

An Examination of the Impact of University Students on Rent Prices, Rough Draft*

Julie Dawkins[†]

April 23, 2024

Abstract

I am going to write this last.

* Acknowledgements here, if any.

[†]Department of Economics, University of Oklahoma. E-mail address: juliedawkins@ou.edu

1 Introduction

Across the United States, large cities are facing a housing crisis (U. S. Government Accountability Office, 2023). Much attention has been paid to high rents and high demand in cities such as New York and San Francisco, challenges closely tied to homelessness rates (Horowitz et al., 2023).

However, regions that have historically faced less scrutiny are college towns. These towns, such as Norman, Oklahoma, are dominated by a university and the large student population who attend the school. These students flow in and out of the region seasonally – anecdotally, Norman is a ghost town during the summer and winter breaks. They also help shape the rental market; many apartments are marketed specifically toward student housing. Meanwhile, individuals who are permanent residents in a collegiate area may face some competition from students for properties.

In this paper, I seek to understand how college campuses and higher education institutions impact county-level median rent. Unlike other literature that examines individual housing prices as a function of proximity to universities, my approach examines the effects of number of students within a county and students as a percentage of the total population on median rent prices of various unit sizes (one- through four-bedroom). Understanding how student populations affect rent has major policy implications, particularly for the Norman region and the University of Oklahoma (OU). The last three freshman classes have each broken records for the largest class in the university's history, and with the upcoming transition from the Big 12 to the SEC, this growth trend will likely continue. Understanding the potential impacts of these changes is critical for both the resident population of a college town and the student population.

By using balanced county-level panel data from 2018 to 2022, I conduct multiple fixed-effects multinomial regressions comparing counties with students to counties without, as well as examin-

ing counties with students to understand potential causal impacts on rent among college counties.
[eventually my results will go here too]

2 Literature Review

Some research has been done on the direct effects of student populations on rental markets, though more has been done to examine the largest determinants of rents generally. It is generally agreed within the literature that the presence of universities increases rent prices; other significant influences include

Smith (2004) discusses the effects of “studentification” on the housing market. Studentification, a type of gentrification, describes how the increase in largely-transient, single-person renters (i.e., students) results in a shift of investments away from (or even repurposing of) single-family units toward multi-unit housing, reducing the overall supply of single-family units. This discussion remains largely theoretical and speaks to the larger cultural implications of this shift. Rugg et al. (2000) further emphasizes the development of a unique student-focused rental market, but notes that the impact on the local market varies widely depending on the culture of the region. For example, they note that local tenants may just avoid “student” areas but not have challenges finding housing elsewhere, and in some instances, students may themselves have less bargaining power than other tenants.

Ogur (1973) addresses the question more directly in an empirical model studying 62 New York counties from the 1960 US Census to examine the impacts of proximity to a university on the local rents. He finds that college/university presence (measured by the percentage of the population enrolled in college) has a positive significant impact on rental prices – i.e., prices increase as a

result of student presence. Similarly, Rivas et al. (2019) use ZIP-code data as well as individual-house level data to examine the impacts of proximity to both hospitals and universities on median housing prices and rents. They use no formal econometric model but study correlations and the significance of those correlations. Though they find positive correlations between house price and proximity to a university, they find slightly smaller correlations between rent and university distance. Lastly, Yilmaz et al. (2022) examines data in the UK and specifically highlights the seasonal nature of rentals in university towns – i.e., demand is higher during the summer before the year starts than during the winter. This seasonal nature of rent is also something to consider when looking at the relationship between student population and rents – there may be a slightly lagged effect as student rentals adapt their prices in between the end of a school year and the beginning of the next.

Understanding other determinants of rent prices is also critical to developing a robust model. Sirmans & Benjamin (1991) provides a robust overview of the literature and highlights important factors such as the “physical attributes” of the property (i.e., size, amenities, etc.), property management, and vacancy rates. Unfortunately, I am unable to access any of these data points, as I am examining data on the county-level.

3 Data

To construct the data for my project, I make use of four separate data sources across the years 2018 to 2022. Prior to 2018, information on the distance learning population is unavailable.

First, I gather data from the National Center for Education Statistics (NCES) to gain information on enrollment in institutions across the country. I then download median rental data from the

Housing and Urban Development department. Next, I gather data on group living estimates from the American Community Survey, which captures information on individuals living in places such as college residences and military quarters; and lastly, I make use of Dr. Tyler Ransom's extensive county-level data to gain controls such as median income, the percentage of various racial demographics, the unemployment rate, and health/poverty indicators.

Within the NCES data, I merge institutional-level data (containing information such as the school's name and location) with data containing the unduplicated student count by form of distance learning. I subtract students who exclusively distance learn from each institution's total student count to gain an estimate of the students present in the region. I then group by the county code and get an estimate of the number of students within each county. Critically, across the 5 years, not every county appears in each year. I therefore used linear interpolation for counties that had data in the before/after period and mean interpolation for counties that were missing either four of the five years or were missing data at the beginning or end of the dataset. Mean interpolation was more logical than linear interpolation in many cases because rarely did a county appear in one year, then missed a year, then appeared again.

A notable component of this data is that a "year" is a school year, meaning 2022 is the unduplicated headcount across the school year 2021-2022. To adapt to this, I use rolling averages for all controls and housing data, so that "2018" represents an average of 2017-2018 and so on. Following this cleaning, I merged this NCES data, county-level housing data, group living data, and control data by year and county code. Lastly, I created several variables, including the percentage of a county population made up of students and the log of the rent prices, population, and student population. I also created a binary variable equal labeled COLLEGE equal to one if student population resided in the county. This binary variable, as well as the number of students and percentage

of students, equaled 0 if I could not find data for that county.

Summary statistics comparing averages in counties with college students and without can be found in Appendix A.

4 Empirical Methods

Because I am seeking to understand causal inference, I will take an econometric approach and look at a few multinomial regression models. Because I have panel data, I will also try a fixed effects approach, which demeans each variable in order to control for unobservable county specific trends.

I will develop a few models:

- $\log(4\text{-bedroom}) = \text{COLLEGE} + X + E$, X is a matrix of regressors including median income, pct white, pct black, population, and unemployment, and population density
- $\log(4\text{-bedroom}) = \text{pctstudent} + X + E$, X is a matrix of regressors
- $\log(4\text{-bedroom}) = \text{pctstudent} + \text{numstuds} + \text{grouphousing} + X$
- etc i still need to play with the modeling
- i will probably look at the different rents of bedroom too
- also all of this will get properly formatted in equation mode

5 Research Findings

This section will discuss the results of my findings. I will find out what these are soon. I have not had the opportunity to run these models since (nearly) finalizing my data.

6 Conclusion

This section will compare my results with the larger literature and, from there, make a policy recommendation with a focus on Norman, OK.

References

i cannot get these bibs to load for the life of me. i will fix it for the final

Figures and Tables

Figure 1

this figure will likely be a visualization similar to what i created in PS6

Table 1

this will be summary statistics based on COLLEGE == 1 or not

Table 2

this will be an output of my regressions; i will probably end up having multiple tables