column for the main specification weighted by total enrollment (Column 3). As recommended, I have also explored weighting by undergraduate enrollment which gives similar results to total enrollment. However, I hesitate to include weights as the default for two main reasons.

First, the weights, unfortunately, do not tighten the results to a significant degree. Columns 2 and 3 of Table 2 show the preferred specification, with the latter weighted by total enrollment. As displayed in Panel A, the standard errors and point estimates are nearly identical for alcohol offenses, exhibiting a 26.5% and 27.8% decrease from the mean respectively. Similarly in Panel B, sexual assaults exhibit a 20% decrease (statistically insignificant) and a 16% decrease (10% significance) in unweighted and weighted estimations. While the weighting enhances the sexual assault results to a small degree, I would prefer to be as conservative as possible, especially when discussing the effect alcohol offenses, which is the strongest and most robust result of the paper.

Second, while I am pleased that the results are consistent, and therefore the model likely specified well, I am apprehensive that weighting all specifications will invite skepticism of the results. Hence, I believe that including the main specification, weighted and unweighted

as shown in Columns 2 and 3, is a great solution to convince readers of the results and, in addition, follows the empirical suggestion of reporting both estimates as outlined in Solon, Haider, and Wooldridge (2015). However, upon request, I am more than happy to report the main results (weighted) from Table 2 in an online appendix table.

2.3 Suggestion #2: Leverage IFC Variation

> To interact the treatment variable with the share of students who are in IFC fraternities. Perhaps I'm missing something, but I would expect the effect size to be essentially proportional to the proportion in IFC fraternities. To be clear, I'm suggesting replacing Moratorium (in equation 1) with Moratorium * Fraction_IFC

3 Response to comments from Referee #4

3.1 Major Suggestion 1: Leverage IFC Variation

> Currently, the main specification treats all moratoriums the same. However, presumably a moratorium at a school with 1.3% of the body in IFC fraternities is less impacted by a moratorium than one with 10.2% IFC membership. Can you provide heterogeneity analysis that probes this difference in intensity of treatment? This could be accomplished by interacting moratorium with share IFC or looking at quartiles of share IFC etc.

3.2 Major Suggestion 2: Separating Effect of Triggering Event

> One interesting result is that impacts of moratoriums seem larger following a student death (Figure 8). This begs the question of whether the incident (death) changes behavior for a short time, or whether the moratorium changes behavior. I would like to see whether a death without a moratorium causes a change in alcohol and sexual assault violations. Can you test this hypothesis by considering the 64 days (avg moratorium length) after these deaths as treated? Note – this suggestion is distinct from the test conducted in Appendix C7. I want to know whether a death causes a treatment effect absent a moratorium. Your test in C7 shows a related but separate idea—that moratoriums still cause a treatment effect when compared to colleges that have deaths but no moratoriums. Given that a large portion of the observations for a school with only a death would still be considered "untreated" (days prior to the event or days after the treatment period ended) even if it had a moratorium, the test in C7 may be a weak test, and I would encourage you to use more conservative language in describing the results.

3.3 Major Suggestion 3: Representativeness

It would be helpful to place the sample in context relative to the Greek-life ecosystem to comment about generalizability. For instance, what is a typical share IFC across all 4-year US colleges and universities and how does this differ from the sample? Some small private colleges have very large Greek presence in percentage terms (e.g. Depauw University is 70

4 References

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3): 414–27. https://doi.org/10.1162/rest.90.3.414.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. "What Are We Weighting For?" *J. Human Resources* 50 (2): 301–16. https://doi.org/10.3368/jhr.50.2.301.

5 Tables

Table 1: Effect of Moratoriums on Alcohol Offenses and Sexual Assaults: Comparison of Daily Crime Logs and Campus Safety and Security (OLS)

		Campus Safety and Security			
	Daily Crime Logs	Displinary Actions/Reported Crime		Arrests	
	All Reports	All Reports	Residence Halls	All Reports	Residence Halls
	(1)	(2)	(3)	(4)	(5)
Panel A: Alcohol Offenses					
In Moratorium	-0.134*	0.297**	0.270**	-0.022	-0.025
	(0.077)	(0.118)	(0.125)	(0.056)	(0.040)
Observations	220	222	222	222	222
Mean of Dependent Variable	131.861	362.978	343.616	55.961	24.280
FE: Year	X	X	X	X	X
FE: University	X	X	X	X	X
Panel B: Sexual Assaults					
In Moratorium	-0.013	-0.046	-0.033**		
	(0.011)	(0.039)	(0.014)		
Observations	220	222	222		
Mean of Dependent Variable	14.099	28.732	14.444		
FE: Year	X	X	X		
FE: University	X	X	X		

Note:
Standard errors are clustered by university and each offense is defined as offense per-25000 enrolled students per-calendar year. Recall that Daily Crime Logs are the primary source of data used in prior analysis. In this model, the In Moratorium treatment variable is defined as the number of calendar-days that experienced a moratorium in a calendar-year. All Reports columns include the entire Daily Crime Logs/Campus Safety and Security Data (CSS), while Residence Halls is a subset of the CSS. All Reports in the CSS data contains both off-campus and on-campus reports. CSS data does not necessarily need to be reported to the university police and hence, may not show up in the Daily Crime Logs. Columns 2 and 3 refer to disciplinary actions for liquor law violations and reported crime for sexual assaults. Columns 4 and 5 refer to arrests for liquor law violations. Fixed effects include university and year fixed effects
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Effect of Moratoriums on Alcohol Offenses and Sexual Assaults (OLS)

					Specification (2)	
	(1)	(2)	(3)	(4)	Weekends (5)	Weekdays (6)
Panel A: Alcohol Offenses						
In Moratorium	-0.125**	-0.123**	-0.129**	-0.131***	-0.238**	-0.038
	(0.047)	(0.051)	(0.050)	(0.046)	(0.106)	(0.026)
Observations	55115	55115	55115	55115	23643	31472
Mean of Dependent Variable	0.464	0.464	0.464	0.464	0.828	0.190
Wild Bootstrap P-Value	0.004	0.010	0.014	0.001	0.014	0.160
Panel B: Sexual Assaults						
In Moratorium	-0.009**	-0.010	-0.008*	-0.007	-0.017*	-0.004
	(0.004)	(0.006)	(0.005)	(0.006)	(0.010)	(0.006)
Observations	55115	55115	55115	55115	23643	31472
Mean of Dependent Variable	0.049	0.049	0.049	0.049	0.058	0.042
Wild Bootstrap P-Value	0.016	0.123	0.097	0.251	0.089	0.504
FE: Holiday	X	X	X	X	X	X
FE: Game Day	X	X	X	X	X	X
FE: Semester (Spring/Fall)	X	X	X		X	X
FE: University	X					
FE: Academic Year	X					
FE: University by Academic Year		X	X		X	X
FE: University by Academic Year by Semester				X		

Note:
Estimates are obtained using OLS. Standard errors shown in parenthesis are clustered by university (37 clusters) and each offense is defined as per-25000 enrolled students. P-values from 1000 wild cluster bootstrap iterations are shown for the In Moratorium coefficient as suggested by Cameron, Gelbach, and Miller (2008) in cases with a small number of clusters (typically lower than 30). This analysis is near, but not below this threshold. Game Day controls consist of university football games within each university. Weekends include Friday-Sunday while Weekdays include Monday-Thursday. Column 2 is the preferred specification due to the flexibility of the fixed effects and the conservativeness of the estimates. Column 3 is specification 2 weighted by total enrollment. Column 3 represents the only weighted specification. Significance stars correspond to clustered standard errors.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$