FORECASTING THE LAND VALUE AROUND COMMUTER RAIL STATIONS USING HEDONIC PRICE MODELING

Mohammed Ali Berawi[1*], Nyoman Suwartha[1], Kahfi Kurnia[1], Gunawan[1], Perdana Miraj[2], Abdur Rohim Boy Berawi[1]

[1]*Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia*
[2]*Department of Civil Engineering, Faculty of Engineering, Universitas Pancasila, DKI Jakarta 12640, Indonesia*

ABSTRACT

The value of land is determined by the relationship between supply and demand in the land market and the land's location, physical structure, and surrounding area. During infrastructure development, changes in land use and its characteristics are possible and may directly affect the land price significantly. This research aims to generate a prediction model of land market values in response to transportation infrastructure development and elaborates a correlation among factors affecting the land price. A hedonic pricing model is used through multiple linear regression to seek contributing factors affecting the land price. The result shows that the proposed hedonic model can predict the price of land using the model LnLand Price = 16.991 + 0.203 Shape of land + 0.084 Economy of Residents − 0.719 LogStation − 0.405 LogAge. This estimated formula could be used as part of a land value capture mechanism for future development.

Keywords: Hedonic price modeling; Land price; Land value capture; Railway

1. INTRODUCTION

Land value capture (LVC) is an effort to monetize infrastructure capital investment in order to capitalize land value. LVC consists of three components: the value uplift, the laws concerning the interpretation of how an area contributes to the value uplift, and the conversion of the increased land price into taxes (Smolka, 2007). The framework of LVC collects the documented value created, identifies the stakeholder who benefits from the built infrastructure, analyzes the LVC mechanism, and proposes a recommendation.

Railway Transit potentially reduces the cost of travel and induces a higher value of property (Berawi & Akbar, 2017; Rahman et al., 2018). The location of a place in a transportation system determines its accessibility and affects the land values (Tolley & Turton, 1995). Studies conducted in relation to the Helsinki Metro Rail, Finland, indicate that property prices within a walkable distance (500–750 m) of a metro station are 7.5% higher than those beyond that distance (Laakso, 1992). Transit-oriented development in San Diego, California, has found that the further the location of land from a station, the lower its value (Duncan, 2011). In a similar study in Portland, results showed that within 500 meters of an LRT Station, land prices increased by about $4,323, while outside the 500-meter radius, the price was relatively stable

(Al-Mosaind et al., 1993).

Similar findings were also found in Wuhan, China, where the government invests heavily in transportation infrastructure, with regard to commercial property prices. The research focused primarily on a radius of 800 meters from the transport station. The study concluded that the price of property was significantly affected within a distance of 400 meters, with the price gradually decreasing every 100 meters (Xu et al., 2016). Location, which plays a significant role in land value, can be divided into two factors: environment and accessibility. Distance to the Central Business District (CBD) may have an impact on land values, depending on the geographical location of a particular property and the economic condition of that area (Haider & Miller, 2000). Other factors contributing to location attributes are the width of the road and the distance from the property to the main road. The good quality of the road will determine whether people are tempted to occupy the land for its access to modes of transportation. A better road also provides greater accessibility for people to conduct their daily activities.

The structural attributes of a particular land area are among the main factors of a land value increase. Land located near the city center has a higher land price, but it is hard to occupy a large amount of land there. By contrast, land far away from the city center is cheaper and less attractive (Siswanto, 2007). The condition of the property also plays a part. The age and the number of rooms in a property contribute to property values. One of the main reasons for this is that people need particular amenities when buying a property and are willing to pay more for better amenities. An older property tends to require repairs, which may decrease the property's value and the interest of buyers.

Environmental attributes indirectly affect the land value market, but through hedonic pricing, a correlation can be found. Property values increase for housing whose owners have a high salary rather than a low salary. People with a high salary can afford an elite property with facilities such as 24-hour security or a swimming pool, which make a property more expensive (Nelson, 1992)

Unfortunately, the price of land is sometimes inaccurate when based on the variables affecting the land price and the characteristics of the surrounding areas. Instead, land prices can be the result of speculation by developers and the transaction between buyers and sellers. This research aims to obtain a model for land price prediction that can be implemented to evaluate land prices in areas near to where railway projects are constructed (Wijito & Sindang, 2012).

2.   METHODS

The research combined qualitative and quantitative approaches. The qualitative approach used in-depth interviews with 92 landowners through a questionnaire survey. The population sample consisted of property owners living within a radius of 800 meters of the Duren Kalibata Station in South Jakarta, Indonesia.

Duren Kalibata is part of a transit-oriented development area, strategically located along the railroad line close to Rawajati Road. Duren Kalibata Station is clustered with commercial properties, such as Kalibata city and Swiss Bell residence; residential property; and shopping centers, such as Kalibata mall.

Each 100-meter interval boundary is located on a specific road as a benchmark for a clustered population sample. The results obtained are the transaction price and asking price of the land. The data will be benchmarked using previous similar research from Wuhan, China (Xu et al., 2016).
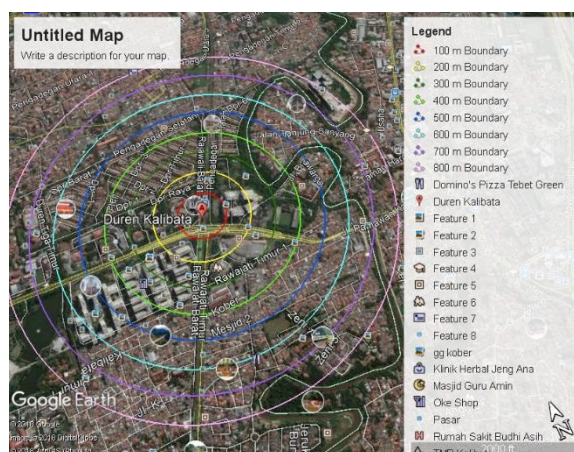
Figure 1 Study area of research and boundaries

The quantitative approach involved measuring the distance of properties to the CBD, the station, and the main road using the Geographical Information System (GIS) software ArcGIS. Variables such as the price of land, area, number of rooms, and the age of the building were obtained through interviews (Karim et al, 2007; Berawi et al., 2016), and the ordinal variables, such as the shape of the land, elevation of the land, and the socio-economic status of residents, were obtained by field observation. All the variables in the research and definition are summarized in Table 1.

Table 1 Dependent, physical, accessibility, and environmental variables

Variable	Description	Expected Sign
Price	Land price of the property (Rupiah)	DV
Land Area	The area of land (in square meters)	-
CBD	Distance to Central Business District (in meters)	-
Station	Distance to station (in meters)	-
Main Road	Distance to main road	-
Shape	Shape of land (ordinal scale 1–5)	+
Elevation	Elevation of land (ordinal scale 1–5)	+
Economy	Social status of residents in the area (ordinal scale 1–5)	+
Environment	Environmental quality of residence (ordinal scale 1–5)	+
Rooms	Number of rooms in the property	+
Age	Age of the building (in years)	-
Width	Width of the road in the area (in meters)	+

Note: DV indicates a Dependent Variable; (+) increasing/ (-) reducing the land price

The price variable is the asking price or the transactional data price of the land of the property expressed in Rp and is a dependent variable. In the proposed model, the variable of the price of land is explained by the following independent variables:

- Land Area: The size of the land (structural attributes)
- CBD: Distance to the CBD, represented by Sudirman Central Business District (accessibility attributes)
- Station: Distance to Duren Kalibata Station (accessibility attributes)
- Main_Road: Distance to the main road represented by Jalan Raya Kalibata (accessibility attributes)
- Shape: The shape of the land, an ordinal variable scaled from 1 to 5, 5 being a regular square and 1 an irregular shape (physical attributes)

- Elevation: The observed slope of the land scaled from 1 to 5, 5 being a flat slope and 1 being more than 40 degrees (physical attributes)
- Economy: The socio-economic status of the residents, varying from 1 to 5, 5 being residents with a high salary and 1 being unemployed residents (environmental attributes)
- Environment: The quality of the environment, scaled from 1 to 5, the larger space of green area, the highest score of environment.
- Rooms: The number of rooms in the building on the land; both bedrooms and bathrooms are counted (physical attributes)
- Age of Building: The age of the building from the time it was built until 2018 (physical attributes)
- Width: The width of the road in front of the property, giving access to the property (accessibility attributes)

A statistical test was used to evaluate the credibility of the model when the data collection was finished, as shown in Figure 2. First, the data were analyzed using multiple linear regression to find the correlation among the variables. A normality residue test was also performed to expose the differences between the model estimate and the observed value. A t-test was conducted to eliminate variables that had no partial correlation to the dependent variables. In addition, a multicollinearity test was conducted to generate correlated variables, and an f-test was used to determine which variables were optimal for the model. Finally, a heteroscedasticity test was used to ensure that the model was homoscedastic and constant.

2.1. Hedonic Price Modeling

Hedonic price modeling (HPM) relates the asking price of land with its accessibility, environmental, and physical attributes. It is a result of independent variables that can be monetized and calculated by observation of price differences of land with the same attributes. HPM consists of a regression analysis, which is a statistical analysis used to correlate different data points (Duncan, 2011).

The result of the analysis can be used to predict future values. It is an instrument to determine how the dependent variable is affected by the independent variables. The correlation is shown as a coefficient, which is the result of the multiple linear regression (Monson, 2009). The result of the variable prediction model after the multiple regression analysis is shown below:

$$Y = a + b_1.x_1 + b_2.x_2 + b_3.x_3 …. \tag{1}$$

where;
 Y : Independent Variable
 a : Constant
 b : Regression Coefficient
 x_1 : First Independent Variable
 x_2 : Second Independent Variable
 x_3 : Third Independent Variable

In this study, the relationship between the asking price and the characteristics of the land is taken as a log-log specification. The log of the dependent variable is regressed against the log of the explanatory variables that have value. Meanwhile, the ordinal variables are not transformed as logs, and the regression is expressed as follows.

$$\ln(Y) = a + b_1\log(x_1) + b_2\log(x_2) + b_3\log(x_3) + … \tag{2}$$

The logarithmic model enables us to calculate and measure the impact of the independent variables related to the dependent variables in relative terms. The coefficient (bn) represents the

elasticity value of the independent variables, while (a) is the constant of the variables when all the independent variables are equal to zero.
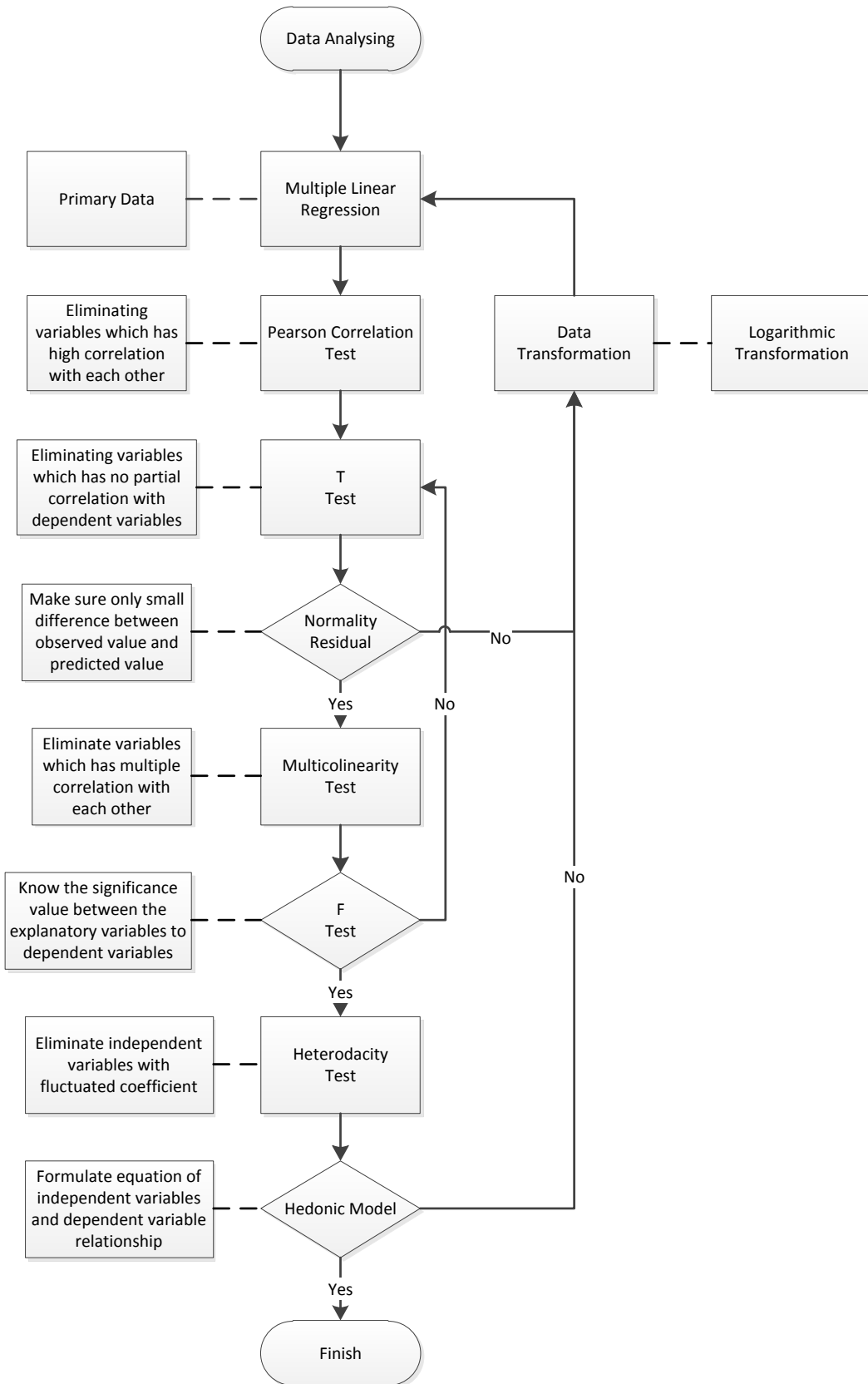
Figure 2 Multiple regression statistical test

3. RESULTS AND DISCUSSION

Table 2 presents the results of the multiple linear regression of the 11 variables. The independent variables that are statistically below the 0.05 significant level are the area of land, elevation of land, quality of the environment, distance to the CBD, distance to the main road, number of rooms, and width of the road. Variables that are significant have the appropriate signs.

The elevation of land and the quality of the environment will be excluded from the final model because the value of their Pearson correlation coefficients is 0.915 and 0.713, respectively. Each of these is above 0.5, which means that the variables correlate with each other.

Table 2 Log-log model estimated by multiple linear regression

Model	Unstandardized B	Coefficient Std. Error	Standardized Coefficients Beta	t	Sig.
Constant	16.991	0.355		47.816	0.000
Shape	0.203	0.037	0.298	5.432	0.000
Economy	0.084	0.026	0.189	3.239	0.002
LogAge	-0.405	0.058	-0.260	-6.957	0.000
LogStation	-0.719	-0.104	-0.386	-6.934	0.000
Adjusted R^2	0.810				
F-Statistics	94.001				0.000
Observation	92				
Normality Residuals	0.086				0.089

Note: Each coefficient is significant at the 0.05 level; Explanatory variables are in log form except for the ordinal variable; The dependent variable is in logarithmic scale

The final model includes all the explanatory variables that have partial significance to the dependent variables. The statistical test reveals that the distance to the station has the most significant explanatory power. It is followed by the age of the building, the shape of the land, and the socio-economic status of the residents, respectively. The explanatory variables are statistically significant at the 0.05 significance level. Variables that have insignificant correlation to the land price have been eliminated, and the regression was repeated with variables to generate the expected significant level.

The set of explanatory variables makes it possible to explain 81% of the dependent variables comprising variation in the asking price of land. The F ratio test has a value of 94 with a significance of 0.000 at the significance level. The explanatory variables model contained no heteroscedasticity; the explanatory variables show significance above 0.05 when the dependent variables are replaced with residuals. It is statistically significant and showing the homoscedasticity essential for the hedonic model due to the stabilized coefficients. As for the residual normality test, it confirms that the model has residual normality with a significance value above 0.05. This means that the difference of the observed value and the predicted value is tolerable and forms residuals along the best fit line. The outcome of the multiple linear regression resulted in the equation shown below.

$$\text{LnPrice of Land} = 16.991 + 0.203\ \text{Shape} + 0.084\ \text{Economy} - 0.719\ \text{LogStation} - 0.405\ \text{LogAge} \quad (3)$$

- The constant coefficient value is stated at 16.991, which has a significance level at 0.05. This means that with other independent variables equal to zero, the value of land is 16.991

in logarithmic form. When converted to normal values, the value of land is Rp. 23,938,533, which also acts as a y-intercept of the multiple linear regression.

- An increase in land shape of 1 point on the scaled ordinal assessment will result in an increase in the asking price of as much as 20.3%. This is a better estimate than that of previous research with the same variables, which obtained a coefficient of 14.7% . The calculation can be used as follows. Say that a property has a land value of Rp. 2,000,000 with an assessed value of 5 for its square shape. If there is a change in its assessment to 3, the value of the land will decrease by as much as 63.52%. This can be calculated by reducing 100% by 20.3% = 79.7 % for the reduction assessment scale to 4, and for the reduction assessment to 3, the 79.7% is multiplied by 20.3% and the result is 16.18%, which is used to reduce the 79.7% to obtain 63.52%. Thus, the price of land is corrected to Rp. 1,270,400.

- The socio-economic status of the residents contributes 8.4% to the increased price of land, with an assessment scale from 1 to 5. Residents with a higher economic status will purchase land that is more expensive and will push developers to create residential properties in a specific location where people with high salaries will be interested in purchasing them.

- In relation to distance from Duren Kalibata Station, the price drops by 71.9% over the 100-meter intervals. This research is in line with the Wuhan research, which used 100-meter intervals for the distance to the station. It shows that the price of a property decreases for up to 400 meters. The Wuhan research used distance intervals as a research variable and used nominal scales. The research found that distance from the station above 400 meters was not significant because the variables did not affect the dependent variable. This indicates that land near to stations provides accessibility for nearby residents, and the residents in Duren Kalibata prefer a land location closer to access to public transportation (Xu et al., 2016). This research is also in line with research conducted by Weinberger (2000), which focused on the effect of the distance from a station on property price. It found that property located within a quarter of a mile from the station saw an increase in price of 13%. On the other hand, a property between a quarter and half a mile from the station has an increase by as much as 7%, and property between a half a mile and three-quarters of a mile only has an increase price of as much as 1%. On the other hand, property located between three-quarters of a mile and a mile away saw no significant difference, which supports the idea that long-distance walking and easy access to public transport have a significant effect on land price.

- The age variable contributes over 40.5% to the decrease in land value, since it has a negative coefficient, resulting from depreciation of the building on the land as the age increases and degradation of land below it, which needs more money for repairs. The hedonic model established in Hong Kong had a negative coefficient of 42.3% with a margin of difference of 1.8%. As we can see, people tend to buy land with newer property on and with newer structures. Older structures are less preferred and demanded by homebuyers, leading to a decrease in land price (Tse, 2002).

Further research needs to be conducted with regard to the variables that characterize the features of an area and to add more samples and different case studies. A different area may produce a different outcome for the linear regression formula and therefore generate results that are more comprehensive.

4.   CONCLUSION

On the basis of its model, this research generates the following equation for the market price of land in the Duren Kalibata Station catchment area of 800 meters:

LnPrice of Land = 16.991 + 0.203 Shape + 0.084 Economy – 0.719 LogStation – 0.405 LogAge

This equation has a constant value of 16.991. This is the value of the land when all of the variables are zero and the price of the land is Rp. 23,938,533. The highest elasticity of the independent variable is the distance to the station, which has 71.9% elasticity. The age of the building decreases the land value by as much as 40.5%, making age the second-most elastic of the independent variables within the model. The price of the land increases by as much as 20.3% when the shape of the land is regular or rectangular; this is the second lowest in elasticity among the variables in the model. The social status of residents is the variable that contributes the least to the price of the land, and with an elasticity of 8.4% has the lowest elasticity within the model.

The model has an R squared of 0.81, which means that the model accounts for 81% of the relationship between the independent and dependent variables. Only 19% of the relationship between the independent and dependent variables cannot be explained by the model.

5.   ACKNOWLEDGEMENT

6.   REFERENCES

Al-Mosaind, M.A., Dueker, K.J., Strathman, J.G., 1993. Light-rail Transit Stations and Property Values: A Hedonic Price Approach. *Transportation Research Record,* No. 1400, pp. 90–94

Berawi, M.A., Miraj, P., Berawi, A.R.B., Silvia, S., Darmawan, F., 2016. Towards Self-sufficient Demand in 2030: Analysis of Life-cycle Cost for Indonesian Energy Infrastructure. *International Journal of Technology*, Volume 7(8), pp. 1445–1454

Berawi, A.R.B., Akbar, F., 2017. A Benchmark Study for Indonesia's High Speed Train Considering Technology Selection. *Advanced Science Letters*, Volume 23(7), pp. 6343–6346

Duncan, M., 2011. The Impact of Transit-oriented Development on Housing Prices in San Diego, CA. *Urban Studies*, Volume 48(1), pp. 101–127

Haider, M., Miller, E.J., 2000. Effects of Transportation Infrastructure and Location on Residential Real Estate Values: Application of Spatial Autoregressive Techniques. *Transportation Research Record*, No. 1722(1), pp. 1–8

Karim, S.B.A., Rahman, H.A., Berawi, M.A., Jaapar, A., 2007. A Review on the Issues and Strategies of Stakeholder Management in the Construction Industry. *In:* Meetings and Conference on Management in Construction and Researchers Association (MICRA), 28-29 August 2007, Shah Alam, Selangor, Malaysia

Laakso, S., 1992. Public Transport Investment and Residential Property Values in Helsinki. *Scandinavian Housing and Planning Research*, Volume 9(4), pp. 217–229

Monson, M., 2009. Valuation using Hedonic Price Models. *Cornell Real Estate Review*, Volume 7, pp. 62–73

Nelson, A.C., 1992. Effects of Elevated Heavy-rail Transit Stations on House Prices with Respect to Neighborhood Income. *Transportation Research Record,* No. 1359, pp. 127–132

Rahman, H.Z., Berawi, M.A., Susantono, B., Miraj, P., Petroceany, J.S., Maya, R., 2018. Investigation of an Operation and Maintenance Framework in the Railway Industry: A Case Study of the Makassar-Parepare. *International Journal of Technology*, Volume 9(3), pp. 549–557

Siswanto, E., 2007. Kajian Harga Lahan dan Kondisi Lokasi Lahan Permukiman di Kecamatan Arga Makmur, Kabupaten Bengkulu Utara (*Study of Land Prices and Settlement Site Conditions in Arga Makmur Sub-District, North Bengkulu Regency*). Master's Thesis, Post Graduate Program of Urban and Regional Development, Diponegoro University, Semarang

Smolka, M.O., 2007. *Implementing Value Capture in Latin America: Policies and Tools for Urban Development.* Lincoln Institute of Land Policy, Cambridge

Xu, T., Zhang, M., Aditjandra, P.T., 2016. The Impact of Urban Rail Transit on Commercial Property Value: New Evidence from Wuhan, China. *Transportation Research Part A: Policy and Practice*, Volume 91, pp. 223–235

Tolley, R.S., Turton, B.J., 1995. *Transport Systems, Policy, and Planning: A Geographical Approach.* Longman Scientific and Technical, England

Tse, R.Y.C., 2002. Estimating Neigbourhood Effects in House Prices: Towards a New Hedonic Model Approach. *Urban Studies*, Volume 39(7), pp. 1165–1180

Weinberger, R.R., 2000. Commercial Property Value and Proximity to Light Rail: Calculating Benefits with a Hedonic Pricing Model. *In*: Transportation Research Board 79th Annual Meeting, Washington, D.C. January 9–13

Wijito, L., Sindang, E., 2012. Penerapan Model Hedonic dalam Penentuan Nilai Tanah sebagai Referensi dalam Penilaian Barang Milik Negara dan Harga Limit Lelang (*Hedonic Model in Land Value Estimation to Assess State Property and Auction Price Limit).* Academic Study, Financial Education and Training Agency, Ministry of Finance, Indonesia (Unpublished)