The Cupola
Scholarship at Gettysburg College

Student Publications | Student Scholarship

Spring 2019

College Crime and Retention Rates

Abigail R. Hauer
Gettysburg College

College Crime and Retention Rates

Abstract
Increased media attention on college crime has led to greater prioritization of campus safety when selecting a college to attend. This, coupled with society's view of higher education as a necessity to succeed in the labor market, creates a potential tradeoff between safety on campus and future job success. To analyze such tradeoff, I examine whether college crime affects retention rates at four-year American institutions. While literature has focused on college crime and factors that affect the decision to begin attending a college, no study has solely focused on the college crime and the decision to continue attending a college. Using data from the US Department of Education, I estimate the effect of college crime and changing college crime expectations on retention rates from 2009 to 2016 for four-year institutions using linear and nonlinear OLS regressions. Such results have implications for college policies to combat crime on campus not only to keep students safe, but to prevent students from transferring or dropping out. Using an instrumental regression with a proxy for average state temperature, along with fixed effects and interaction terms, I find that college crime expectations and college crime overall have a negative, statistically insignificant effect on retention rates.

Keywords
Retention, College Crime, Econometrics

Disciplines
Behavioral Economics | Higher Education | Social Control, Law, Crime, and Deviance

Comments
Written for ECON 350: Econometrics

College Crime and Retention Rates

Abigail Hauer
Gettysburg College
haueab01@gettysburg.edu

May 1, 2019

Abstract:
Increased media attention on college crime has led to greater prioritization of campus safety when selecting a college to attend. This, coupled with society's view of higher education as a necessity to succeed in the labor market, creates a potential tradeoff between safety on campus and future job success. To analyze such tradeoff, I examine whether college crime affects retention rates at four-year American institutions. While literature has focused on college crime and factors that affect the decision to *begin* attending a college, no study has solely focused on the college crime and the decision to *continue* attending a college. Using data from the US Department of Education, I estimate the effect of college crime and changing college crime expectations on retention rates from 2009 to 2016 for four-year institutions using linear and nonlinear OLS regressions. Such results have implications for college policies to combat crime on campus not only to keep students safe, but to prevent students from transferring or dropping out. Using an instrumental regression with a proxy for average state temperature, along with fixed effects and interaction terms, I find that college crime expectations and college crime overall have a negative, statistically insignificant effect on retention rates.

Acknowledgements:
The author would like to thank Kenzie Horst for helpful suggestions and assisting in data analysis. The author would like to give additional thanks to the referee who provided feedback in the draft process and those who gave helpful comments after the author's presentation.

Honor Code Statement
I affirm that I have upheld the highest principles of honesty and integrity in my academic work and have not witnessed a violation of the Honor Code.

Introduction

The Jeanne Clery Disclosure of Campus Security Policy and Campus Crime Statistics Act, known as the Clery Act, requires colleges that receive federal funding to report campus crime statistics to the US Department of Education (DOE) each year (Gregory and Janosik, 2002). Since the implementation of Clery Act by Congress in 1990, 78 million college students have reported criminal victimization (National Crime Victims' Rights Week, 2017). Of the 78 million students reporting, "a few violent campus incidents highlighted by the media have drawn a spotlight to college and university campuses has created the impression that campuses are increasingly dangerous [places]" (Fisher, 1995). Although the media has portrayed college campuses as increasingly unsafe, society has progressively viewed college education as a necessity to succeed in the American labor market. The safety of campus and societal necessity of a degree forces students to decide whether to continue attending an unsafe school, transfer to another school, or drop out entirely. The following paper investigates the relationship between college crime and retention rates at four-year colleges and universities in the US.

Existing literature has focused on college selection, using survey data to determine the factors that students' value when selecting a college, such as academic programs, location, and cost (Pampaloni, 2010). Students are found to place less importance on campus safety and security than their parents (Warwick and Mansfield, 2003). While students do not prioritize college crime when selecting a college to attend, survey data shows how college crime does impact their behavior on campus based upon their perception of safety (Morrall et al, 2010; Patton and Gregory, 2014), as well as their knowledge of crime on campus from formal and informal crime reports (Gregory and Janosik, 2002; Janosik and Gehring, 2003). However, there are a lack of studies that use national data and solely focus on college crime and its influence on the decision to *begin* attending a school and to *continue* attending a school.

To examine the effect of college crime on retention rates at four-year American institutions, I use national data from the Department of Education (DOE) College Scorecard and Campus Safety and Security datasets to determine the effect of college crime on retention rates at four-year institutions, using both overall annual college crime numbers and changing crime expectations, from 2009 to 2016. I utilize ordinary least squares (OLS) regressions, both linear and nonlinear, controlling for factors that influence college crime and relate to retention rates, specifically average cost of attendance, transfer rates, share of female students, average family income, number of undergraduates, and admission rates to prevent omitted variable bias. I also use school and year fixed effects to capture unobserved institutional characteristics that do not change over time, limiting the effects of omitted variables that are not quantified. I also use an average state temperature ranking instrumental variable to limit any other potential biases.

While college crime has pertinent effects on college policy to keep students safe, such crime may also influence whether students continue attending a school, transfer, or drop out. If a school cannot keep its students safe due to crime on campus, it likely already faces backlash in the media and from its students. However, if a school cannot keep its students safe due to crime on campus *and* if such crime leads to lower retention rates, schools may lose recognition and prestige, diminishing the number of applicants. Over time, high college crime and low retention rates may result in a school losing donors and funding, potentially causing a school to shut down. Overall, college crime affecting retention rates may perpetuate preexisting issues that schools have from college crime, potentially prompting faster and more aggressive institutional policy responses to combat crime and keep students safe.

Initially, I find statistically significant results with my log-log model with school and year fixed effects, specifically for college crime expectation, meaning that students who had an expectation of crime before arriving on campus were likely to transfer or drop out. However, counterintuitive to literature's depiction of the negative effects of college crime, retention rates increase as college crime increases. After implementing an instrumental variable of annual average state temperature rankings, both college crime expectations and college crime overall have a negative effect on retention rates, although such effects are statistically insignificant. Such results depict that students who do and do not have an expectation of college crime before arriving on campus were negatively impacted by such crime, making them likely to transfer or drop out—reducing retention rates.

Literature Review

Literature relevant to college campus crime and retention rates falls into four categories: college selection, college safety/crime, the Clery Act, and college retention. All papers focusing on college selection use survey data, questioning high school students, college students, and parents of college-bound students about the factors that influence the decision to select a college. Pampaloni (2010) finds that students prioritize academic programs, location, and cost when analyzing institutional characteristics. Galotti (1995) finds that males and females weigh the importance of certain school criteria differently: males focus more on prestige traits, specifically reputation, success of graduates, and academic challenge, while females focus more on aesthetic traits, specifically campus appearance and dorms/residence halls. Unable to control for location, my models control for average cost of attendance as a proxy for cost.[1]

[1] Location in the literature is typically interpreted as rural, urban, or suburban.

While there is difference among gender in Galotti's (1995) study, Warwick and Mansfield (2003) find that students and their parents place importance on many similar financial, physical, functional, social, and psychological characteristics. Warwick and Mansfield (2003) find no statistical difference between student and parental weighing of tuition and scholarship (financial), size and location (physical), academics and degrees offered (functional), cultural diversity and social activities (social), and reputation of degree (psychological). However, parents place a statistically significantly higher level of importance on security and safety (Warwick and Mansfield, 2003) and students report they are least satisfied with campus safety and security regarding campus life (Elliott and Healy, 2001). My paper specifically focuses on the influence of campus security and safety on college selection since no individual study has done so.

The second overarching body of literature focuses on campus safety, specifically campus crime. A portion of the literature uses survey data of undergraduate students. Patton and Gregory (2014), using survey data among students of Virginia community colleges, find no statistically significant difference regarding safety between campuses with police departments or security departments—the campus perceived to be the safest had one part-time security guard in the evening. Although most campuses have some form of security, students surveyed report they feel least safe in places they visit alone and do not have a strong security presence, specifically parking lots, walkways, and bathrooms (Patton and Gregory, 2014).

While students find on-campus parking lots unsafe, Cornell (2010), using data from the Virginia State Police, finds that campuses as a whole are safer than other off-campus locations, such as off-campus roads and parking lots. Morrall et al (2010) find from surveying undergraduates in the United Kingdom that women, both victims and non-victims, are most likely to change their behavior to avoid unsafe practices susceptible to crime. Specifically, 45.2% of female victims of crime do not go out after dark, 51.1% avoid going out alone, and 65.4% avoid certain streets on and near campus (Morrall et al, 2010). Although women are more likely to change their behavior to avoid victimization (Morrall et al, 2010), Fisher et al (2000) find from surveying undergraduate women that 27.7 per 1,000 females are raped during an academic year, and that 22.8% of such victims are multiple-rape victims. Despite 2.8% of women raped and 15.5% of women sexually victimized, Fisher et al (2000) finds that less than 5% of such crimes are reported to law enforcement. My models control for the share of female students and determines the influence of college crime and changes in college crime expectations specifically on females by using interaction variables.

While the survey data on college crime speaks volumes, other literature focusing on college crime uses mixed methods of surveys and national data, specifically the Uniform Crime Report (UCR). Volkwein et al (1995) use data from the UCR and Consortium for Higher Education Campus Crime Research (CHECCR) to measure crime and College Board Survey data to measure student body community and organizational characteristics. Volkwein et al (1995) find that the student characteristics (percentages of applicants accepted, students receiving financial aid, males, African Americans, foreign students, and students living in residence halls) explain the greatest amount of variance in violent crime on campus and that the 23 community, organizational, and student variables explain 79% of the variance in property crime. In terms of student characteristics, my models control for share of female students.

While location is one of the community variables Volkwein et al (1995) uses, Morriss (1993) finds that location is not positively related to college crime rates. Morriss (1993), by using the UCR and questionnaires sent to institutional research offices and campus police forces, finds that campuses with higher crime rates are more accessible to people, as determined by square footage of campus per campus population and accessibility to automobiles and public transportation. Morriss (1993) also finds that campuses with higher crime rates utilize more deterrents, such as a larger campus police force relative to campus population and more police involvement with the community, than institutions with lower crime rates. Although wealthier campuses (measured using tuition cost, ratio of total university operating expenditures to campus population, percentage of campus applicants not admitted, percentage of faculty holding tenure, and ratio of students to faculty) can provide more deterrents, Morriss (1993) finds that wealthier campuses offer more opportunities and targets for individuals to commit crime. My models control for such variables, specifically cost.

The third body of literature focuses on student knowledge of the Clery Act and how such knowledge changes student behavior. Using survey data of students at "community colleges, comprehensive colleges, and research universities," Gregory and Janosik (2002) find that 71% of respondents are unaware of the Clery Act, and that 99% of males and 94% of females do not use Clery Act data when making an enrollment decision. Janosik and Gehring (2003) find similar results, with 78% of their two and four-year college student survey respondents not knowing about the Clery Act and only 8% being influenced by Clery Act report data when selecting a college to attend. Students attending private and smaller institutions are 10% more likely to use Clery Act report data when selecting a college (Janosik and Gehring, 2003).

Many surveyed read less formal crime reports, such as flyers and newspaper articles—55% of women and 48% of men surveyed by Gregory and Janosik (2002), and 60% of all surveyed by Janosik and Gehring (2003), read the less formal reports. While Clery Act report data may not be frequently read or significantly influence college selection, less formal reports on the crime data *are* read but have not been studied to determine if such reports influence college selection. My models use crime expectations as an independent variable since changes in college crime may invoke a change in behavior, potentially by transferring or dropping out since some students have a preconceived notion of college crime from reading formal crime reports.

The fourth body of literature focuses on factors that affect college retention rates. Chapman (1986) finds that the decision to matriculate to the next academic year is typically made during the final months of the spring semester. Chapman (1986) outlines how some factors, such as changed family or personal circumstances, or admittance from the wait list of a school highly preferred by the student, may affect a student's decision to continue attending a school. Allen et al (2008), using a previous study's survey data but focusing on third-year students, find that first-year GPA has a strong relationship with whether a student stayed at the school or dropped out (r=0.56). Allen et al (2008) also find that college commitment/motivation and social connectedness have statistically significant positive effects on whether a student stayed or dropped out. My models control of admission rate as a proxy for college commitment.[2]

Although studies have used national data and surveys, no study has *solely* used national data on college crime and characteristics. It may be difficult to control for some of the variables that researchers hypothesize influence crime, such as "the proportion of students living in on-campus dormitories… the number of national fraternities and sororities on the campus… [and] academic quality" as such data is unavailable on Department of Education's (DOE) College Scorecard or other national databases (Fisher, 1995). However, my paper advances the literature on the effects of college crime by being the first to look at national data and determining if such crime affects retention rates, controlling for many other variables and proxies of variables that influence both college crime and retention rates.

[2] A school with a lower admission rate is likely more rigorous and thus requires a greater amount of college commitment.

Theory and Methodology

Literature depicts how college crime changes student behavior and contributes to the decision to *begin* attending a college. College crime, among other factors, also likely affects a student's decision to *continue* attending a college as opposed to substantially changing their behavior by transferring or dropping out. When a student intends to go to a college, they may have an expectation of crime that occurs on the campus and how it will affect them, whether this be from viewing Clery Act data reports or less formal crime reports.[3] If college crime during a student's first or second year on campus differs from college crime when they made the decision to attend the college, such change may affect them differently than they expected, potentially causing the student to transfer or drop out.

I use standard ordinary least squares (OLS) models, checked for robustness, to calculate the effect of change in college crime expectations on retention rates. I model individual four-year college *i*'s retention rate as a function of change in crime expectations:

$$retention\ rate_i = \beta_0 + \beta_1 \Delta crime\ expectations_i + \varepsilon_i \quad (1)$$

The dependent variable, $retention\ rate_i$, measures first-time, full time retention rates at four-year institutions *i*. The independent variable, $\Delta crime\ expectations_i$, captures the change in crime relative to one's crime expectation at four-year institutions *i*.

Some students may not go into college with a preconceived notion as to what crime on campus will be like or how it will affect them, potentially because they did not read crime reports, formal or informal.[4] College crime impacts the retention of these students in a different manner

[3] Janosik and Gehring (2003) find that 8% of students surveyed were influenced by Clery Act report data when selecting a college to attend, and that 60% read less formal crime reports, such as newspaper articles or flyers.
[4] Gregory and Janosik (2002) find that only 6% of female students surveyed and 1% of male students surveyed were influenced by Clery Act report data when selecting a college to attend, and that 55% of female students and 48% of male students read less formal crime reports.

than in equation 1—these students are impacted by college crime that they do not anticipate or have an expectation of when making the decision to attend their school. For example, a student who comes from an area with little crime may not account for how crime will affect her in college since crime has not affected her life pre-college. Once she arrives on campus as a first-year student, crime may impact her in a way that causes her to transfer or drop out.[5] For example, she may be a victim or witness of a crime, or she may know someone who was a victim or a witness of a crime.[6] While equation 1 assumes that students have expectations of crime on campus, I also model individual four-year college *i*'s retention rate as a function of crime to account for those students who do not have an expectation of crime when going into college but are still impacted by such crime when they are on campus:

(2) $$retention\ rate_i = \beta_0 + \beta_1 crime_i + \varepsilon_i$$

where $crime_i$ measures the amount of crime on campus *i* instead of the change in expectations of crime.

I control for other college characteristics to mitigate the effects of omitted variable bias and potential overestimation or underestimation of the effects of change in crime expectations on retention rates. γX_i captures other college characteristics that relate to change in crime expectations and affect retention rates.[7] I add such controls to equations 1 and 2 in equations 4 and 5, respectively.

(3) $$\gamma X_i = \beta_1 cost_i + \beta_2 transfer_i + \beta_3 female_i + \beta_4 income_i + \beta_5 size_i + \beta_6 admis_i$$

[5] For example, a student who is socially disconnected once she arrives on campus as a first-year may transfer or drop out if she did not anticipate or expect such social disconnectedness. See Allen et al (2008) for further discussion of how social connectedness affects whether a student stays or drops out.
[6] In Morrall et al (2010), 22% of 866 undergraduate students surveyed were victims of crime.
[7] I control for wealth with average cost of attendance and average family income. I also control for transfer rate which makes up retention rate, as well as gender breakdown of student population, number of undergraduates, and admission rate of the college. I also include school and year fixed effects. All data comes from the Department of Education's College Scorecard dataset from 2009/2010-2016/2017.

(4) $$retention\ rate_i = \beta_0 + \beta_1 \Delta crime\ expectations_i + \gamma X_i + \varepsilon_i$$

(5) $$retention\ rate_i = \beta_0 + \beta_1 crime_i + \gamma X_i + \varepsilon_i$$

My main model incorporates the independent variables from equation 1 and equation 2, accounting for the changes in crime expectations and the amount of crime on campus. Such model accounts for students who have expectations about crime on campus and how it will affect them, as well as students who have no expectations but are still impacted by crime on campus:

(6) $$retention\ rate_i = \beta_0 + \beta_1 \Delta crime\ expectations_i + \beta_2 crime_i + \gamma X_i + \varepsilon_i$$

where all variables measure the same as in model 1 and include the independent variable, $crime_i$, from equation 2.

To specifically model the independent variable in equation 1 and equation 4, I develop a proxy measurement for change in expectations since there is no actual data on expected college crime. I measure the change in expectations of college crime as $(crime_{i,t} - crime_{i,t-2})$ using college crime data from the DOE Campus Safety and Security Data:

(7) $$\Delta crime\ expectations_i = (crime_{i,t} - crime_{i,t-2})$$

(8) $$retention\ rate_{i,t} = \beta_0 + \beta_1 (crime_{i,t} - crime_{i,t-2}) + \gamma X_i + \varepsilon_i$$

where β_1 represents the percentage point change in retention rate when crime expectations change over two years. Such measurement of change in expected crime delineates how an individual bases their expectations of college crime from when they were a senior in high school choosing to attend a college at time $t-2$ and compare such expectations to crime as a college sophomore at time t. Using a lag of $t-2$ years allows students to have been on campus for at least one full year and gained comprehension of the crime occurring around them.[8]

To model the overall crime rate variable in equation 2 and equation 5, I use the raw number of annual college crimes from the DOE's Campus Safety and Security Data. However, it may be more appropriate to model the independent variable, raw number of annual college crimes, as a percentage change since students are more impacted by a percentage increase in crime rather than *one* additional crime per year:

(9) $$retention\ rate_i = \beta_0 + \beta_1 \ln(crime_i) + \gamma X_i + \varepsilon_i$$

where β_1 would measure the percentage point change in retention rate for campus *i* when college crime increases by 100%. Such nonlinearity within the model may better approximate the data if equation 5 presents a nonlinear relationship between raw college crime and retention rates.

It may also be more appropriate to measure change in crime expectations as a percent change because students are impacted more by a percentage increase in crime rather than an increase in *one* crime over a two-year period:

(10) $$retention\ rate_i = \beta_0 + \beta_1 \ln(\Delta crime\ expectations_i) + \gamma X_i + \varepsilon_i$$

where β_1 represents the percent change in retention rate percentage point when crime expectations increase by 100% over two years for campus *i*. Just as with crime overall, such nonlinearity within the model may better approximate the data if equation 4 presents a nonlinear relationship between college crime expectations and retention rates.

[8] While Allen et al (2008) focuses on retention rates of college students at the beginning of their third year, I choose to use a time lag of 2 years. For most schools, the earliest one can decide where to attend college is December of their senior year and the latest one can decide is May 1 (National College Decision Day). A lag of 2 years would account for students at the end of their second-year fall semester (at the earliest) up to the end of their second-year spring semester. This is off by 1 semester at most to Allen et al (2008) data on students at the beginning of their third year.

It may also be more appropriate to measure both main independent variables as percent changes within the same model. Incorporating the nonlinearity in equation 9 and equation 10, I model equation 6 as:

$$(11) \qquad retention\ rate_i = \beta_0 + \beta_1 \ln(\Delta crime\ expectations_i) + \beta_2 \ln(crime_i) + \gamma X_i + \varepsilon_i$$

where β_1 represents the percent change in retention rate percentage point when crime expectations increase by 100% over two years and β_2 represents the percentage point change in retention rate for campus i when college crime increases by 100%. Again, such nonlinearity may better approximate the relationship between the independent variables, change in college crime expectations and percentage change in college crime, and the dependent variable, retention rate.

It may also be appropriate to measure retention rates as a percentage change, along with the two main independent variables being represented as percentage changes. Incorporating such log-log nonlinearity, I also model equation 6 as:

$$(12) \qquad \ln(retention\ rate_i) = \beta_0 + \beta_1 \ln(\Delta crime\ expectations_i) + \beta_2 \ln(crime_i) + \gamma X_i + \varepsilon_i$$

Although I have accounted for potentially nonlinear relationships between retention rate and college crime, the results of my models are still subject to biases. I control for certain variables that relate to college crime and affect retention rates to prevent omitted variable bias. There are some variables that I am unable to control for because of a lack of data. Specifically, I am unable to control for the location of the institution (rural, suburban, urban) because such data was "NULL" in the DOE College Scorecard dataset, unable to be used.[9]

[9] Volkwein et al (1995) and Morriss (1993) look at the relationship between campus location and college crime. Morriss (1993) discusses how campuses that are more accessible to automobiles and public transportation have higher crime rates, potentially because such campuses are in urban areas that have more public transportation.

While I am unable to control for location, thus overestimating the effect of campus crime on retention rates, I do control for other variables to mitigate such overestimation or underestimation caused by omitted variable bias. Using data from the DOE Scorecard data from 2009/2010-2016/2017, I control for campus wealth via proxies of average cost of attendance at an institution and average family income.[10] I also control for transfer rates for first-time, full time students within 150% of expected graduation (6 years) since transfers and dropouts make up retention rates (or reduce them). I also control for admission rate since a school with a higher admission rate is likely more rigorous, affecting retention rates and relating to college crime.[11] I control for the number of undergraduate students at an institution as the size of the institution may affect crime prevention.[12] I also use individual school fixed effects since such unobservable factors may make a campus more or less susceptible to crime and also affect retention rates.

Females are likely the ones to remove themselves from the sample of colleges as females are disproportionately affected by certain types of crimes, and females are more likely to change their behavior in response to crime—potentially by dropping out of college.[13] To prevent gender from causing an overestimation of the effect of college crime on retention rates, I control for the share of females in the student body at each institution. I also include an interaction term between the share of female students and crime expectation, as well as the share of female students and crime overall:

[10] Morriss (1993) finds that wealthier campuses offer more opportunities and targets for individuals to commit crime, and wealthier campuses may have higher retention rates if students realize they cannot continue to afford tuition.
[11] A school with a higher admission rate is likely more rigorous, which may decrease college crime since students have less time to commit crime. A school with a higher admission rate may also increase retention rate since schools with higher admission rates likely have strong reputations, meaning students will likely want to continue to attend the acclaimed school (Galotti 1993).
[12] Morriss (1993) describes how the ratio of full-time police officers to campus population may affect deterrence of college crime and how a larger campus population is more accessible and visible to criminals. Campus size also affects retention rates as students may attend a school and transfer if they wish to attend a larger/smaller school.
[13] Morrall et al (2010) finds that females are affected more by rape and fondling crimes than men.

$$(13)\ \ln(retention\ rate_i) = \beta_0 + \beta_1 \ln(\Delta crime\ expectations_i) + \beta_2 \ln(crime_i) + \beta_3 female_i * \Delta crime\ expectations_i + \beta_4 female_i * crime_i + \gamma X_i + \varepsilon_i$$

Measurement error may be present within my models. Combining the Campus Safety and Security and College Scorecard datasets may create bias due to measurement error as the Campus Safety and Security Dataset reports crime statistics for the *calendar* year for each campus, while the College Scorecard dataset reports statistics for the *academic* year for each campus. I combine Scorecard data from the year starting the academic year with the year of the Campus Safety and Security data; for example, I merge the 2009/2010 Scorecard data to the 2009 Campus Safety and Security data. However, since I merge consistently with each academic and calendar year from 2009 to 2016, it is unlikely that the difference in time periods for the two datasets causes an overestimate or underestimate the effect of college crime on retention rates.

Total college crime and retention rates may simply be trending over time without the former causing the latter, thus there may be no causal relationship between the two, rather their relationship may be driven by time.[14] To prevent such spurious correlation in my panel data and the underestimation of the effect of college crime on retention rates, I include a year fixed effect in my regressions that include control variables (γX_i).

Retention rates may be impacting college crime rather than college crime impacting retention rates, thus overestimating the effect of college crime on retention rates.[15] To prevent such reverse causality and other omitted variable bias I am unable to control for, I use an average state temperature ranking instrumental variable. Such variable, accessed from the National

[14] Crime may be increasing due to a larger number of students attending college, and retention rates may be increasing due to increased pressure to obtain a college degree to succeed in the employment market.

[15] Low retention rates may represent a school with students' possessing a low commitment. Such students with low commitment to their studies may be more likely to engage in criminal behavior since their time is not being spent doing work or studying.

Oceanic and Atmospheric Administration (NOAA) National Climate Reports from 2009 to 2016, ranks each state's average annual temperature relative to its previous annual average temperatures, with the ranking periods running from 1895 to the year of the climate report. Average state temperature affects college crime as literature speaks to higher temperatures correlating to higher rates of violent and property crime due to more individuals being outside (Field 1992). Average state temperature does not affect retention rates except through college crime since nearly all high school students research the weather of the schools they anticipate attending. All students have an expectation of temperature on campus and such temperature does not drastically vary from year to year, meaning that it is unlikely that temperature and expectations of temperature affect retention rates. Although the rankings on average state temperature are not the raw data on state temperature that I desire, this is the best measure of average state temperature to prevent endogeneity in the error term ε_i.

I model my instrument as:

$$(14) \qquad \ln(\Delta crime\ \widetilde{expectations_i}) = \pi_0 + \pi_1 ranktemp_i + \gamma X_i + v_i$$

$$(15) \qquad \ln(retention\ rate_i) = \beta_0 + \beta_1 \ln(\Delta crime\ \widetilde{expectations_i}) + \beta_2 female_i * \Delta crime\ expectations_i + \gamma X_i + \varepsilon_i$$

$$(16) \qquad \ln(\widetilde{crime_i}) = \pi_0 + \pi_1 ranktemp_i + \gamma X_i + v_i$$

$$(17) \qquad \ln(retention\ rate_i) = \beta_0 + \beta_1 \ln(\widetilde{crime_i}) + \beta_2 female_i * crime_i + \gamma X_i + \varepsilon_i$$

where equations 14 and 16 represent the first-stage instrumental regressions and equations 15 and 17 represent second-stage least-squares results. In each instrumental regression, I include the female and crime interaction term for the respective independent variable. I also include the same controls in my first-stage instrumental regressions in equations 14 and 16 that I include in my previous regressions that include control variables γX_i.

Although there may be bias within my models that I am unable to account for, causing me to overestimate or underestimate the effect of college crime on retention rates, I am able to minimize such bias by controlling for outside factors that relate to college crime and affect retention rates, using school and year fixed effects, and using an instrument. I also run robust regressions since my regression is not likely homoskedastic, making the t values of my coefficients more conservative. Overall, my attempts to limit bias within my models allow me to best estimate the effect of the change in college crime, whether it be change in expectations or change in crime, on retention rates at four-year institutions in the US.

Data

I utilize data from the DOE College Scorecard and Campus Safety and Security Data Analysis Cutting Tool. The College Scorecard data includes panel data on many variables for each institution in the US regarding academia. Such variables are collected after each academic year and provided through federal reporting from institutions, specifically institutions that receive federal financial aid dollars. The Campus Safety and Security Data includes panel data on college crime for each institution in the US broken down by campus, collected annually and provided through federal reporting from institutions per the Clery Act that requires colleges that receive federal funding to report college crime statistics to the US DOE each year (McCallion 2014). For this paper, I utilize College Scorecard data from the 2009-2010 academic year to the 2016-2017 academic year, and Campus Safety and Security Data from 2009 to 2016.

From the College Scorecard dataset, I select identifier variables, including the unit ID for each institution, the institution name, and the city and zip code of the institution. I also choose variables that allow me to use data for the main campuses of four-year institutions, eliminating other campuses of each institution, such as abroad campuses, and non-four-year institutions, such as community colleges and technical schools. Of the 2,146 variables, I choose 22 variables that affect retention rates and relate to college crime.[16] Since numerous variables are not included every year or privacy suppressed, I restrict my regression analysis to exclude such values.

From the Campus Safety and Security Dataset, I select similar identifier variables, including the unit ID, the institution name, and the state and zip code of the institution. I also eliminate non-four-year institutions and non-main campuses, including the sector of the institution labeling each school as a public or private two-year or four-year institution. I choose variables to measure crime on each campus, specifically the number of murders, negligent manslaughters, rapes, fondling instances, incestuous instances, statutory rapes, robberies, aggravated assaults, burglaries, motor vehicle thefts, and arsons.[17] I sum the values of each crime variable to generate a total crime variable.[18] I then create a two-year lag variable and subtract it from the total crime variable to represent change in crime expectations, overall using 13 variables from the dataset.[19]

I then merge the College Scorecard and Campus Safety and Security Data Analysis set by using each institutions' unit ID and the corresponding year for each data point. The merged dataset contains 1,902 observations (each a college-year pair) with 66 variables, including identification variables such as unit ID and institution name.[20] I specifically focus on 10 variables in my regressions: retention rate, crime expectation, total crime, average cost of attendance, transfer rate,

[16] Such variables include: type of school; state; admission rate; number of undergraduate; number of white, black, Hispanic, Asian, American Indian, Pacific Islander, 2+ races, non-resident Aliens, or unknown race undergraduates; average cost of attendance; transfer rate; number of undergraduate men; number of undergraduate women; average SAT score; average age of entry; share of female students; share of first generation students; average family income.
[17] Data on rape, fondling, incest, and statutory rape were not available until 2014.
[18] The total crime variable includes murder, negligent manslaughter, robbery, aggravated assault, burglary, motor vehicle theft, and arson from 2009-2013. The total crime variable includes murder, negligent manslaughter, rape, fondling, incest, statutory rape, robbery, aggravated assault, burglary, motor vehicle theft, and arson from 2014-2016.
[19] The lag data was only available from 2011-2016 since 2009 and 2010 could not be lagged with a 2-year lag.
[20] Some of the college-year pairs were not for the entire 2009-2016 span of the dataset due to missing either missing data from the College Scorecard dataset or the Campus Safety and Security Dataset.

share of female students, average family income, state, and number of undergraduates. I then run OLS regressions, both linear and nonlinear, some with school and year fixed effects.

Table 1 presents the descriptive statistics for the variables of interest, with the first variable being the dependent variable and the following two variables being the primary independent variables. The average annual retention rate is 0.7135561, or 71.35561%, but the standard deviation of such annual retention rate is 0.1746543, or 17.46543%. The relatively high variation in retention rates may be due to numerous factors discussed in the literature, but it may be due to college crime, either in the form of changed college crime expectations or changed college crime overall. Table 1 also shows that college crime expectations have an average of 0.3762557, meaning that college crime increases by 0.3762557 crimes every two years. However, college crime expectations also have a relatively large standard deviation of 6.760393 that is greater than the average, meaning that college crime may *decrease* over two years. Table 1 also shows that overall college crime has an average of 5.309745, meaning that a four-year college in the US has an average of approximately five crimes on college each year from 2009 to 2016. Again, the standard deviation for college crime, 8.65957, is larger than its average. The large standard deviations among the dependent and two independent variables indicates that regressions are needed to determine the effect of college crime, both expectations and overall, on retention rates.

Results

Table 2 reports linear regression results, with column 1 regressing the independent variable, crime expectations, and the dependent variable, retention rates, without any controls (see equation 1). In column 1, the coefficient β_1 represents how an increase in one crime over a two-year period increases retention rates by 0.0016502 percentage points. This coefficient, capturing students who had an expectation of college crime before arriving on campus, means that as college crime increases by one crime relative to college crime two years ago, retention rates increase—counterintuitive to the literature that speaks to the negative effects of college crime on student behavior on campus. Although β_1 is significant at the 10 percent level, this is likely due to the lack of control variables (i.e. other omitted factors)—ε_i, the catch-all error term, is likely accounting for the factors aside from change in crime expectations that influence retention rates. The lack of controls results in the covariance of change in crime expectations and ε_i being a positive, non-zero value, thus biasing the results.

Column 2 regresses my second independent variable, college crime, and retention rates without any controls (see equation 2). In column 2, β_1, capturing students who had an expectation of college crime before arriving on campus, represents how an increase in college crime by one crime increases retention rates by 0.0029588 percentage points—positive and again counterintuitive regarding the negative effects of college crime on student behavior. Although β_1 is significant again at the 1 percent level, the results are likely again biased due to ε_i capturing the variables that affect retention rates that are not being controlled for. Thus, ε_i is again a positive, non-zero value biasing the results.

Figure 1 plots the linear relationship between change in crime expectations and retention rates, and the relatively small R^2 value of 0.0044 indicates the weak correlation between the change in college crime expectations and retention rates. Figure 2 plots the linear relationship between college crime overall and retention rates, and the similarly relatively small R^2 value of 0.0212 indicates the correspondingly weak correlation between college crime and retention rates. Since less than 3 percent of the variation in retention rates is explained by the two models, the relationship between the independent variables and dependent variable may be nonlinear and/or influenced by external factors. For example, there may be sample selection bias since females are disproportionately affected by certain types of crimes and thus are more likely to change their behavior in response to crime, potentially by dropping out or transferring—affecting retention rates. Campus wealth may also affect retention rates and relate to college crime since a wealthier campus may be too expensive for students to afford but may have more deterrents of college crime in place. These omitted variables are potentially some of the factors captured by ε_i and causing an overestimate of the relationship between college crime, both changed expectations and overall, and retention rates.

To prevent such omitted variable biases causing the regressions in column 1 and 2 to overestimate the effect of college crime on retention rates, I implement control variables that affect retention rates and relate to college crime, and also use school and year fixed effects (see equation 3). I then re-run the regression in column 1, including the control variables in γX_i , along with school and year fixed effects (see equations 4 and 5). Column 3 in Table 2 regresses crime expectations on retention rates, including the control variables and fixed effects (see equation 4). The coefficient β_1 represents how an increase in one crime over a two-year period decreases retention rates by 0.0001209 percentage points. The coefficient, aligning with the literature's depiction of college crime having negative impacts, delineates how an increase in college crime relative to college crime two years ago decreases retention rates. This coefficient also represents how students who have an expectation of crime before arriving on campus are likely to change their behavior in response to such crime by transferring or dropping out, thus decreasing retention rates. Although not statistically significant, R^2 increases to 0.3279, relatively greater than the previous regression that excluded controls.

I also re-run the regression in column 2, including the control variables in γX_i (see equation 3). Column 4 in Table 2 regresses overall college crime on retention rates, including the control variables (see equation 5). The coefficient β_1 represents how an increase in college crime by one crime decreases retention rates by 0.0002195 percentage points. The coefficient, again aligning with the literature, depicts how those without a college crime expectation before arriving on campus also reduce retention rates by changing their behavior in response to such crime by transferring or dropping out. Although not statistically significant, R^2 increases again to 0.1746, relatively much greater than the previous regression excluding controls.

The statistical insignificance and small magnitude of the independent variable coefficients in columns 3 and 4 may be due to omitted variable bias for data not available, such the presence and/or number of social fraternities and/or sororities on campus. The small and statistically insignificant coefficients may also be because both college crime expectations and college crime overall need to be included in the same regression in order to account for students who have expectations about crime on campus and how it will affect them, as well as students who have no expectations but are still impacted by crime on campus.

I then regress both college crime expectations and overall college crime in column 5 (see equation 6). The coefficient β_1 represents how an increase in one crime over a two-year period increases retention rates by 0.0000468 percentage points, and the coefficient β_2 represents how an increase in college crime by one crime decreases retention rates by 0.0003857 percentage points. β_1, capturing the students who have an expectation of crime, is again counterintuitive like in the regression in column 1 where an increase in college crime by one crime relative to college crime two years ago increases retention rates. β_2, representing the students who do not have an expectation of crime, is intuitive in terms of the negative impacts of college crime on student behavior as an increase in crime decreases retention rates A subsequent F-test between crime expectations and college crime show that the coefficients of the variables likely equal 0, meaning students with and without crime expectations may not jointly contribute to retention rates.

Natural Log of Independent Variables

In response to the counter-intuitive results from column 5 and the low F-statistic, I model overall college crime and crime expectations as natural logs in attempt to more appropriately model and better capture the relationship between college crime and retention rates. Table 3 presents my OLS regressions with non-linear terms results, with column 1 regressing the natural log of overall college crime and retention rates, including the previous control variables (see equation 9). The coefficient β_1 represents how a 100% increase in college crime increases retention rates by 0.0011352 percentage points. The coefficient, capturing those without an expectation of college crime, represents that a 100% increase in college crime increases retention rates. The positive coefficient is again counterintuitive like in the previous linear regressions of college (see Table 2 column 2) and the coefficient is statistically insignificant and small in magnitude.

Column 2 in Table 3 regresses the natural log of college crime expectations on retention rates (see equation 10). The coefficient β_1 represents how a 100% increase in college crime over a two-year period decreases retention rates by 0.0102814 percentage points. The coefficient, capturing students with an expectation of crime, represents that a 100% increase in crime relative to college crime two years ago decreases retention rates due to students transferring or dropping out. Although negative and thus intuitive with the literature describing the negative impacts of college crime, the coefficient is statistically insignificant and small in magnitude. Such statistical insignificance again may be due to needing to regress both the nonlinear variables of college crime expectations and college crime overall.

I regress both the natural log college crime expectations and the natural log of overall college crime, including control variables and fixed effects, in column 3 of Table 3 (see equation 11). The coefficient β_1, capturing students with a college crime expectation, represents how a 100% increase in college crime over a two-year period decreases retention rates by 0.0154352 percentage points, and the coefficient β_2 represents how a 100% increase in college crime overall increases retention rates by 0.0240867 percentage points. These results mean that students with an expectation of college crime are negatively impacted by an increase in college crime and are likely to transfer and drop out when college crime increases relative to their expectation of college crime formed during their senior year of high school. Students without an expectation of crime are positively impacted by an increase in college crime, increasing retention rates. While the linear regression of both crime expectations and college crime overall produce a positive coefficient for crime expectation and a negative coefficient for college crime overall, the nonlinear version of the variables produces the opposite with crime expectation having a negative coefficient and college crime overall having a positive coefficient. A subsequent F-test between crime expectations and college crime shows that the coefficients of the variables may jointly equal 0, meaning both may not contribute to retention rates in the model with non-linear variables.

Log-Log Models and Interaction Terms

Again, in response to the counterintuitive results from column 3 of Table 3, I model retention rate as a natural log along with college crime expectations and college crime as natural logs (see equation 11). Table 4 presents my log-log OLS regression results. In column 1, β_1 represents how a 100% increase in college crime expectations decreases retention rate percentage points by 2.59779%, and β_2 represents how a 100% increase in college crime increases retention rate percentage points by 3.64419%. While β_2 is counterintuitively positive and statistically insignificant just as it was when the dependent variable was linear, β_1 is negative and statistically significant at the 10% level. The coefficients mean that those with an expectation of crime are negatively impacted by crime and are likely to transfer or drop out to reduce retention rates, while those without an expectation of crime are not negatively impacted by college crime and increase retention rates. An F-test between college crime expectations and college crime shows that again college crime expectations and college crime overall may not contribute to retention rates.

In my log-log model, I also include interaction terms between the share of female students and both crime expectations and college crime overall to mitigate any sample selection bias (see equation 13).[21] In column 2 of Table 4, I regress the natural log of both crime expectations and college crime overall on the natural log of retention rates, including a female and total college crime interaction term. β_1 represents how a 100% increase in college crime expectations decreases retention rate percentage points by 1.88815%, and β_2 represents how a 100% increase in college crime decreases retention rate percentage points by 4.33022%. Although not statistically significant and having an F-test value unable to rule out both coefficients being jointly 0, both coefficients are intuitively negative, meaning that students that have an expectation of college crime and those who do not have an expectation are negatively impacted by college crime and thus are likely to transfer or drop out, reducing retention rates.

The female interaction term with college crime overall represents how retention rates decrease by an amount smaller than 4.33022% as the share of female students increases (specifically, 0.0078006 multiplied by the share of female students). Counterintuitive to the

[21] Females are likely the ones to remove themselves from the sample of colleges as females are disproportionately affected by certain types of crimes, and females are more likely to change their behavior in response to crime. See the negative coefficient of the share of female students, significant at the 5% level, in column 1 of Table 3.

literature, the interaction term depicts that a greater number of female students at an institution decreases the magnitude of the reduction in retention rates.

In column 3 of Table 4, I regress the natural log of both crime expectations and college crime overall on the natural log of retention rates with female and total crime expectation interaction term. β_1 represents how a 100% increase in college crime expectations decreases retention rate percentage points by 4.7221% significant at the 5% level, and β_2 represents how a 100% increase in college crime decreases retention rate percentage points by 1.37488%. With one of my two main independent variables being significant and the joint F-test being significant at the 10% level, these results show the most statistically significant results so far. These results show that both students with and without college crime expectations are negatively impacted by such crime and reduce retention rates, with those having an expectation of crime being significantly likely to transfer or drop out.

The female interaction term with college crime expectations represents how retention rates decrease by an amount smaller than 4.7221% as the share of female students increases (specifically, 0.0139123 multiplied by the share of female students). Again, counterintuitive to the literature, the interaction term depicts that a greater number of female students at an institution decreases the magnitude of the reduction in retention rates.

I also include a regression of both independent variables—college crime expectations and college crime overall—and both interaction terms—female and college crime expectations, and female and college crime overall—in column 4 of Table 4 (see equation 13). β_1 represents how a 100% increase in college crime expectations decreases retention rate percentage points by 3.93592%, and β_2 represents how a 100% increase in college crime decreases retention rate percentage points by 3.34811%. Although statistically insignificant and unable to rule out the coefficients of both independent variables jointly equaling 0 via an F-test, both coefficients are again intuitively negative, meaning that students with and without crime expectations are negatively impacted by college crime and reduce the retention rate.

Both female interaction terms represent how retention rates decrease by an amount less than 3.93592% and 3.34811%, respectively, as the share of female students increases (specifically, 0.0106083 multiplied by the share of female students for college crime expectations, and 0.0030962 multiplied by the share of female students for college crime overall). Since the three regressions in columns 2 through 4 depict the female interaction term having a positive coefficient, it is likely that although female students are more likely to change their behavior in response to crime, such change in behavior is not represented by female students transferring or dropping out to reduce retention rates.

Instrumental Regressions

Both linear and non-linear versions of college crime expectations and college crime present statistically insignificant and counterintuitive results delineating a positive effect of either college crime expectations or college crime overall on retention rates. Such counterintuitive results may be driven by bias that is not being captured by γX_i or the school or time fixed effects and is thus being captured by ε_i. To prevent such bias in further regressions, I implement an instrumental variable: ranking of annual average state temperatures.[22] I implement my instrument on my log-log regressions and my linear regressions, including both female interaction terms, in Table 5.

Column 1 of Table 5 regresses the natural log of college crime expectations on the natural log of retention rates with the annual average state temperature ranking instrument (see equation 15). β_1 represents how a 100% increase in college crime expectations decreases retention rate

[22] See equations 14-17 and the accompanied theoretical description for the use of the ranking of annual average state temperatures.

percentage points by 11.7293%, meaning that students with an expectation of college crime before arriving on campus are likely to transfer or drop out, reducing retention rates. Although the coefficient is not statistically significant and the F-test is less than 10 (F=1.79), meaning the annual average state temperature ranking instrument is relatively weak, the results are consistent with literature and previous regressions (see Table 4, columns 2 through 4) that college crime has a negative impact on students, causing them to transfer or drop out.

The female interaction term represents how retention rates decrease by an amount smaller than 11.7293% as the share of female students increases (specifically, 0.0310292 multiplied by the share of female students). Again, the positive interaction terms represents how the change in female student behavior is not represented by female students transferring or dropping out to reduce retention rates. Overall, although the instrument is relatively weak, the results of the log-log instrumental regression are similar in terms of students with an expectation of college crime before arriving on campus being negatively impacted by increases in college crime relative to their expectation made during their senior year of high school, causing them to transfer or drop out and thus reduce retention rates.

Column 2 of Table 5 regresses the natural log of college crime overall on the natural log of retention rates with the annual average state temperature ranking instrument (see equation 17). β_1 represents how a 100% increase in college crime overall decreases retention rate percentage points by 25.18799%, showing again that students with an expectation of college crime before arriving on campus are negatively impacted by such crime, making them likely to transfer or drop out. Again, although the R^2 decreases to 0.0688 and the coefficient is not statistically significant and the F-test is less than 10 (F=1.03), meaning the annual average state temperature ranking instrument is relatively weak, the results are again consistent with the negative impact of college crime on student behavior in terms of retention.

The female interaction term again represents how retention rates decrease by an amount smaller than 24.28621% as the share of female students increases (specifically, 0.0302396 multiplied by the share of female students). Again, although the literature speaks to how females are more likely to change their behavior in response to crime, the two instrumental regressions in Table 5 depict the female interaction term having a positive coefficient, meaning that the change in female student behavior is not represented by female students transferring or dropping out to reduce retention rates.

Overall, although my instruments are weak as depicted by the low first-stage F-statistics, and my coefficients are statistically insignificant, the results depict the same signs and interpretation. When accounting for controls, interaction terms, and an instrument, college crime expectations and college crime overall reduce retention rates, meaning that both students who do and do not have expectations of college crime and how it will affect them are negatively impacted by college crime on campus, making them more likely to transfer or drop out.

Conclusion

While college crime presents issues for college administrators regarding student safety, such crime may also present issues regarding retention rates. Although occasionally counterintuitive and with coefficients jointly potentially equaling 0, my instrumental regressions with female interaction terms depict that retention rates decreased when college crime expectations and college crime overall increased, meaning that college crime negatively impacts both students who do and do not anticipate college crime before arriving on campus, causing them to transfer or drop out.

Such results call for an increase in college administrative efforts to combat crime on campus by increasing campus police force presence, improving lighting around campus at night, and installing other safety measures to prevent crime. Overall, since the causality between college crime and retention rates cannot be ruled out definitively, these results should encourage administrators of higher education to more quickly and aggressively combat college crime, both for the safety of students *and* the future success of the institution.

References

Allen, Jeff, Steven B. Robbins, Alex Casillas, and In-Sue Oh. 2008. "Third-Year College Retention and Transfer: Effects of Academic Performance, Motivation, and Social Connectedness." *Research in Higher Education* 49 (7): 647–64.

Chapman, Randall G. 1986. "Toward a Theory of College Selection: A Model of College Search and Choice Behavior." *ACR North American Advances* NA-13. http://acrwebsite.org/volumes/6497/volumes/v13/NA-13.

Cornell, Dewey. 2010. "Threat Assessment in College Settings." *Change* 42 (1): 8–15.

Elliott, Kevin M., and Margaret A. Healy. 2001. "Key Factors Influencing Student Satisfaction Related to Recruitment and Retention." *Journal of Marketing for Higher Education* 10 (4): 1–11. https://doi.org/10.1300/J050v10n04_01.

Field, Simon. 1992. "The Effect of Temperature on Crime." *British Journal of Criminology* 32: 340–51.

Fisher, Bonnie S. 1995. "Crime and Fear on Campus." *The Annals of the American Academy of Political and Social Science* 539: 85–101.

Fisher, Bonnie S, Francis T Cullen, and Michael G Turner. 2000. "The Sexual Victimization of College Women," December, 47.

Galotti, Kathleen M. 1995. "A Longitudinal Study of Real-Life Decision Making: Choosing a College." *Applied Cognitive Psychology* 9 (6): 459–84. https://doi.org/10.1002/acp.2350090602.

Gregory, Dennis E, and Steven M Janosik. 2002. "The Clery Act: How Effective Is It? Perceptions from the Field – The Current State of the Research and Recommendations for Improvement." *Stetson Law Review*, 53.

Janosik, Steven M. (Steven Michael), and Donald D Gehring. 2003. “The Impact of the Clery Campus Crime Disclosure Act on Student Behavior.” *Journal of College Student Development* 44 (1): 81–91. https://doi.org/10.1353/csd.2003.0005.

McCallion, Gail. 2014. “History of the Clery Act: Fact Sheet.” *Fact Sheet*, 3.

Morrall, P., P. Marshall, S. Pattison, and G. Macdonald. 2010. “Crime and Health: A Preliminary Study into the Effects of Crime on the Mental Health of UK University Students.” *Journal of Psychiatric and Mental Health Nursing* 17 (9): 821–28. https://doi.org/10.1111/j.1365-2850.2010.01594.x.

Morriss, Susan B. 1993. “The Influences of Campus Characteristics on College Crime Rates,” *Annual Forum of the Association of for Institutional Research*.

National Crime Victims’ Rights Week. 2017. “School & Campus Crime.” *Office for Victims of Crime.*

National Oceanic and Atmospheric Administration. National Centers for Environmental Information. *National Climate Report: Annual 2009-2016*. https://search.usa.gov/search?utf8=%E2%9C%93&affiliate=ncdc&query=national+climate+report+annual

Pampaloni, Andrea M. 2010. “The Influence of Organizational Image on College Selection: What Students Seek in Institutions of Higher Education.” *Journal of Marketing for Higher Education* 20 (1): 19–48. https://doi.org/10.1080/08841241003788037.

Patton, Robert C., and Dennis E. Gregory. 2014. “Perceptions of Safety by On-Campus Location, Rurality, and Type of Security/Police Force: The Case of the Community

College." *Journal of College Student Development* 55 (5): 451–60. https://doi.org/10.1353/csd.2014.0049.

United States Department of Education. Office of Postsecondary Education. *Campus Safety and Security 2009/2010-2016/2017*. https://ope.ed.gov/campussafety/#/datafile/list

United States Department of Education. *College Scorecard Data 2009-2016*. https://collegescorecard.ed.gov/data/

Volkwein, J. Fredericks, Bruce P. Szelest, and Alan J. Lizotte. 1995. "The Relationship of Campus Crime to Campus and Student Characteristics." *Research in Higher Education* 36 (6): 647–70.

Warwick, Jacquelyn, and Phylis M. Mansfield. 2003. "Perceived Risk in College Selection: Differences in Evaluative Criteria Used by Students and Parents." *Journal of Marketing for Higher Education* 13 (1/2): 101–25. https://doi.org/10.1300/J050v13n01_07.

Tables

Table 1
Summary of Statistics of Independent, Dependent, and Control Variables

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Retention Rate	1,484	0.7135561	0.1746543	0	1
Crime Expectation	1,095	0.3762557	6.760393	-67	46
Total Crime	1,724	5.309745	8.65957	0	91
Average Cost of Attendance	1,081	29,774.93	12,827.43	3,368	66,045
Transfer Rate	1,615	0.1025132	0.189466	0	1
Share of Female Students	1,460	0.5643445	0.1375096	0.0669145	0.9719626
Average Family Income	1,671	54,725.37	26,635.36	3,000.464	144,807.90
Number of Undergraduates	1,729	1,115.311	1,729.802	0	24,061
Admission Rate	1,252	0.6478022	0.2283397	0	1
Annual Average State Temperature Ranking	1,729	81.40197	34.64713	6	122

Notes: Data come from 2009-2016. The number of observations vary due to availability of data. Total crime includes rape, fondling, incest, and statutory rape only for 2014 and 2016. Crime expectation is measured as a 2-year lag, thus there is no lag data from 2009 or 2010.

Table 2
OLS Estimates (Linear) of College Crime and Retention Rates, 2009-2016

Dependent Variable: Retention Rate

Regressor	**1**	**2**	**3**	**4**	**5**
Crime Expectation	0.0016502* (0.0008942)		-0.0001209 (0.004573)		0.0000468 (0.000631)
Total Crime		0.0029588*** (0.0005995)		-0.0002195 (0.0002953)	-0.0003857 (0.0009456)
Average Cost of Attendance			6.49e-08 (2.23e-06)	9.42e-07 (1.71e-06)	1.15e-07 (2.20e-06)
Transfer Rate			-0.0445032 (0.1058321)	-0.0082136 (0.053441)	-0.0464742 (0.1065766)
Share of Female Students			0.1091484 (0.2295849)	0.0592878 (0.256543)	0.1020449 (0.2290608)
Average Family Income			-1.23e-06 (1.13e-06)	-9.59e-07 (8.42e-07)	-1.25e-06 (1.15e-06)
Number of Undergraduates			7.27e-06 (0.0000383)	-0.0000232 (0.00003)	7.58e-06 (0.0000383)
Admission Rate			0.0460858 (0.1024079)	0.0358425 (0.0402863)	0.0461244 (0.1023514)
School Fixed Effect	No	No	Yes	Yes	Yes
Year Fixed Effect	No	No	Yes	Yes	Yes
Intercept	0.7150725*** (0.0057905)	0.6977569*** (0.0055002)	0.7348753*** (0.1534721)	0.7662703*** (0.1506491)	0.740687*** (0.1553503)
R-Squared	0.0044	0.0212	0.3279	0.1746	0.3351
Observations	887	1,480	320	729	320

Notes: * represents a coefficient's significance at the 10% level. ** represents a coefficient's significance at the 5% level. *** represents a coefficient's significance at the 1% level. Robust standard errors are in parentheses under the coefficients. Regressor: retention rates.

Table 3
OLS Estimates (Nonlinear) of College Crime and Retention Rates, 2009-2016

Dependent Variable: Retention Rate

Regressor	1	2	3
ln(Crime Expectation)		-0.0102814 (0.0077575)	-0.0154352 (0.0094457)
ln(Total Crime)	0.0011352 (0.0041582)		0.0240867 (0.0199168)
Average Cost of Attendance	-9.87e-07 (1.71e-06)	3.66e-06 (4.12e-06)	1.17e-06 (4.33e-06)
Transfer Rate	-0.0071378 (0.0576195)	-0.0292375 (0.0849227)	0.0248127 (0.091064)
Share of Female Students	0.2742775** (0.1266679)	-0.3834016 (0.2549435)	-0.4006061 (0.2578942)
Average Family Income	-8.80e-07 (6.30e-07)	-1.28e-06 (1.17e-06)	-1.14e-06 (1.22e-06)
Number of Undergraduates	0.0000483** (0.0000228)	0.0001713** (0.0000771)	0.0001666** (0.0000749)
Admission Rate	-0.0343934 (0.0375806)	-0.0645891 (0.0912071)	-0.1054078 (0.0929969)
School Fixed Effect	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes
Intercept	0.6198126*** (0.0991102)	0.7090389*** (0.1625603)	0.7951179*** (0.1527307)
R-Squared	0.0674	0.3761	0.3946
Observations	536	131	131

Notes: * represents a coefficient's significance at the 10% level. ** represents a coefficient's significance at the 5% level. *** represents a coefficient's significance at the 1% level. Robust standard errors are in parentheses under the coefficients. Regressor: retention rates.

Table 4
Log-Log OLS Estimates of College Crime and Retention Rates, 2009-2016

Dependent Variable: ln(Retention Rate)

Regressor	1	2	3	4
ln(Crime Expectation)	-0.0259779* (0.0148528)	-0.0188815 (0.0158404)	-0.047221** (0.0218432)	-0.0393592 (0.032414)
ln(Total Crime)	0.0364419 (0.0303749)	-0.0433022 (0.0749029)	-0.0137488 (0.0554004)	-0.0334811 (0.0780481)
Average Cost of Attendance	1.96e-06 (6.99e-06)	5.60e-06 (8.05e-06)	3.82e-06 (7.23e-06)	4.83e-06 (7.97e-06)
Transfer Rate	0.0294137 (0.1520195)	-0.1096537 (0.20111129)	-0.0434923 (0.1732997)	-0.0813766 (0.2110622)
Share of Female Students	-0.6203646 (0.3987538)	-0.6830524* (0.4119308)	-0.6606259 (0.4130421)	-0.6759463 (0.4158835)
Average Family Income	-1.59e-06 (1.89e-06)	-2.08e-06 (1.94e-06)	-2.10e-06 (1.91e-06)	-2.17-06 (1.92e-06)
Number of Undergraduates	0.0002981** (0.0001328)	0.0002725** (0.0001168)	0.0003389** (1.91e-06)	0.0003191** (0.0001406)
Admission Rate	-0.2049344 (0.160046)	-0.1277042 (0.1604885)	-0.189203 (0.1494387)	-0.1622847 (0.153949)
Female*Crime Expectation			0.0139123 (0.009544)	0.0106083 (0.0128366)
Female*Total Crime		0.0078006 (0.0059368)		0.0030962 (0.0084744)
School Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
Intercept	-0.2943231 (0.2488589)	-0.270508 (0.2465933)	-0.2946908 (0.23667908)	-0.2851508 (0.239488)
R-Squared	0.4048	0.4336	0.3397	0.4453
Observations	131	131	131	131

Notes: * represents a coefficient's significance at the 10% level. ** represents a coefficient's significance at the 5% level. *** represents a coefficient's significance at the 1% level. Robust standard errors are in parentheses under the coefficients. Regressor: natural log of retention rates.

Table 5
Log-Log OLS Instrumental Regression Estimates of College Crime and Retention Rates, 2009-2016

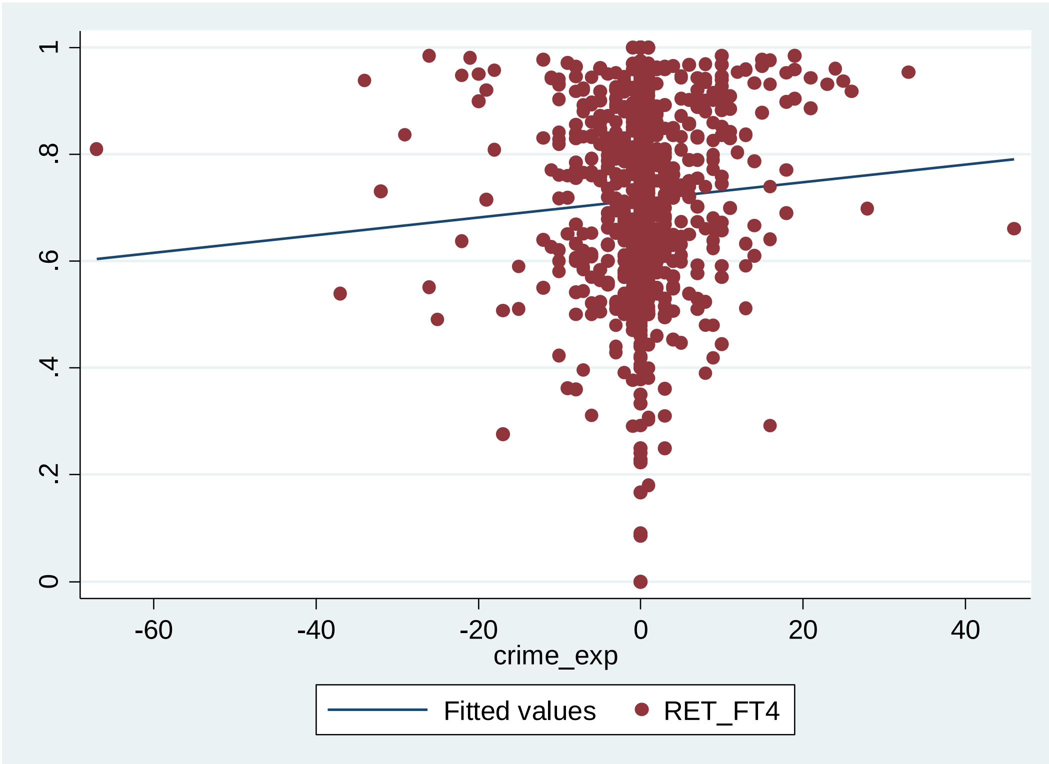
Dependent Variable: ln(Retention Rate)
Instrument: Annual Average State Temperature Ranking

Regressor	**1**	**2**
ln(Crime Expectation)	-0.1117293 (0.1362255)	
ln(Total Crime)		-0.2438621 (0.3098388)
Average Cost of Attendance	3.89e-06 (6.84e-06)	-2.45e-06 (4.53e-06)
Transfer Rate	-0.0478886 (0.1455495)	-0.0348639 (0.1020763)
Share of Female Students	-0.7810026* (0.4768086)	0.336661 (0.2864268)
Average Family Income	-2.24e-06 (2.08e-06)	-4.72e-06 (1.80e-06)
Number of Undergraduates	0.0003984*** (0.0001451)	-0.0000158 (0.0001274)
Admission Rate	-0.3322336 (0.2727446)	0.0605296 (0.212875)
Female*Crime Expectation	0.0310292 (0.0397025)	
Female*Total Crime		0.0302396 (0.0389647)
School Fixed Effect	Yes	Yes
Year Fixed Effect	Yes	Yes
Intercept	-0.2776344 (0.284331)	-0.1295499 (0.538445)
First-Stage F-Statistic	1.79	1.03
R-Squared	0.2494	0.0688
Observations	134	549

Notes: * represents a coefficient's significance at the 10% level. ** represents a coefficient's significance at the 5% level. *** represents a coefficient's significance at the 1% level. Robust standard errors are in parentheses under the coefficients. Regressor: natural log of retention rates.

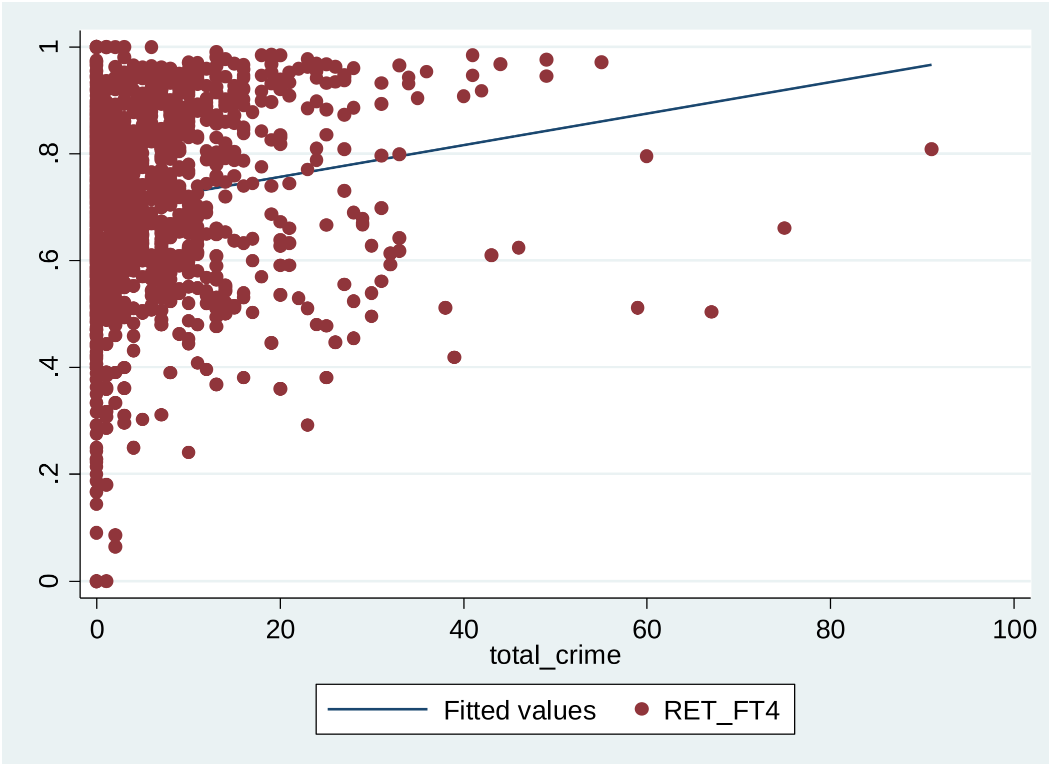
Figures

Figure 1
Change in College Crime Expectations and Retention Rates



Note: R^2= 0.0044. Data correspond to regression in Equation 1.

Figure 2
Change in College Crime and Retention Rates



Note: R^2= 0.0212. Data correspond to regression in Equation 2.

Appendix A
Table A
First-Stage Instrumental Regression Estimates of Annual Average State Temperature Ranking and College Crime, 2009-2016

Dependent Variable: ln(Total Crime) [1], ln(Crime Expectation) [2]

Regressor	**1**	**2**
Rank Temperature	0.000727 (0.0007168)	0.003779 (0.0028254)
Average Cost of Attendance	-7.59e-06 (0.0000144)	-7.58e-06 (0.0000494)
Transfer Rate	-0.0096495 (0.387304)	-0.3684589 (0.7599995)
Share of Female Students	-0.3280922 (0.9190485)	0.5083703 (3.534219)
Average Family Income	3.06e-06 (4.50e-06)	-2.79e-06 (0.0000146)
Number of Undergraduates	-0.0003829** (0.0001525)	0.0006328 (0.0008152)
Admission Rate	0.6355552 (0.2947623)	-2.070069** (1.028158)
Female*Crime Expectation		0.2832292*** (0.0446176)
Female*Total Crime	0.1245503*** (0.0225829)	
School Fixed Effect	Yes	Yes
Year Fixed Effect	Yes	Yes
Intercept	1.559979** (0.6852026)	0.2824916 (2.363875)
First-Stage F-Statistic	1.03	1.79
R-Squared	0.2172	0.4184
Observations	560	135

Notes: * represents a coefficient's significance at the 10% level. ** represents a coefficient's significance at the 5% level. *** represents a coefficient's significance at the 1% level. Robust standard errors are in parentheses under the coefficients. Regressor: ln(Total Crime) [column 1], ln(Crime Expectation) [column 2].