Statistical Analysis of Campus Safety Factors

By

Thomas Bode

&

Loren Snow

Presented to the Department of Economics, University of Oregon, in partial fulfillment of the requirements for honors in Economics

Prepared under the supervision of Dr. William Harbaugh

Abstract: We attempt to determine the marginal effect campus police and institutional characteristics have on the assault rate of university and college campuses in the United States using ordinary least squares and two-stage least squares regression analysis. We find that although we are unable to unable to determine with confidence the effects of campus police officers on assault rates, there appears to be spillover effects from the environment surrounding the campus on university assault rates.

Approved: _____

Prof. William T. Harbaugh Date

Table of Contents

INTRODUCTION

Within the past ten years, there have been a tragic number of incredibly violent shootings at universities and colleges across this country. We wanted to study this phenomenon from a statistical economic perspective with the ultimate goal of measuring the effect of policy variables on the likelihood of campus shootings. Specifically, we wanted to investigate the affect of money spent on student counseling services and campus security services, and find the marginal effects of each on the incidence of violence on campus. But from this grand vision, there were several confining factors that made our ultimate project something other than what we first imagined.

Fortunately, school shootings do not occur all that often, so there is not a large enough data set for regression analysis. What is available is a set of data on campus crime from the FBI that includes numbers for murders, assaults, rape and robbery, as well as several property crimes. We decided to shift the focus of our statistical analysis to the incidence of assaults on college campuses. Assaults are a highly violent crime, and we suggest it can be a proxy for the likelihood of a more violent shooting to occur. The incidence of murder is probably better proxy for severely violent crime, but again the murder rate on college campuses is so low that meaningful regression analysis would not be possible.

From the beginning this project was intended to include a policy analysis component, so that school administrator could more effectively understand how to protect their schools and their students. We had hoped to look at the effect of rates of counseling available for students on

campus as something that affects crime. However, an extensive period of search revealed no existing public data on the number or budget of counseling centers for individual schools. We settled on substituting a distant proxy: the student services budget for universities. This data was available from the IPEDS database and is part of the mandatory financial disclosure that universities must do. This budget line item includes campus fixtures such as the health center, the career center, and athletic programs. While the budget for counseling is included in this line item, the definition was too broad for it to be a satisfactory substitute.

We had more success gathering data for the campus security programs. The FBI collects the number of sworn officers and the number of civilians for the campus security departments of many schools.

LITERATURE REVIEW

While there is no precedent economic analysis of crime on college campuses, there is a well-established economic theory of crime and a consistent methodology for analyzing crime rates as a function of environmental and policy variables in larger population units such as large cities, counties, and states. Authors use both OLS and two-stage least squares to conduct their regressions. Interestingly, there is no consensus for the statistical significance of police levels on crime. There is more consensus that other non-policy variables, such as ethnic and gender demographics significantly influence crime.

Bukenya, James O. (2005). Crime Trends and Socio-economic Interactions: A County-level Analysis. *Criminal Justice Studies,* 18(4), 365-378.

Bukenya (2005) conducts a straightforward statistical analysis of crime in Alabama by county as a function of police expenditures, demographic figures, and environmental conditions. His model appeals to the Routine Activity Theory which says that three conditions affect the probability of a crime: motivated offenders, suitable targets, and the absence of a capable guardian. Additionally, he hypothesizes that "economic development and unidimensional crime interventions such as increase in law enforcement personnel is not enough to ensure constant crime decline" (372). This is supported by his conclusions that county-level police expenditures do not have a statistically significant effect on crime. He offers the common explanation that police levels are endogenous to crime. Factors he found to be statistically significant in predicting crime rates were age, education, and economic conditions.

Gius, Mark. (1999). The Economics of the Criminal Behavior of Young Adults: Estimation of an Economic Model of Crime with a Correction for Aggregate Market and Public Policy Variables. *American Journal of Economics and Sociology*, 58(4), 947-957.

Guis (1999) uses data from a longitudinal survey of youth that captures individual demographic information, economic data, and self-reported criminal history to measure by proxy the relative influence of individual variable and structural variables on propensity for crime. He concludes that while sex, race, and peer pressure are statistically significant factors for crime of all types, "police levels have no statistical deterrent effect on the criminal levels of young adults" (954).

Levitt, Steven. (1997). Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime. *The American Economic Review*, 87(3), 270-290.

Levitt (1997) uses the innovation that police rates in large cities are sensitive to mayoral and gubernatorial election cycles. He finds that "the mean percentage change in sworn police officers is 2.1 percent in gubernatorial election years, 2.0 percent in mayoral election years, and 0.0 percent in nonelection years" (271). He uses this phenomenon to evade the problems of endogeneity and simultaneity that have caused other regression analyses of police effects on crime to find a positive or zero correlation with crime. Using a two-stage least squares regression, he concludes that, in addition to other crimes, "large negative impacts of police are also observed for robbery, aggravated assault, and auto theft" (284).

Carr, J. L. (2005). *American College Health Association campus violence white paper.* Baltimore, MD: American College Health Association.

Carr (2005) conducts a diverse review of studies on college campus crime. It is clear from his research that campus crime is not a random event, but rather can be correlated with certain specific characteristics of the student body and the college campus. He finds that race and sex have a highly significant effect on the likelihood of victimization in campus violence and that 65% of violent acts against students go unreported and that "students were under the influence of alcohol or other drugs in 64% of physical assaults" (10).

Witt, Robert and Ann Dryden Witte. (1998). Crime, Imprisonment, and Female Labor Force Participation: A Time Series Approach. ***National Bureau of Economic Research working paper 6786.*** **Cambridge, MA.**

The authors examine the effects of the labor force participation of women as a proxy for different family and neighborhood structures on crime rates. They find that labor force participation is "highly significant" and suggest that it may represent the incidence of single-parent households, and unsupervised children, "especially teenagers" (11, 14). In addition, a higher labor force participation rate creates depopulated neighborhoods during the work day, which decreases likely apprehension and interruption of neighborhood criminal behavior.

THEORETICAL ANALYSIS

Economic theories of crime claim that the level of crime in a location is the function of a number of social and economic factors – this is well established in the literature. It also suggests that criminals operate rationally, so that increased disincentives such as greater chance of being caught or harsher punishments will drive down the crime rate. This is the deterrence effect. Following the assumption that people commit crime because it is positive, making available more positive experiences will cause potential criminals to substitute away from crimes. This is the substitution effect. Additionally, economic analysis of crime assumes that some individuals are more prone to crime that others, so that reducing the opportunity for the more criminal

people to commit crimes will decrease the crime rate. This is the incapacitation effect. However, the empirical effect of police on crime rates is less clear: while common sense suggests that an increased police presence activates both the deterrence and incapacitation effects and would drive down crime levels, regression analyses of crime have trouble with the inherent endogeneity of crime levels and police levels.

A policy variable for which there is less empirical studies is student services, which we hypothesize provides students (potential criminals) with alternatives to crime and encourage non-criminal behavior by showing good examples. This hypothesis is supported by the established negative effects of marriage and unemployment rates on crime, which both provide activity alternates to crime.

Regressions of crime that include police levels are conducted using both Ordinary Least Squares (Bukenya 2005) and two-stage least squares (Leavitt 1997). OLS regressions face the common problem of the endogeneity of the police levels. This problem can be theoretically be worked around by including sufficient control variables to capture the factors that do cause crime, which would reveal the crime reducing effect of police. Two stage least squares offers a possible alternative, if an instrument for the first regression can be found that is unrelated to crime. However, it can be difficult to find an instrument that accurately predicts police levels while remaining unrelated to crime.

DESIGN

In designing our model we knew from the start that we would have an endogeneity problem with police levels and crime level being positively correlated. We attempted to find the factors that cause campus assaults by putting together our model in pieces. In the first piece we created a model for city crime with officer levels and control variables such as demographic and economic conditions to explain assault rates for the city locations of universities. We then applied those explanatory variables to see how well they explained the campus assault rate. After determining how campus crime rates were determined by city variables we added campus specific control variables similar to the control variables used to explain the city assault rate. We theorize that there should be some spillover effects from the city environment onto a campus, but that these could not explain campus crime in full.

We have focused our attention on campuses located in cities with populations of 200,000 or less. We chose to do this because campuses that are in smaller towns and not in large metropolitan areas will reduce spillover effects from their surroundings and allow us to get a better picture of how campus characteristics affect the campus assault rate. This will also make it easier to compare data from other campuses in smaller towns that we do not have in our observations and for estimating purposes makes the task of explaining campus assaults much more simple.

In addition to the police per citizen variable used to explain the city assault rate, we also included explanatory variables such as racial demographics, income, and education level. Bukenya (2005) found that crime levels decrease as income goes up and found similar effects for

increased education levels. We used the number of officers per citizen, median income, percent married, age, percent completed high school, percent in the labor force, percent African-American, percent Asian, percent Native American, and the percentage of citizens identifying themselves as "other" on the race question of the 2000 US Census. We also included a measure of median age. Because the relationship between crime and age is parabolic in nature we used a variable of median age and also median age squared. We expect a population to commit more crimes as it grows to a certain age and less as they grow older. We expect more people in the labor forces to correlate with higher levels of crime. We expected police officers, median income, percent married, and percent completed high school to be negatively correlated with assault rates.

After determining the model for city assaults we used the same explanatory variables but used campus assaults as our explained variable. Here we explored the effect of a campus's surroundings on its assault rate. We expected to find the similar significance and sign direction, but lower magnitude, for city variables on campus assaults as existed for city assaults. After analyzing this model to observe the difference in explanatory power between city and campus assault rates, we put campus officers in the model to see how well this policy variable explained the level of assaults on a campus. We expect to find that campus officers are statistically significant, but positively correlated with campus crime because they are endogenous. This is the basic problem with our regression model.

To solve our endogeneity problem we added more campus-specific control variables to our model that will help to explain campus assaults. We added to our model percentage of African-American students, percentage of Hispanic students, percentage of Asian students, percentage of Native American students, percent of students that are men, the 25^{th} percentile

ACT score for the student body, the dollars per student spent on student support services, the number of bars per student within one mile of campus, and dummy variables for whether the campus is primarily non-residential or highly residential. We expect ACT scores, support money, and a non-residential campus to have negative coefficients while we expect bars and high campus residence to have positive coefficients.

DATA

The Integrated Postsecondary Education Data System (IPEDS), available from the National Center for Education Statistics (NCES), provides a wealth of information about postsecondary education institutions. From the 2006 universe of institutions, we restricted our selection to the following: located in the United States; public, four-year or above; private not-for-profit, four year and above; and private for-profit, four year and above. This resulted in a set of 2720 institutions. For this set of schools, we gathered identification information on the school, including address and name; the geographic region of the school's location; fall term enrollment; racial and gender demographics for the fall enrollment cohort; the 25th percentile for the ACT score of first year enrolled students; school expenditures on student services[1]; and the number of students receiving athletic related financial aid. These data were collected for 2004 and 2005 to correspond with the crime data available from the FBI Uniform Crime Reports. The geographical region information and the school control (public/private) were transformed into

[1] These include "admissions, registrar activities, and activities whose primary purpose is to contribute to students' emotional and physical well-being and to their intellectual, cultural, and social development outside the context of the formal instructional program" (Data Dictionary).

dummy variables. The demographic data were divided by school enrolment to create percentage variables, as was the absolute number of students receiving athletic financial aid. The student service and total institution revenue data were also divided by the student enrolment to create per capita variables. The data from IPEDS was manually matched with the FBI Uniform Crime Report data by the name of the school.

The demographic data on marriage rates, labor force participation rates, educational attainment, and median age were collected from the US Census. These particular data were available from the 2006 American Communities Survey. The American Community Survey collects data in geographical areas with a population of 65,000 or more, including counties and cities. The county and city data was matched with the university location using zip codes and cities. Except for median age, these variables were then transformed into percents of the total population. Marriage rate is derived from the number of residents in a married household (of any size); labor force participation rate is derived from the number of size of the labor force made of residents age 16 and older divided by the population of residents 15 and older (there was no disparate category for 16 years old); and the high school completion rate is derived from the number of residents 25 and older who earned a high school diploma or achieved higher education.

For the level of student residence of a campus we used a variable from the IPEDS Peer Analysis System. We used the Carnegie Classification of 2005: Size and Setting variable to describe the proportion of students living on campus. The size and setting classification divides campuses into very small, small, medium, and large classifications based on the number of students enrolled full-time at the campus. Very small campuses have an enrollment of fewer than 1,000 students. Small campuses have enrollments between 1,000 and 2,999 students. Medium

campuses have between 3,000 and 9,999 students enrolled. Large universities have 10,000 or more students enrolled. These size classifications are further broken down into three categories based on the proportion of enrolled students living in a university or college facility. The categories are primarily non-residential, primarily residential, and highly residential. Campuses that are primarily not residential have fewer than 25 percent of enrolled students living in campus residence facilities. Campuses that are primarily residential have between 25 and 49 percent of enrolled students in campus residence, while campuses that are highly residential have 50 percent of more of their students in campus housing. Campus housing is defined as institutionally owned, controlled, or affiliated housing. We used this classification to create dummy variables for primarily non-residential, primarily residential, and highly residential campuses. We obtained classification data for 755 observations; of which 286 are primarily non-residential, 340 are primarily residential, and 129 are highly residential.

We also used IPEDS for our revenue variable. We got core total revenues for our observations from the IPEDS Peer Analysis System. We then converted that into a dollar amount per student enrolled for use in our regressions by dividing total core revenues by the number of students enrolled.

To get data for the crimes committed on a campus and the crimes committed in the city that the campus is located we downloaded tables from the Federal Bureau of Investigation website. The data was downloaded from the Uniform Crime Report that the FBI makes available to the public on an annual basis. The tables from the Uniform Crime Report we used were tables eight and nine. Table eight is offenses know to law enforcement by state by city. Table nine is offenses known to law enforcement by state by university and college. We got data for campuses and their respective locations for the years 2004 and 2005. The data is separated into two

different categories of crime, violent crime and property crime. Violent crimes are murder, rape, assault, and robberies. Property crimes are burglary, theft, car theft, and arson. We were unable to obtain two years of data for every university or city due to differences in reporting by the FBI in the years 2004 and 2005. We were able to obtain 796 observations for violent crime on campus with 634 observations for the violent crimes of cities. For property crimes we were able to obtain 796 observations for campuses and 638 observations for cities. To make these variables comparable to other universities or cities of different sizes we converted the variables by dividing the number of campus crimes by the enrollment of the university or city crimes by the city population for use in our regressions.

To obtain the number of officers for a campus police force or city police department we used the FBI Uniform Crime Reports as well. We used the years 2004 and 2005 to match with our crime data. The FBI states their definition of an officer as such: "The UCR Program defines law enforcement officers as individuals who ordinarily carry a firearm and a badge, have full arrest powers, and are paid from government funds set aside specifically for law enforcement." The data we utilized from this part of the UCR were tables 78 and 79. Table 78 is full-time law enforcement employees by state by city and table 79 is full time law enforcement employees by state by university and college. Like the data for crime, we were unable to get two years of data for all schools or cities, but obtained two years for most. These datasets contain data for the number of officers, civilians, and the total of the two employed by a university or city. We obtained 779 observations for officers and civilians of campus police departments and 711 observations for officers and civilians of city police departments.

For city control variables and characteristics we used data from the United States Census of 2000. For income measurements we used the median income of a city and the poverty rate of

a city. These variables were taken from the American FactFinder on the Census Bureau website, using the one-in-six sample of the 2000 Census. These observations were collected from a one-in-six sample of the population and weighted to meet the total population. Poverty rate was created by dividing the number of people under the poverty line in the city by the total population in the city. We obtained 762 observations for poverty rate and median income. Data for race characteristics for a city were obtained from the American FactFinder as well using the 100 percent data of the 2000 Census. The Census categorizes races between White, Black or African-American, American Indian and Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, other, and two or more races combined. These separate race categories were changed into percentages by dividing each race by the total population of the city.

For the bars variable we obtained data using Google Maps. We first located the general address for the university or college on Google Maps. If there was no general address for the campus found we used the address of the admissions office. We then used the "Find Businesses" and searched for "category: bars" within a mile of the address for the university or college. The current findings were used for both years of observations assuming there would be little significant change between years and the bars in business today represent an approximation of the bars in business for 2004 and 2005. We then converted the number of bars into a variable for bars per student enrolled in the university or college by dividing the number of bars found by the total enrolment. We found 788 observations for the number of bars per student within a mile of the address.

Statistical distribution of the raw variables is available in Appendix 2.

EMPIRICAL RESULTS

Below are the results from our estimation of city assault rates. Only the city officers and statistically significant explanatory variables are shown. As expected, more city officers per citizen in the city have a negative effect on assaults, but are insignificant statistically. Median income and the percentage of high school graduates are significant at the ten percent level and are negatively correlated, as we predicted. Race was a statistically significant factor, which we would expect based on other statistical analysis of crime. The coefficient of percent in the labor force is positive, which is consistent with past literature (Witt & Witte 1998).

	Coefficient	P-Value
City Officers per Citizen	-0.11413	(0.618)
Median Income	-3.57e-08	(0.079)*
High School Grad %	-0.00504	(0.061)*
% In Labor Force	0.01382	(0.000)***
% African-American	0.00997	(0.000)***
% Other	0.00823	(0.001)***
Observations	347	
R-squared	0.432	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on city assault rate

The R-squared value of .432 tells us that we have a somewhat decent model for predicting assaults in a city. We then used this model to see how well it did at predicting assaults on campus by running the same regression using campus assault rates instead of city assault rates as the dependent variable. The results for that regression are below.

	Coefficient	P-Value
City Officers per Citizen	0.16982	(0.085)*
Median Income	-1.56e-08	(0.075)*
Observations	349	
R-squared	0.034	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus assault rate using city variables

In this regression we have only two significant variables, median income and city police. Also, these two variables are only significant at the ten percent level. We begin to see the effects of endogeneity in this regression, as the coefficient for city police officers is a positive effect on campus assaults. However, income still affects crime negatively. We can see a slight amount of spillover from income to the campus in this model. However, the city model is clearly not very good at explaining campus crime as it has an R-squared value of just .034. In the next regression we included campus police officers as an explanatory variable to try to see what effect they may have on campus assault rates.

	Coefficient	P-Value
Campus Officers per Student	0.13312	(0.000)***
City Officers per Citizen	0.16194	(0.068)*
Median Income	-2.34e-08	(0.003)***
Percent African-American	-0.00107	(0.015)**
Observations	344	
R-squared	0.254	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus assault rate using city variables and campus police rates

In this regression we found that campus officers certainly do have an effect on crime, but it is the opposite effect that we theorized. The jump in R-squared values from .034 to .254 tells

us that campus police have a large effect on crime relative to city characteristics. Here is where we really run into our problem of endogeneity. We know logically that neither city nor campus police cause assaults on students, so we need to find a way to solve this problem of endogeneity. The problem is that campuses with a high rate of crime will hire more police to act as deterrents. Of course our simple statistical model tells us that police cause crime, not that police deter the crime. Our first step in explaining this was to add university variables as explanatory variables to try to explain the assault rate better. If we account for the things that really do cause crime, perhaps we could see the true effect that campus police have on assault rates.

In our next model we add in our control variables for the university, similar to the control variables for the city along with university characteristics. The table following shows the statistically significant independent variables with city and campus police variables.

	Coefficient	P-value
Campus Officers per Student	-0.01938	(0.120)
City Officers per Citizen	0.07229	(0.070)*
Bars per Student	-0.07462	(0.015)**
% Male Students	0.00072	(0.015)**
% African-American Students	0.00057	(0.000)***
City Median Income	-1.29e-08	(0.002)***
City % Married	-0.00228	(0.046)**
City % Native American	-0.00109	(0.069)*
Observations	215	
R-squared	0.334	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus assault rate using city and campus variables

In this regression we now have the coefficient of the campus police negative, but it is statistically insignificant. Also, the coefficient for city officers is positive and statistically significant, suggesting high endogeneity in this estimate. Percent students African-American and

median income are the two most significant variables. The percentage of people married in the city is also statistically significant at the five percent level. This and the median income significance suggest that there is some spillover effect of city variables affecting campus crime. Because we do not know from the data whether students or non-students are committing crime on campus, this is possible evidence that not only students commit crime on campus, or that campus environments and student behavior are affected at least a little by the general surroundings of the campus. We can also see the men tend to cause assaults as well. This is in line with what we predicted. The R-squared goes up to .334, so the extra explanatory variables significantly increased the explanatory power of our model.

Another interesting result we get from this regression is the coefficient on the bars variable. According to our estimate, the more bars per student that are within one mile of campus, the fewer assaults there will be on campus. This contradicts our theory, as studies show that violence increases when individuals have been drinking (Carr 2005). There is clearly some other effect taking place here. It could be that the bars have an incapacitating effect on students. When potential criminals drink at off-campus bars, they are removed from the campus and do cannot commit crimes there without traveling back. Without the bars, students remain on campus more, and the propensity for violence goes up. We see the same effect in student support spending per student. While we predicted this would have a negative coefficient, it is actually positive, though insignificant. This is potentially due to the fact that as support dollars go up, more students are on campus more of the time and have more potential to commit assaults. We had hypothesized that increased opportunity for non-criminal activity, such as provided by the student support budget, would substitute potential criminals away from crime. This appears not to be the case with student support, although it could be happening with bars.

However, the research showing the increase in crime that accompanies alcohol consumption makes this unlikely. We conclude that bars create a significant incapacitating effect on students.

We still needed to try to solve our endogeneity problem in some way. Other researches use a two-stage least squares regression to avoid the endogeneity problem (Leavitt 1997). We can use instruments to determine an estimated amount of police in one regression, and use this estimate in the second regression to try to determine a better model for explaining campus assaults. However, it is difficult to find instruments that estimate police that are unrelated to crime. We chose total college revenues per student. We would expect that as a university earns more money per student that they would hire more police. A simple correlation test provides rough evidence for this, with a correlation of .36 between campus officers and revenue per student, while assault rates and revenue per student have a correlation value of just -.0032. Doing a simple regression with officers as the dependent variable and revenue per student as the explanatory variable shows how well revenue estimates the amount of officers on a campus.

	Coefficient	P-Value
Revenue per Student	0.00000	(0.000)***
Constant	0.00178	(0.000)***
Observations	717	
R-squared	0.129	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus police rates using per-student revenues

With an R-squared value of just .129 this estimate is far from perfect, but without any better alternative instruments we proceed with the regression anyway. Campus and city officers along with statistically significant variables are shown below.

	Coefficient	P-value
Campus Officers per Student	0.01025	(0.947)
City Officers per Citizen	0.13067	(0.072)*
Bars per Student	-0.03982	(0.067)*
City Median Income	-1.28e-08	(0.039)**
City % Married	-0.00319	(0.047)**
Observations	256	
R-squared	0.192	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Two-stage least squares regression on campus assaults using per-student revenue as the instrument.

The two-stage least squares regression has a lower r-squared value than our model that used the true values for campus police rates. It also gives a positive value for the coefficient on campus police rates, and they are much less significant. This model is significantly worse than the OLS model that uses the true campus police rates. This could be expected, though, because of the low explanatory power of the per-student revenue on campus police rates.

One of the main problems we had is that assault rates are not variable across campuses. Over half of our observations had two assaults or less. Assaults on campus are random events, and it is impossible to get any sort of conclusion out of our data. On last thing we tried was to use our model to explain property crime instead of assaults. Although property crime is relatively low on campuses compared to cities, it gives us more variety in the data. Also, property crime is not totally independent of assaults. The correlation between assaults and property crime on

campuses is .4185. The campus and city officer coefficients along with significant results from our property crime regression are below.

	Coefficients	P-values
Campus Officers per Student	1.59781	(0.000)***
City Officers per Citizen	3.37479	(0.000)***
Primarily Non-Residential	-0.00260	(0.009)***
% Men on Campus	0.0231	(0.001)***
% Students African-American	0.01462	(0.000)***
% Students Hispanic	0.01393	(0.079)*
% Students Asian	0.03011	(0.091)*
% Students Native American	0.09089	(0.026)**
City Median Income	-2.37e-07	(0.013)**
% City Completed HS	0.02129	(0.050)**
% City Native American	-0.02521	(0.063)*
Constant	-0.04104	(0.258)
Observations	215	
R-squared	0.541	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus property crime

With property crimes the number of bars per student becomes insignificant, while the residence characteristic becomes significant at the 99 percent level. Many burglaries that may happen in dorms or other campus residences on high residence campuses don't have the chance to happen on a primarily non-residential campus. Of course our endogeneity problem still exists with the officers. We tried using two-stage least squares for property crime also, but we got the same results as with our first two-stage least squares regression. Full tables with all results for the regressions discussed in this section can be found in the appendix.

CONCLUSION

It is difficult to create a useful regression model for the effect of police on crime. We did not escape the endogeneity problem that has faced other researchers, and our attempt to evade it through a two-stage least squares regression failed because we could not find a sufficiently explanatory instrument. In addition to the endogeneity problems that plague regression analysis of crime and police, there are characteristics unique to crime on college campuses that make this type of analysis more difficult: the overall low rate of crime on campuses lacks sufficient variance to support a strong regression; high rates of spillover from the city to a campus and a campus to the surrounding city make analysis difficult; and unusually low rates of reported crime create large errors.

Even with our relatively large data sample (n=796), half of the observations had less than two assaults per year. With such low numbers of assaults, it is difficult to attribute them to a general pattern of crime because they are more likely to be arbitrary events. In contrast, the fiftieth percentile for city rates of assaults is 264. The higher incidence of crime in the city is one reason why our regression of city crime, though simpler than our full model for campus crime, had more explanatory power. The generally low rate of assaults on college campuses casts a high error on all of our results. A related problem to the low rates of crime on campuses is the immense underreporting that occurs for campus crime. One author cites that only 35% of "acts of violence against students" are reported (Carr 2005). The social and legal pressures that influence a student's decision to report violent crime are complex and it is unsafe to assume that the crimes that do get reported are representative of the total incidence of crime on campus.

College campuses are often uniquely situated within as city as an open community center, but one with its own resident population. The degree to which a college is integrated within a community can vary greatly, from an urban commuter college to an isolated liberal arts enclave. Crime from a city can spill over onto campus, when students become victims to non-students; but campus crime can also be drawn off campus, as when students leave the campus to drink. These spillover effects make it necessary to include both city and college characteristics in a regression, but the inability to measure integration makes it difficult to capture with precision these spillover effects. Our regression results indicated a high significance for some city characteristics that were also highly significant for city assault rates, such as the median income of the city. Other variables that were highly significant in the city regression did not spill over onto campus crime, such as the percent of the population in the labor force.

To further pursue the explanation of campus assaults it would be prudent to separate city and campus. Detailed crime statistics such as whether the perpetrator or victim was in fact a university student would go a long way in discerning what the magnitude of the spillover effects from the surrounding city are. From our research it is unclear whether the spillover effects are citizens of the city causing crime on campus or university students causing crime on campus being influenced by city characteristics. It would also be useful to expand the area of analysis into any residential areas surrounding the campus that are heavily populated by students. Also, more years of data and more observations per year would help in determining whether changes in officer staffing lead to changes in the assault rate.

APPENDIX

Full regression results for empirical analysis:

	Coefficients	P-values
City Police Rate	-0.11413	(0.618)
Median Income	-3.57e-08	(0.079)*
% Married	-0.00440	(0.557)
Median Age	0.00052	(0.324)
Age^2	-0.00001	(0.328)
% Completed HS	-0.00504	(0.061)*
% In Labor Force	0.01382	(0.000)***
% African-American	0.00997	(0.000)***
% Native American	0.00287	(0.314)
% Other	0.00823	(0.001)***
% Asian	-0.00603	(0.108)
Constant	-0.01232	(0.185)
Observations	347	
R-squared	0.432	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on city assault rate

	Coefficients	P-values
City Police Rate	0.16982	(0.085)*
Median Income	-1.56e-08	(0.075)*
% Married	-0.00390	(0.222)
Median Age	0.00012	(0.587)
Age^2	-1.43e-06	(0.666)
% Completed HS	-0.00004	(0.971)
% In Labor Force	0.00023	(0.886)
% African-American	-0.00026	(0.584)
% Native American	-0.00093	(0.449)
% Other	-0.00026	(0.807)
% Asian	0.00110	(0.494)
Constant	-0.00159	(0.691)
Observations	349	
R-squared	0.034	
p values in parentheses		
* significant at 10%; ** significant at 5%;		

*** significant at 1%		

Regression on campus assault rate using city variables

	Coefficients	P-values
Campus Officer Rate	0.13312	(0.000)***
City Officer Rate	0.16194	(0.068)*
Median Income	-2.34e-08	(0.003)***
% Married	-0.00256	(0.362)
Median Age	0.00015	(0.442)
Age^2	-2.19e-06	(0.458)
% Completed HS	-0.00070	(0.496)
% In Labor Force	0.00013	(0.928)
% African-American	-0.00107	(0.015)**
% Native American	-0.00114	(0.294)
% Other	-0.00094	(0.316)
Constant	-0.00133	(0.707)
Observations	344	
R-squared	0.254	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus assault rate using city variables and campus police rates

	Coefficients	P-values
Campus Officer Rate	-0.01938	(0.120)
City Officer Rate	0.07229	(0.070)*
Bars per Student	-0.07462	(0.015)**
Highly Residential	-0.00008	(0.438)
Primarily Non-Residential	-0.00004	(0.365)
% Men	0.00072	(0.015)**
% African-American, Campus	0.00057	(0.000)***
% Hispanic, Campus	0.00012	(0.727)
% Asian, Campus	-0.00021	(0.789)
% Native American, Campus	0.00220	(0.220)
25^{th}% ACT Score	-0.00001	(0.324)
Support Dollars per Student	5.98e-08	(0.218)
Median Income, City	-1.29e-08	(0.002)***
% Married, City	-0.00228	(0.046)**

Median Age, City	0.00006	(0.517)
Age^2, City	-5.50e-07	(0.670)
% Completed HS, City	-0.00002	(0.974)
% In Labor Force, City	-0.00094	(0.222)
% African-American, City	-0.00005	(0.795)
% Native American, City	-0.00109	(0.069)*
% Other, City	-0.00023	(0.700)
% Asian, City	0.00034	(0.749)
Constant	0.00019	(0.907)
Observations	215	
R-squared	0.334	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus assault rate using city and campus variables

	Coefficients	P-Values
Campus Officer Rate	0.01025	(0.947)
City Officer Rate	0.13067	(0.072)*
Bars per Student	-0.03982	(0.067)*
Highly Residential	0.00023	(0.422)
Primarily Non-Residential	-0.00003	(0.711)
% Men	0.00026	(0.508)
% African-American, Campus	0.00035	(0.237)
% Hispanic, Campus	-0.00014	(0.776)
% Asian, Campus	-0.00052	(0.599)
% Native American, Campus	0.00172	(0.527)
25^{th}% ACT Score	-0.00001	(0.263)
Median Income, City	-1.28e-08	(0.039)**
% Married, City	-0.00319	(0.047)**
Median Age, City	0.00003	(0.804)
Age^2, City	-8.46e-09	(0.900)
% Completed HS, City	0.00034	(0.668)
% In Labor Force, City	0.00025	(0.749)
% African-American, City	-0.00021	(0.468)
% Native American, City	-0.00106	(0.242)
% Other, City	0.00032	(0.685)
% Asian, City	0.00176	(0.248)
Constant	-0.00029	(0.894)
Observations	256	
R-squared	0.192	
p values in parentheses		
* significant at 10%;		

** significant at 5%; *** significant at 1%		

Two-stage least squares regression on campus assaults using per-student revenue as the instrument.

	Coefficients	P-values
Campus Officer Rate	1.59781	(0.000)***
City Officer Rate	3.37479	(0.000)***
Bars per Student	0.31549	(0.646)
Highly Residential	0.00129	(0.594)
Primarily Non-Residential	-0.00260	(0.009)***
% Men	0.0231	(0.001)***
% African-American, Campus	0.01462	(0.000)***
% Hispanic, Campus	0.01393	(0.079)*
% Asian, Campus	0.03011	(0.091)*
% Native American, Campus	0.09089	(0.026)**
25th% ACT Score	0.00024	(0.353)
Support Dollars per Student	1.46e-06	(0.183)
Median Income, City	-2.37e-07	(0.013)**
% Married, City	-0.01294	(0.615)
Median Age, City	0.00138	(0.486)
Age^2, City	-0.00002	(0.509)
% Completed HS, City	0.02129	(0.050)**
% In Labor Force, City	-0.01070	(0.539)
% African-American, City	-0.00093	(0.844)
% Native American, City	-0.02521	(0.063)*
% Other, City	0.00638	(0.633)
% Asian, City	-0.00012	(0.996)
Constant	-0.04104	(0.258)
Observations	215	
R-squared	0.541	
p values in parentheses		
* significant at 10%; ** significant at 5%; *** significant at 1%		

Regression on campus property crime

APPENDIX 2

Statistical distribution of raw variables

Variable	Obs	Mean	Std. Dev.	Min	Max
City Population	639	291131	825211	1092	8101321
Violent Offenses, City	634	2628.992	6748.847	0	55688
Murders, City	638	35.51724	85.30329	0	570
Rapes, City	638	114.2085	213.1918	0	1428
Robberies, City	638	1008.027	2852.685	0	24373
Assaults, City	635	1462.093	3652.098	0	29317
Property Crime, City	638	13624.79	26288.14	0	171188
Burglary, City	638	2776.884	5193.445	0	27541
Theft, City	638	8791.994	17132.35	0	124016
Car Theft, City	638	2055.911	4481.105	0	29973
Arsons, City	581	106.4079	260.5109	0	2229
Total Law Enforcement Employees, City	711	1055.291	4538.789	1	52335
City Officers	712	781.5801	3116.568	1	35513
City Civilians	711	272.8664	1433.184	0	16822
Enrollment	793	12386.85	10738.49	290	52261
Violent Offenses, Campus	796	5.777638	7.068901	0	58
Murders, Campus	796	0.015075	0.121929	0	1
Rapes, Campus	796	1.183417	1.819261	0	16
Robberies, Campus	796	1.552764	3.063846	0	30
Assaults, Campus	796	3.026382	4.066808	0	28
Property Crime, Campus	796	197.4925	204.5101	0	1358
Burglaries, Campus	796	25.70854	37.62721	0	387
Theft, Campus	796	165.2588	171.9901	0	1110
Car Theft, Campus	796	6.525126	12.27292	0	105
Arson Campus	731	0.971272	1.96977	0	20
Total Law Enforcement Employees, Campus	779	33.66752	27.70184	1	158
Campus Officers	779	21.58151	15.45903	1	89
Campus Civilians	779	12.08601	14.75138	0	86
25th % ACT	580	20.16207	3.300904	5	31
Student Service	656	1.23E+07	1.25E+07	0	1.03E+08
Student Service	119	2.40E+07	3.69E+07	53212	2.60E+08
Total Enrollment, Campus	794	12530.38	10825.34	74	51612
Total Enrolled, Men	795	5610.128	5226.015	5	25960
Total Enrollment, White	794	8164.555	7598.462	2	37622

Total Enrollment, African-American	795	1230.054	1624.343	4	11943
Total Enrollment, Hispanic	795	987.5484	2135	0	20567
Total Enrollment, Asian	795	815.0465	1675.314	0	11614
Total Enrollment, Native American	795	114.1132	251.9033	0	2701
Total Enrollment, Unknown	794	643.2657	984.4358	0	7533
Total Enrollment, Resident Alien	794	574.8123	835.1774	0	4650
Median Age, City	644	35.37314	3.190335	25	42.8
Percent Married, City		0.184116	0.030909	0.095627	0.46239
Percent With HS Diploma, City	644	0.866717	0.062054	0.595495	0.988724
Percent in Labor Force, City	642	0.832848	0.049094	0.622422	0.905308
Median Income, City	763	18974.27	6814.942	7573	71867
Bars within One Mile of Campus	793	9.300126	27.44118	0	291
Carnegie Size Classification	756	13.49735	2.106123	6	17
City Population White	765	141909.2	386950.4	50	3576385
City Population African-American	766	66906.25	221225.2	13	2129762
City Population Native American	766	1654.74	4927.939	0	41289
City Population Asian	766	17203.14	82318.65	4	787047
City Population Hawaiian	765	292.1856	894.8909	0	5915
City Population Other	766	30238.21	130185.2	0	1074406
City Population No Response	765	9651.656	40626.28	4	393959
Total Yearly College Revenue	736	3.35E+08	5.44E+08	1.00E+07	6.27E+09
Primarily Non-Residential Campus	798	0.359649	0.480199	0	1
Primarily Residential Campus	798	0.426065	0.494814	0	1
Highly Residential Campus	798	0.161654	0.368364	0	1