

Land Use Efficiency: How Unit Count and Stories Impact Lot Value*

Ziheng (Tony) Fang

Abstract

For properties on the real estate market, one may expect that the larger the lot size, the larger the market value of the property. In the real world, is that true? Would the relationship between lot size and market value instead be impacted by the total number of units the property contains and the number of stories the property hosts? Some in the real estate industry would wonder if having more residential units or a higher building would increase or decrease the value of the land the property is built on. Data regarding total assessed value of different residential properties was studied from real estate data provided by the Assessor's Office in San Francisco. Methods utilized included building a regression model, studying the interaction effects, and running a hypothesis test for the interaction effects.

Introduction

In many urban areas across the country, the population is extremely dense, making land among one of the most scarcest and valuable resources for real estate developers. In the eyes of a developer, how efficiently land is utilized, whether it's having multiple units in one property or multiple stories in a unit, significantly impacts the city's housing supply. In cities like San Francisco, where housing is limited and prices are up, land use is a significant problem that city leaders want to address. On top of supply, understanding how lot area would interact with the number of units inside a property or the number of stories a unit has can provide additional insights on how efficient land use can actually have an impact on property value in the eyes of a potential homeowner. From this our hypothesis test in this study, we can also see that contrary to the belief that increases in lot area would increase the assessed value, the two variables are actually not related.

*Project GitHub Repository: <https://github.com/TFang96/MATH261A-Project2>

Prior research often examined how the prices of property are predicted by direct predictors such as the lot area, room count, bedroom count, etc. However, these studies very rarely address how variables regarding efficient land use, such as having multiple units in a building and multiple stories in a unit could impact how our predictors affect property value. In reality, many of these variables are interdependent whereas while we could utilize land more efficiently by having a residential unit with multiple floors and having multiple residential units in one building, in the eyes of a property owner, the value per unit of lot size may also change significantly.

This paper uses data from the San Francisco Assessor-Recorder office to see how property values are impacted not only by the lot area, but by how the number of units inside a property and the number of stories per unit impacts how lot area affects property values. A multiple linear regression model was utilized to examine both the direct effect of lot area on the assessed value and how the unit count and story count affects the impact lot area has on the assessed value through the interaction effect.

By identifying how vertical expansion and having multiple units inside a building affects how a unit change in lot area affects property value, this paper provides additional insight for developers when considering land use. Insights from this analysis can help developers and policymakers better understand trade-offs when it comes to land use efficiency and real estate value.

The Data section gives further details on data that was utilized for this paper. Details on the statistical methods utilized are listed in the Methods section. The Results section includes our statistical findings. The conclusion we drew can be found in the Discussion section.

Data

The data set used for this paper, Assessor Historical Secured Property Tax Rolls (*Assessor Historical Secured Property Tax Rolls 2025*), provides details such as the assessed values, lot area, number of units, number of stories, and more for properties within the city of San Francisco since 2016. The data was compiled by the city assessor's office. We have a total of 3.72M observations in the original dataset where each observation represents a real estate property. Variables that used for the response variable in our study included `Assessed.Fixtures.Value`, `Assessed.Improvement.Value`, `Assessed.Land.Value`, and `Assessed.Personal.Property.Value`, which were summed up into the `AssessedValue`. Variables used as our predictor variables included `Lot.Area`, `Number.of.Units`, and `Number.of.Stores`. To allow numerical computations, all commas were removed. Since clients look for very different features when on the market for residential and commercial properties, we only kept data on single family residential and multiple family residential units.

Table 1: Summary statistics for assessed property values, unit count within each building and story count within a building.

	Value
Mean_Assessed_Value	7.628087e+05
Assessed_Value_SD	3.012549e+06
MinAssessedValue	0.000000e+00
MaxAssessedValue	5.387552e+08
Mean_LotArea	3.708348e+03
LotArea_SD	1.611940e+05
Min_LotArea	1.200000e-01
Max_LotArea	5.800145e+07
Mean_NumberOfUnits	2.145812e+00
NumberOfUnits_SD	8.095218e+00
Min_NumberOfUnits	0.000000e+00
Max_NumberOfUnits	9.820000e+02
Mean_NumberOfStories	1.633178e+00
NumberOfStories_SD	5.258824e+00
Min_NumberOfStories	0.000000e+00
Max_NumberOfStores	9.940000e+02

Methods

We started off by fitting a linear regression model based on the dataset using the `lm()` function in R (R Core Team 2024). The original model is as follows:

$$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} X_{i1} + \beta_3 X_{i1} X_{i2} + \epsilon_i$$

Results

Discussion

References

- Assessor Historical Secured Property Tax Rolls*. 2025. Assessor’s Office, City of San Francisco. <https://data.sfgov.org/Housing-and-Buildings/Assessor-Historical-Secured-Property-Tax-Rolls/wv5m-vpq2>.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.