


RESEARCH

Open Access



The impact of colleges and hospitals to local real estate markets

Ryan Rivas^{1*} , Dinesh Patil¹, Vagelis Hristidis¹, Joseph R. Barr² and Narayanan Srinivasan³

*Correspondence:

rriva002@ucr.edu

¹ Department of Computer Science and Engineering, University of California, 900 University Ave., Riverside, CA 92521, USA

Full list of author information is available at the end of the article

Abstract

This paper studies how the presence of universities and hospitals influences local home prices and rents. We analyze the data on ZIP code level and on the level of individual homes. Our ZIP code-level analysis uses median home price data from 13,105 ZIP codes over 21 years and rent data from 15,918 ZIP codes over 7 years to compare a ZIP code's appreciation, volatility and vacancies to the size of a university or hospital within that ZIP code. Our home-level analysis uses data from 2,786,895 homes for sale and 267,486 homes for rent to study the impact of the distance from the nearest university or hospital to individual home prices. While our results generally agree with our expectations that larger, closer institutions yield higher prices, we also find some interesting results that challenge these expectations, such as positive correlations between volatility and university/hospital size in some ZIP codes, a positive correlation between rent and distance from a hospital for some homes, and lower correlations of rent vs. distance from a university compared to price vs. distance.

Keywords: Home prices, Rent, Appreciation, Volatility, University proximity, Hospital proximity

Introduction

Home price is made of two parts: price of land and the cost of the house. Land value is derived from its location which often, especially in urban areas, accounts for the lion's share of overall home price. The value of the land is subject to the laws of supply and demand and in turn depends on the land's scarcity. Indeed, decoupling price of land from price of construction has been extensively researched [1, 2]. Many factors are baked into land price, including proximity to amenities and land's inherent quality (e.g., proximity to a shoreline, the mountains, etc.). Unique and somewhat subjective home characteristics like view as well as proximity to ocean, lake, etc. are known to influence home price [3]. Conversely, land price may be adversely affected by proximity to sources of noise and pollution (airports, major highways, etc.) [4, 5]. Unlike building material, labor and capital, land is a "finite," or "non-renewable," resource, often limited by stringent geographic and topographical constraints. Amenities pertain to proximity and accessibility to things like opportunities in employment, education, transportation, entertainment, retail, cultural, recreational, etc.

This analysis focuses on universities and hospitals as "opportunity hubs," which encapsulate "packaged amenities" in terms of those listed above. It studies the impact of these

institutions on both home sale price and rent. Both types of institutions attract a “stable,” educated and mobile workforce, a mix of demographics and incomes, and various amenities. Unpacking amenities isn’t altogether simple, and is a somewhat subjective art. For example, a neighborhood’s school rating is on one hand a reflection of the neighborhood and its characteristics, demographics and economics. Conversely, the rating of a neighborhood’s schools affects its home prices, primarily via the value of the land on which each home in the neighborhood is built. Throughout this paper, we will unravel pricing substructure via the correlations between home value and proximity to said institutions as well as their “idiosyncratic rhythm.”

This study focuses on US homes. We perform two types of analysis. In *ZIP code-level analysis*, we use the median home price per ZIP code and study how the containment of a university or a hospital affects (correlates to) home prices by comparing against ZIP codes that do not contain such institutions. As real estate is always local, we looked for a wide availability of data at a very local possible level and ZIP code level data fit our requirements [6]. In *home-level analysis*, we consider the prices of individual homes with respect to their exact distance from institutions. The goal of this analysis is to study how the distance of a property to an institution affects its price or rent and test our assumption that universities and hospitals generally increase the sale prices and rent of nearby homes. Our basis for this assumption includes both prior work that has analyzed the effect of real estate prices in relation to proximity of various features as described above and in the Related Work section as well as intuition; for example, one would expect that homes close to universities have higher rent as students without cars prefer them. In addition to computing the price and rent correlations with the distance from a university or hospital, we study how rent or price appreciation and volatility as well as vacancies change over time and how they correlate with the size of a university or hospital.

We collected data for home listings and historic home rent and price trends from public listing sources. We also collected ZIP code populations, hospitals and vacancies data from government sources. We built a university dataset by crawling and combining data from online sources (US News and Wikipedia).

Our results show several correlations. In our ZIP code-level analysis, we found the strongest correlations in ZIP codes with a population below the national ZIP code average. These correlations were between appreciation and hospital size, volatility and university size, volatility and hospital size, and vacancies and university size. In our home-level analysis we found significant correlations between rent and distance from a university (especially private universities), and rent and distance from a smaller hospital.

Related work

Real estate prices have been a frequent subject of analysis. Cesa-Bianchi et al. compared house prices in advanced and emerging economies between 1990 and 2012 and found that house prices in emerging markets experience faster growth, more volatility, less persistence and less synchronized than house prices in advanced economies [7]. Favara and Imbs found that housing prices increased in response to the expansion of mortgage credit [8]. Muehlenbachs et al. found that shale gas development has a negative effect on property values in areas dependent on groundwater, but a positive effect on property values in areas with piped water [9]. Waddell et al. drew several conclusions from their

analysis of residential property values in Dallas County, Texas: including a significant but fairly localized central business district price gradient; improvements to modeling the price effect of proximity to employment centers and other nodes of activity; amenities such as highways, retail, universities, and hospitals had a significant effect on modeling housing values; and a significant influence of race on housing prices [10]. Nau and Bishai found that life expectancy within communities predicted increases in home price indexes [11]. Otto and Schmid analyzed real estate prices in Germany using spatiotemporal models and found that urban regions with higher population density and higher per-capita disposable income have higher land prices than rural areas, shocks in regional real estate prices “ripple out” and affect the whole economy, and population density had an increasing impact on real estate prices [12].

Several papers have evaluated the impact of nearby points of interest on home prices. Rascoff and Humphries found that homes within a quarter mile of Starbucks locations appreciated more quickly than the overall rate of nationwide home appreciation [13]. Turner found that a several of points of interest, including supermarkets, restaurants and movie theaters, increase nearby home values in three neighborhoods in the San Francisco Bay Area [14]. Bolitzer and Netusil found that open spaces such as parks and golf courses increased nearby home prices in Portland, Oregon [15]. Similarly, Anderson and West’s analysis of the Minneapolis-St. Paul metropolitan area found that open spaces provide more value to homes in certain neighborhoods, such as those near the central business district or with many children [16]. Debrezion et al. found that real estate prices in three Dutch metropolitan areas are affected more by the most frequently chosen railway station in an area than the nearest station, and this effect is more pronounced in more urbanized areas [17].

Other work has analyzed economic statistics in populations near universities and hospitals. Moore and Sufrin concluded that large nonprofit institutions such as universities and hospitals can generate employment and personal income through interregional trade [18]. Beeson and Montgomery found that employment growth rates and income are higher in areas with higher-ranked universities, the probability of being employed as a scientist or engineer increases with local university research and development funding, and the probability of being employed in a high-tech industry increases with the number of local university graduates [19]. Hedrick et al. found that university commercial activities reduce private employment in small counties, particularly in the areas of finance, insurance, and real estate, but university enrollment and spending increase local employment, leading in a net positive effect on employment [20]. Moore’s analysis of the State University of New York university system found that per capita income generation in counties with a university is negatively correlated with per capita personal income; in other words, the greatest impact on income generation per capita is found in counties with lower personal incomes [21].

Methods

Data collection

Median ZIP code home price and rent

Zillow maintains a dataset of home and rental data for public use [22]. For our ZIP code home price analysis, we used the ZIP Code Zillow Home Value Index (ZHVI) data for

May 2017, which lists the median home price in 13,105 ZIP codes for each month from no earlier than April 1996 to May 2017. For our ZIP code rent analysis, we used the May 2017 ZIP Code Zillow Rent Index (ZRI) data, which lists the median rent in 15,918 ZIP codes for each month from no earlier than November 2010 to May 2017. Apartments were not included in the ZRI calculation, thus our statements regarding rent in our ZIP code-level analysis refer only to the rent of houses. In our ZIP code-level analysis, we use the terms “home price” and “rent” to refer to ZHVI and ZRI, respectively. Note that these amounts are computed based on Zillow’s estimate of market price and rent.

ZIP code population data

For population data, we used the 2010 census data provided by the United States Census Bureau [23]. We used two datasets extracted from this data for our analysis: a list of ZIP code tabulation area (ZCTA) populations, and a list of ZCTA population densities, where population density is given by the average number of people per square mile. We assume that the population and population density of each ZCTA are equal to the population and population density of the ZCTA’s corresponding ZIP code.

Universities

We collected university details via a twofold approach restricted to universities in the United States. The first step consists of collecting details about the universities in the United States from Wikipedia [24]. This data source provides many crucial details about the university such as name, number of enrolled students, location and university-type to name a few. The second step includes finding rankings for these universities, the data for which is collected from US News and World Report’s ranking and is restricted to the first 200 ranked universities, while the others are unranked [25]. For our ZIP code-level analyses of price, rent and vacancies over time, we use four subsets of ZIP codes based on the number of students enrolled in a university in each ZIP code as described in Table 1. This distribution was selected to give each subset relatively similar sizes between ZIP codes with home price data and ZIP codes with rent data. Each ZIP code that contains more than one university is assigned to the subset corresponding to the university in that ZIP code with the most enrolled students.

Hospitals

The Centers for Medicare and Medicaid Services (CMS) provide data used by the Medicare.gov website, including data on hospitals and physicians [26, 27]. Using the hospital

Table 1 Distribution of ZIP codes with home price data based on the number of enrolled students

Number of enrolled students	ZIP codes with home price data	ZIP codes with rent data
0 (No university)	12,473	15,153
Fewer than 10,000 Students	501	611
10,000–20,000 students	73	85
20,000 or more students	58	69

Table 2 Distribution of ZIP codes with home price data based on the number of affiliated doctors

Number of affiliated doctors	ZIP codes with home price data	ZIP codes with rent data
0 (No hospital)	10,819	13,009
Fewer than 100 doctors	309	539
100–500 doctors	1496	1837
500 or more doctors	481	533

Table 3 Home Listings details: Number of records, start and end date of record collection

Home data type	Number of records	Start date	End date
For rent	267,486	2017-04-03	2017-04-15
For sale	2,786,895	2017-05-03	2017-05-26

data, we determined which ZIP codes contained a hospital. To determine the number of doctors each hospital has, we used the affiliated hospitals listed for each physician in the physician data. For our ZIP code-level analyses of price, rent and vacancies over time, we use four subsets of ZIP codes based on the number of doctors affiliated with a hospital in each ZIP code as described in Table 2. As above, this distribution was selected to give each subset relatively similar sizes between ZIP codes with home price data and ZIP codes with rent data. Each ZIP code that contains more than one hospital is assigned to the subset corresponding to the hospital in that ZIP code with the most affiliated doctors.

Home listings

We collected data related to homes available for rent and sale from an online listings source that provides various details related to each home such as rent/sale price, home address, number of bedrooms and bathrooms, ZIP code and exact location (latitude and longitude). Apartments account for 7% of the rental data. We used each home's latitude and longitude to calculate the distance from any universities in the same ZIP code or neighboring ZIP codes. We cleaned the data, which includes removing entries with no details about the location, rent/sale price and number of bedrooms. In our home price analysis, we use the term “home price” to refer to the listed sale price and “rent” to refer to the listed rent price. A quantitative summary of the data of homes for rent and for sale is shown in Table 3.

Home vacancy

The US Department of Housing and Urban Development (HUD) provides home vacancy data [28]. This dataset includes the vacancy statistics for homes and businesses within each census tract. Census Tracts are “small, relatively permanent statistical subdivisions of a county or equivalent entity that are updated by local participants prior to each decennial census as part of the Census Bureau's Participant Statistical Areas Program” [29]. We mapped the census tract data to ZIP codes by

using the Tract-ZIP code mapping provided by HUD and assuming a uniform distribution of vacant homes in each tract. The vacancy details include statistics such as the count of vacant homes, count of homes, and periods of vacancy.

Table 4 shows a summary of the types of data used and their use in either or both our ZIP code-level analysis and home-level analysis.

ZIP code-level analysis

We used two metrics to analyze median home price and rent in each ZIP code. The first of these is average annual appreciation, which is the average difference in median home price or rent in a ZIP code compared to twelve months prior. To calculate this, we sampled the median home price and rent for May of each year. The second metric is volatility. Given P_z , a list of median home price or rent over time in ZIP code z , we define volatility as σ/μ , where σ is the standard deviation of the values in P_z and μ is the mean of the values in P_z .

We also analyzed the percentage of vacancies in each ZIP code. For our analysis, we averaged the ratio of vacant homes over the four most recent quarters in our data (Q3 2016, Q4 2016, Q1 2017 and Q2 2017) for each ZIP code. This was done to account for changes in the vacancy ratio over the course of a year due to homes with seasonal vacancies (e.g. vacation homes).

For ZIP codes that contain a university or a hospital, we analyzed each of these metrics as a function of the size of the university or hospital to observe their correlations. We define size as the number of students enrolled in a university or the number of doctors affiliated with a hospital. We calculate these correlations using the Pearson correlation coefficient. For random variables X and Y , the Pearson correlation coefficient is defined as $\rho_{X,Y} = \text{cov}(X,Y)/\sigma_X\sigma_Y$, where $\text{cov}(X,Y)$ is the covariance of X and Y , σ_X is the standard deviation of X and σ_Y is the standard deviation of Y [30].

In addition to analyzing all ZIP codes together, we also partitioned the ZIP codes across various dimensions and analyzed each partition separately. Table 5 shows these dimensions and the threshold used to split the ZIP codes into two partitions. We also analyzed subsets of ZIP codes in metropolitan areas or non-metropolitan areas.

Table 4 Data source description and usage summary

Type	Description	Usage
Median ZIP code home price and rent	Monthly median rent and sale prices by ZIP code	ZIP code-level analysis
ZIP code population data	Demographic data of each ZIP code in USA	Both
University details	Statistical and locational details of 1991 universities	Home-level analysis
Hospital details	Statistical and locational details of 4691 hospitals	Both
Vacancy details	Quarterly statistical details of home vacancies in various ZIP codes in USA	ZIP code-level analysis
Home details	Details of homes for rent and sale throughout USA	Home-level analysis

Table 5 Dimensions and thresholds to partition the ZIP codes into two sets

Splitting dimension	Splitting threshold
Number of students (ZIP codes with universities)	20,000
Number of doctors (ZIP codes with hospitals)	1000
Median ZIP code home price/rent	National ZIP code average
ZIP code population	National ZIP code average
ZIP code population density	National ZIP code average
Metropolitan area ZIP code, as listed in Zillow data	Yes/no

Home-level analysis

The home level analysis focuses on the impact of distance to a university or hospital on the home price or rent. This impact is gauged from the correlation of the home price or rent with the distance of the home from the nearest university or hospital. As in ZIP code-level analysis, we explore the Pearson correlations for various subsets of the homes, defined across various dimensions, such as the number of bedrooms or population of their ZIP code, university types and number of doctors in hospitals.

To partition the home data across such dimensions, a key step is to join the university (or hospital) and home data, as described in Table 4, based on the nearest university (or hospital) decided by the home-university (or hospital) distance. Specifically, for each home, we store its closest university (or hospital) if there is more than one institution within the home's vicinity. The result is a pool of homes which are within a defined vicinity range from a university (or hospital). We create separate data pools for home rent and sale price data. Here the defined maximum vicinity is ten miles from the location of the university or hospital. In the analysis we also consider reducing the vicinity ranges, to see if the correlation is stronger if we focus on homes that are very close to the institutions.

Note that the research and analysis did not use any data from HomeUnion.

Results

ZIP code-level analysis

Home price and rent over time

We grouped ZIP codes into four subsets: ZIP codes with no university, ZIP codes that have a small university (fewer than 10,000 students enrolled), ZIP codes that have a medium university (at least 10,000 but fewer than 20,000 students enrolled) and ZIP codes with a large university (20,000 or more students enrolled). We then compared the average of the median home price and rent over time for each of these subsets. For brevity, we refer to these as “average home price” and “average rent,” respectively. This comparison is shown in Fig. 1 for both home price and rent, where we see that *the average home price and rent are higher in ZIP codes with a university than those without, and highest in ZIP codes with a medium university*. The pairwise significance of the most recent values (May 2017), calculated using a one-tailed heteroscedastic Student's $-t$ -test, is shown in Table 6 for home prices and Table 7 for rent.

Similarly, we compared the average home price and rent over time for four ZIP code subsets based on hospitals. This comparison was between ZIP codes with no hospital,

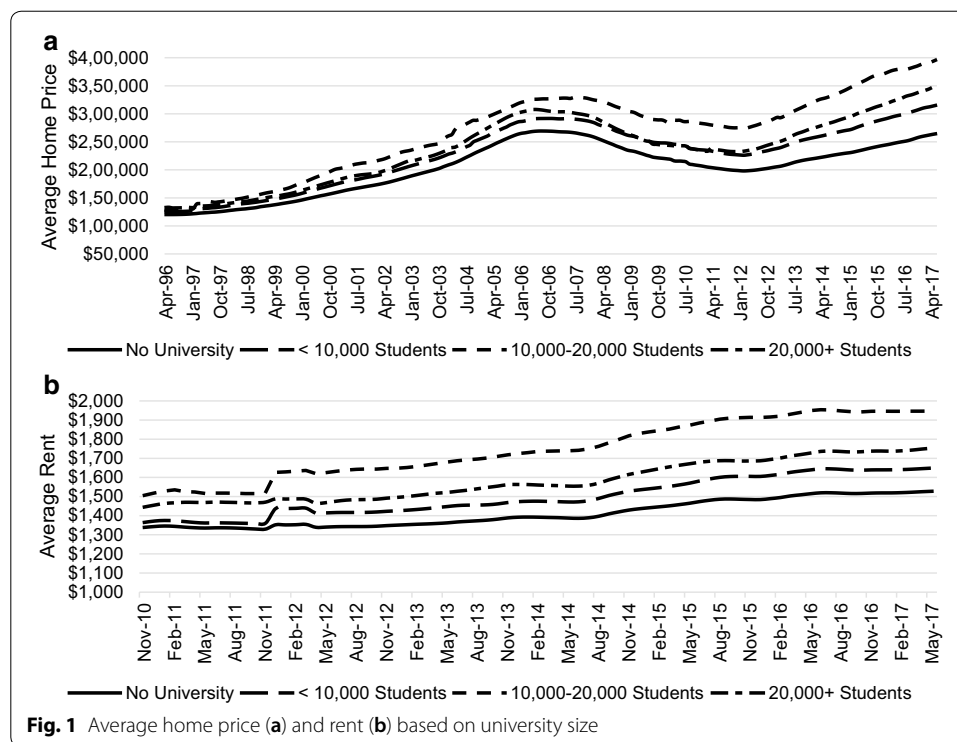


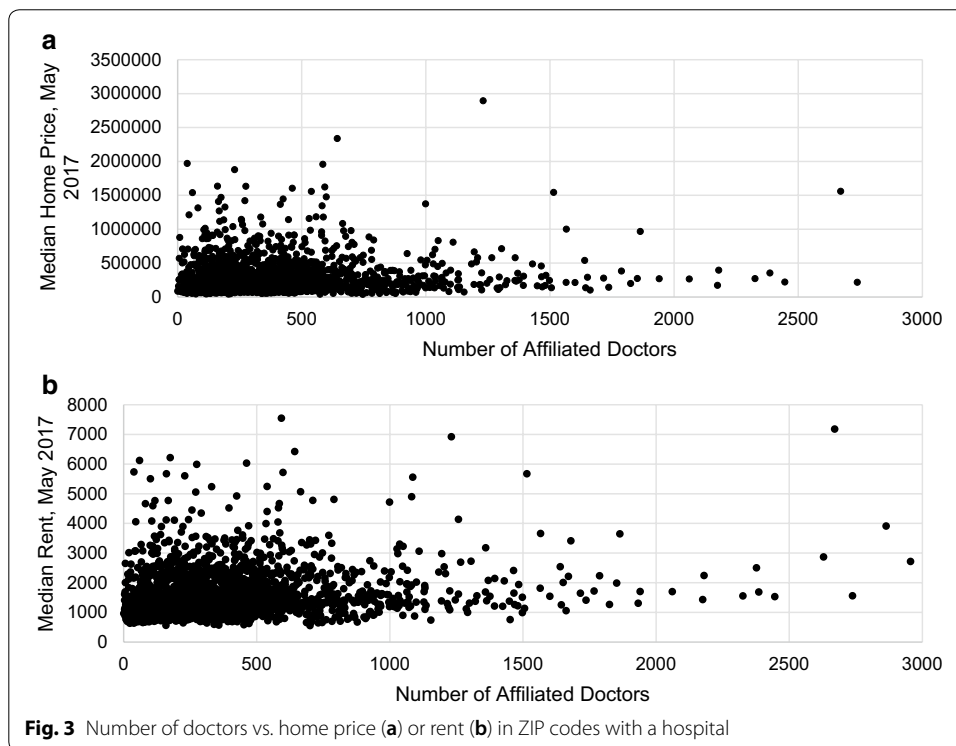
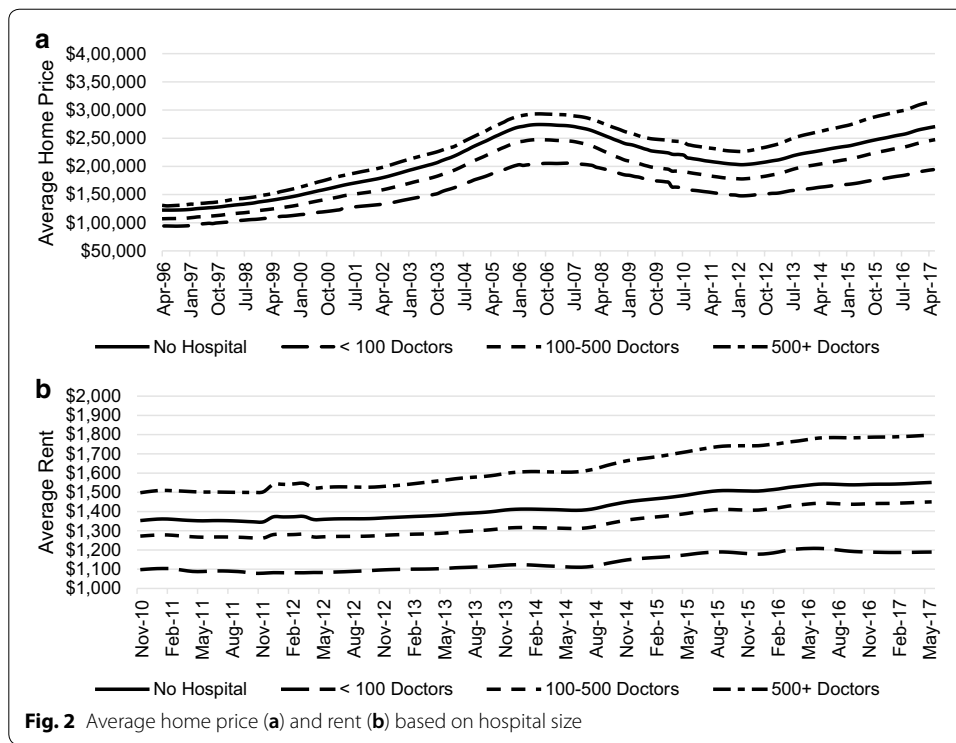
Table 6 Pairwise significance (p-values) of average May 2017 home prices in ZIP codes by university size

	No university	Fewer than 10,000 students	10,000–20,000 students
Fewer than 10,000 students	0.00361	–	–
10,000–20,000 students	0.00676	0.0733	–
20,000 or more students	0.0132	0.203	0.239

Table 7 Pairwise significance (p-values) of average May 2017 rent in ZIP codes by university size

	No university	Fewer than 10,000 students	10,000–20,000 students
Fewer than 10,000 students	0.00615	–	–
10,000–20,000 students	0.00118	0.0189	–
20,000 or more students	0.00258	0.128	0.106

ZIP codes with a small hospital (fewer than 100 affiliated doctors), ZIP codes with a medium hospital (at least 100 but fewer than 500 affiliated doctors) and ZIP codes with a large hospital (500 or more affiliated doctors). This comparison is shown in Fig. 2 for both home price and rent, where we see that ZIP codes with larger hospitals have higher average home price and rent than those with smaller hospitals, while only ZIP codes with large hospitals have higher average home price and rent than ZIP codes



with no hospital. Figure 3 shows the correlations between the number of doctors affiliated with a hospital and both home price (Pearson correlation 0.154) and rent (Pearson correlation 0.261). The p -value for both correlations is less than 1×10^{-5} .

Table 8 Pairwise significance (p-values) of average May 2017 home prices in ZIP codes by hospital size

	No hospital	Fewer than 100 doctors	100–500 doctors
Fewer than 100 doctors	4.02×10^{-10}	–	–
100–500 doctors	7.27×10^{-5}	2.85×10^{-5}	–
500 or more doctors	0.000637	1.94×10^{-11}	2.35×10^{-6}

Table 9 Pairwise significance (p-values) of average May 2017 rent in ZIP codes by hospital size

	No hospital	Fewer than 100 doctors	100–500 doctors
Fewer than 100 doctors	1.71×10^{-43}	–	–
100–500 doctors	4.11×10^{-9}	4.84×10^{-20}	–
500 or more doctors	2.72×10^{-9}	2.63×10^{-35}	4.79×10^{-15}

Table 10 Correlations between the number of students enrolled in a university and home price appreciation

Subset	Pearson correlation	p-value
ZIP codes with a university	0.132	0.000864
ZIP codes with a university and home prices below the national ZIP code average	0.136	0.0063
ZIP codes with a university and home prices above the national ZIP code average	0.115	0.0822
ZIP codes with a university and a population above the national ZIP code average	0.134	0.00126
ZIP codes with a university and population density above the national ZIP code average	0.107	0.0359

Table 11 Correlations between the number of students enrolled in a university and rent appreciation

Subset	Pearson correlation	p-value
ZIP codes with a university	0.124	0.000575
ZIP codes with a university and rent below the national ZIP code average	0.142	0.00232
ZIP codes with a university and a population above the national ZIP code average	0.117	0.00218
ZIP codes with a university and population density below the national ZIP code average	0.144	0.00793

The pairwise significance of the most recent home price and rent values (May 2017), calculated using a one-tailed heteroscedastic Student's *t*-test, is shown in Tables 8 and 9 for rent.

Appreciation of home price and rent

We found several very weak correlations between the number of students enrolled in a university and average annual home price and rent appreciation in ZIP codes with a university. These correlations are listed in Table 10 for home price appreciation and Table 11 for rent appreciation. For hospitals, we found *a weak positive correlation between the number of doctors affiliated with a hospital and average annual home price*

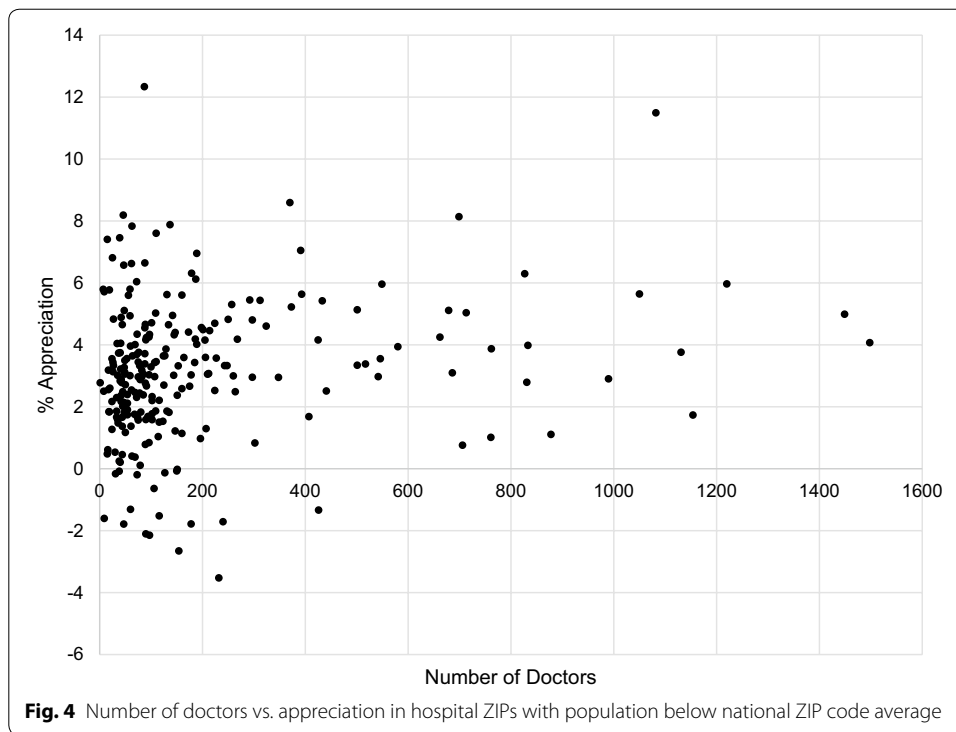


Table 12 Correlations between the number of doctors affiliated with a hospital and rent appreciation

Subset	Pearson Correlation	p-value
ZIP codes with a hospital	0.197	$< 1 \times 10^{-5}$
ZIP codes with a hospital and rent below the national ZIP code average	0.173	$< 1 \times 10^{-5}$
ZIP codes with a hospital with fewer than 1000 affiliated doctors	0.17	$< 1 \times 10^{-5}$
ZIP codes with a hospital and a population below the national ZIP code average	0.156	0.000832
ZIP codes with a hospital and a population above the national ZIP code average	0.167	$< 1 \times 10^{-5}$
ZIP codes with a hospital in a metropolitan area	0.166	$< 1 \times 10^{-5}$

appreciation in ZIP codes with a hospital and a population below the national ZIP code average (Fig. 4; Pearson correlation 0.203, p -value 0.0016). We also found a very weak correlation between the number of doctors affiliated with a hospital and average annual home price appreciation in all ZIP codes with a hospital (Pearson correlation 0.107, p -value $< 1 \times 10^{-5}$) in addition to several very weak correlations between the number of doctors affiliated with a hospital and average annual rent appreciation in ZIP codes with a hospital. These correlations are listed in Table 12.

Volatility of home price and rent

We found a weak positive correlation between the number of students enrolled in a university and home price volatility in ZIP codes with a university and a population below the national ZIP code average (Fig. 5a; Pearson correlation 0.296, p -value 0.0299) as well as several very weak correlations between the number of students enrolled in a

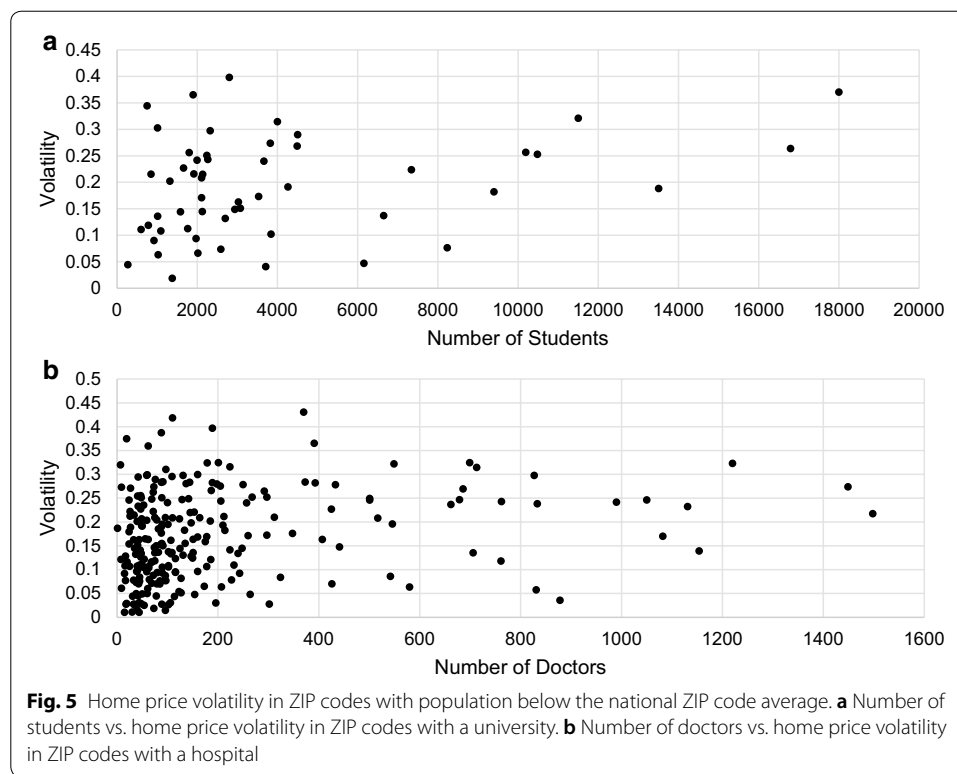


Table 13 Correlations between the number of students enrolled in a university and home price volatility

Subset	Pearson correlation	p-value
ZIP codes with a university	0.139	0.000449
ZIP codes with a university and home prices below the national ZIP code average	0.1749	0.000434
ZIP codes with a university and a population above the national ZIP code average	0.129	0.00193
ZIP codes with a university and population density below the national ZIP code average	0.124	0.0518

university and home price volatility in ZIP codes with a university. These correlations are listed in Table 13. For hospitals, we found a weak positive correlation between the number of doctors affiliated with a hospital and home price volatility in ZIP codes with a hospital and a population below the national ZIP code average (Fig. 5b; Pearson correlation 0.244, p-value 0.000134). We also found several very weak correlations between the number of doctors affiliated with a hospital and home price and rent volatility in ZIP codes with a hospital. These correlations are listed in Table 14 for home price volatility and Table 15 for rent volatility.

Vacancies

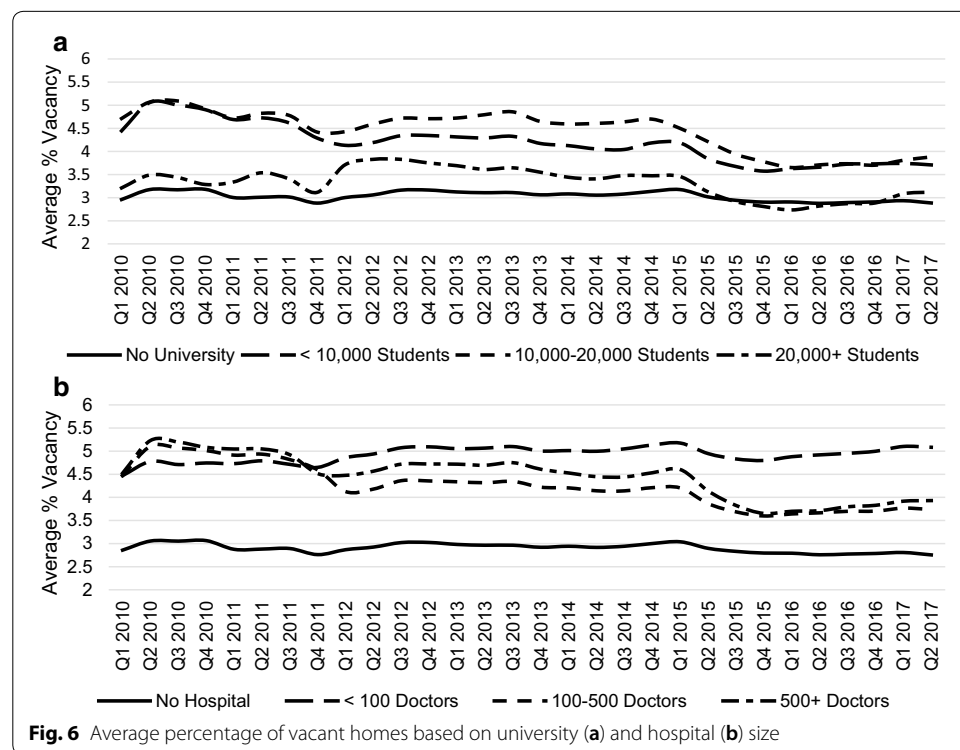
We again grouped ZIP codes into four subsets for both universities and hospitals to compare the average percentage of vacant homes between subsets. These comparisons are shown in Fig. 6. Among ZIP codes with a university, we see that the average

Table 14 Correlations between the number of doctors affiliated with a hospital and home price volatility

Subset	Pearson Correlation	p-value
ZIP codes with a hospital	0.135	$< 1 \times 10^{-5}$
ZIP codes with a hospital with fewer than 1000 affiliated doctors	0.144	$< 1 \times 10^{-5}$
ZIP codes with a hospital and home prices below the national ZIP code average	-0.152	$< 1 \times 10^{-5}$
ZIP codes with a hospital and home prices above the national ZIP code average	-0.108	0.0052
ZIP codes with a hospital and a population above the national ZIP code average	0.103	$< 1 \times 10^{-5}$
ZIP codes with a hospital and population density below the national ZIP code average	0.113	7.6×10^{-5}
ZIP codes with a hospital and population density above the national ZIP code average	-0.106	0.000537

Table 15 Correlations between the number of doctors affiliated with a hospital and rent volatility

Subset	Pearson correlation	p-value
ZIP codes with a hospital	0.105	$< 1 \times 10^{-5}$
ZIP codes with a hospital and a population below the national ZIP code average	0.147	0.00165



percentage of vacant homes is highest in ZIP codes with medium universities and lowest in ZIP codes with no university, while ZIP codes with small universities have a higher average percentage of vacant homes than ZIP codes with large universities. Among ZIP codes with a hospital, we see that the average percentage of vacant homes is highest in ZIP codes with small hospitals and lowest in ZIP codes with no hospital, while ZIP codes

with large hospitals have a higher average percentage of vacant homes than ZIP codes with medium hospitals. The pairwise significance of the most recent values (Q2 2017), calculated using a one-tailed heteroscedastic Student's $-t$ -test, is shown in Table 16 for ZIP codes grouped by university size and Table 17 for ZIP codes grouped by hospital size.

We found a weak positive correlation between the number of students enrolled in a university and the percentage of vacant homes in ZIP codes with a university and a population below the national ZIP code average (Fig. 7; Pearson correlation 0.285, p -value 0.0368). However, we also found a very weak negative correlation between the number of students enrolled in a university and the percentage of vacant homes in ZIP codes with population density below the national ZIP code average (Pearson correlation -0.134 , p -value 0.0361). Among ZIP codes with a hospital, we found very weak correlations between the number of doctors affiliated with a hospital and the percentage of vacant homes in ZIP codes with home prices above the national ZIP code average (Pearson correlation 0.14, p -value 0.000296) and rent above the national ZIP code average (Pearson correlation 0.129, p -value 0.000134).

University rankings

We found *weak negative correlations between university ranking and both home price and rent*. The Pearson correlation is -0.269 for university ranking vs. home price with a p -value of 0.021. The Pearson correlation is -0.327 for university ranking vs. rent with a p -value of 0.00271. These results are not surprising as the negative correlations imply that real estate prices tend to be higher in ZIP codes with higher ranked universities.

Home-level analysis

Our goal is to determine whether there exist subsets of the data, partitioned across the dimensions of the university, hospital and home data, in which the distance to a university or hospital is significantly correlated to home price or rent. A few possible avenues for finding these subsets are along the features of the data such as distance ranges, number of bedrooms, types of university and number of doctors affiliated with hospitals. We try to filter data layer by layer by using the feature filters to arrive at a particular high-correlation subset of data.

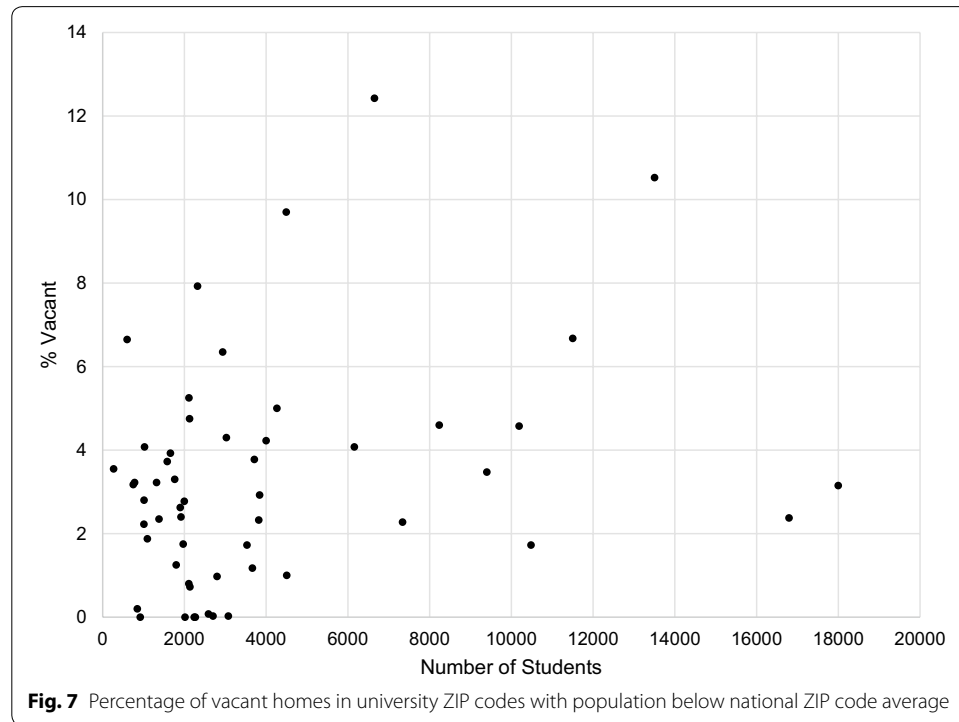
As discussed in “Methods” section, each entry in the data table consists of details for homes within ten miles of a university along with that university's details. If a home is near multiple universities, only the entry with the shortest distance from a university

Table 16 Pairwise significance (p-values) of average home vacancy percentage in ZIP codes by university size

	No university	Fewer than 10,000 students	10,000–20,000 students
Fewer than 10,000 students	1.89×10^{-11}	–	–
10,000–20,000 students	0.0013	0.43	–
20,000 or more students	0.362	0.00802	0.0246

Table 17 Pairwise significance (p-values) of average home vacancy percentage in ZIP codes by hospital size

	No hospital	Fewer than 100 doctors	100–500 doctors
Fewer than 100 doctors	1.3×10^{-66}	–	–
100–500 doctors	8.55×10^{-48}	1.13×10^{-21}	–
500 or more doctors	4.9×10^{-15}	3.07×10^{-11}	0.206



is considered. We applied the same scheme to home-hospital data. Table 18 shows the average number of homes for sale and for rent within ten miles of a university or hospital.

Analysis of university and hospital proximity on home price and rent

As a preliminary analysis, we analyzed the effective distance range to which the presence of a university affects home prices and rent. Table 19 shows the correlations between home price/rent and distance from the nearest university based on different maximum distances. Although all such correlations are very weak, we observed slightly higher correlations for both home price and rent among homes within two miles of a university.

Table 18 Average number of homes near a university or hospital (within ten-mile radius)

Home type	University	Hospital
For rent	165.028	497.483
For sale	107.686	315.817

Table 19 Distribution of homes by distance from nearest university and home price/rent-distance correlation

Miles from a university	Homes with home price data	Pearson correlation (home price-distance)	Homes with rent data	Pearson correlation (rent-distance)
< 2	129,102	− 0.149	115,340	− 0.087
< 3	228,185	− 0.122	153,668	− 0.076
< 4	326,055	− 0.101	183,404	− 0.055
< 5	414,769	− 0.095	205,884	− 0.038
> 5 and < 10	290,713	− 0.017	61,602	0.016

Table 20 Home price/rent-distance from hospital correlations for various distance ranges up to ten miles

Home price-distance analysis			Rent-distance analysis		
Miles from a hospital	Number of homes	Pearson correlation	Miles from a hospital	Number of homes	Pearson correlation
Any	877,067	− 0.0495	Any	300,768	0.079
< 3 miles	303,313	− 0.081	< 2 miles	123,313	− 0.013
> 3 miles	573,754	0.006	> 2 miles	177,455	0.094

Therefore, unless mentioned otherwise, all further experiments related to homes near universities limit the dataset to homes that are within two miles of the nearest university.

Next, we examined at the effect of proximity to a hospital on home prices and rent. As a preliminary experiment, we calculated the home price-distance and rent-distance correlations based on different maximum distances and found the best correlations by partitioning the home data at three miles for home prices and two miles for rent. As seen in Table 20, *homes within a three-mile radius of a hospital have higher correlation between home price and distance from a hospital*. In the remainder of this section, we consequently focus on other data filters based on the number of bedrooms in a home and the number of doctors affiliated with a hospital to find correlations between home price and distance from a hospital. Table 20 also shows that for rent data, the highest correlation between rent and distance from the nearest hospital exists beyond a two-mile radius from the hospitals. Interestingly, the correlation is positive, that is, the rent is higher for homes farther from a hospital.

Analysis of university/hospital proximity: price and rent analysis by number of bedrooms

For these experiments, we partitioned the home data based on the number of bedrooms. In the first of these experiments, we analyzed the correlations between home price/rent and distance from a university within two miles of a university. We found that *two-bedroom homes have the highest correlation between home price and distance from a university* (Pearson correlation -0.319). This was followed closely by one-bedroom homes. The correlation was very weak for homes with more than two bedrooms. We also found a *weak correlation between rent and distance from a university for one-bedroom homes* (Pearson correlation − 0.191). The Pearson correlations for various numbers of bedrooms and the home counts for each such category are shown in Table 21.

Table 21 Home price/rent-distance from university correlations based on distribution by number of bedrooms

Number of bedrooms	Homes with home price data (< 2 miles)	Pearson correlation (home price-distance)	homes with rent data (< 2 miles)	Pearson correlation (rent-distance)
1	13,663	−0.219	32,473	−0.191
2	31,050	−0.319	40,371	−0.142
3	44,028	−0.225	26,369	−0.149
4	24,206	−0.212	11,566	−0.056
>4	16,155	−0.119	4563	0.073

Table 22 Home price/rent-distance from hospital correlations based on distribution by number of bedrooms

Number of bedrooms	Homes with home price data (< 3 miles)	Pearson correlation (home price-distance)	Homes with rent data (> 2 miles)	Pearson correlation (rent-distance)
1	20,524	−0.223	24,102	−0.008
2	61,831	−0.140	55,195	0.015
3	114,657	−0.100	63,904	0.055
4	68,914	−0.095	26,151	0.114
>4	37,837	−0.048	2935	0.186

Next, we analyzed the correlations between home price/rent and distance from a hospital. As shown in our earlier experiments, the set of homes less than three miles away from the nearest hospital is a good candidate for analyzing the effect of proximity to a hospital on home prices, while the set of homes more than two miles away from the nearest hospital is a good candidate for analyzing the effect of proximity to a hospital on rent. For home price data, Table 22 shows that single-bedroom homes have a higher correlation between home price and distance from a hospital (Pearson correlation −0.223) than the other bedroom categories. In the next subsection, we shall thus focus only on these one-bedroom homes. Among the correlations between rent and distance from the nearest hospital for each category, the correlations get stronger as the number of bedrooms increases, as shown in Table 22. The strongest of these is a weak positive correlation for homes with more than four bedrooms (Pearson correlation 0.186).

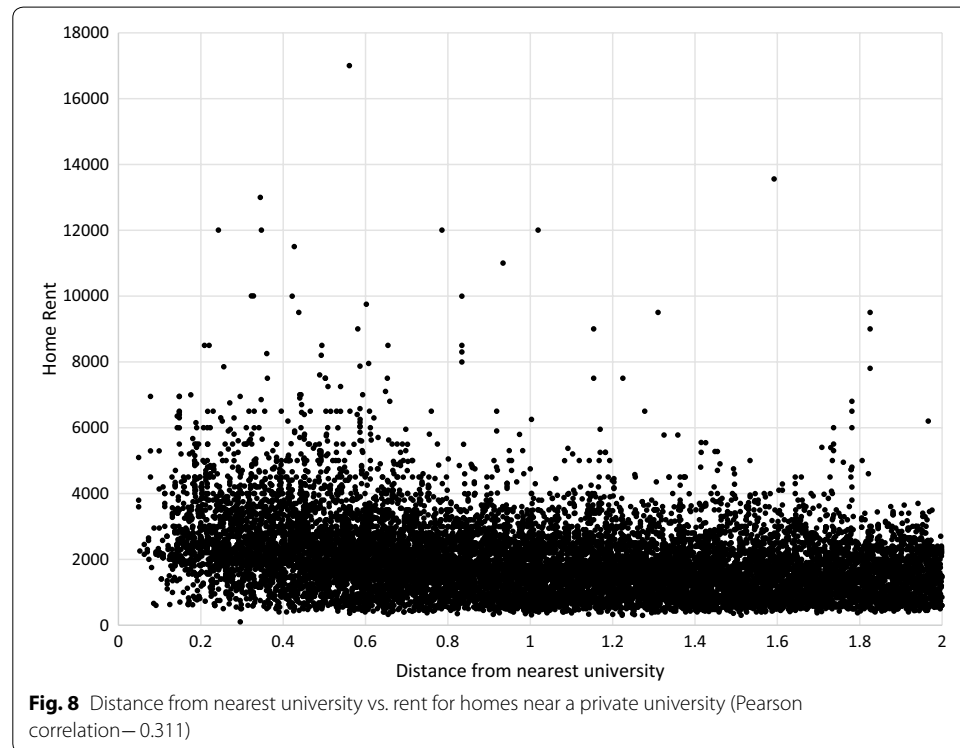
Analysis of university proximity: price and rent analysis by type and rank of university

For home price analysis within two miles of a university, we classify the universities into the following three types: public, private and other. Also, as observed in previously, two-bedroom homes near universities provide a good enough correlation to be explored further. Table 23, which compares the correlations between home price and distance for these types of universities, shows that *two-bedroom homes have a weak negative correlation between home price and distance from a private university within a two-mile radius*.

For rent analysis within two miles of a university, we again classify the universities into the three types mentioned previously and limit our experiment to one-bedroom homes due to the higher correlations found with those homes. In Table 23, we found a *negative correlation between rent and the distance from a university for one-bedroom homes*

Table 23 Home price/rent-distance correlations based on university type distribution

University type	University count	Homes with home price data (< 2 miles, 2 bedrooms)	Pearson correlation (home price-distance)	Homes with rent data (< 2 miles, 1 bedroom)	Pearson correlation (rent-distance)
Private	1068	16,680	− 0.368	16,992	− 0.311
Public	548	7812	− 0.220	6463	− 0.203



near a private university (Fig. 8; Pearson correlation − 0.311). We also observed a weaker correlation for one-bedroom homes near public universities. Note that we have omitted results for homes near “other” universities since there were very few of these universities compared to the other two types and they yielded very small correlations.

We also considered the rank of universities in our analysis. The university rankings provided by US News and World Report provide data for only the top 200 schools. However, we found no significant correlations in our experiments involving university rankings as can be seen in Table 24.

We then checked for any interesting correlations between home price/rent and distance from university on the filters of ranked or unranked universities for homes within a two-mile radius. Further, on filtering over one-bedroom homes for rent and two-bedrooms for home price, we found similar correlation for ranked as well as unranked universities. Results for these experiments are shown in Tables 25 and 26. Hence, as per our analysis, the ranking of a university does not play a crucial role in the dynamics of real estate prices of nearby homes.

Table 24 Correlation of distance from university to home price/rent for top ranked/unranked universities

	Distance-rent correlation	Distance-price correlation
Top 200 ranked universities	− 0.007	− 0.055
Unranked universities	− 0.005	− 0.039

Table 25 Correlation of university distance (within two miles) to home price/rent for top ranked/unranked universities

	Distance-rent correlation	Distance-price correlation
Top 200 ranked universities	− 0.052	− 0.137
Unranked universities	− 0.105	− 0.160

Table 26 Correlation of university distance (within two miles) to two-bedroom home price/one-bedroom rent for top ranked/unranked universities

	Distance-rent correlation (one-bedroom homes)	Distance-price correlation (two-bedroom homes)
Top 200 ranked universities	− 0.230	− 0.298
Unranked universities	− 0.180	− 0.327

Table 27 Home price/rent-distance correlations by number of doctors affiliated with the nearest hospital

Number of doctors	Homes with home price data (1 bedroom, < 3 miles)	Pearson correlation (home price-distance)	Homes with rent data (> 4 bedrooms, > 2 miles)	Pearson correlation (rent-distance)
< 500	12,976	− 0.196	5168	0.214
> 500	7548	− 0.293	2935	− 0.085

Analysis of hospital proximity: price and rent analysis by number of affiliated doctors

For home price data, we consider only single-bedroom homes as they exhibited the highest correlation between home price and distance from the nearest hospital. Table 27 shows that *single-bedroom homes near larger hospitals (more than 500 doctors) have a higher distance-home price correlation compared to those near smaller hospitals.*

For rent data, we restricted our analysis to homes with more than four bedrooms at distance of over two miles from the nearest hospital. We then categorize this data into two subsets of fewer than 500 or more than 500 doctors affiliated with the nearest hospital. Table 27 shows that *larger homes (more than four bedrooms) near a smaller hospital (fewer than 500 doctors) had a significantly higher rent-distance correlation as compared to homes near a larger hospital.*

Discussion

In our analysis of average ZIP code median home price and median rent over time (“average home price” and “average rent”), we found that the average home price and rent are higher in ZIP codes with a university than those without, and highest in ZIP codes with a medium-sized university (10,000–20,000 students). One possible explanation for this observation is that public universities tend to have a more positive effect on home price and rent, as most medium-sized universities in our analysis are public. We also found that ZIP codes with larger hospitals have higher average home price and rent than those with smaller hospitals, while only ZIP codes with large hospitals have higher average home price and rent than ZIP codes with no hospital. One possible reason why ZIP codes with small and medium hospitals have lower home price and rent than ZIP codes with no hospital is that smaller hospitals tend to be in more remote areas with lower real estate prices. In general, these measures were positively affected by the presence of a university and negatively affected by the presence of a hospital (this should not be confused by the impact of the hospital distance of individual homes within a ZIP code). Note that the existence of a large (or small) university in a ZIP code does not imply the existence of a large (or small) hospital or vice versa (Table 28).

The strongest ZIP code-level correlations discovered in this study were found for smaller ZIP codes (population below the national ZIP code average). The reason may be that institutions have a higher impact in smaller ZIP codes as they are one of the main employers or drivers of economic activity. Specifically, we found that *in smaller ZIP codes with at least one hospital, there is a positive correlation (0.203) between the number of affiliated doctors and home price appreciation*. This result, along with several weaker correlations we found between home price/rent and appreciation, agrees with our expectation that appreciation is higher near larger institutions. Our analysis of volatility in smaller ZIP codes showed that for ZIP codes with at least one university, there is a positive correlation (0.296) between the number of enrolled students and home price volatility, and for ZIP codes with at least one hospital there is a positive correlation (0.244) between the number of affiliated doctors and home price volatility. *These results on volatility are opposite from what we expected, as larger universities or hospitals generally imply more job security for the area, and hence one would expect lower price volatility as well*. We also found that smaller ZIP codes with at least one university have a positive correlation (0.285) between the number of students enrolled and the percentage of vacant homes. This agrees with our expectation that the vacancy rate is higher near larger universities, as many students leave for the summer.

Table 28 Number of ZIP codes with both a university and a hospital for each subset

	No university	Fewer than 10,000 students	10,000–20,000 students	20,000 or more students
No hospital	N/A	305	43	32
Fewer than 100 doctors	290	16	2	1
100–500 doctors	1340	122	24	10
500 or more doctors	404	58	4	15

Our analysis of homes near universities or hospitals based on the number of bedrooms in homes showed several interesting correlations. We found that the correlation between home price and distance from a university is strongest for two-bedroom homes (-0.319), while the correlation between rent and distance from a university is strongest with one-bedroom homes (-0.191). *That is, smaller homes are of higher demand closer to universities.* This conclusion seems logical as most of the occupants within a two-mile radius from a university would be students and not big families.

Similarly, we found that the correlation between home price and distance from a hospital is strongest for one-bedroom homes (-0.223), which could imply high demand for single bedroom homes near hospitals. In contrast, the correlation between rent and distance from a hospital was strongest for homes with more than four bedrooms (0.186), which implies that larger families may prefer to live farther from a hospital.

We found negative correlations between the price of a two-bedroom home and distance from a private university (-0.368) or a public university (-0.220). We also found negative correlations between the rent for a one-bedroom home and the distance from a private university (-0.311) or a public university (-0.203). *A probable cause for the difference in correlations between public and private universities is that private university students may be willing to pay more rent to be closer to the university.* These results also show that renting a home near a university has a slightly lower correlation compared to the sale of a home, implying a higher demand for buying a home. This may be accounted for by sales to investors for the purposes of renting out these homes. This possibility may be a subject for future research.

As noted in previous sections, economic laws as viewed from the lens of homes lying in the proximity of universities and hospitals act in subtle ways. What seems to be true near a university may not be true near a hospital. Indeed, one should not be altogether surprised by those findings. Although both universities and hospitals are magnets of highly educated workforce, universities have students while hospitals generally do not (with the exception of teaching hospitals, which are by definition universities). Demand for housing is a function of multiple factors which aren't altogether easy to decouple—variations in demand differ according to factors that would appeal to different demographic and economic strata. For example, students fuel demand for inexpensive housing lying in close proximity to a university campus. On the other hand, hospitals' professional staff, some highly paid (doctors, senior nurses and senior management), are adult, mostly with families that compete for larger homes, in neighborhoods having amenities commensurate to their needs and desires. Clearly the differences between those two demographical strata are stark. That said, there are many examples of universities that are situated in what may be considered as "inner city" and those include some of the finest universities in this country, e.g. University of Pennsylvania and Temple University (both in Philadelphia), University of Southern California (Los Angeles), Wayne State University (Detroit), etc., where this analysis would prove wrong. More often, there are many examples of hospitals situated in what one would consider a "bad" part of town, where the professional staff does not live; where doctors, nurses, management, etc. drive to work, sometime for an hour one-way, "put their time" and drive back to their home in a middle- or upper-middle-class suburbia. It is also interesting to note that "job security" plays a secondary role, if that. Indeed, those "old" notions of job security do not

seem to play prominently into the economic calculus, especially as it manifests in real estate terms. However, as expected, what is confirmed by the analysis is the notion that demand for modest rent housing is high near an employer promising job security.

Limitations

When considering the distance between homes and universities/hospitals, we used the geographical distance without regard to elevation or roads. The Google Maps API could be used to account for these, but the API rate limits imposed by Google made this impractical. The university rankings provided by US News and World Report provide data for only the top 200 schools. For that, we generally study them in two groups, ranked and unranked.

The CMS hospital data includes smaller medical centers in addition to traditional hospitals. These medical centers tend to have very few affiliated doctors, which may affect our calculations involving subsets of ZIP codes that contain these medical centers. However, these medical centers are often in small cities with no other hospital nearby, thus we believe they are appropriate for our analysis.

Two limitations apply to the Zillow data. First, the prices/rent are based on listed prices/rent and not actual sale prices/rent. Second, the median monthly home price and rent data provided by Zillow had 1 or more months of data missing for some ZIP codes. To account for a ZIP code has one or more consecutive months of missing data between months with data, we assume the change in home price or rent is linear during the months with missing data. If a ZIP code's first month of data is after the first month of Zillow data (April 1996 for home prices and November 2010 for rent), that ZIP code is not included in our calculation of average median home price/rent for months before that ZIP code's first month of data, and our calculation of appreciation and volatility for that ZIP code are made using only the range of months for which we have data for that ZIP code.

We assume a ZIP code containing a university or a hospital contained that institution throughout the entire range of dates used in calculations for that ZIP code; however, some universities or hospitals may have been built after the start of their containing ZIP codes' ranges of home price/rent data.

As discussed above, many factors affect the demand—and therefore the price—of housing. While our study focuses on a select few factors, our home price and rent data may be affected by one or more other variables that we do not consider.

Conclusions

We analyzed several measures of real estate valuation near universities and hospitals based on both individual home sales and ZIP code level aggregates. In our ZIP code-level analysis, we found that ZIP codes with universities tend to have above average median home price and median rent, especially those with medium-sized universities, while ZIP codes with hospitals tend to have below average median home price and median rent, with the exception of those with large hospitals, and that less populated ZIP codes have positive correlations between the number of doctors affiliated with a hospital and home price appreciation, and between the number of enrolled university students and home vacancy rate. Notably, less populated ZIP codes also have positive correlations between

home price volatility and both the number of enrolled students (in ZIP codes with a university) and the number of affiliated doctors (in ZIP codes with a hospital), which is surprising given that one would expect these institutions to have a stabilizing effect on home prices. In our home-level analysis, we found that the home price and rent for smaller homes tend to be the most affected by distance from a university, while distance from a hospital has greater effect on both the price of one-bedroom homes as well as on the rent of large homes. Of particular interest is our finding of a positive correlation between rent and distance from a hospital beyond two miles, suggesting that renters prefer homes in areas without a hospital nearby.

The findings point at complex interactions between demand and supply in the ZIP codes and homes under study. There is little doubt that supply–demand curves should be stratified by price points and possibly additional factors. This is clearly demonstrated in the city of Irvine, California, (ZIP code 92618) where two large healthcare facilities, Kaiser and Hoag hospitals, employ a large staff at a diverse income levels: from board-certified surgeons at the higher end, to nurse assistants and orderlies at the other. As one may readily check on Zillow or similar websites, there is little, if any “affordable” housing in the vicinity of ZIP code 92618, presumably necessitating low-income hospital staff to seek housing in lower-rent areas. An overall theory to explain behavior of real estate in the vicinity of a university or a hospital may prove complex as it should take into account myriad hard-to-measure factors. We will take this kind of analysis in a subsequent study, specifically the effects of interactions between economics, demographics, and amenities, to further explore how all the effects interact with the metrics we normally associate with real estate and potentially develop a machine learning model based on these analyses.

Abbreviations

CMS: Centers for Medicare and Medicaid Services; HUD: US Department of Housing and Urban Development; ZCTA: ZIP code tabulation area; ZHVI: Zillow Home Value Index; ZRI: Zillow Rent Index.

Authors' contributions

RR performed the ZIP-code level analysis and added the methods and results relevant to that analysis to the manuscript. DP performed the home-level analysis and added the methods and results relevant to that analysis to the manuscript. VH conceived the study and provided coordination and guidance in the experiments and writing of the manuscript. JRB provided guidance in the experiments and assisted in writing the Introduction and Discussion sections of the manuscript. NS provided guidance in the experiments and the writing of the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Computer Science and Engineering, University of California, 900 University Ave., Riverside, CA 92521, USA. ² Barr Analytics, Irvine, CA, USA. ³ HomeUnion, Inc., 2010 Main St. Ste. 250, Irvine, CA 92614, USA.

Acknowledgements

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The median ZIP code home price (ZHVI) and rent (ZRI) datasets are available from Zillow [22].

The ZIP code population data are available from Splitwise, which includes ZIP code population data [31] and ZIP code population density data [32]. These datasets were derived from the US Census Bureau [23].

The university data was generated from Wikipedia [24] and US News and World Report's National University Rankings [25].

The hospital data was generated from two datasets provided by the Centers for Medicare and Medicaid Services: Hospital General Information [26] and the Physician Compare National Downloadable File [27].

The home listings data is from a public home listings website.

The home vacancy dataset is available from the US Department of Housing and Urban Development [28].

Funding

This work was partially supported by NSF grants IIS-1619463, IIS-1746031, IIS-1447826 and IIS-1838222.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 7 November 2018 Accepted: 18 January 2019

Published online: 31 January 2019

References

- Bostic RW, Longhofer SD, Redfearn CL. Land leverage: decomposing home price dynamics. *Real Estate Econ*. 2007;35(2):183–208.
- Diewert WE, Haan JD, Hendricks R. Hedonic regressions and the decomposition of a house price index into land and structure components. *Econometric Rev*. 2015;34(1–2):106–26.
- Benson ED, Hansen JL, Schwartz AL, Smersh GT. Pricing residential amenities: the value of a view. *J Real Estate Finance Econ*. 1998;16(1):55–73.
- van Praag B, Baarsma BE. The shadow price of aircraft noise nuisance (No. 01-010/3). Tinbergen Institute Discussion Paper. 2001.
- Ozdenerol E, Huang Y, Javadnejad F, Antipova A. The impact of traffic noise on housing values. *J Real Estate Pract Educ*. 2015;18(1):35–54.
- Barr JR, Ellis EA, Kassab A, Redfearn CL, Srinivasan NN, Voris KB. Home price index: a machine learning methodology. *Int J Semantic Comput*. 2017;11(01):11–33.
- Cesa-Bianchi A, Cespedes LF, Rebucci A. Global liquidity, house prices, and the macroeconomy: evidence from advanced and emerging economies. *J Money Credit Banking*. 2015;47(S1):301–35.
- Favara G, Imbs J. Credit supply and the price of housing. *Am Econ Rev*. 2015;105(3):958–92.
- Muehlenbachs L, Spiller E, Timmins C. The housing market impacts of shale gas development. *Am Econ Rev*. 2015;105(12):3633–59.
- Waddell P, Berry BJ, Hoch I. Residential property values in a multinodal urban area: new evidence on the implicit price of location. *J Real Estate Finan Econ*. 1993;7(2):117–41.
- Nau C, Bishai D. Green pastures: do US real estate prices respond to population health? *Health Place*. 2018;49(1):59–67.
- Otto P, Schmid W. Spatiotemporal analysis of German real-estate prices. *Ann Reg Sci*. 2018;60(1):41–72.
- Raschoff S, Humphries S. Zillow talk: the new rules of real estate. New York: Grand Central Publishing; 2015.
- Turner J. The impact of walkability on home values: findings from neighborhoods in three Bay Area cities. *Digital Commons @ Cal Poly*. 2017. <http://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=1178&context=crpsp>. Accessed Sept 2017.
- Bolitzer B, Netusil NR. The impact of open spaces on property values in Portland, Oregon. *J Environ Manag*. 2000;59(3):185–93.
- Anderson ST, West SE. Open space, residential property values, and spatial context. *Reg Sci Urban Econ*. 2006;36(6):773–89.
- Debrezion G, Pels E, Rietveld P. The impact of rail transport on real estate prices: an empirical analysis of the Dutch housing market. *Urban Stud*. 2011;48(5):997–1015.
- Moore CL, Sufrin SC. The impact of a nonprofit institution on regional income. *Growth and Change*. 1974;5(1):36–40.
- Beeson P, Montgomery EB. The effects of colleges and universities on local labor markets. Massachusetts: National Bureau of Economic Research; 1990.
- Hedrick DW, Henson ST, Mack RS. The effects of universities on local retail, service, and F.I.R.E. employment: some cross-sectional evidence. *Growth Change*. 1990;21(3):9–20.
- Moore GA. Local income generation and regional income redistribution in a system of public higher education. *J High Educ*. 1979;50(3):334–48.
- Zillow data. <http://www.zillow.com/research/data/>. Accessed June 2017.
- United States Census Bureau data. <http://www.census.gov/data.html>. Accessed June 2017.
- Wikipedia. <http://en.wikipedia.org>. Accessed June 2017.
- National university rankings. <http://www.usnews.com/best-colleges/rankings/national-universities>. Accessed June 2017.
- Hospital general information. <http://data.medicare.gov/Hospital-Compare/Hospital-General-Information/xubh-q36u>. Accessed June 2017.
- Physician compare national downloadable file. <http://data.medicare.gov/Physician-Compare/Physician-Compare-National-Downloadable-File/mj5m-pzi6>. Accessed June 2017.
- HUD aggregated USPS administrative data on address vacancies. <http://www.huduser.gov/portal/datasets/usps.html>. Accessed July 2017.
- Geographic terms and concepts - census tract. http://www.census.gov/geo/reference/gtc/gtc_ct.html. Accessed January 2018.
- Pearson K. Notes on regression and inheritance in the case of two parents. *Proc R Soc London*. 1895;58(1):240–2.
- The 2010 US census population by zip code (totally free). <http://blog.splitwise.com/2013/09/18/the-2010-us-census-population-by-zip-code-totally-free/>. Accessed June 2017.
- Free US population density and unemployment rate by zip code. <http://blog.splitwise.com/2014/01/06/free-us-population-density-and-unemployment-rate-by-zip-code/>. Accessed June 2017.