**On the market: The longitudinal relationship between housing market valuation and job market postings in California**

**1. Introduction**

**1.1 Question**

What is the temporal relationship between job postings and housing prices in California, -- specifically we will investigate the strength and effect of the lags between these two factors, and also looking for its dependence on job categorization and education level, which we think are two most important demographic data that will influence the characteristics of the job market.

**1.2 Significance of Our Question**

California is the largest economy in the United States, and is a major provider of jobs in the technology and healthcare sector (Table 7). However with rapidly increasing housing prices in California nowadays, even people who work in highly paid industries feel great pressure living there. Will this eventually stop the growth of California’s economy considering companies there will have to pay much higher salaries than other parts of US to cover employee’s living expense and thus there would be less job openings from these companies? This question will be satisfyingly answered with the analysis of data of housing prices and job openings within a certain time period.

Job postings data are easily available online and could be regularly scraped from the web using ordinary methods. These postings could be used as an important source of data for predicting market trends, including real estate values (including predicting the values of unlisted properties). Because the temporal ordering of causal influence is unclear, it is important to assess the time specificity of the relationship between job growth and real estate values.

**1.3 Data Sets Utilized**

Four of the provided datasets were used for the analysis:

1. jobs
2. real\_estate
3. education
4. geographic

Because of hardware constraints, special measures were taken to reduce the size of the analysis dataset. We’ll specifically focus on the data from the most recent 5 years which can reflect the situation nowadays most accurately.

**2. Non-Technical Executive Summary**

Our main finding in the cross-sectional analysis was that for every $100,000 increase in home value over the period of 2011-2015 in California, the number of job postings increased by about 2 per thousand persons (Table 1). This effect was robust to adjustment for variables that could potentially confound the relationship (Table 2).

Additionally, the effect was very significant, suggesting that home values could be a consistent predictor of job posting activity. Interestingly, education was not strongly associated with job postings, suggesting that the expectation for migration of new hires might heavily influence how positions are filled (this could also reflect high competition for jobs.)

**3. Technical Executive Summary**

**3.1 Data Selection and Manipulation**

Special measures were taken to reduce the size of the dataset to make the analysis more feasible for a highly limited time period. The Job dataset was first filtered to only jobs in California, and variables of interest. This dataset was then merged with similarly subsetted data sets on home value and education.

The education dataset was summarized as one variable - years of education - by calculating a weighted average of all of the years of education variable. This was done by assigning discrete values to each category, e.g. 1st grade is one year while a Bachelor’s Degree is 14 years. This greatly simplifies our analysis of education factor while at the same time captures the most intrinsic property of education--length.

The home values dataset was transposed so that each row represented a time-place combination (replacing the original wide format of the dataset). This yielded an analysis dataset where each observation represented a unique time-place-job category combination. There were 128,607 total jobs posted for California across 174 cities. The study period of 2011-2015 was chosen based on the date range of data from the education dataset, so there were 60 time periods representing 1 month each.

Finally, the **geographic** data set was chosen for the creation of heat maps.

**3.2 Methodology**

Linear models were used as our starting point, as in most social-economic problem, to assess the cross-sectional association between home value and number of jobs. Specifically, a Blundell-Bond longitudinal model was used to assess the strength of lags between variables at different time points. The relationship was also assessed using graphical methods (Table 3).

**4. Appendix**

Table 1. Crude linear model of number of job postings per capita as a function of home value.

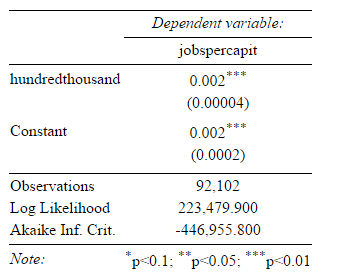


Table 2. Model of home value adjusting for education (in years) of the city.

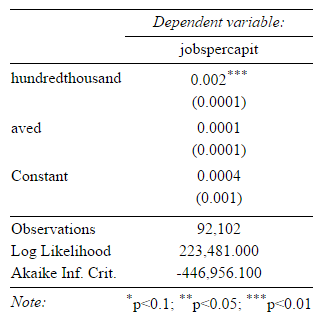


Table 3. results from Blundell-Bond longitudinal model (jobs as outcome)

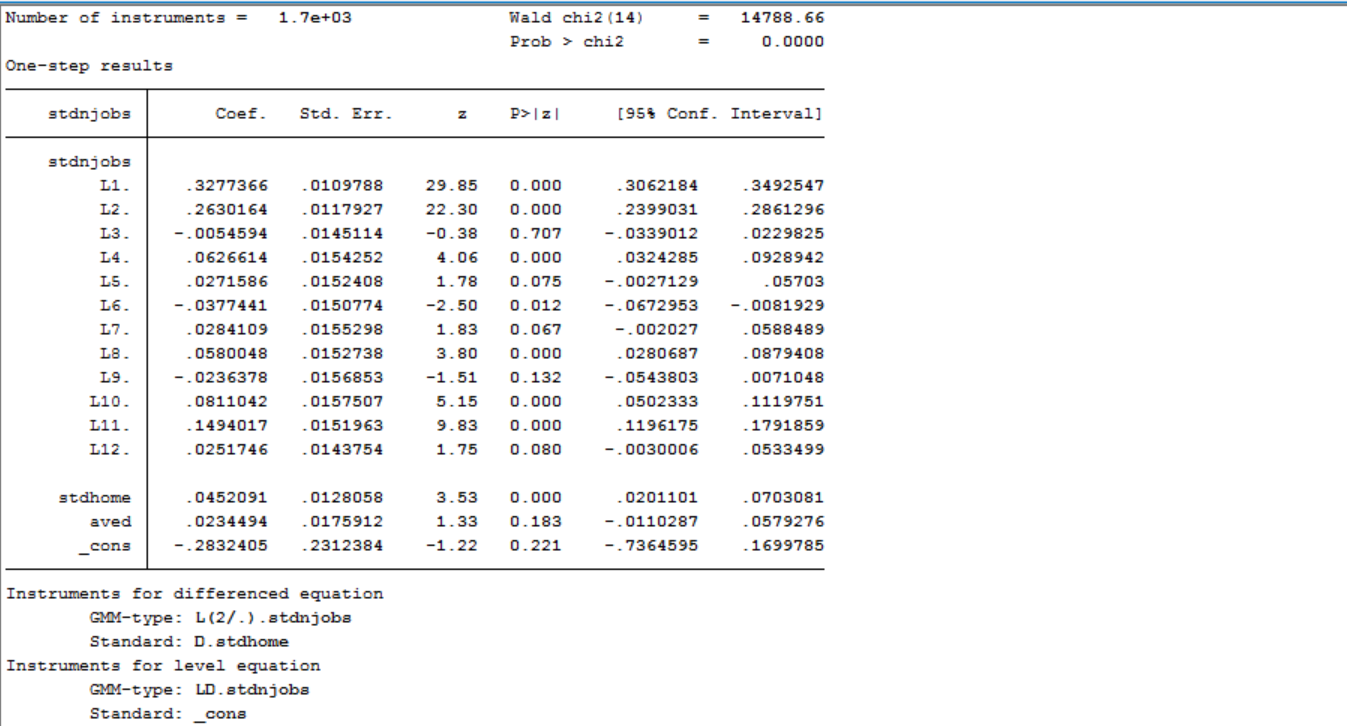


Table 4. results from Blundel-Bond longitudinal model (home value as outcome)

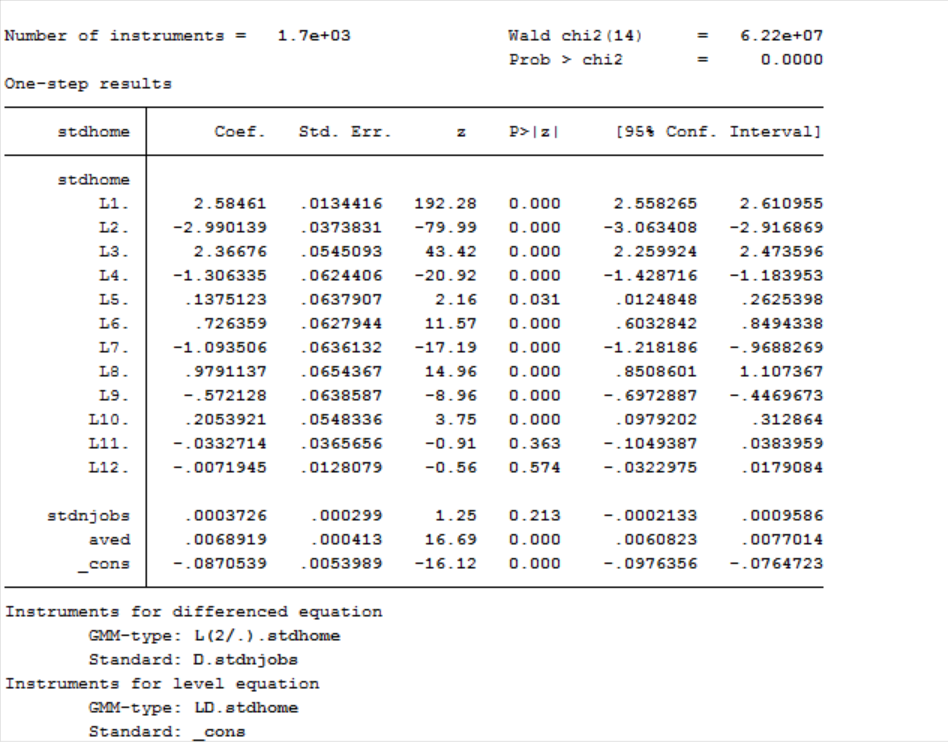


Table 5. Distribution of home values in California in 2011-2015

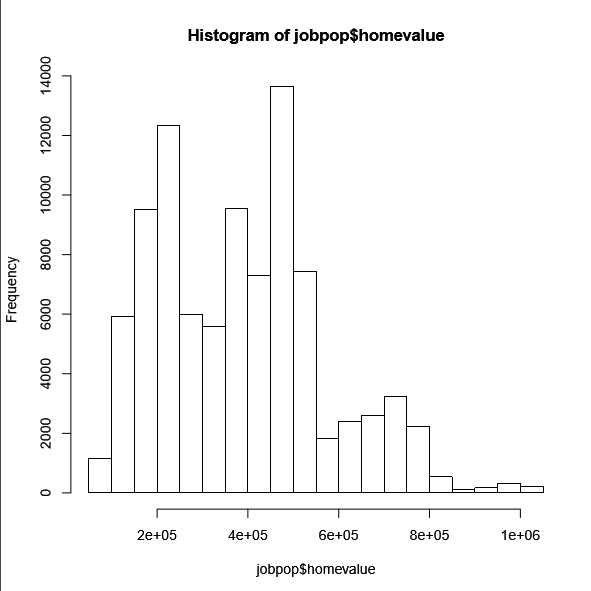


Table 6. Distribution of job per capita in California in 2011-2015

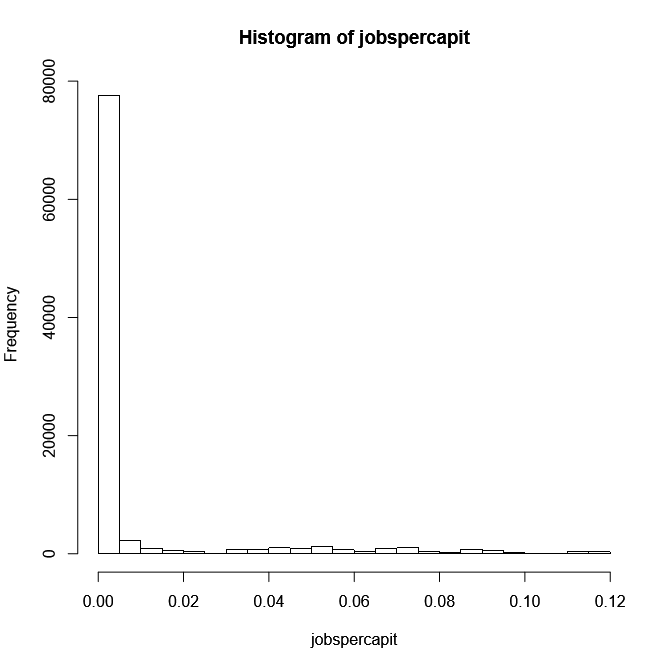


Table 7. Jobs postings in California in 2011-2015

|  |  |  |
| --- | --- | --- |
|  | Job Category | Frequency |
| 1 | Missing | 9376 |
| 2 | Account Management | 578 |
| 3 | Accounting & Finance | 2615 |
| 4 | Admin & Clerical | 591 |
| 5 | Administrative Assistant | 279 |
| 6 | Advertising | 83 |
| 7 | Aerospace, Aviation & Defense | 75 |
| 8 | Agriculture, Forestry & Fishing | 10 |
| 9 | Architecture | 10 |
| 10 | Arts, Media & Publishing | 437 |
| 11 | Auditing | 56 |
| 12 | Automotive | 465 |
| 13 | Bank Teller | 100 |
| 14 | Banking & Financial Services | 603 |
| 15 | Banquet, Catering & Events | 69 |
| 16 | Biological Sciences | 168 |
| 17 | Bookkeeping | 72 |
| 18 | Business Development | 1113 |
| 19 | Channel Sales | 73 |
| 20 | Child Care | 20 |
| 21 | CNAs, Aides, MAs, Home Health | 308 |
| 22 | Computers & Hardware | 47 |
| 23 | Concierge & Guest Services | 193 |
| 24 | Construction & Skilled Trade | 735 |
| 25 | Consultants & Freelance Opportunities | 268 |
| 26 | Credit | 18 |
| 27 | Customer Service | 1773 |
| 28 | Data Entry | 47 |
| 29 | Database Administrator | 4 |
| 30 | Direct Sales | 196 |
| 31 | Education & Training | 325 |
| 32 | Engineering & Architecture | 380 |
| 33 | Engineers | 1 |
| 34 | Executive Assistant | 59 |
| 35 | Executive Management | 35 |
| 36 | Finance Management | 52 |
| 37 | Financial Services | 145 |
| 38 | General Management & Business | 1897 |
| 39 | Government | 40 |
| 40 | Green | 358 |
| 41 | Health & Medical | 6244 |
| 42 | Healthcare Management & Finance | 300 |
| 43 | Healthcare Office & Finance | 24 |
| 44 | Healthcare Support Services | 4 |
| 45 | Healthcare Technologists & Technicians | 46 |
| 46 | Hospitality & Travel | 286 |
| 47 | Hotel Housekeeping & Maintenance | 8 |
| 48 | Hotel Management | 1 |
| 49 | HR Benefits & Compensation | 84 |
| 50 | HR Management | 100 |
| 51 | Human Resources | 443 |
| 52 | Inside Sales | 283 |
| 53 | Insurance | 116 |
| 54 | Intern / New Graduate | 513 |
| 55 | Internet | 32 |
| 56 | Inventory | 134 |
| 57 | IT Operations | 162 |
| 58 | Job Fairs | 126 |
| 59 | Lab Technician | 133 |
| 60 | Law Enforcement & Security | 1318 |
| 61 | Legal | 45 |
| 62 | Library | 9 |
| 63 | Life, Physical, and Social Science | 14 |
| 64 | Logistics | 246 |
| 65 | Maintenance & Repair | 1847 |
| 66 | Management Consulting | 71 |
| 67 | Manufacturing & Operations | 1059 |
| 68 | Marketing | 896 |
| 69 | Medical & Dental Practitioners | 13 |
| 70 | Medical Records & Health IT | 74 |
| 71 | Mortgage & Loan | 54 |
| 72 | Network Administrator | 216 |
| 73 | Nonprofit & Volunteer | 5 |
| 74 | Nursing | 5884 |
| 75 | Office Manager | 795 |
| 76 | Oil, Gas & Utilities | 216 |
| 77 | Operations | 442 |
| 78 | Other Healthcare | 257 |
| 79 | Pharmacy | 21 |
| 80 | Physician | 8 |
| 81 | Plant Management | 144 |
| 82 | Product Marketing | 202 |
| 83 | Public Relations | 25 |
| 84 | Publishing | 186 |
| 85 | Purchasing | 192 |
| 86 | Radiology & Imaging | 286 |
| 87 | Real Estate | 3 |
| 88 | Receptionist | 161 |
| 89 | Recruiting | 146 |
| 90 | Research | 33 |
| 91 | Restaurant & Food Service | 6455 |
| 92 | Retail | 14888 |
| 93 | Sales | 549 |
| 94 | Sales & Business Development | 6414 |
| 95 | Sales Engineers | 72 |
| 96 | Sales Rep | 1278 |
| 97 | Salon/Spa/Fitness | 243 |
| 98 | Science, Pharmaceutical & Biotech | 410 |
| 99 | Shipping/Receiving | 135 |
| 100 | Social Services & Mental Health | 218 |
| 101 | Software Architecture | 158 |
| 102 | Software Development | 1017 |
| 103 | Software, Gaming & Web Developers | 9 |
| 104 | Supply Chain & Logistics | 449 |
| 105 | System Admininistrator | 118 |
| 106 | Teaching | 51 |
| 107 | Tech Management | 1596 |
| 108 | Tech Quality Assurance | 2055 |
| 109 | Technical Design | 1267 |
| 110 | Technical Support | 400 |
| 111 | Technology | 1784 |
| 112 | Telecommunications | 28 |
| 113 | Therapy & Rehab | 559 |
| 114 | Trading | 23 |
| 115 | Training & Instructor | 680 |
| 116 | Transportation | 2748 |
| 117 | Travel & Tourism | 6 |
| 118 | TV, Film & Video | 6 |
| 119 | UI Design | 34 |
| 120 | Veterinary & Animal Health | 3 |
| 121 | Warehouse Management | 29 |
| 122 | Warehousing | 406 |
| 123 | Web Design | 73 |
| 124 | Web Development | 261 |
| 125 | Writing & Editing | 99 |
|  |  | 92102 |

Figure 1. scatterplot of home value (x-axis) versus number of jobs (y-axis)

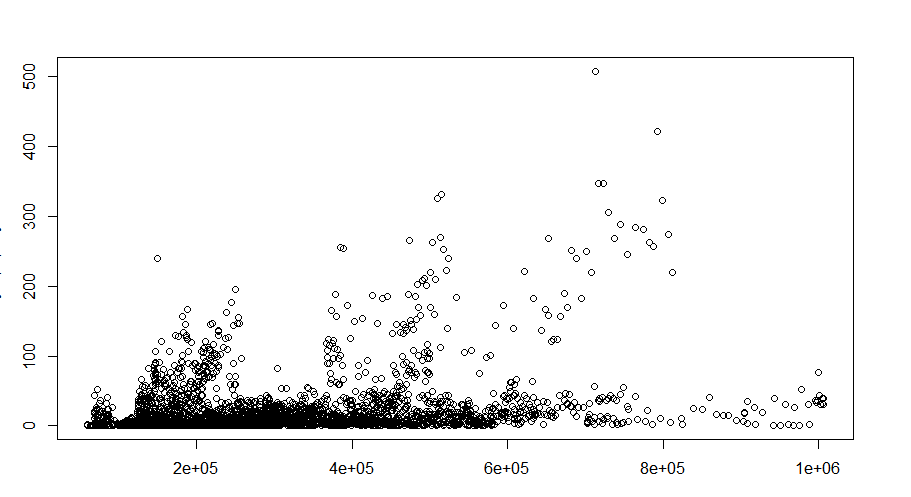
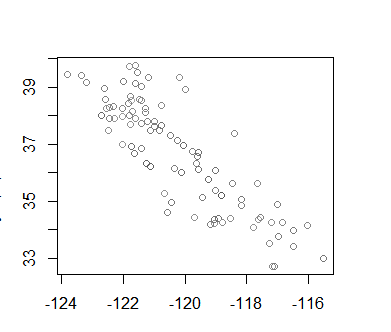


Figure 2. Spatial distribution of jobs in California.



sum

import excel "C:\Users\anpas\Documents\Datathon\jobpop.xlsx", sheet("jobpop") firstrow

sum

hist stdhome

hist stdnjobs

xtset city dates

clear

import excel "C:\Users\anpas\Documents\Datathon\jobpop.xlsx", sheet("jobpop") firstrow

xtset city1 dates

gen city12=city1+100

xtset city12 dates

clear

import excel "C:\Users\anpas\Documents\Datathon\jobpop.xlsx", sheet("jobpop") firstrow

xtset city1 dates

xtdpdsys stdnjobs stdhome aved lags(12)

xtdpdsys stdnjobs stdhome aved, lags(12)

xtdpdsys stdnjobs stdhome aved, lags(58)

xtdpdsys stdhome stdnjobs aved, lags(12)