Data

Data

The main analysis uses data from a variety sources. In particular, I construct a novel crime data set that links incidence-level crime from university police departments, fraternity moratorium dates, and university characteristics over a six-year period (2014-2019).

Daily incidences of offenses are collected from Daily Crime Logs maintained by the 38 university's police departments resulting in approximately 500000 distinct incidences. The Daily Crime Log is an incidence-level source of information; each crime log contains the date occurred, date reported, time occurred, time reported, a short summary of the incident, the general location of the incident, and a distinct case number (see Figure **??** for an example).[1] The Daily Crime Log contains the universe of incidents that are reported by (or to) the university-specific police department. Hence, each of the incidences listed in these logs represent incidences that occurred on or nearby university property.[2] There are two main advantages of the Daily Crimes Logs over readily available crime data sources such as the National Incidence-Based Reporting System (NIBRS), Uniform Crime Reporting System (UCR), and the Campus Safety and Security Data (CSS). First, each university police department is mandated under the Clery Act to maintain and make available a Daily Crime Log. Crime logs must be kept for seven years, although this mandate is subject to each university's interpretation.[3] Hence, only one university is missing data from a complete calendar-year.[4] Second, the Daily Crime Logs contain all daily incidences of alcohol offenses, drug offenses, and sexual assault offenses reported to or by the police—the primary outcomes used in the main analysis. This is a major advantage as the UCR does not contain alcohol offenses and the NIBRS only contains alcohol violations that end in arrests. Since most violations of underage drinking at universities do not end in arrests, the NIBRS data would under-report the prevalence of alcohol misuse. While the CSS data includes similar information as the Daily Crime Logs,[5] the CSS data is aggregated to the calendar-year level which makes the effect of moratoriums difficult to study given their short-lived nature. See Table **??** for more details on the advantages of the Daily Crime Logs.

The existence of a fraternity moratorium is identified using Google and Lexis Nexis searches, in addition to documents from various fraternity associations. In total, there are 45 fraternity moratoriums in the sample. Importantly, these do not represent the universe of fraternity moratoriums that occurred from 2014-2019. In particular, there are five schools that are known to have experienced a moratorium in this time frame, but are excluded due to data issues or their definition of a moratorium.[6] Each moratorium's start and end dates

[1] While the date occurred is technically mandated under the Clery Act to include each of these categories, only 33 of the 38 universities contained the date occurred. However, these five schools contained the date reported.

[2] Sometimes, university police may respond to calls slightly outside of university property. Based on conversations with university police, this is usually when a student is involved.

[3] For instance, if a crime log from 2014 is requested in year 2021, most police departments will have this information as it falls within 7 years. However, some police departments may consider seven-years to be inclusive of their current year, and hence, may only contain records for 2015-2021.

[4] The one university is Rollins College and it is missing data from 2014. North Carolina State University is also missing data, although their missing data spans from January 2014-August 2014.

[5] There are important differences between these two sources. The CSS provides data on liquor and drug violations that occur in residence halls that may not be reported to the police and therefore not appear in the Daily Crime Logs. Hence, an aggregated Daily Crime Log should not (and will not) match the CSS exactly.

[6] Miami University is excluded due to being unable to verify the end-date of their moratorium. Pennsylvannia State University is excluded because they would not digitally release their Daily Crime Logs. University of Texas at Arlington is excluded because the crime logs are scanned images that can not be read reliably by any computer software. Cal State Northridge is excluded because it is unclear whether the moratorium includes a ban on alcohol. University of North Florida is excluded because of a discrepancy between public records information and newspaper articles: newspaper articles claim there is a moratorium

are obtained through public records requests, conversations with Fraternity and Sorority Life advisers, and school newspaper articles. All start and end date are verified by at least one of these sources.[7]

University characteristics such as total enrollment, student demographics, and academic calendars are obtained through the Integrated Postsecondary Education Data System (IPEDS) or directly from the university. However, not all academic calendars for each year in the sample are available. Therefore, only the most current academic calendar found on a university's website is utilized. To account for small changes in academic calendars year-to-year, a seven-day window is added to each start and end date of a semester.[8]

Matching and Harmonization

One of the challenges of using the Daily Crime Logs is their uniqueness to each university. While all crime logs contain daily reports of incidences, each university police department describes their incidences differently. For example, Indiana University's crime log describes driving under the influence as "driving under the influence" while Cal Poly San Luis Obispo's describes this as "dui". As such, there is a lack of harmonization between the crime logs—incidences do not have a standardized way of being reported between university police departments. To mitigate this issue, I use regular expressions to match on typical words, phrases, and abbreviations seen in each crime log for descriptions relating to alcohol offenses, drug offenses, and sexual assaults.[9] For each offense, I use the following definitions for matching the incident descriptions:

- **Alcohol Offense** - Any incident description that refers to a public intoxication, underage drinking, or drinking in an unlawful manner. For instance, public drunkenness, a minor in possession, and driving while intoxicated refer to each of these definitions respectively.
- **Drug Offense** - Any incident description that relates to the possession or use of an illegal drug. Common descriptions may include a "drug incident" or "possession of marijuana".
- **Sexual Assault** - Any incident description that refers to a sexual assault or sex crime including rape and fondling. This corresponds to the types of sex crimes that are reported in the CSS data: rape, statutory rape, incest, and fondling. However, incest sex crimes are omitted as these are infrequent and less likely to be associated between college students.

Table **??** shows the corresponding words, phrases, and abbreviations used to match each incident description to it's corresponding offense. Importantly, each of these phrases are only portions of an incident's description. For instance, the word "sex" is used as a word to match on sexual assaults. The advantage to this method is that the word "sex" will be matched to descriptions such as "sexual assault" or "sex offense" since the word "sex" appears in each of these descriptions. While this is an imperfect method, it is conservative. It is likely that this method is under-counting the true amount of offenses in each category since there are instances in incidence descriptions where words are misspelled (e.g., "aclohol" vs. "alcohol"). Table **??** shows a snapshot of the results of this matching process with the most frequent descriptions matched to each offense.

Descriptive Statistics

Table **??** summarizes the characteristics of the 38 universities and their corresponding ubiquity of offenses and fraternity moratoriums. Panel A shows descriptive statistics of the universities' demographics. On average, the universities are large with total enrollment exceeding 28000. Undergraduates are the majority

beginning 12/4/17, but the public records department says this is untrue. There may exist other universities that experienced a moratorium, but may not have had any sort of news coverage-these are also be excluded from the sample.

[7]There is one exception to this which is the first moratorium at San Diego State University. While the start date has been verified by a newspaper article, the exact end date is a little ambiguous. However, evidence shows that the moratorium ended before the start of the 2015 spring semester, and hence, this is the date used in the analysis. The newspaper article showing this evidence can be seen here: https://newscenter.sdsu.edu/sdsu_newscenter/news_story.aspx?sid=75357.

[8]To define the start of a semester, the first day of instruction is used. For the end of a semester, the finalized grade date is used.

[9]In particular, I found all unique descriptions of incidences in each Daily Crime Log, and then independently analyzed which descriptions matched to each offense.

population with 62% being white. Graduation rates vary substantially between schools and there is particularly large variation in the selectively of each university. For instance, graduation rates and the fraction of students admitted range between 39-95% and 14-94% respectively. Panel B shows summary statistics of the three primary outcome measures: alcohol offenses, drug offenses, and sexual assaults. Each of these outcomes is measured as per-25000 enrolled students per-academic-calendar day. Therefore, the average amount of alcohol offenses per-25000 enrolled students in an academic-calendar day is approximately 0.5. Lastly, Panel C describes characteristics of the 45 moratoriums in the sample. On average, each university experiences approximately one moratorium, although universities can experience up to three. Furthermore, the moratoriums persist for an average of 64 academic-calendar days. Notably, there is significant variation in the length of the moratoriums. In particular, the minimum length of a moratorium is only 6 academic-calendar days while the maximum is 541. Due to this large range, it is important to note that a median moratorium lasts for 46 academic-calendar days (approximately 1.5 months).