

Commercial Real Estate Valuation Using Hedonic Pricing Models

Overview:

Buildings can be compared to a bundle of goods sold in a market, where each of the building characteristics combined equate to the expected overall transaction value. By collecting data on many different buildings a regression analysis can be used to determine the correlation (relationship) of each characteristic to the transaction price —e.g. physical characteristics and other external influencing elements that may add or subtract from the building value. Each of these correlations can be measured to determine a degree of confidence (i.e. significance) and then subsequently be used to build a hedonic pricing model. Hedonic pricing models can be useful to determine the intrinsic value of each attribute, as well as to predict transaction prices. This can be particularly useful when traditional discounted cash flow models fall short because of the absence of a market, when no comparable buildings exist, and for non-income generating buildings.

Introduction & Purpose

Current financial sector woes have left many real estate owners and investment fund managers wondering how to properly value real estate assets. Recent mark-to-market valuations based on comparable properties and area cap rates are driving down the market value of many assets even though the underlying fundamentals have not changed.

Many now wonder what a proper approach is to value potential future acquisitions and developments; in addition, some claim that traditional discounted cash flow (DCF) valuation techniques fail because of the broad-based assumptions used (e.g. discount rates, going-out cap rates, the effect of market cycles on disposition values, etcetera).

At the height of the latest real estate cycle, proper underwriting fundamentals seem to have been lost because of what some call irrational exuberance fueled by inexpensive capital chasing deals. As a result, the market value of real estate far outweighed the true, or intrinsic value. Understanding the intrinsic value of a real estate asset and the characteristics that contribute to its potential transaction price (market value) is imperative for proper valuation and can only be calculated by fastidious underwriting.

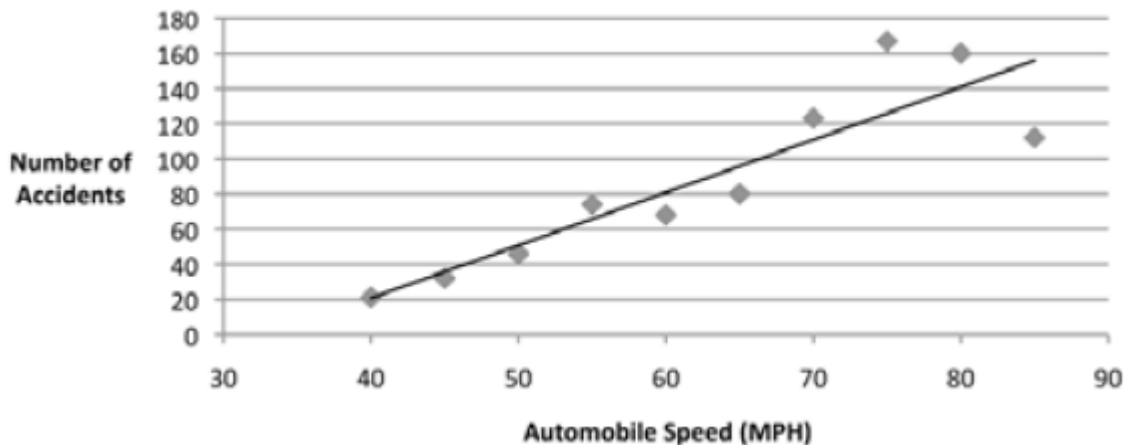
The intent of this paper is to introduce regression analysis and an alternative valuation technique called hedonic price modeling. The theory behind these tools will be discussed, as well as how they are constructed and their potential uses, benefits, and weaknesses. Regression analysis can be used to determine the correlation between building characteristics and the transaction price and hedonic modeling can be used to predict future transaction prices. Three case studies are presented in which this valuation methodology has been implemented.

Statistical Models

Statistics is a mathematical science pertaining to the collection, analysis, interpretation or explanation and presentation of data. Statistics is used in many disciplines to solve problems where numerous variables are present. Industries in science, such as pharmacology, rely heavily on statistics to determine the safety and effectiveness of experimental drugs and the probability of potential outcomes. Because of fixed locations, variation in geographical regions, and the unique physical characteristics associated with real estate, statistics can be an invaluable tool in determining building's intrinsic value. Statistical tools have been developed to assist in determining such values, and include regression analysis and hedonic modeling.

Regression Analysis

Regression analysis is a statistical technique used to determine correlation between different data points. The results of the regression can then be used to predict future results. It is a tool used to determine how different independent variables influence a single dependent variable. These influencing variables are called explanatory or independent variables. The most basic regression, a least squared method, measures and plots the correlation between a single dependent (Y) variable and a single independent (X) variable. A theoretical example is to measure the correlation of vehicle speed to highway accidents. We could establish traveling speed as our independent variable (X) and the number of accidents as our dependent (Y) variable. Using historical documentation of car accidents on highways and the speed traveled when the accident occurred, one could plot a simple graph – Figure 1.



One possible explanation of these results could be as automobile speed on highways increase the number of accidents increase, as represented by the trend line in Figure 1. This, however, explains only one of a myriad of other influencing factors—time of day, inclement weather, road quality—that could be attributed to the true cause of any single accident. Multiple regression analysis is a technique used to determine the contributing effect of several independent variables on the dependent variable. The mathematical result of a regression produces correlation coefficients for each of the variables.

This methodology can be used in the context of real estate to help explain and account for value. Malpezzi, Ozanne, and Thibodeau (1980) compare housing to a bag of groceries. Each bag may be large or small and contain a variety of different items (oranges and apples), each of which contribute to the overall price. By collecting information about many different bags and the items contained therein, regression analysis can be used to determine the contributing affect (or correlation) each of those items has on the overall price. This technique also helps to determine which item(s) most significantly impact the price.

According to their theory, real estate is no different. A building may be large or small and have a number of items contributing to the overall price; and, regression analysis provides a method to measure how much each of those items affect the transaction price. One of the strengths of regression analysis is that it allows one to measure items qualitative in nature in quantitative terms.

An example of this is an office building located near a train station. While the train station is not physically part of the building, it is a trait that may positively influence a higher transaction price due to the added convenience for employees who work within the building. Other similar intangible characteristics include proximity to a central business district and an airport. Variables such as these are often referred to as internal characteristics. External characteristics include items such as square footage, building finishes, parking availability, number of restrooms, and etcetera. Theoretically, there are indefinite numbers of internal and external characteristics within a building that add value. What an investor/developer may be most interested in is which of these characteristics will most highly affect the transaction price. Regression analysis is a tool which helps to measure such variables if sufficient current and historical data is available.

Two cases will be provided later in the paper where multiple regression is used to explain what tangible and intangible characteristics—our independent variables—contribute significantly to a property's likely transaction price—our dependent variable. A third case will then be presented to test the predictability of a hedonic model.

Hedonic Pricing Models

The underlying goal when generating a hedonic pricing model is to create an accurate predictive model. Unlike groceries, the value of individual features within a building cannot be directly observed (e.g. the value of a building's atrium relative to the overall value of the building). Hedonic pricing models, however, can be used to measure the influencing affect of these characteristics on the overall transaction price. These models are developed by using the coefficients generated from a regression analysis. This relationship can be described as "market price is a function of each tangible & intangible building characteristic and other outside influencing factors." This is illustrated in the following equation:

Market Price = f(tangible & building characteristics, other influencing factors)

Other influencing factors include assets sold in a non-fair market and will not be discussed herein.

Hedonic models can also be instrumental in answering the following questions:

- How should real estate be valued in the absence of a liquid active market?
- How do intangible characteristics—such as proximity to a train station—affect price?
- How should one account for an asset's option value for potential future redevelopment scenarios?

Answers to these questions can be determined by collecting data on buildings that have known transaction prices and are similar in nature to the subject building being analyzed. A regression analysis can then be calculated to determine the correlation for each of the characteristics measured against the transaction price. Those correlation measurements are then used to create a hedonic pricing model which will help determine the expected price of the subject property. Additional guidance is given below in the Structuring an Analysis section.

As an alternate real estate valuation method, hedonic modeling can be used by developers, corporate real estate groups, owners, and operators to determine which building characteristics add significant value to the potential transaction price. The results produced can provide important information for future decisions and help each party better understand the economics surrounding each asset, thus improving asset underwriting.

Structuring an Analysis

One must determine what characteristics are important in analyzing. To a retail developer preparing to build a new shopping center, characteristics to analyze may include geographic location, size of the buildings, and typical parking ratios. Other characteristics which can be analyzed are listed in the following section, Elements to Consider, with a brief definition as to their usefulness.

One must then determine which characteristics can be measured or quantified. Of those that can be quantified through the collection of data, one must also ensure a large enough sample size can be obtained. There are several companies which collect and provide data on individual buildings such as CoStar, SNL, The Warren Group, LoopNet, and Property Shark. Most of these databases include tools to search for properties by property type, size, specific geographic regions, last sales date, age of building, rental rates, and etcetera. This historical data can then be used to calculate a multiple regression analysis.

Elements to Consider

The following is a list of typical building/asset characteristics which may be considered when developing a regression analysis. Notes have been included to support a regression analysis.

- **CPP Index:** Useful in considering the effect of the economic conditions on commercial real estate transaction prices if the data used spans several economic cycles. The index number used is that of when the last sales transaction of the building took place.
- **Total Square Feet:** total number of square feet considered in the transaction. This variable should be squared when running the regression to account for the non-linear relationship between size of the building and price. I.e. increasing the size of the building won't continue to increase the price indefinitely, but rather will have a diminishing effect at some point on price. Squaring the data will help correct for this effect.
- **Green (Energy Star or LEED certified):** established as a 'yes/no' variable represented by a 1 for 'yes' or 0 for 'no'.
- **Fitness Center:** established as a 'yes/no' variable represented by a 1 for 'yes' or 0 for 'no'.
- **Food Service:** usually established as a 'yes/no' variable represented by a 1 for 'yes' or 0 for 'no'. However, a more detailed analysis on the type of food service can be done. E.g. restaurant v. vending v. catering.
- **Atrium:** established as a 'yes/no' variable represented by a 1 for 'yes' or 0 for 'no'. This is often an important attribute in office buildings and hotels.
- **Land (Acres):** Can be established as the number of acres included in the transaction price. It is recommended to separate land cost from the building cost in the transaction. This is done to ensure an apples-to-apples comparison.
- **Typical Floor Square Feet:** it is important to distinguish how this is measured to ensure consistency. E.g. buildable v. rentable square footage.
- **Parking Ratio:** this can be measured as the number of parking spaces per 1000 square feet of building space.
- **Stories:** this factor may also be indirectly attributed by establishing a characteristic for 'suburban v. urban', as buildings with more stories tend to be in more urban settings.
- **Year Renovated:** typically a better indicator of value than 'age of building' because many older buildings recently renovated transact at prices similar to new buildings.

- Building Park: established as a 'yes/no' variable represented by a 1 for 'yes' or 0 for 'no'. Used to indicate whether or not the building is located on a campus. Near to train: can be established as a 'yes/no' variable represented by a 1 for 'yes' or 0 for 'no'. This, however, is a subjective measure to be determined by the model creator. In urban environments, it is reasonable to suggest that a building less than 1 mile from a train stop would be considered 'near'.
- Urban: used to distinguish whether the building is in a suburban or urban environment
- Year Built: entered in directly as the year built.
- Class: can be used to distinguish A, B, C or any other building measure that distinguish a building. The '1 & 0' method can be used up to 3 variables. E.g. if a building is an 'A' it would receive a 1 and 'B' factors would be left as 0's, similar in nature to a 'B' class building. If the building is a 'C' class building, both 'A' and 'B' would receive 0's.
- Construction Type: similar in nature to 'Class', up to 3 different types of construction can be used by implementing the 1 & 0 methodology (a.k.a. dummy variables).
- CPI: Consumer Price Index-an important factor to consider in using data across different geographic regions. E.g. collecting data on industrial buildings across the United States. Economic conditions for different transactions can be captured by recording the current consumer price index tied to the region in which the real estate is located.
- Tenancy: an example of this could be to consider if the building is occupied by a single or multiple tenants.

Data Interpretation

Once the regression analysis is run and a hedonic model is created, a proper interpretation of the pricing model results should be stated as follows:

“Based on the historical data used in the regression, it can be expected that the transaction price of the tested real estate in this model would transact for “x” dollars.”

To interpret the individual value of each characteristic (i.e. independent variable(s)), proper interpretation would be stated as follows:

“Based on the historical data used in the regression, with all other characteristics holding constant, the value of the characteristic on the overall transaction price is “y” dollars.”

It should be noted that other unaccounted factors not included in the regression may also play an important role in determining the transaction price, such as other influencing factors noted

above. Many statisticians agree that econometrics is a field comprised of investigating the whole picture; thus, the results in the regression cannot always be explained without looking at other factors. For example, in the sample regression shown below in Figure 2, fitness center has a correlation coefficient equal to negative eighty-nine thousand. Can this be interpreted to mean that including a fitness center into a building will lower the overall transaction price by this amount? Possibly this is true. However, another explanation could be that those buildings in the data set transacted at a point in time when fitness centers were not highly sought after by building owners and viewed as lost leasable space. Or, older dilapidated buildings had fitness centers built into them at a later point in time in an effort to increase their attractiveness and value. So, additional investigation may be needed for a more clear explanation of the results.

An example of a regression analysis and hedonic model is shown below.

Note that the characteristics analyzed and their associated correlation coefficients are highlighted within a box.

The algorithm created to predict the price of \$658,938 in Figure 4 is derived from the correlation coefficients produced from the regression in Figure 2. Each of the correlation coefficients can be tested to determine if they are statistically significant in predicting the final price, i.e. if the coefficients influence on price didn't happen by chance. This can be done by observing the t-Stat or P-value produced for each of the variables and measuring them against a predetermined threshold. In the case above, the threshold of significance is 5%.

Case Examples

Three cases have been provided below. Case #1 and #2 are used to show the usefulness of regression analysis in explaining the correlation between building characteristics and transaction price. Case #3 is provided to determine how hedonic modeling is used to predict future transaction prices.

Case #1: Residential Development in South Boston

Developers of Olmsted Green, a townhouse and condominium development located near Jamaica Plain in Boston, Massachusetts, have recently completed phase 1 of 5. With a wide swing in economic conditions since conceptualization of the project, owners of the project have been faced with several tough decisions regarding the quality of finishes to include in each condominium (e.g. countertops, appliances, light fixtures) as well as the community amenities (e.g. fitness center, swimming pool, club house). Several different strategies could be implemented for the future phases of the development to correct for current economic conditions, including eliminating the high-grade finishes and lowering the price point to attract a wider customer base. They could also add additional amenities (perks) not currently in the project to maintain sales volume. In an effort to determine what decisions to pursue, a regression analysis and hedonic model was created based on similar property types in the

The independent variables chosen for the condominiums at Olmsted Green were characteristics found in townhomes and condominiums in the area. These variables include the following:

- Square footage
- Number of bathrooms and bedrooms
- If the units have an attached garage
- If there is an affordable component in the complex
- If complex is within walking distance to a train
- Home owner association fees
- If the complex has a fitness center
- If the complex has a swimming pool
- If the units have a fireplace
- If there is an outdoor courtyard
- If the units are ADA accessible
- If the units have private outdoor space (i.e. porch, deck, etc)
- If the building is LEED certified
- If the complex is a mixed-use development (i.e. retail component to the building or complex)
- If units have security systems
- The number of units in the complex
- Appliance quality
- If laundry hookups are in the unit
- Countertop quality
- Year building was constructed
- If basements are included
- If there is extra storage in or around the building

Regression & Hedonic Results

Of the independent variables measured and tested in the regression analysis, the following proved statistically significant: attached garage, swimming pool, private outdoor space, security systems, and extra storage space. In other words, the probability that these characteristics influenced price purely by chance is less than 5%. It is reasonable to suggest that based on the historical data used to populate this database, these significant variables contributed to transaction price. The hedonic model predicted a price very similar to the current asking price for the condominiums. For example, a 1,850 square foot, three bedroom, and two bathroom unit had a predicted price of \$552,000 and the current asking price of the unit was \$549,000. The developer was certainly in the ball park on price for the type of unit produced but incidentally this price point was higher than what the area demand was willing to pay. By analyzing the hedonic model results, the developer could determine which building attributes (variables) should or should not be used on future phases. As in the case

above, he/she may want to consider whether the impact of the significant variables were positive or negative and then choose which of these to build into the project.

Case #2: Office Building in Peoria, Illinois

Google, Microsoft, HP, Fidelity, SAS, and Wal-Mart are similar when it comes to their corporate-owned real estate. Each owns and occupies campus-style office complexes in various locations throughout the U.S., often in areas that are a considerable distance from a major metropolitan area. These firms develop new buildings to fulfill employee and business operation needs. At times these organizations have been criticized by investors because of the high expense associated with developing these unique buildings rather than leasing out space. In addition, resale transactions of these buildings are far and few between due to their unique locations and customized designs which aren't always desirable for another owner—e.g. Wal-Mart's headquarters located in Bentonville, Arkansas. Luxury amenities are often included in these campuses, and consist of fitness centers, cafeterias with gourmet menus, outdoor pavilions, volleyball and basketball courts, ergonomic furniture, and child day care. Justification for such amenities may include that these help boost employee morale and productivity. While this may be true, such statements are often difficult to quantify and measure relative to the bottom line. Part of the difficulty associated with determining a 'true' value for these buildings is because they are held indefinitely within the firm and they lack a resale market. Because of these factors, typical discounted valuation methods render useless. As an example, a traditional discounted cash flow model, or income model, is difficult to populate because these buildings generate no income, holding periods are not usually defined, and these buildings are not developed/owned for investment purposes.

Due to the sensitive competitive nature of the building campus test case the facts and figures have been changed. The building to be analyzed is a class A, 350,000 square foot office building located in Peoria, Illinois, about 165 miles outside of Chicago and was constructed in 2005 for \$185 million. The location was chosen by the company because of the City of Peoria's land grant, which was intended to incentivize business growth. In addition, the firm is aiming to create an environment for its employees where they can live in close proximity to work for a reasonable cost. The building includes state-of-the-art electronics and telecommunications, raised floors, high-tech conference space, large auditoriums, a cafeteria with several restaurant choices, a large atrium space, outdoor activity space, walking paths, and fitness centers with 24/7 access. Because of the location and the unique nature of the building, it is reasonable to suppose that a comparable property doesn't exist. In addition, because this building is used exclusively by the owner, it generates no income thus making the income approach moot.

Regression & Hedonic Results

Over 280 office properties were collected and analyzed in a 120 mile radius surrounding Peoria. The variables analyzed included the following:

- CPP index

- Whether the building was ‘green’
- Existence of food services
- Typical floor size
- Parking ratio
- Number of stories
- Year renovated and built
- Proximity to a train
- In an urban or suburban location
- Class of the building

The regression showed that the following characteristics were statistically significant: Total building square footage, CPP index, ‘green’, year renovated, and class of the building.

The hedonic model produced an expected market value of \$150 million, a predicted price one could consider reasonable based on the bundle of goods the building contains. This is much lower than the construction price tag, indicating that the firm paid a premium for the building relative to the price the market would have paid in that geographic location for the bundle of goods. Certainly this doesn’t provide additional justification for the cost associated with the construction in 2005. This does, however, give additional insight to the firm as to which variables contributed greatest marginally to the overall price and which building characteristics could be scaled down in future building developments in order to keep price in check.

Case #3: Multi-family Condominium Units in Reston, Virginia

A database of multi-family condominium units in Reston, Virginia was collected in an effort to determine what a developer could reasonably expect to sell each unit for. Units that were sold in which we know the actual transaction price and unit/building attributes will be inserted into the hedonic model to see how accurately the model comes to predicting correctly.

Regression & Hedonic Results

Those units recorded were only transactions in the last two years and within a 1 mile radius of the target development site. The total number of units analyzed was 154 and the variables analyzed included the following:

- Square footage and square footage squared
- Bedrooms
- Bathrooms
- Year building was built
- Number of stories of the building

Three properties were then inserted into the hedonic model created from the regression results. The output, listed as Predicted Price within each figure below, is then compared to the actual transaction price.

For the condominium unit shown below in Figure 5, the actual transaction price as recorded in September 2008 was \$1,196,513, a \$169,416 difference.

	Intercept	Square Footage	SF^2	Bedrooms	Bathrooms	Year Built	Stories
Building Inputs	1	2,800	7,840,000	3	4	2005	15
Correlation Coefficients	-9175186.322	130.8187048	0.016683993	-4476.116207	69473.77416	4604.996192	16162.76843
Predicted Price	\$ 1,027,097						

For the condominium unit shown below in Figure 6, the actual transaction price as recorded in September 2008 was \$349,921, a \$37,757 difference.

	Intercept	Square Footage	SF^2	Bedrooms	Bathrooms	Year Built	Stories
Building Inputs	1	859	737,881	1	1	2005	4
Correlation Coefficients	-9175186.322	130.8187048	0.016683993	-4476.116207	69473.77416	4604.996192	16162.76843
Predicted Price	\$ 312,164						

And lastly, for the condominium unit shown below in Figure 7, the actual transaction price as recorded in May 2008 was \$749,900, a \$14,799 difference.

	Intercept	Square Footage	SF^2	Bedrooms	Bathrooms	Year Built	Stories
Building Inputs	1	1,510	2,280,100	2	2	2006	19
Correlation Coefficients	-9175186.322	130.8187048	0.016683993	-4476.116207	69473.77416	4604.996192	16162.76843
Predicted Price	\$ 735,101						

The results suggest a fairly accurate prediction of pricing for each of the three test units relative to their actual transaction price, with an average difference of 10%.

Conclusion

Regression analysis and hedonic modeling are valuable tools for real estate professionals in determining the correlation between building characteristics and the transaction price, as well as to predict future transaction pricing. Using these techniques can facilitate building valuation when traditional discounted cash flow models cannot be populated, which can often be the case if an asset is located in a remote location and/or its structure is unique to other building assets. The results from such a study can also produce answers to development decisions such as what building attributes to include in an effort to generate the highest value on a parcel of land.