African Journal of Business Management Vol. 6(14), pp. 4948-4956, 11 April, 2012
Available online at http://www.academicjournals.org/AJBM
DOI: 10.5897/AJBM11.2781
ISSN 1993-8233 ©2012 Academic Journals

Full Length Research Paper

Waiting period for agency-brokered housing transactions

Jeng-Hsiang Lin

Hwa Hsia Institute of Technology, No.111, Gongjhuan Rd., Jhonghe District, New Taipei City 23568, Taiwan. E-mail: hsiang@cc.hwh.edu.tw. Fax: 886-2-29454892.

Accepted 9 December, 2011

Prospective sellers of housing often commission housing agencies to broker transactions. The relationship between the waiting period for sellers (from the date of commission to the date of transaction) and housing variables is a topic of considerable interest to real estate investors, housing agents, and researchers. This study examines the effect of these factors on the waiting period. Data was collected on 4,256 housing transactions brokered by agencies from 2007 to 2011 in New Taipei City, Taiwan. Parametric survival analysis was performed to produce a probability model for the transaction waiting period for sellers and establish a relationship between the waiting period and housing variables. Empirical results show that the transaction waiting period for sellers follows the Weibull probability distribution.

Key words: Housing transactions, waiting period, housing agents, survival analysis.

INTRODUCTION

In recent years, increasingly active trading in the real estate market has led a growing number of prospective sellers of housing to commission agencies to broker transactions. The means to effectively analyze the relationship between various features and attributes of real estate and the transaction waiting period, and use this relationship to predict the waiting period for sellers for future transactions, are topics of keen interest to all parties involved in the real estate market.

The trends in real estate prices and how quickly real estate transactions are concluded can be seen as indicators of economic health. Most studies in the domain of real estate have concentrated on estimation of real estate prices, developed analysis models and methods to target these issues, and derived specific and constructive research conclusions. Previous researchers have generally used two approaches when studying factors that influence housing prices. One approach is to combine macroeconomic conditions and the supply-demand principle of microeconomics to observe factors that may affect fluctuations in housing prices (Giannias, 1998; Thalmann, 1999). The second is the hedonic price approach (Lancaster, 1966), which focuses on various features and attributes of real estate, and uses the consumer utility theory to analyze the influence of the implicit price of these attributes on actual product price (Singell and Lillydahl, 1990; Evans, 1973; Ha and Weber, 1994; Laurice and Bhattacharya, 2005; Giannias, 1998). Most previous studies, however, have applied these techniques to processing price data. However, very few studies have been conducted on the subject of waiting time for agency-brokered housing transactions.

To understand how various features and attributes of real estate influence the length of time required to successfully broker housing transactions, data were collected on 4,256 housing transactions brokered by agencies from 2007 to 2011 in New Taipei City, Taiwan. Survival analysis (Fleming and Harrington, 1991; Hosmer, 1999; Lawless, 1981) was then performed to produce a probability model for the transaction waiting period, and a parametric method was used to establish the relationship between the waiting period and housing variables. For investors, predicting the waiting period and the probability of exceeding it provides an important reference for investment decisions and financial planning. Housing agents can also use this information to manage time costs, control management quality, and enhance service efficiency.

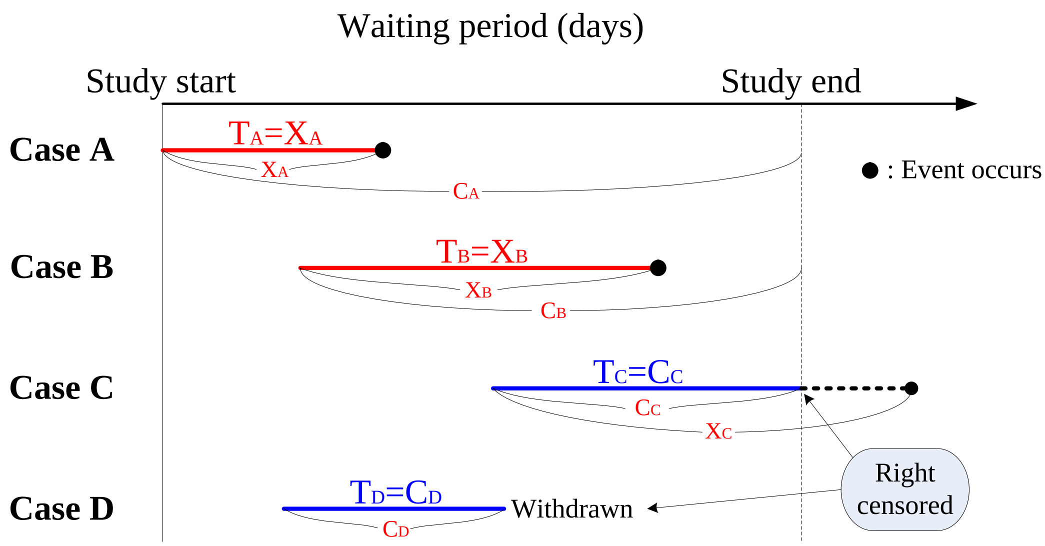

Figure 1. Various types of data collected.

METHODOLOGY

This study used survival analysis to develop a probability model for the waiting period of agency-brokered housing transactions. The parametric method was used to establish the relationship between waiting period and various features and attributes of housing; research inferences were then made based on this relationship. The objective of survival analysis is to analyze the random distribution of events in time (Hosmer, 1999). When conducting survival analysis of temporal data, we must first define the event and event time. This study defined the conclusion of a housing transaction as the event. The length of time from the date the case was commissioned to the date of successful transaction (known as the waiting period) was defined as event time, and counted in days.

The observed values of event time are related to the definition of the event, the starting point of observation (date of commission) and the ending point of observation (date of transaction). Of course, the waiting period of housing transactions is always closely associated with the state of the real estate market, and the market economy fluctuates over time. In the process of statistical analysis and inference, a large sample number indicates more reliable results. The commonly-used SAS package software (Der and Everitt, 2002) was used for all statistical analysis in this study.

Data on waiting period and housing attributes

The collected research data may include incomplete data. In statistics, such incomplete data is known as right-censored. As the main objective of survival analysis is to observe event time, right-censored data can also be called right-censored time. This means that the starting point of observation is known, but at the end point of observation, the defined event has either not occurred or it is impossible to ascertain when the event occurred. Figure 1 shows the various types of data collected. As shown in Figure 1, the waiting times TA and TB are not censored; the waiting times TC and TD are censored because we can say only that the waiting times for the cases C and D are at least TC and TD, respectively.

In addition to creating a probability model for the waiting period, this study also used multiple regressions to establish the relationship between waiting period (response variable) and housing attributes (explanatory variables). The purpose was to determine the significance of the marginal effects of these factors on transaction waiting period. The housing variables considered in this study include internal variables and external variables. Internal variables include building age, floor area of the house, building height (low-rise (1 to 3 stories), mid-rise (4 to 7 stories), or high-rise (8+ stories)), number of rooms, type of construction material (reinforced concrete/steel/steel reinforced concrete), apartment house type (apartment houses with and without elevators). External variables include proximity to schools, MRT stations or bus stops, parks, and markets. The definition of "in proximity" was within 500 m. For example, if a house was designated in proximity to a park, this indicated that the distance between the house and the park was within 500 m.

The housing variables used in this study include both quantitative and qualitative variables. Using regression analysis, qualitative variables were converted into dummy variables. The number of dummy variables was equal to the number of variants minus 1 (Kleinbaum et al., 2007). For example, the variable "building height" had three variants: high-rise, medium-rise, and low-rise buildings. Thus, this variable was converted into two dummy variables. The variable "proximity to parks" had two variants (in proximity/not in proximity) and was therefore converted into one dummy variable.

The incorporation of censored data is essential to survival analysis, as it still includes useful information. To enhance the comprehensiveness of data analysis, this study integrated censored time data into the survival analysis, thereby considering all observations. In other words, the likelihood of both complete data and right-censored data was considered when estimating the likelihood function.

Likelihood function of waiting period

Consider the waiting period *X* and censored time *C* are two non-negative continuous random variables. *X* and *C* are mutually independent and defined as $[0,\infty)$.

The variable X has the probability density function $f(x)$ and distribution function $F(x)$; the variable c has the probability density function $g(x)$ and distribution function $G(x)$. This study considered *n* samples (x_1, c_1),....., (x_n, c_n). Under random censoring, (T_i, δ_i), $i=1,2,.....,n$, were observed. T_i

and δ_i can be expressed as follows:

$$T_i = \min.(X_i, C_i) = \begin{cases} X_i, & if \quad X_i \le C_i \\ C_i, & if \quad X_i > C_i \end{cases} \tag{1}$$

$$\delta_i = I[X_i \le C_i] = \begin{cases} 1, & if \quad X_i \le C_i \\ 0, & if \quad X_i > C_i \end{cases} \tag{2}$$

where $I[.]$ is the indicator function and δ_i is used to express the state of T_i data. As $\delta_i = 1$, this shows that T_i is waiting period; as $\delta_i = 0$, T_i is censored time.

Set (T_i, δ_i), $i = 1, 2, \ldots, n$, as *n* number of independently and identically distributed observations. The likelihood function of transaction waiting period can be expressed as follows (Dabrowska and Doksum, 1988):

$$L = \prod_{i=1}^{n} [f(T_i)\bar{G}(T_i)]^{\delta_i} [g(T_i)\bar{F}(T_i)]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} [f(T_i)^{\delta_i} \bar{F}(T_i)^{1-\delta_i}][\bar{G}(T_i)^{\delta_i} g(T_i)^{1-\delta_i}] \tag{3}$$

where

$$\bar{F}(T_i) = 1 - F(T_i) \tag{4}$$

$$\bar{G}(T_i) = 1 - G(T_i) \tag{5}$$

Survival function and hazard function of waiting period

Two primary quantities desired from the perspective of survival analysis are survival rates and hazard rates. The survival function $S(x)$ expresses the probability that the transaction will not be concluded until time x. The hazard function $h(x)$ expresses the rate of occurrence in which the transaction is concluded precisely at time x, given that the waiting period is greater than or equal to x. The survival function can be expressed as follows (Hosmer, 1999):

$$S(x) = \Pr(X > x) = \int_x^{\infty} f(t)dt \tag{6}$$

It is worth noting that $S(0) = 1$, and $S(\infty) = \lim_{x\to\infty} S(x) = 0$ are functions that increase with time and then monotonically decreases to zero.

The relationships among the survival function, probability density function $f(x)$, and distribution function $F(x)$ are as follows:

$$F(x) = \Pr(X \le x) = 1 - S(x) \tag{7}$$

$$f(x) = \frac{dF(x)}{dx} = -\frac{dS(x)}{dx} \tag{8}$$

The hazard function is an important function in survival analysis. The hazard function $h(x)$ is defined as follows:

$$h(x) = \lim_{\Delta x \to 0} \frac{\Pr(x \le T < x + \Delta x | T \ge x)}{\Delta x} \tag{9}$$

The relationships among the hazard function, probability density function $f(x)$, and survival function $S(x)$ are as follows:

$$h(x) = \frac{f(x)}{S(x)} = -\frac{d}{dx}\ln[S(x)] \tag{10}$$

Another important function in survival analysis is the cumulative hazard function $H(x)$, which is defined as follows:

$$H(x) = \int_0^x h(t)dt = -\ln[S(x)] \tag{11}$$

The survival function can be expressed as

$$S(x) = \exp[-H(x)] = \exp[-\int_0^x h(t)dt] \tag{12}$$

The hazard function and the cumulative hazard function include important information related to the probability model of the waiting period. Both functions are particularly useful in the creation of the probability model. Commonly used hazard functions include constant function, incremental function, decreasing function, bathtub-shaped function, and hump function. The parametric method of survival analysis has been widely applied to solve problems in the domains of biomedicine and food science; however, few studies have used survival analysis to predict the waiting period for agency-brokered housing transactions. The following parametric probability distributions are frequently used when survival analysis is applied to problem-solving: exponential distribution, log-normal distribution, and Weibull distribution. Table 1 shows the theoretical relationships between the cumulative hazard functions and waiting period for these three models.

As shown in the table, if the probability distribution of the waiting period is exponential, then there is a linear relationship between cumulative hazard function and waiting period. If the probability distribution is Weibull, there is a linear relationship between the natural logarithm of the cumulative hazard function and the natural logarithm of waiting period. If the probability distribution is log-normal, the relationship between the cumulative hazard function and waiting period is more complex. However, if the cumulative hazard function is appropriately converted, there will be a linear relationship between the Probit function values and the natural logarithm of the waiting period.

Probability distribution of the waiting period

The survival function of the variable of interest can be estimated using the well-known product-limit estimator (Kaplan and Meier, 1958), which provides a nonparametric estimate. In this study, the probability distribution of waiting period was assumed to be a two-parameter Weibull distribution (Weibull, 1951), thus the probability distribution function, hazard function, cumulative hazard function, and survival function of waiting period can be given as follows:

$$f(x) = \alpha\beta x^{\beta-1} \exp[-\alpha x^{\beta}] \tag{13}$$

$$h(x) = \alpha\beta x^{\beta-1} \tag{14}$$

Table 1. Theoretical relationships between the cumulative hazard functions and waiting period.

Probability distribution f (x)	Cumulative hazard function H (X) = -log [S(X)]	Linear relation
Exponential	$H(X) = \lambda X$	$\hat{H}$ versus X
Weibull	$H(X) = \alpha X^{\beta}$	$\ln \hat{H}$ versus $\ln X$
Log-normal	$H(X) = -\ln\left(1 - \Phi\left[\frac{(\ln X - \mu)}{\sigma}\right]\right)$ $\Phi^{-1}(.)$: Probit function	$\Phi^{-1}\left[1 - \exp\left(-\hat{H}\right)\right]$ versus $\ln X$

$$H(x) = \int_0^x h(t)dt = \alpha x^{\beta} \quad (15)$$

$$S(x) = \exp[-H(x)] = \exp[-\alpha x^{\beta}] \quad (16)$$

where α and β are the parameters of Weibull distribution. α is the scale parameter and β is the shape parameter. Equations (15) and (16) can be rewritten as follows:

$$\ln[H(x)] = \beta \ln x + \ln \alpha \quad (17)$$

$$H(x) = -\ln[S(x)] = \alpha x^{\beta} \quad (18)$$

The linear relationship between $\ln[H(x)]$ and $\ln x$ can be used to ascertain whether the assumption that the probability distribution of the waiting period is appropriate. It should be noted that the estimated cumulative hazard function can be obtained once the survival function is estimated (Equation 18).

Parametric regression model of waiting period

In most instances, the waiting period may depend on housing variables. To reduce the bias in the estimation, the effects of these variables on the waiting period must be considered. Parametric models can be used to predict the distribution of the waiting period to an event from a set of housing variables. This study used the well-known Weibull regression model to conduct survival analysis of agency-brokered housing transactions. The model can be expressed as follows:

$$\ln T = C_0 + \sum_{i=1}^{m} C_i Z_i + \varepsilon \quad (19)$$

where Z_i($i = 1, 2, \ldots, m$) are the housing variables, C_i is the regression coefficient that corresponds to Z_i, m is the number of variables, C_0 is the intercept, and ε is the error term. The density function of $\ln T$, given Z_i($i = 1, 2, \ldots, m$), has an extreme value distribution.

Estimation of parameters

All parameters in the parametric model can be estimated by the maximum likelihood method. Assume that there are p parameters ($\theta_1, \theta_2, .., \theta_p$) to be considered in the model. Let $\Theta = (\theta_1, \theta_2, .., \theta_p)$ be denoted as the parameter vector. Considering a group of n randomly selected samples $(X_1, X_2, .., X_n)$, the likelihood function $L(\Theta)$ can be expressed as:

$$L(\Theta) = \prod_{i=1}^{n} f(X_i, \Theta) \quad (20)$$

The maximum likelihood estimates of the parameters must satisfy Equation 21, and the Hessian for parameter vectors must be negative-definite:

$$\frac{\partial}{\partial \theta_j} L(\Theta) = 0 \qquad j = 1, 2, ..., p \quad (21)$$

Numerical methods are often used to obtain the maximum likelihood estimates of parameters because a closed-form solution cannot be obtained for the majority of these parameter estimates. Maximum likelihood estimates can be determined through iterations using the Newton-Raphson algorithm. The main advantage of this method is that in addition to the maximum likelihood estimates for parameters, this method simultaneously solves the Fisher information matrix. The inverse of the Fisher information matrix is the covariance matrix of estimators. This matrix is particularly useful when making inferences about parameters, including the processes of deriving confidence intervals and conducting hypothesis testing (Hosmer, 1999).

Evaluation of goodness of fit

The most intuitive approach involves observing the model graphically (Arjas, 1988). This study examined the relationship between waiting period and the estimated value of the cumulative hazard function, to determine whether it matched the theoretical relationship expressed in Equation (17). In addition, the significance of each parameter of the regression model was also tested. The purpose was to evaluate the goodness-of-fit of the hypothesized model.

The parameter estimates and their estimated covariance matrix are available in an output SAS data set and can be used to construct additional tests or confidence intervals for the parameters (Andersen, 1982; Schoenfeld, 1980). Alternatively, tests of parameters can be based on likelihood ratios (Vuong, 1989). It is believed that likelihood ratio tests are generally more reliable in small samples than tests based on the information matrix (Cox and Oakes, 1984). The statistic of likelihood ratio test follows chi-square distribution with *p* degrees of freedom.

Analysis process

The analytical procedure is briefly summarized as a step-by-step format as follows:

i) Collect data on housing attributes and the waiting period for agency-brokered housing transactions;
ii) Identify housing attributes or variables that influence waiting period;
iii) Test the correlations between waiting period and all housing variables; variables with low correlations were eliminated from the regression model. SAS software provides two test methods: the log-rank test and the Wilcoxon test.
iv) Hypothesize that the probability distribution of waiting period was a specific type of parametric probability distribution: Weibull distribution;
v) Conduct survival analysis of the data collected in Step 1;
vi) Estimate regression coefficients and establish the relationships between transaction waiting period and housing attributes;
vii. Evaluate the model's goodness of fit.

By following the steps described, we can select parametric models with a better fit, which can be used as the basis for further statistical inference. Two main procedures executed in SAS software were PROC LIFETEST and PROC LIFEREG.

EMPIRICAL ANALYSIS

This study used survival analysis to create a probability model for the waiting period for agency-brokered housing transactions. A parametric survival method was used to establish the relationship between the waiting period and the factors associated with housing. Housing agencies provided the data used in this study, covering 4,256 agency-brokered housing transactions, from 2007 to 2011 in New Taipei City. The 4,256 housing cases collected for this study included 2,307 transactions that had been concluded and 1,949 cases that had not been concluded. Due to the limitations of the research deadline and the fact that the contracts for some of these cases were terminated prior to completion, we were unable to ascertain the waiting period for 1,949 cases. As a result, the research data collected may have included right-censored data.

This study considered both internal and external variables. Internal variables refer to housing attributes, while external variables refer to the attributes of the surrounding environment. Both quantitative and qualitative variables were included. Internal variables include building age (AGE), floor area of the house (AREA), building height (HEI), number of rooms (ROOM), type of construction material (MAT), apartment house type (TYP). External variables include proximity to schools (SCH), MRT stations or bus stops (STAT), parks (PAR), and markets (MAR). Descriptive analysis was performed on these variables.

The results show that average building age was 16.5 years (standard deviation=11.2 years); the relative frequency distribution for building age is shown in Figure 2a. The average floor area of housing was 100.94 M^2 (standard deviation=42.88 M^2); the relative frequency distribution for floor area is shown in Figure 2b. As shown in Figure 2b, the histogram highly skewed to the right. A few very high areas create the skewness in the right tail. There are no areas above 400 M^2; most of the areas are often below 150 M^2. Table 2 shows the descriptive statistics for the variables. Approximately 53% of buildings were high-rise buildings, 46% were mid-rise buildings, and 1% were low-rise buildings. A high proportion (51 %) of houses had three rooms. 64 % of apartment houses had elevators and 36 % were lower-rise apartment houses without elevators. Reinforced concrete was the most common construction material (approximately 96%), followed by steel reinforced concrete (3%). Houses built with steel structures accounted for only 1% of housing cases in research data. 70% of houses were near schools, 62% near MRT stations or bus stops, 55% near parks, and 71% near markets.

Figure 3 illustrates the survival curve for agency-brokered housing transactions. This curve was estimated using the Product-Limit method. The x-axis in Figure 3 indicates the waiting period, and the y-axis indicates the exceedance probability of the waiting period. The figure shows that from 2007 to 2011, the waiting period for 75% of agency-brokered housing transactions in New Taipei City exceeded 21 days; 50% exceeded 101 days, and25% exceeded 272 days. This study used the Wilcoxon and Log-Rank methods to test the correlations between housing attributes and waiting period. The results showed that the correlation between waiting period and AREA was not significant. Therefore, AREA was eliminated from the regression model. The qualitative variables SCH, STAT, PAR, MAR, HEI, MAT, and TYP were all converted to dummy variables during regression analysis. The relationships between transaction waiting period and housing attributes can be expressed as follows:

$$\ln T = C_0 + C_1 * \mathrm{AGE} + C_2 * \mathrm{BTYP1} + C_3 * \mathrm{BTYP2} + C_4 * \mathrm{ROOM} + C_5 * \mathrm{SCH} + C_6 * \mathrm{STOP} + C_7 * \mathrm{PAR} + C_8 * \mathrm{MAR} + C_9 * \mathrm{MAT1} + C_{10} * \mathrm{MAT2} + C_{11} * \mathrm{TYP} \quad (22)$$

or

$$T = \exp\{C_0 + C_1 * \mathrm{AGE} + C_2 * \mathrm{BTYP1} + C_3 * \mathrm{BTYP2} + C_4 * \mathrm{ROOM} + C_5 * \mathrm{SCH} + C_6 * \mathrm{STOP} + C_7 * \mathrm{PAR} + C_8 * \mathrm{MAR} + C_9 * \mathrm{MAT1} + C_{10} * \mathrm{MAT2} + C_{11} * \mathrm{TYP}\} \quad (23)$$

The variable "HEI" included three possibilities: high-rise, mid-rise, and low-rise buildings. Therefore, two dummy

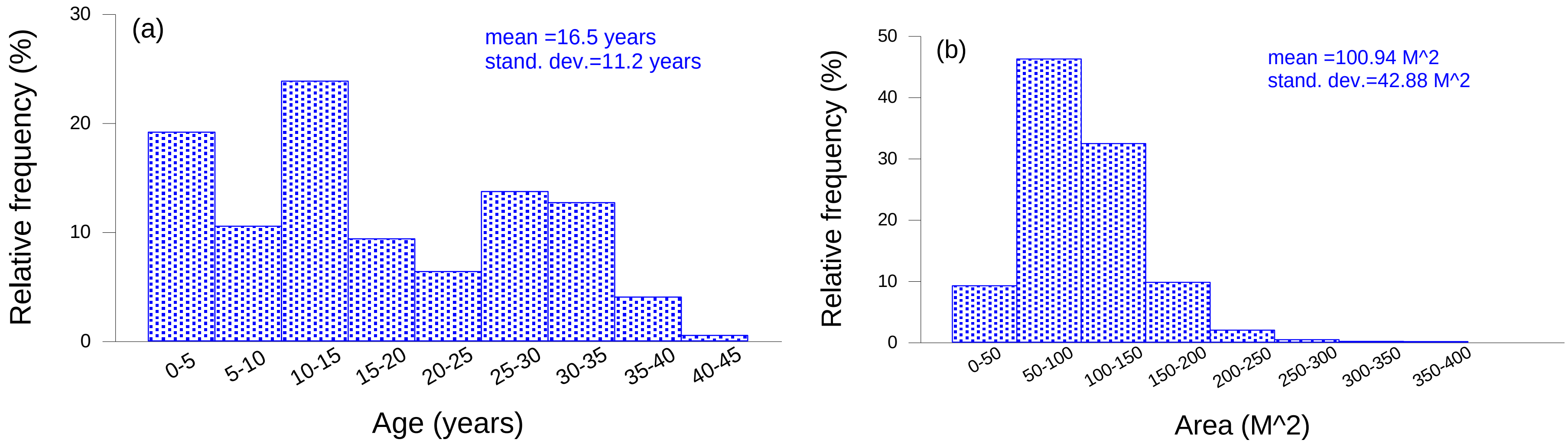

Figure 2. Histogram of housing variables: (a) age and (b) area (sample size =4256).

variables, BTYP1 and BTYP2, were established for this variable. When BTYP1=0 and BTYP2=0, this indicates cases in which the buildings are low-rise. When BTYP1=1 and BTYP2=0, this indicates that the buildings are mid-rise. When BTYP1=0 and BTYP2=1, this indicates that the buildings are high-rise.

If SCH=1, the house in question is within a distance of 500 M from a school. If SCH=0, the house does not satisfy the condition of "in proximity to a school". Likewise, when STOP, PAR, and MAR equal 1, this indicates that the house in question is near MRT stations or bus stops, parks, and markets, respectively. Conversely, when STOP, PAR, and MAR equal 0, the house does not satisfy the conditions of "in proximity to MRT stations or bus stops /parks/markets."

The variable "MAT" included three possibilities: steel, reinforced concrete, and steel reinforced concrete. Two dummy variables, MAT1 and MAT2 were established for this variable. If MAT1=0 and MAT2=0, the construction material is reinforced concrete. If MAT1=1 and MAT2=0, the construction material is steel. If MAT1=0 and MAT2=1, the construction material is steel reinforced concrete. The variable "TYP" represented apartment houses with or without elevators. When TYP=1, this indicated that the apartment house is equipped with an elevator.

Figure 4 illustrates the relational curve between the natural logarithm of the estimated cumulative hazard function, $\ln \hat{H}(x)$, and the natural logarithm of the waiting period, $\ln x$, based on the data from the 4,256 housing cases. The figure shows that the relationship is nearly linear, which complies with the theoretical relationship expressed in Equation 17. This demonstrated the validity of the Weibull distribution model.

Table 3 shows the results of regression analysis. The first column lists all the variables that were included in the regression model. The second column shows the coefficients of the variables, while the third column lists the estimated coefficient values. The chi-square test results in the table show that 6 of the 12 regression coefficients were significant at the 5% level. The coefficient of the variable AGE was C_1=-0.027, meaning that with each year a building increased in age, the waiting period for housing transactions is reduced by $\exp(-0.027)=0.973$ times. This indicates that houses in older buildings have a shorter transaction waiting period.

The waiting period for mid-rise buildings was $\exp(-0.307)=0.736$ times that of low-rise buildings. The waiting period for high-rise buildings was $\exp(-0.316)=0.729$ times that of low-rise buildings. This indicates that higher-rise buildings correspond to a shorter waiting period.

With each unit increase in the number of rooms, waiting period increased by $\exp(-0.1716)=0.842$ times, indicating that

Table 2. Descriptive statistics for the housing variables.

Variable	Description	Outcome	Frequency	Relative frequency (%)
BTYP	Building type	Low-rise building (1~3 stories)	38	1
		Mid-rise building (4~7 stories)	1948	46
		High-rise building (8+ stories)	2270	53
ROOM	Number of room	One room	655	15
		Two rooms	812	19
		Three rooms	2170	51
		Four rooms	619	15
MAT	Construction material	Reinforced concrete	4099	96
		Steel	39	1
		Steel Reinforced concrete	118	3
TYP	Property type	With elevator	2716	64
		Without elevator	1540	36
SCH	Proximity to schools?	Yes	3002	71
		No	1254	29
STOP	Proximity to MRT stations or bus stops?	Yes	2636	62
		No	1620	38
PAR	Proximity to parks?	Yes	2362	55
		No	1894	45
MAR	Proximity to markets?	Yes	3032	71
		No	1224	29

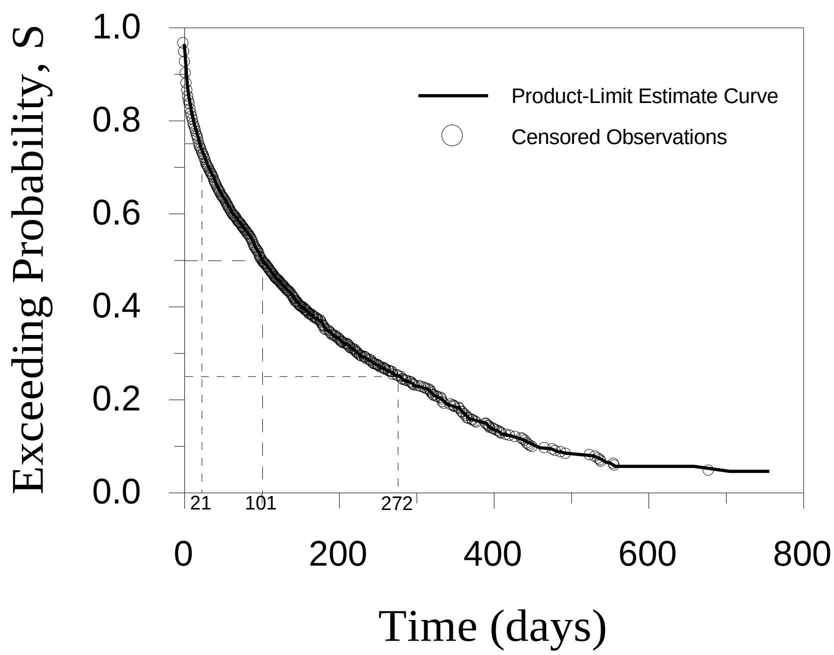

Figure 3. Exceedance probability of waiting period.

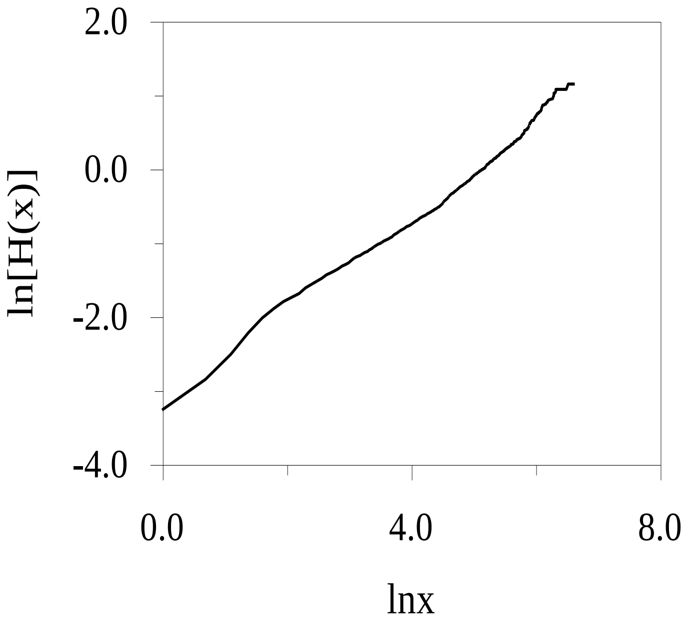

Figure 4. Relationship between the natural logarithm of cumulative hazard function and the natural logarithm of waiting period.

Table 3. Regression variables and the estimated values of regression coefficients.

Variable	Coefficient	Estimate
Intercept	C_0	5.604* (221.90)
AGE	C_1	-0.027* (44.19)
BTYP1	C_2	-0.307 (0.77)
BTYP2	C_3	-0.316 (0.80)
ROOM	C_4	-0.172* (22.31)
MAT1	C_5	-0.153 (0.18)
MAT2	C_6	0.215 (1.01)
TYP	C_7	-0.013 (0.01)
SCH	C_8	0.082 (0.83)
STOP	C_9	-0.165** (4.05)
PAR	C_{10}	0.773* (106.09)
MAR	C_{11}	0.493*(28.66)

*, ** indicate significance at the 0.1 and 5% level, respectively. The figures in parentheses are the values of chi-square statistic.

houses with a higher number of rooms have a shorter transaction waiting period. The waiting period for houses built from steel was $\exp(-0.153)=0.858$ times that of houses built from reinforced concrete. The waiting period for houses built from steel reinforced concrete was $\exp(0.215)=1.24$ times that of houses built from reinforced concrete. In other words, houses built from steel demonstrated the shortest waiting period, followed by houses built from reinforced concrete. Residences constructed from steel reinforced concrete showed the longest waiting period. If houses were located in apartment buildings with elevators, the waiting period was reduced by $\exp(-0.013)=0.987$ times.

As shown in Table 3, the waiting period for houses near schools was $\exp(0.082)=1.085$ times that of houses not in proximity to schools. Thus, agency-brokered housing transactions in school neighborhoods require a longer waiting period. The waiting period for houses close to MRT stations or bus stops was reduced by $\exp(-0.165)=0.848$ times. The waiting period for housing close to parks was $\exp(0.773)=2.166$ times that of housing not in proximity to parks. Thus, the purchase or sale of housing near parks requires a longer waiting period. The waiting period for housing near markets was $\exp(0.493)=1.637$ times that of housing not in proximity to markets. This indicates that brokering transactions of housing near markets requires a longer waiting period.

Conclusion

Based on survival analysis, this study proposed an analysis framework and procedures to predict the waiting period for agency-brokered housing transactions. Empirical results show that the transaction waiting period follows Weibull distribution, and housing features and attributes significantly influence the waiting period for sellers. The results also show that housing with a higher number of rooms in older buildings has a shorter waiting period. Housing in higher-rise buildings indicates a longer waiting period, while housing in apartment buildings with elevators corresponds to a shorter waiting period. With regard to construction materials, the waiting period for houses built with a steel structure was the shortest, followed by housing constructed from reinforced concrete. The purchase or sale of housing constructed from steel reinforced concrete requires the longest waiting period. Housing in proximity to schools, parks, and markets has a longer waiting period, while housing in proximity to MRT stations or bus stops indicates a shorter waiting period.

REFERENCES

Andersen PK (1982). Testing goodness-of-fit of Cox's regression and life model. Biometrics, 38: 67-77.

Arjas E (1988). A graphical method for assessing goodness of fit in Cox's proportional hazards model. J. Am. Stat. Assoc., 83(401): 204-212.

Cox DR, Oakes D (1984). Analysis of survival data. Chapman and Hall, London.

Dabrowska DM, Doksum KA (1988). Partial likelihood in transformation models with censored data. Scand. J. Stat., 15: 1-23.

Der G, Everitt BS (2002). A handbook of statistical analyses using SAS (2nd ed.). Boca Raton, Chapman and Hall/CRC, Fla.

Evan AW (1973). The economics of residential location. Macmillan, London.

Fleming TR, Harrington DP (1991). Counting processes and survival analysis. Wiley, New York.

Giannias DA (1998). A quality of life based ranking of Canadian cities. Urban Stud., 35(12): 2241-2251.

Ha M, Weber MJ (1994). Residential quality and satisfaction: Toward developing residential quality indexes. Home Econ. Res. J., 22(3): 296-308.

Hosmer DW (1999). Applied survival analysis, Wiley series in probability and statistics. Wiley, New York.

Kaplan EL, Meier P (1958). Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc., 53: 457–481.

Kleinbaum DG, Kupper LL, Nizam A, Muller KE (2007). Applied regression analysis and other multivariable methods (4th ed.). Wadsworth, Belmont, Thomson, CA.

Lancaster KJ (1966). A new approach to consumer theory. J. Polit. Econ., 74(2): 132-157.

Laurice J, Bhattacharya R (2005). Prediction performance of a hedonic pricing model for housing. Appraisal J., 73(2): 198-209.

Lawless JF (1981). Statistical models and methods for lifetime data. Wiley series in probability and statistics, Wiley, New York.

Schoenfeld D (1980). Chi-squared goodness-of-fit tests for the proportional hazards regression model. Biometrika, 67(1): 145-153.

Singell LD, Lillydahl JH (1990). An empirical examination of the effect of impact fees on the housing market. Land Econom., 66(1): 88-92.

Thalmann P (1999). Identifying households which need housing assistance. Urban Stud., 36(11): 1933-1947.

Vuong QH (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, 57(2): 307-333.

Weibull W (1951). A statistical distribution function of wide applicability. J. Appl. Mech. Trans. ASME., 18(3): 293–297.