INTERNATIONAL REAL ESTATE REVIEW

2021 Vol. 24 No. 3: pp. 385 – 403

An Empirical Method for Decomposing the Contributions of Land and Building Values to Housing Value

Kuan-Lun Pan
Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan. jackklpan@gmail.com

Hsiao Jung Teng
Anfu Co., Ltd, Taoyuan, Taiwan. karen.teng@anfusolutions.com

Shih-Yuan Lin
Department of Land Economics, National Chengchi University, Taipei, Taiwan. syl@mail2.nccu.tw

Yu En Cheng
Anfu Co., Ltd, Taoyuan, Taiwan. yu.cheng@anfusolutions.com

This paper develops an empirical method that uses two separate housing related components to estimate housing value: land and building. The artificial neural network (ANN) technique is used to iteratively solve for two hedonic models simultaneously by minimizing the difference in the observed total value and the sum of the estimated land and building values. This method enables one to objectively separate housing value into land and building components. Using actual sales transaction data from Taipei City, we estimate the land value as a share of the total housing value. The results show that the land value accounts for a higher share with older properties. The share of the land value of low-rise buildings tends to be higher than that of high-rise buildings. The share of the land value can deviate by 20 percentage points between more or less expensive housing communities within Taipei City.

Keywords

Land Value, Building Value, Housing Value, Apportionment Theory, Artificial Neural Network

1. Introduction

It is known that the factors that influence housing price can be categorized into land or improvement characteristics. Nevertheless, it is also known that separating real estate value into land and building components is still an academic challenge. After their construction, buildings are in the best condition and therefore considered to contribute to a higher price. With time, the price of the building decreases along with the deterioration of the building. Meanwhile, the opposite is true for the price of land, which tends to increase due to the potential for redevelopment . As we can only observe one value of an entire property after its development, the proportion of contribution of building and land to real estate value remains an unresolved puzzle.

In order to separate housing price into land and building components, the appraisal industry commonly applies one of the two approaches: residual apportionment or proportional apportionment. Under the former approach, it is assumed that if land or building value is calculated first, the remainder of the housing value belongs to the other component. For example, after deducting the replacement cost of a building from the property price, land accounts for the remainder of the price. Hendriks (2005) suggests that this is the most commonly used method to apportion value into a building and a land portion. On the other hand, the proportional apportionment approach argues that when land and building are combined as a new joint good, then the price is attributed to both the land and building. There is a relationship between the share of the value of the land and the building. This implied ratio is used to determine the value of both components. Guerin (2000) argues that building replacement or reconstruction cost does not necessarily equal to the market price of the existing building, hence, using the residual apportionment approach to assess the land price will cause biased results. Despite the long history of the use of these two methods and debates about them, there is no empirical research that can separate the two components of value in a robust manner. The challenge in this exercise is that after development, only one value of the entire property, including land and improvement, can be observed. The objective, however, is to estimate the two components of value from this one single observed transaction price. This limitation due to the lack of related empirical research has inhibited the feasibility of separating land value from building value.

Nevertheless, this value apportionment is needed to support many business functions worldwide. The most common situation found are property tax-related issues. Governments tend to separate the assessment value of a property into land and building components for property tax purposes, so improvements are depreciable while land usually gains value (Kutty, 1999). For countries like the United States, the same property tax rate is applied to both land and building. The variation of the apportionment between the value of the two components does not have tangible impact on the total tax amount. However, the impact could be greater if there are different tax rates assessed for the land and building, such as in the cases of Taiwan, Thailand, Turkey, Malaysia, Belarus, etc. Under such a condition, homeowners can face very different tax amounts if the value is biased toward the land or building. Hendriks (2005) argues that, when determining the risk structure of an investment, the owner and the lender need to know the share of their capital that is invested in the land or building. With the different risks in future value increase potential, different risk-adjusted costs of capital need to be used for the share of capital invested in the land vs. the building. Beside, Ilić and Mizdrakovic (2016) point out that accountants need to report the value of the land and building separately for financial reporting purposes. Zhai et al. (2003) also mention that separate values are required to manage and estimate insurance and mortgage. In practice, the total value of a property must be allocated appropriately between land and improvements (Weinberger et al., 2009). However, in real world practices, the separation is generally derived from the opinion of the appraiser, construction cost manuals, or some other intuitive guidelines. There is no strong empirical evidence to support the existing apportionment methods.

In response to this challenge, an empirical method that separates the contributions of land and building to property value is needed which has motivated this research work. Multiple regression, one of the common regression methods, is used to separate the components of property price (Guerin, 2000; Özdilek, 2016; Sunderman and Birch, 2001). When performing a multiple regression for this purpose, there are factors, such as location and time, that affect both land and building prices. However, such factors affect prices with different weights, while it is difficult to distinguish the weights in an ordinary regression model (Gloudemans, 2002). Moreover, the constant derived from the hedonic model is defined as the common value of the combined land and building. It is also difficult to distinguish the share of land and that of building from one ordinary least square (OLS) model (Glodemans, 2002; Özdilek, 2016).

With rapid advancements in numerical computation capability, data mining and implementing machine learning algorithms have been significantly improved over the past two decades. In particular, artificial neural networks (ANNs) have been largely applied in various fields. Abidoye and Chan (2017) report that the first study to use an ANN on real estate is Borst (1991). Since then, a large volume of studies in the literature have used ANNs to analyze real estate prices (Do and Grudnitski, 1992; Kauko, 2003; González et al., 2005; McCluskey et al., 2012; Nguyen and Cripps, 2001; Pagourtzi et al., 2003; Peterson and Flanagan, 2009; Renigier-Bilozor and Wisniewski, 2012; Yacim and Boshoff, 2018). Moreover, it has been found that the performance of mass appraisal models of real estate developed on ANNs are superior to those of conventional multiple regression based hedonic models. Peterson and Flanagan (2009) report that ANNs can effectively conduct nonlinear analyses and are more suitable for multiple regressions that use numerous dummy variables.

The success of ANN applications in establishing mass appraisal models, and the capability of ANNs to simulate and interpret complex functions (Ribeiro et al., 2016; Shrikumar et al., 2017) have motivated us to propose the use of an ANN to further evaluate the feasibility of separating the value of land and building. Actual sales transaction data from Taipei City, Taiwan, are used to empirically estimate the land vs. building value of each transaction. The estimated results are then converted to a ratio of the land value as a proportion of the total housing value.

This study consists of five parts. After the introduction, the models and methods are then discussed. The transaction record data used and the empirical results are elaborated in the third and fourth parts respectively. The final part provides the conclusions and recommendations for future work.

2. Methods

Given that there is only one observed sale price of a property, it is mathematically not possible to uniquely identify the two separate values of land and building that contribute to the housing value.

$$TV_i = LV_i + BV_i + e_i$$

where TV denotes total value, LV is land value, and BV is building value. The objective is to minimize the sum of the squared residual by estimating the LV and BV functions at the same time. Once the optimal solution of the LV and BV functions is estimated, the proportion of the value of the land and different building types can be directly calculated.

In a traditional econometric based hedonic model, the equations above can be written as:

$$\begin{aligned} TV_i &= LV_i + BV_i + e_i \\ &= [\alpha_L + \beta_L XL_i + \varepsilon_{L,i}] + [\alpha_B + \beta_B XB_i + \varepsilon_{B,i}] \\ &= \alpha + \beta_L XL_i + \beta_B XB_i + e_i \end{aligned}$$

where XL_i and XB_i are the attributes relevant to the land and building, respectively; $\alpha = \alpha_L + \alpha_B$ is the intercept of the regression, and β is the coefficient of an individual attribute. Note that $e_i \neq \varepsilon_{L,i} + \varepsilon_{B,i}$ if the land and building equations can be estimated separately. With only TV being observable, only one α can be estimated, which cannot be systematically divided into α_L and α_B, because this is a limitation due to the lack of related empirical research (Özdilek, 2012, 2016). Similarly, if there are factors that are relevant to both land and building values, the β coefficients are also non-separable. This is a property of the infeasibility of separation as in Ely (1922), Ratcliff (1950) and Fisher (1958).

Advancements in numerical computation capability and machine learning techniques have made it possible to solve this simultaneous nonlinear equation system. Specifically, we adopt an ANN to numerically solve the above equation system simultaneously.

The basic principles and concepts of ANNs were pioneered by McCulloch and Pitts (1943). ANNs have now become effective machine learning algorithms, which are then used to model different patterns and processes. ANNs develop algorithms by using brain processing, that is, their architecture is based on interconnected systems, much like a biological neural network. Like the brain, ANNs have neurons, albeit artificial ones. Each connected neuron receives a signal from the other neurons. Then the neuron processes the signals and transmits them to the connected neurons.

Rosenblatt (1958) and Rumelhart et al. (1986) propose different methods to effectively perform ANNs. The former develops a model of how the brain stores and organizes information by using a perceptron, which is defined as a hypothetical nervous system. The function of the signal-receiving neurons is simplified into a linear model. The latter describe a learning procedure called back-propagation. Determining whether to send signals to the connected neurons is done by using a nonlinear function transformation. A method for training the model is also proposed. The overall framework is currently a popular neuron-simulation model.

The structure of ANNs can be input, hidden, and output layers of neurons. The input layer inputs features provided by the data into the network. The hidden layer can freely decide on the number of hidden layers required and the number of neurons in a hidden layer. This enables an ANN model to simulate complicated functions. The output layer represents the output of the final outcomes calculated by the ANN model.

The entire ANN can be expressed by using the following formulas:

$$Z^l = W^l a^{(l-1)} + \theta^l \tag{1}$$

$$a^l = h^l(Z^l) \tag{2}$$

where

Z^l is the output vector after the l^{th}-level linear transformation,
W^l is the parameter matrix of the l^{th}-level linear transformation,
$a^{(l-1)}$ is the output vector at the $(l-1)^{\text{th}}$ level,
θ^l is the offset function at the l^{th} level,
a^l is the output vector at the l^{th} level, and
h^l is the nonlinear transformation function at the l^{th} level.

According to the above ANN framework, all studies that have applied an ANN to predict housing value design the ANN with numerous variables in the input layer, and set multiple hidden layers with different nodes. There is only one node in the output layer, which is the target value (housing value). As housing values are composed of land and building values, therefore, we are the first to the best of our knowledge to design an ANN with two nodes in the last hidden layer which are land value and building value respectively. We add the results of these two nodes to the output layer. The output layer still only has one node, which is the property value. In this study, the model is expressed as follows:

$$f1 + f2 = property\ value \quad (3)$$

where $f1$ are the features that influence the building value function, and $f2$ are those that influence the land value function.

First, the features of real estate are classified into building or land features. The features that influence the building prices comprise transaction date, building area, transaction floor (number of floors of a house), total number of floors in the building, age of the building, building management , first floor, number of rooms, number of bathrooms, top floor , extension or unregistered building, building type, main material used to construct building, main use of building, and distance from each transaction location to the environmental facilities[1]. The features that influence the land value are land area, road width, transaction date, longitude, latitude, zoning, and distance to the surrounding environmental facilities. Transaction time and distance to the surrounding environmental facilities influence house and land values at the same time. The variables and their applications in functions $f1$ and $f2$ are listed in Table 1.

Secondly, the functions $f1$ and $f2$ are two individual ANNs that have an identical structure. The input of the $f1$ function is real estate features and its output is building value, and the input of the $f2$ function is real estate features and its output is land value. The hidden layers of each ANN have three levels. The input value is readjusted to be between -1 and 1 on the basis of the maximum and minimum values. In the calculations, we apply the Xavier weight initialization algorithm in Glorot and Bengio (2010). Meanwhile, the use of batch normalization in Ioffe and Szegedy (2015) is adopted to prevent problems such as gradient exploding or vanishing in the calculations. Finally, each ANN outputs one value, and the sum of these two values is the combined price of the real estate. For nonlinear transformation, the hidden layer uses leaky rectified linear units (ReLUs) as proposed by Maas et al. (2013)

---

[1] The surrounding environmental facilities include parks and green spaces, elementary schools, subway stations, convenience stores, parking space, hospitals, financial facilities, and not in my backyard (NIMBY) facilities.

$$f(x) = x \ \ if \ x > 0\,;\ f(x) = \lambda x \ if \ x \leq 0 \tag{4}$$

where $\lambda > 0$.

The output layer uses a ReLU as follows:

$$f(x) = x \ \ if \ x > 0\,;\ f(x) = 0 \ if \ x \leq 0 \tag{5}$$

As $f(x)$ represents the value of the land or building, in practice, both will not be negative. When the age of the building is beyond its useful life, the value of the building at most equals to zero, but the land is still valuable. Hence, when $x \leq 0,\ f(x) = 0$.

The Adam algorithm proposed by Kingma and Ba (2017) is used to automatically adjust the learning speed. The loss (target) function of the model is the mean absolute percent error (MAPE) for minimizing errors between the output and target values.

Table 1 Summary of Variables

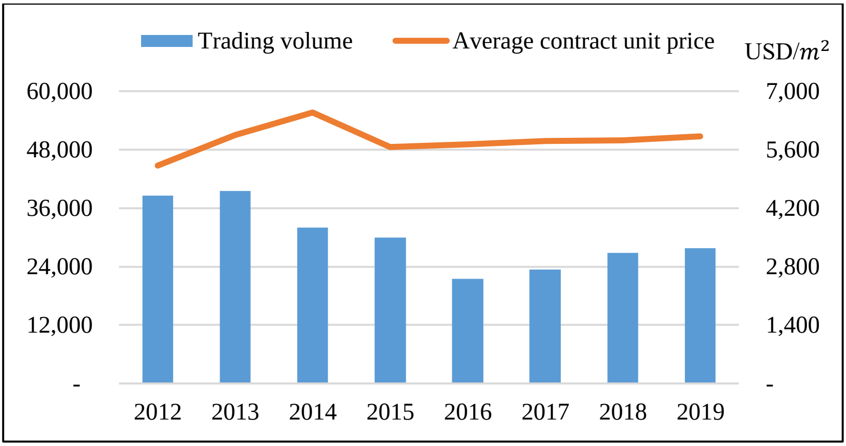
Variable	f1	f2	Variable	f1	f2
Land area		V	Number of bathrooms	V	
Road width		V	Top floor	V	
Transaction date	V	V	Extension or Unregistered building	V	
Position coordinate x		V	Building Type	V	
Position coordinate y		V	Building materials	V	
Zoning		V	Main use of building	V	
Building area	V		Distance to park	V	V
Transaction floor	V		Distance to elementary school	V	V
Building total floor	V		Distance to MRT	V	V
Age of the building	V		Distance to convenience store	V	V
Management	V		Distance to parking	V	V
First floor	V		Distance to NIMBY facilities	V	V
Number of rooms	V		Distance to hospital	V	V

3. Data Analysis

This study uses actual individual transaction data from Taipei City collected between August 2012 and May 2019. Taipei City is a well-developed city in Taiwan with 2.57 million people who are concentrated in an area of 272 square kilometers. Due to the high population and economic concentration, the housing market is characterized by high unit price and low affordability. Virtually all housing properties are in the form of multi-level condominiums with different building heights and age, which offer a large variation in the building density and quality per unit of land in Taipei. The context provides a good opportunity to study the variations in the contribution of the building value to the housing value. As the housing market conditions in Taipei City are very similar to most other Asian cities with a high population concentration, the empirical findings from this paper can be a good benchmark for other high-density cities in the east Asia region, such as Tokyo, Seoul, Singapore, etc.

There were an average of about 3000 real estate transactions each year during the period of 2012 to 2019. As shown in Figure 1, the trading volume gradually declined while the unit price rose during the sample period.

Figure 1 Transaction Volume and Contract Price Trend in Taipei City



Source: Real Estate Information Platform of the Ministry of the Interior

This paper focuses on residential land price and increase in price. Thus, only residential property sales are used for the model estimation. Data cleaning is applied to the sample to remove observations with null values, outliers, and unconventional types of transactions[2]. The remaining observations are further divided into three categories: low-rise buildings (less than five floors, no elevators), mid-rise buildings (less than 10 floors), and high-rise buildings (more than 10 floors). The final sample includes 48,675 sales, of which 25,156 are low-rise units, 10,465 are mid-rise units, and 13,054 are high-rise units. Tables 2 and 3 present the descriptive statistics of the samples.

[2] Unconventional types of transactions include transactions between families or friends or flawed transactions.

Table 2 Descriptive Statistics of Transaction Record Cases for Different Types of Buildings in Taipei City

Variable	count	mean	std	min	median	max
High-rise						
floor area (m^2)	13,054	107.79	76.33	18.82	87.20	646.51
land area (m^2)	13,054	17.74	16.46	1.77	11.96	153.43
age	13,054	15.43	12.43	0.10	10.80	50.00
trans floor	13,054	8.27	4.05	1.00	8.00	32.00
total floor	13,054	14.33	3.42	11.00	14.00	38.00
unit price (USD/m^2)	13,054	6,967	2,207	1,678	6,632	23,066
Mid-rise						
floor area (m^2)	10,465	95.56	54.97	21.83	88.30	397.84
land area (m^2)	10,465	23.88	18.88	3.25	19.81	204.46
age	10,465	20.45	13.54	0.10	21.00	51.00
trans floor	10,464	4.55	2.09	1.00	4.00	10.00
total floor	10,465	7.53	1.36	5.00	7.00	10.00
unit price (USD/m^2)	10,465	6,528	2,071	2,354	6,216	23,540
Low-rise						
floor area (m^2)	25,156	94.61	24.87	30.00	94.00	278.18
land area (m^2)	25,156	31.32	10.88	8.36	29.72	135.50
age	25,156	36.92	7.10	0.20	37.00	59.90
trans floor	25,081	3.02	1.25	1.00	3.00	5.00
total floor	25,156	4.49	0.59	2.00	5.00	5.00
unit price (USD/m^2)	25,156	5,299	1,887	2,008	4,956	47,349

Table 3 Sample Distribution of Transactions Based on Building Type at District Level

District	Low-rise		Mid-rise		High-rise		Total	
	Count	Price	Count	Price	Count	Price	Count	Price
Zhongshan	1,918	6,119	1,707	6,867	2,569	8,127	6,194	6,619
Zhongzheng	932	7,038	613	7,637	850	8,267	2,395	7,374
Xinyi	2,656	6,070	647	6,947	800	8,566	4,103	6,395
Neihu	3,528	4,604	1,174	5,682	1,330	6,510	6,032	5,098
Beitou	2,838	4,470	1,206	5,594	717	5,769	4,761	4,796
Nangang	1,177	4,978	441	5,643	643	6,232	2,261	5,439
Shilin	3,098	5,044	785	6,493	443	6,125	4,326	5,315
Datong	806	4,786	314	5,654	637	6,145	1,757	5,145
Daan	1,665	7,691	1,259	8,873	1,175	9,963	4,099	8,372
Wenshan	3,143	4,284	1,162	5,042	1,343	5,147	5,648	4,585
Songshan	1,799	6,897	823	7,179	1,140	7,547	3,762	7,037
Wanhua	1,596	3,868	334	4,874	1,407	5,617	3,337	4,376

Note: Unit of land price is USD/m^2

4. Empirical Results

Note that to further determine the accuracy and stability of the model, an out-of-sample test is performed. Therefore, we hold 20 percent of randomly selected observations as the test sample, and the other 80 percent of observation data is the training sample. The performance of the ANN model is shown in Table 4. The MAPE of the training sample model is 9.0%; the percentage predicted error of the training sample model within a 10 percentage error is 67.9%, and that within a 20 percentage error is 90.2%. The MAPE of the out-of-sample model is 10.7%; the percentage predicted error of the test model within a 10 percentage error is 59.9%, and that within a 20 percentage error is 86.6%. The prediction results of the model in this study are superior to those of previous studies that use an ANN to predict the housing prices in Taiwan (Lai, 2007; Tsaih et al., 1999). The test sample and training sample performances are consistent (the error between the two is < 3%), thus indicating that the model in this study is sufficiently stable and no significant overfitting problem is observed.

Table 4 Overall Appraisal Error Performance of the Artificial Neural Network Model

	Training sample	Testing sample
MAPE	9.0%	10.7%
Percentage Predicted Error 10%	67.9%	59.9%
Percentage Predicted Error 20%	90.2%	86.6%
Sample Size	38,940	9,735

The results of a refinement analysis of the appraisal performance of the different building types are listed in Table 5. The appraised performance of high-rise buildings is the highest with the lowest MAPE (7.9%), followed by condominiums (8.7%), and apartments which have a larger MAPE (9.8%). In terms of the hit rates, high-rise buildings also have the highest performance followed by condominiums and apartments. The overall performances are higher than those of international standards[3].

As the ratios of factors that influence building and land can be calculated separately by using an ANN analysis, the ratios of land prices to total real estate prices can be calculated. McCain et al. (2003) conclude that different building types result in different land prices. To determine the validity of this statement, we analyze house and land results separately across building type. The proportion of land value for low-rise buildings is 63.4%, and the proportion of

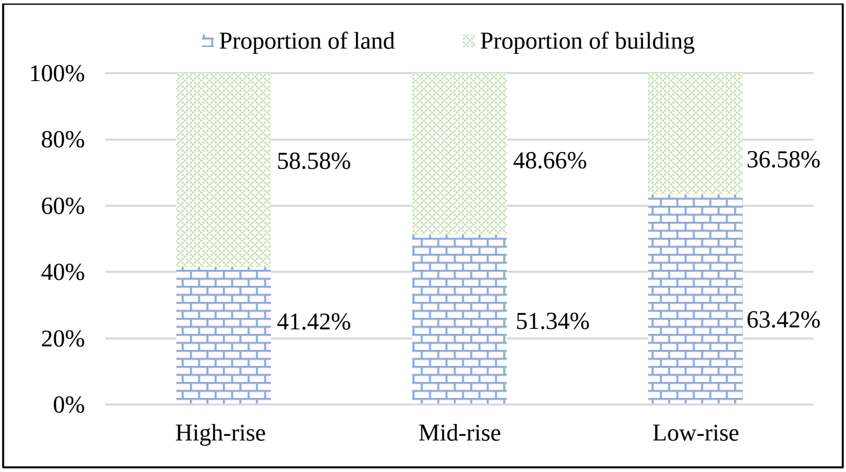
[3] According to international standards for mass appraisal models, a hit rate with a 10% error must reach 40%, and that with a 20% error must reach 70%.

building value is 36.6%. For mid-rise buildings, the proportion of land value is 51.3%, and the proportion of building value is 51.3%. For high-rise buildings, the proportion of land value is 41.4%, and the proportion of building value is 58.6%. It is found that different land area ratios of different building types result in different proportions of land value.

Table 5 Performance of ANN Model for Different Building Types

	Low-rise		Mid-rise		High-rise	
	Training sample	Testing sample	Training sample	Testing sample	Training sample	Testing sample
MAPE	9.8%	11.5%	8.7%	10.5%	7.9%	9.3%
Percentage Predicted Error 10%	64.9%	56.6%	68.6%	59.8%	73.1%	66.2%
Percentage Predicted Error 20%	88.6%	84.3%	91.0%	87.0%	92.8%	90.7%
Sample Size	20,140	5,016	8,348	2,117	10,452	2,602

Figure 2 Proportion of Land and Building Values for Different Building Types



Housing age is also an important factor that influences housing value. As Kutty (1999) states, buildings depreciate with age, yet the value of land continues to increase because of its scarcity. Land is immortal in nature, whereas buildings deteriorate over time. Therefore, land value typically increases as the building ages. Furthermore, this study examines the performance of decomposed land

and building values with respect to building type and age. Intervals of 5 years are used to group buildings into different categories. The differences in proportion of land value with age of the building are examined by using a total of seven intervals.

Table 6 presents the proportion of the land value of different building types as the house age increases. Low-rise buildings that are less than 5 years old have a proportion of building value of 47.7%, which is lower than that of the land value of 52.3%. However, as the house age increases, the proportion of the building value gradually decreases, and the proportion of the land value increases. When a low-rise building is more than 20 years old, the proportion of the building value decreases to less than 40%, and that of land value increases to more than 60%. For a mid-rise that is less than 10 years old, the proportion of the building value is higher than that of the land value. For a mid-rise building that is between 15 and 30 years old, the proportion of the land value converges to approximately 55%. A high-rise building that is less than 5 years old has the highest proportion of building value at 63.2% among all of the combinations. This is consistent with the intuition that buildings with more floors have a higher share of building value because they require higher construction costs and more building materials per unit of land than the mid- and low-rise buildings. Therefore, the building value of a newer high-rise building is worth 72.5% more than its land value. As high-rise buildings typically have superior and more durable building materials that deteriorate at a slower rate, their building value continues to exceed their land value until the buildings are older than 30 years.

Table 6 Proportion of Land Value for Different Building Types and Different Building Age[4]

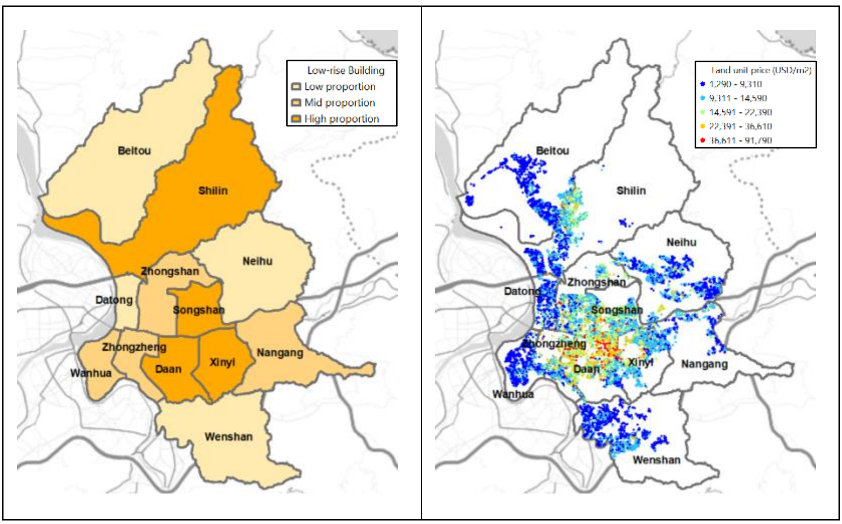
	High-rise	Mid-rise	Low-rise
<5 years	36.8%	43.9%	52.3%
5-10 years	38.3%	46.4%	56.4%
10-15 years	39.1%	51.3%	57.5%
15-20 years	43.3%	54.1%	61.6%
20-25 years	46.1%	52.5%	64.3%
25-30 years	48.6%	56.1%	64.4%
>30 years	50.4%	56.2%	63.5%

Land value is affected by the locations and submarkets. Table 7 shows the proportion of the land value of different building types and administrative

[4] Results are for Taipei City. The same method is applied to other cities in Taiwan, such as Taoyuan City, New Taipei City, and Tainan City. The results show similar land value sensitivity to building age. Those results are available upon request.

districts of Taipei City. The table shows that the top three districts with the highest proportion of land value are Songshan, Shilin and Daan. The proportion of the land value accounts for over two-thirds of the total value. On the other hand, Daton and Beitou have the lowest proportion of land value, which is slightly over half of the total value. For mid-rise buildings, the highest proportion of the land value are observed in Daan, Songshan, Xinyi, and Shilin; all of them with a proportion of land value that exceeds 57%. Daton and Wanhua have the lowest proportion of land value, or less than 40%. For high-rise buildings, the highest proportion of the land value is found in Daan, Xinyi, and Shilin, where the proportion of the land value is over 48%. Again, Daton and Wanhua have the lowest proportion of land value, or less than 31%. The geographical pattern is very similar among the three building types. Districts with buildings that have a higher proportion of land value tend to be more expensive communities. This indicates that the share of the land value tends to be more sensitive to local amenities while that of the building value tends to be less volatile across locations. To verify the rationality of the splitting results of the ANN model, we calculate the unit land price per square meter. Figure 3 shows the land price hierarchy map in Taipei. It can be observed that the areas with higher land prices are concentrated in Daan, Zhongzheng, Zhongshan, Songshan and Xinyi, and other districts with lower land prices. These results are consistent with the housing price relationship of each district.

Figure 3 Proportion of Land Value for Different Districts and Land Value Heat Map



Base map source: Stamen Toner/OSM

Table 7 Proportion of Land Value for Each Building Type in Different Districts

District	Low-rise	Mid-rise	High-rise
Zhongshan	63.1%	45.2%	37.0%
Zhongzheng	62.7%	51.2%	40.8%
Xinyi	65.3%	57.7%	50.2%
Neihu	59.8%	51.9%	43.3%
Beitou	55.3%	44.2%	31.2%
Nangang	62.5%	53.2%	43.8%
Shilin	71.0%	57.1%	48.4%
Datong	53.1%	39.1%	29.8%
Daan	67.8%	61.1%	53.8%
Wenshan	61.1%	47.7%	42.1%
Songshan	72.6%	59.7%	47.6%
Wanhua	63.3%	40.3%	33.3%

5. Conclusions and Recommendations

A number of practical applications require separate estimates of the value of the land and the building of a housing property. Since there is only one observable value of the entire property after development, it is generally infeasible to empirically single out land value from the total value. Therefore, the value separation exercise remains dependent on artificially designed appraisal approaches, such as methods that use the cost of construction. There is very limited data driven empirical research in the existing literature.

This paper has developed an empirical model that can be used to separately estimate housing value by using two components, the value of the land and the building. Specifically, the hedonic factors of housing value are divided into land/location related and building related factors. Thus, two hedonic models are constructed that separately estimate land and improvement values. With the actual sales price of the combined property, an ANN algorithm is used to estimate the two hedonic equations by minimizing the estimation error of the total value. Advancements in numerical computation capability means that ANN algorithms can be calculated within a reasonable amount of time.

Individual housing transaction data of Taipei City during 2012-2019 are used to empirically test the model. The empirical results show that the proportion of the land value increases with age of the house. This is consistent with the findings in the existing literature and industry experience. We also find that the proportion of the land value differs by building type. The proportion of the land value tends to be the highest for low-rise buildings and the lowest for high-rise buildings. This is again consistent with the fact that high-rise buildings tend to be constructed with higher quality and more durable materials. More building space is offered within each unit of land space. Therefore, the proportion of the building value tends to increase with more floors. Finally, the proportion of the building value tends to be higher in less expensive locations. This indicates that the high housing value in expensive communities are mainly driven by the higher land value. Within our Taipei City sample, the proportion of the building value can differ by 20 percentage points between the most and least expensive housing sub-markets. Performance in different administrative districts differs as well.

References

Abidoye, R.B., and Chan, A.P.C. (2017). Artificial Neural Network in Property Valuation: Application Framework and Research Trend. *Property Management*, 35(5), 554-571.

Borst, R.A. (1991). Artificial Neural Networks: The Next Modelling/Calibration Technology for the Assessment Community. *Property Tax Journal*, 10(1), 69-94.

Do, A.Q., and Grudnitski, G. (1992). A Neural Network Approach to Residential Property Appraisal. *Real Estate Appraiser*, 58(3), 38.

Ely, T.R. (1922). Research in Land and Public Utility Economics. *Land Economics, 1*(1), 1–5.

Fisher, E.M. (1958). Economic Aspects of Urban Land Use Patterns. *The Journal of Industrial Economics, 6*(3), 379–386.

Glorot, X., and Bengio, Y. (2010). Understanding the Difficulty of Training Deep Feedforward Neural Networks. In : Proceeding of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS) 2010, Chia Laguna Resort, Sardinia, Italy.

Gloudemans, R.J. (2002). An Empirical Analysis of the Incidence of Location on Land and Building Values. Lincoln Institute of Land Policy Working Paper.

Guerin, B.G. (2000). MRA Model Development Using Vacant Land and Improved Property in a Single Valuation Model. *Assessment Journal,* 7(4), 27-34.

Hendriks, D. (2005). Apportionment in Property Valuation: Should We Separate the Inseparable? *Journal of Property Investment & Finance, 23*(5).

Ilić, D., and Mizdrakovic, V. (2016). Risks of Property Value Apportionment. Conference Finiz 2016. 10.15308/finiz-2016-125-129.

Ioffe, S., and Szegedy, C. (2015). Ioffe, S., and Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In : Proceeding of the 32nd International Conference on Machine Learning, Lille, France.

Kauko, T. (2003). On Current Neural Network Applications Involving Spatial Modelling of Property Prices. *Journal of Housing and the Built Environment, 18*(2), 159.

Kingma, D.P., and Ba, J. (2017). Adam: A Method for Stochastic Optimization. International Conference on Learning Representations

Kutty, N. (1999). Determinants of Structural Adequacy of Dwellings. *Journal of Housing Research*, 10(1), 27-43.

Lai, P.P.-Y. (2007). Applying the Artificial Neural Network in Computer-assisted Mass Appraisal. *Journal of Housing Studies,* 16(2), 43-65.

Maas, A.L., Hannum, A.Y., and Ng, A.Y. (2013). Rectifier Nonlinearities Improve Neural Network Acoustic Models. *30th International Conference on Machine Learning, 28. Atlanta, Georgia.*

González, M.A.S., G., Soibelman, L., and Carlos Torres, F. (2005). A New Approach to Spatial Analysis in CAMA. *Property Management*, 23(5), 312-327.

McCluskey, W., Davis, P., Haran, M., McCord, M., and McIlhatton, D. (2012). The Potential of Artificial Neural Networks in Mass Appraisal: The Case Revisited. *Journal of Financial Management of Property and Construction,* 17(3), 274-292.

McCulloch, W.S., and Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biology,* 52(1-2), 99-115; discussion 173-197.

Nguyen, N., and Cripps, A. (2001). Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks. *The Journal of Real Estate Research,* 22(3), 313-336.

Özdilek, Ü. (2012). An Overview of the Enquiries on the Issue of Apportionment of Value Between Land and Improvements. *Journal of Property Research*, 29(1), 69-84.

Özdilek, Ü. (2016). Property Price Separation between Land and Building Components. *The Journal of Real Estate Research*, 38(2), 205-228.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., and French, N. (2003). Real Estate Appraisal: A Review of Valuation Methods. *Journal of Property Investment & Finance, 21*(4), 383-401.

Peterson, S., and Flanagan, A.B. (2009). Neural Network Hedonic Pricing Models in Mass Real Estate Appraisal. *The Journal of Real Estate Research, 31*(2), 147-164.

Ratcliff, R.U. (1950). Net Income Can't Be Split. *The Appraisal Journal,* 18(April), 168–172.

Renigier-Bilozor, M., and Wisniewski, R. (2012). The Impact of Macroeconomic Factors on Residential Property Price Indices in Europe. *Folia Oeconomica Stetinensia,* 12(2), 103-125.

Ribeiro, M., Singh, S. and Guestrin, C. (2016). '"Why Should I Trust You?": Explaining the Predictions of Any Classifier', *Conference: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations,* 97-101.

Roger A. McCain, Jensen, P., and Meyer, S. (2003). Research on Valuation of Land and Improvements in Philadelphia. *Department of Economics and International Business, LeBow College of Business Administration, Drexel University, Philadelphia*.

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological review,* 65(6), 386-408.

Rumelhart, D., Williams, R., and Hinton, G. (1986). Learning Representations by Back-Propagating Errors. *Nature*, 323, 533-536.

Shrikumar, A., Greenside, P., and Kundaje, A. (2017). Learning Important Features Through Propagating Activation Differences. In. Ithaca: Cornell University Library, arXiv.org

Sunderman, M. A., and Birch, J. W. (2001). Valuation of Land Using Regression Analysis, US, Springer.

Tsaih, R.-H., Kao, M.-C., and Chang, C.-O. (1999). Neural Networks Technique for Residential Property Appraisal in Taipei. *Journal of Housing Studies,* 8, 1-20.

Weinberger, A. M., JD, and LeGrand, Lauren E, JD. (2009). Property's Total Value Must Be Allocated Appropriately between Land and Improvements. *The Appraisal Journal*, 77(2), 93.

Yacim, J.A., and Boshoff, D.G.B. (2018). Impact of Artificial Neural Networks Training Algorithms on Accurate Prediction of Property Values. *The Journal of Real Estate Research*, 40(3), 375-418.

Zhai, G., Fukuzono, T., and Ikeda, S. (2003). Effect of Flooding on Megalopolitan Land Prices: A Case Study of the 2000 Tokai Flood in Japan. *Journal of Naural Disaster Science,* 25(1), 23–36.