

Modeling home property listings' time-on-market duration and listing outcome using copula-based competing risk method

Yicong Liu (corresponding author)
University of Toronto
nora.liu@mail.utoronto.ca

Saeed Shakib
University of Toronto
saeed.shakib@mail.utoronto.ca

Eric J. Miller
University of Toronto
eric.miller@utoronto.ca

Khandker Nurul Habib
University of Toronto
khandker.nurulhabib@utoronto.ca

Abstract: Modeling housing market dynamics is an important component of land use and transport interaction (LUTI) models, particularly for microsimulation models and how they handle the market clearance mechanism. However, most of these models include key assumptions not derived or validated through empirical testing, such as when and what action a seller would take if a property could not be sold within an expected time. However, these are key decision elements of the housing market clearance process. To fill this research gap, this study uses real estate sale listing data to investigate the factors influencing a property listing's time-on-market (TOM) duration, listing outcome, and correlation. A copula-based structure is developed to jointly estimate the TOM and listing outcome through a competing hazard duration model and a nested logit model. The results show statistically significant and positive correlations between the TOM of terminated listings and termination choices (i.e., whether the terminated listing will be withdrawn from the market, converted to a lease, or re-listed as a sale). This implies that the unobserved factors that may increase a seller's probability of terminating a listing would decrease its TOM duration until the termination. It is also found that an increase in the asking price of a property listing can significantly increase its TOM duration and probability of being terminated. The copula-based joint model can be integrated into a LUTI microsimulation framework to parameterize the maximum TOM duration of each simulated property for sale in the housing market, improving its market-clearing process to represent real-world behavior better.

Keywords: land-use and transportation interaction (LUTI) model, housing market, copula-based joint model

Article history:

Received: October 28, 2023

Received in revised form:
May 29, 2024

Accepted: June 18, 2024

Available online: July 29,
2024

1 Introduction

Residential relocation decisions and housing market dynamics have been critical components of land-use and transport interaction (LUTI) models. A two-way interaction is believed to exist between the transportation and land-use systems through accessibility,

Copyright 2024 Yicong Liu, Saeed Shakib, Eric J. Miller & Khandker Nurul Habib

<https://doi.org/10.5198/jtlu.2024.2447>

ISSN: 1938-7849 | Licensed under the [Creative Commons Attribution – Noncommercial License 4.0](#)

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

which can influence households' travel behavior and their choices of where and when to relocate. In addition, the dynamics of the housing market and the interaction between housing buyers and sellers can also significantly impact a household's residential location choice. In general, the housing market module of a LUTI framework models various aspects of households' relocation process, from mobility decisions and search processes to relocation choices and clearing the housing market (Habib, 2009).

Several approaches have been developed to model residential location choices and market-clearing with the LUTI framework, including the microsimulation approach (Farooq & Miller, 2012). This approach attempts to mimic the bidding process between housing buyers and sellers and clear the market under certain assumptions. It is often considered a more realistic representation of housing market behavior, as it simulates the interactions between buyers and sellers (Miller, 2018). However, most studies utilizing the microsimulation approach primarily focus on modelling the households' residential location choices or their willingness to pay through econometric models. In contrast, other key assumptions of the housing market-clearing process are not studied empirically due to the lack of data.

Such key assumptions include how long a residential property listing would stay active in the market before it is sold or withdrawn and the seller's decision of what to do if the property cannot be sold within a certain time. For LUTI microsimulation models, they can impact the results of the housing market-clearing process. However, the determination of such assumptions is often not explained or justified in the existing literature. It is suspected that they are assumed purely based on the modelers' judgment about the market-clearing behavior. This study aims to fill this gap by exploring the sellers' behavior on when and what actions to take if a residential property listing cannot be sold within an expected timeframe. The findings of the study show that there are significant differences between the sellers' behavior and the oversimplified assumptions used in existing microsimulation models. This study can show how to represent real-world housing market dynamics more realistically through microsimulation models.

This study analyzes the time-on-market (TOM) durations of residential property listings and their outcomes in the Greater Toronto Area (GTA) in 2021. Specifically, the focus is solely on property listings intended for sale rather than lease. The time-on-market duration is the days between the listing date and the sale or termination date of a residential property (Benefield & Hardin, 2015). The listing date is when the residential property is listed on the market for potential buyers to view and bid for. The sold or termination date is when the listing becomes inactive in the market, given that it is sold or withdrawn. This study considers four possible outcomes for each property sale listing depending on whether the property is sold. The first possible outcome is selling the property and becoming inactive in the market.

However, suppose the property is not sold and the seller terminates the current listing. In that case, there are three potential choices: 1) withdraw the property from the market and do not re-list, 2) re-list the property as a lease, and 3) re-list the property for sale. This study utilizes multinomial logit (MNL) and nested logit (NL) models to investigate the factors that affect sellers' choices regarding the four possible outcomes of a residential property listing. In addition, a hazard duration model is developed to model the listings' TOM durations and understand the influential factors.

Since two competing events may result in a property listing becoming inactive in the market (i.e., sold or terminated), a competing hazard approach is employed to account for the interdependence between the hazard of the competing outcomes. Last, a copula-based modelling approach describes the dependence structure between the discrete choice model for listing outcomes and the hazard duration model for listings' TOM.

Comparisons between the copula-based joint model and the independent models are made to reveal any potential unobserved correlations between the joint outcomes.

The remainder of the paper is organized as follows. Section 2 reviews the literature on modelling TOM and outcomes of residential property listings, and the use of copula-based models in the transportation field. Section 3 presents the methodological framework of the proposed model. Section 4 describes the dataset utilized in this study. Section 5 presents the results of the empirical models. Section 6 discusses the implications of the study, followed by Section 7 which concludes the study and provides recommendations for future research.

2 Literature review

This literature review section provides an overview of the existing literature on three topics that are relevant to this study: the modelling of residential property listings' TOM and probability of sale, the application of copula-based joint modelling in transportation research, and the housing market-clearing mechanisms in LUTI models.

2.1 Modelling residential property listings' time-on-market and probability of sale

Although modelling residential property listing's TOM and probability of sale has not received much attention from studies on LUTI models, it has been a popular topic in real estate research. Hazard models have always been applied to estimate the time until properties are sold in the research field of real estate economies. Using housing market data from New Orleans, Das (2007) developed a Weibull hazard model to investigate the relationship between uncertainty in the sale and TOM duration of a property. It was concluded that vacant housing units had a positive and higher rate of duration dependence than occupied housing units (Das, 2007). Similarly, Allen et al. (2018) utilized a hazard model in accelerated failure-time form to examine the effect of Multiple Listing Services (MLS) information-sharing intensity on the TOM durations of housing listings.

It was discovered that several variables were negatively related to TOM, including MLS information sharing, number of bedrooms, number of garage parking spaces, age of the housing unit, etc. Bian et al. (2016) applied parametric and semi-parametric hazard models to estimate residential properties' TOM. They found that factors that increase brokers' search and bargaining costs positively affect TOM durations. A thorough review of techniques utilized to model properties' TOM in real estate research can be found in Benefield et al. (2014).

Several studies that investigated factors influencing TOM also examined the effects of such factors on the properties' probability of sale (Allen et al., 2015; Allen et al. 2018; Bian et al., 2016). Most real estate studies only consider binary outcomes for a property transaction – sold or not sold during the marketing period. Therefore, binary logit and probit models are commonly used by such studies to examine the determinations of the probability of sale. Bian et al. (2016) and Allen et al. (2015) estimated both binary logit and probit models for property probability of sale and concluded that both techniques would reveal similar implications. Allen et al. (2015) discovered that virtual tours and photographs increase the probability of sale, whereas public open houses may decrease such probability. Allen et al. (2018) applied a binary probit model and found that a housing unit's age, size, and price can negatively impact its probability of sale.

In contrast, the number of garage parking spaces can positively impact such probability. In addition to using the traditional methods to model property listings' TOM

and probability of sale, some studies also attempted to develop new techniques. Caudill et al. (2022) formulated a generalized geometric hazard model that can estimate the marginal effects of explanatory variables on TOM and the probability of sale from one regression model.

2.2 Copula-based joint modelling in transportation research

Over the last decade, copula distributions have been utilized in transportation research for various topics. A copula is defined as a multivariate joint distribution of random variables that can be obtained from the marginal distributions of each random variable (Bhat & Eluru, 2009). It is a popular method in the transportation field to be applied for joint discrete-continuous models. Bhat and Eluru (2009) provide overviews of commonly used copula structures and applied bivariate copula models to analyze the dependency structure between residential neighborhood choice and daily household vehicle miles of travel (VMT). Spissu et al. (2009) developed a copula-based joint multinomial discrete-continuous model to examine the correlation between vehicle type choice and VMT. In contrast, Golshani et al. (2018) applied a similar modelling structure to investigate the correlation between travel mode and departure time decisions. Ozonder and Miller (2021) developed a copula model of work episode start time and duration.

2.3 Housing market-clearing mechanisms in LUTI models

The housing market-clearing mechanism constitutes a substantial component of a microsimulation LUTI model. Prior to the market-clearing process, the buyers' residential location choices and their willingness to pay for the chosen dwellings are generated through a location choice model, while the dwellings for sale are also determined through a housing supply model. Subsequently, the market-clearing mechanism simulates the matching process between active buyers and dwellings in the market. There are variations between the mechanisms applied in different LUTI models. For example, the UrbanSim model posits buyers as price-takers, and they would be matched with the dwelling providing the highest utility. In case the dwelling has been occupied, then the buyer would be forced to choose the dwelling with the second highest utility (Waddell, 2010). Conversely, the ILUTE model assumes that buyers and sellers are non-cooperative agents aiming to maximize utilities and profits. In its market-clearing mechanism, a dwelling for sale would be randomly selected and then assigned to its highest bidder, providing the bid exceeds the seller's minimum acceptable price (Rosenfield et al., 2013). A similar approach is employed by the SimMobility model, wherein dwellings for sale are also allocated to their highest bidders (Zhu et al., 2018). A more comprehensive review of the housing market-clearing mechanisms in different LUTI models can be found in Liu et al. (2023).

Most microsimulation LUTI models adopt a dynamic disequilibrium approach, whereby housing supply and demand may not be equal. Instances may arise where dwellings for sale cannot be matched with any potential buyers in the current simulation cycle. How the LUTI models address such unsold dwellings in their market-clearing processes appears ambiguous in the literature. For instance, the UrbanSim model assumes that the market is cleared after all the buyers are matched with a dwelling for sale, without providing a clarification on the unmatched dwellings (Waddell, 2010). However, it does utilize the current vacancy rate to adjust the land prices. Similarly, the ILUTE model does not delineate a specific protocol for handling unsold dwellings (Rosenfield et al., 2013). It is presumed that in these models, the dwellings for sale would remain active

in the market until sold, lacking the option for withdrawal prior to a sale completion. Conversely, the SimMobility model, which conducts daily market clearance simulations, assumes that any dwelling remaining unsold after a period of 210 days is withdrawn from the market (Zhu et al., 2018).

The above review of LUTI models suggests that there is a research gap in the housing market-clearing process, particularly regarding the sellers' behavior on terminating listings that are not sold within the expected timeframes. This study addresses this gap by investigating the determinations and correlations between the TOM and possible outcomes of a residential property listing. Additionally, it explores the potential integration of the proposed model into a LUTI framework to provide a market exit option to residential property sellers. Although relevant studies on TOM and probability of sale exist in real estate research, none can be directly applied in a LUTI framework. These studies mostly focus on the impacts of brokers' behavior and marketing strategies on the probability of sale, which are not the focus of LUTI models. Instead of factors related to real estate marketing strategies and brokers' behavior, this study examines the effects of transportation accessibility on listings' TOM and outcomes. To the best of our knowledge, this is the first study that applied competing risk and copula-based modelling structures to investigate the differences in determinations behind sold and terminated property listings and the correlations between TOM and sellers' actions regarding terminated listings. The findings of this study can provide empirical evidence for several key assumptions in the housing market-clearing process of LUTI models.

3 Data description

The real estate listing data used in this study were collected from the Toronto Regional Real Estate Board (TRREB). These are a sample of residential property listings in the GTA in the year 2021. Only properties initially listed for sale in 2021 are of interest to this study and are included in the dataset. The dataset is not in a panel format because this study only focuses on the outcome of the initial listing of each property for sale. The dataset contains 18,272 residential properties initially listed for sale, among which 50% were successfully sold in the initial listing. For the properties whose initial listings were terminated before a successful sale, there are three possible subsequent actions that the sellers can choose from, resulting in a total of four possible outcomes for the initial listing of each property:

- The property is sold in the initial listing (herein referred to as "sold").
- The initial listing is terminated, and the property is withdrawn from the market (herein referred to as "terminate & withdraw").
- The initial listing is terminated, and the property is re-listed as a lease (herein referred to as "terminate & lease").
- The initial listing is terminated, and the property is re-listed as a sale (herein referred to as "terminate & resale").

The distribution of percentages of properties with the four different outcomes is presented in Figure 1. It is observed that for properties that were not sold in the initial listings, most sellers chose to re-list the property and attempt to sell again, whereas very few of them chose to lease out the property instead.

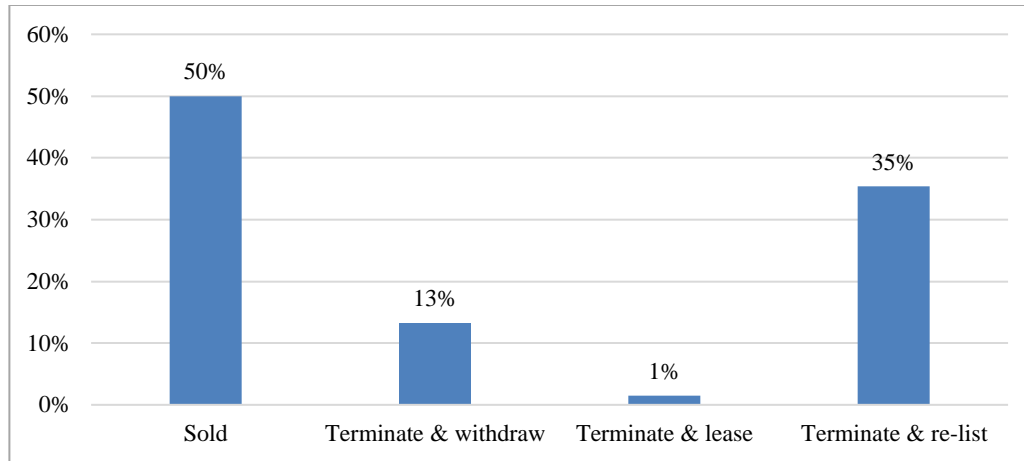


Figure 1. Distributions of the outcomes of a residential property's initial listing

Moreover, the relative frequencies of the listings' TOM durations are calculated at an interval of five days and plotted by listing outcomes, as shown in Figure 2. It is observed that the distributions of TOM durations are quite different across the four listing outcomes. Generally, the on-market days of the properties sold in the initial listings are smaller than those terminated. Over 40% of the sold listings were on-market for less than five days. In contrast, most terminated listings were on the market for at least ten to fifteen days. It is intuitive to observe these differences because, for the unsold properties, it is possible that their initial TOM durations exceeded the sellers' expectations, leading to their decision to terminate the initial listings and adjust for subsequent actions.

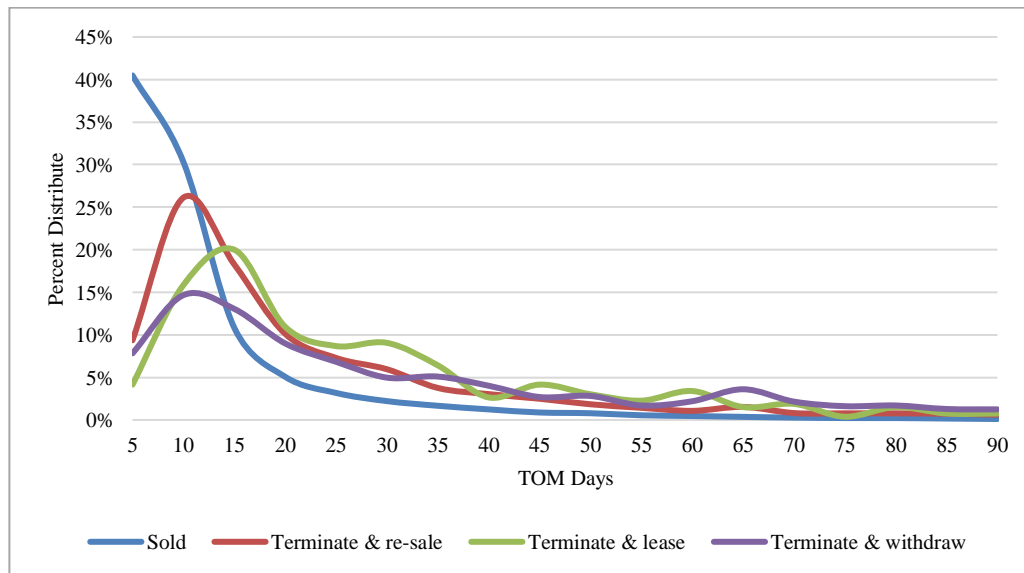


Figure 2. Relative frequencies of TOM durations by listing outcomes

Variables that can be obtained from the listing data are dwelling attributes, including property location, dwelling type, listing price, dwelling area, lot size, number of bedrooms, washrooms, parking spots, etc. Figure 3 presents the distributions of some key dwelling attributes of the dataset. The distributions suggest a higher representation of

detached houses and a lower representation of dwellings in Toronto. This is because the samples are collected to cover all DAs in the GTA instead of stratifying based on population density. Since most of the condo dwellings are in Toronto, it is not surprising to observe such distributions.

Since the listing data does not provide socio-economic information about the sellers and transportation accessibility measurements of the dwellings, additional sources are utilized to derive such information. The 2021 Canadian Census data are used to generate key socio-economic variables of the dissemination area (DA) where the dwelling is located, such as median household income, average age of the residents, and percentage of movers and non-movers in the DA. In addition, the number of subway stations, bus rapid transit (BRT) stations, and streetcar stops within a 500-meter radius of each listed property is computed and used as isochrones-based accessibility measures for public transit. Figure 4 presents a map of five regional municipalities in the study area, along with the subway routes, BRT routes, and streetcar routes.

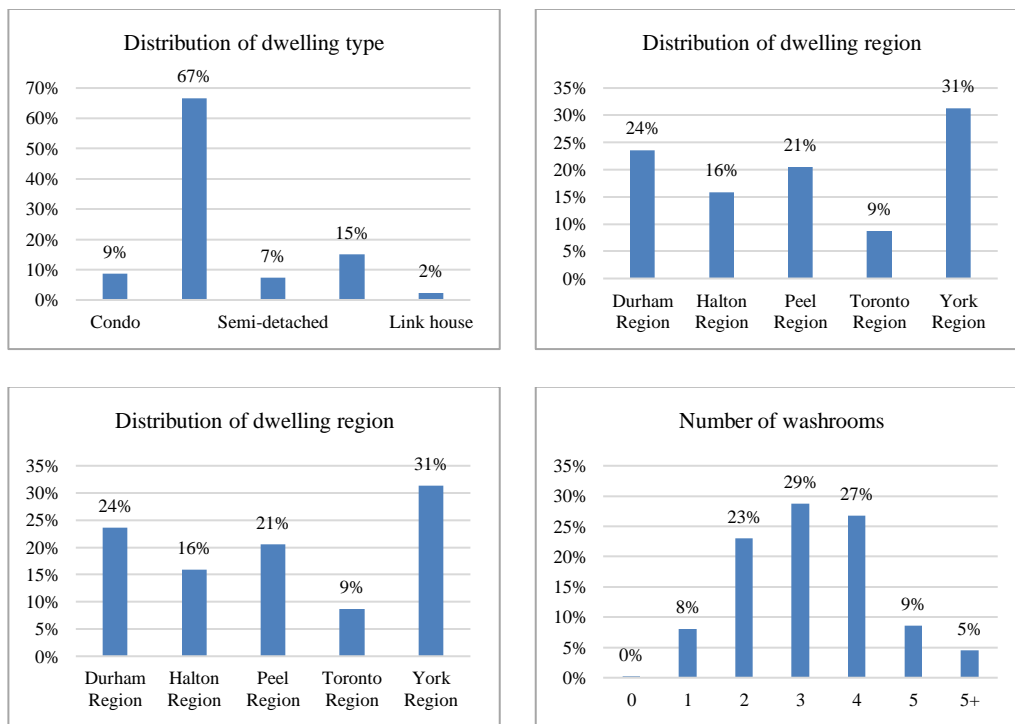


Figure 3. Distributions of key dwelling attributes

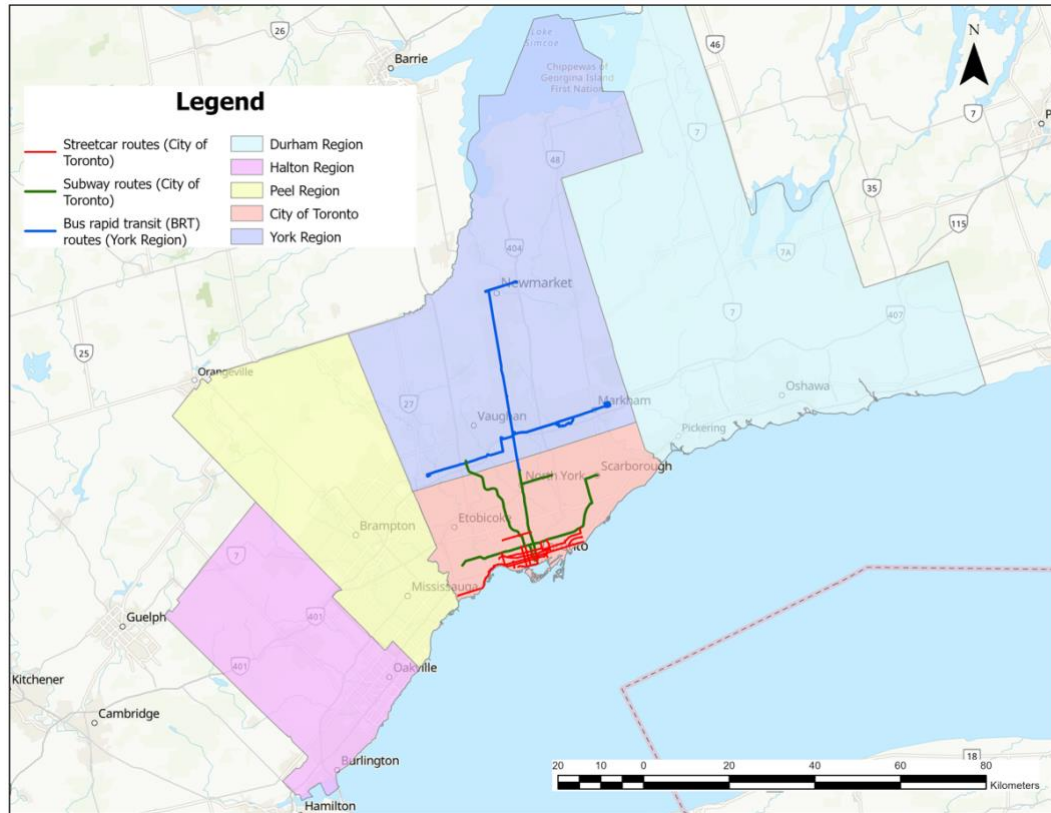


Figure 4. Study area map

4 Methodology

4.1 Housing market-clearing mechanisms in LUTI models

This study examines multinomial logit (MNL) and nested logit (NL) formulations to model the factors influencing the four outcomes of property listings. These two common discrete choice models estimate the probabilities of selecting specific alternatives from a choice set (Liu et al., 2022). Each choice alternative has a utility function with the following formulation:

$$U_j = V_j + \varepsilon_j = \beta_j x_j + \varepsilon_j \quad (1)$$

where V_j is the systematic utility of choice alternative j , β_j is a vector of estimated parameters for explanatory variables x_j , and ε_j is the random error term accounting for the utility from unobserved factors. For logit models, the error term is assumed to follow a standard Type-I extreme value distribution. The equation to compute the probability of choosing alternative j ($j = 1, \dots, k$) in an MNL model is presented below:

$$P_j = \frac{\exp(V_j)}{\sum_j \exp(V_k)} \tag{2}$$

In an MNL model, the alternatives are assumed to be independent and irrelevant, which means proportional substitution exists between the alternatives (Train, 2009). However, there may be situations in which correlated variations may exist between subsets of alternatives and violate the assumption. In this study, it is suspected that the three choice alternatives followed by the termination of initial listings may be correlated and can be viewed as a choice bundle. Therefore, an NL model is also developed to accommodate listing termination as an upper-level nest. The nesting structure of the NL model is illustrated in Figure 5.

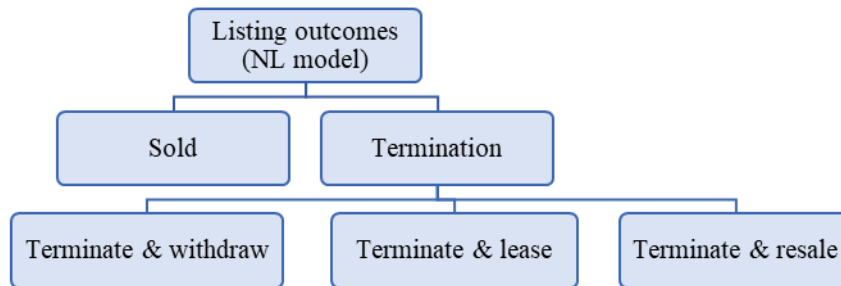


Figure 5. Nesting structure of the NL model

The probability of choosing an upper-level alternative s ($s = 1, \dots, q$) that is not in a nest (i.e., Sold) is formulated in Equation 3. The probability of choosing a nest n ($n = 1, \dots, m$) and the condition probability of choosing an alternative k ($k = 1, \dots, h$) within the nest given that the nest has been chosen can be calculated using Equation 4 and 5, respectively.

$$P_s = \frac{\exp(V_s)}{\sum_S \exp(V_q) + \sum_N \exp(V_m)} \tag{3}$$

$$P_n = \frac{\exp(V_n)}{\sum_S \exp(V_q) + \sum_N \exp(V_m)} \tag{4}$$

$$P_{k|n} = \frac{\exp(\mu V_k)}{\sum_K \exp(\mu V_h)} \tag{5}$$

where μ is the estimated scale of the lower-level conditional choices, and $V_n = \frac{1}{\mu} \ln(\sum_K \exp(\mu V_k))$. The scale of the upper-level choices is assumed to be 1; therefore, the lower-level scale *needs* to be greater than 1 for the nesting structure to be valid.

4.2 Hazard duration model

This study utilizes a continuous hazard duration model to estimate the time until a residential property listing becomes inactive in the market, namely the TOM duration. Hazard duration models are often used to estimate the time and probability of the occurrence of an event. The rate of the event occurring at time t given that it has not

happened until time t can be represented by the following equation (Golshani et al., 2018; Mohammadian & Rashidi, 2007):

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (6)$$

where $h(t)$ is the hazard function, $f(t)$ is the probability density function, and $F(t)$ is the cumulative density function that shows the probability of the event occurring until time t . Each hazard function has a corresponding survival function, which represents the probability of surviving from the occurrence of the event until time t and can be formulated as the following:

$$S(t) = 1 - F(t) \quad (7)$$

This study uses an accelerated time hazard formulation, which allows time to be expressed as a function of covariates and makes interpreting the estimated parameters' effects easier (Golshani et al., 2018; Habib & Miller, 2006). The accelerated time hazard model expresses the time duration as a log-linear regression function of the covariates (Kiefer, 1988):

$$\ln(t_j) = \alpha_j W_j + \psi_j \quad (8)$$

where $\ln(t_j)$ is the natural logarithm of the TOM duration of a listing with outcome event j , α_j is a vector of estimated parameters for covariates W_j , and ψ_j is the random error term representing unobserved factors. If the error term ψ_j is assumed to be a normal distribution, then the model becomes a lognormal hazard model with the following survival function (Habib & Miller, 2006):

$$S(t_j) = 1 - \Phi\left(\frac{\ln(t_j) - \alpha_j W_j}{\sigma_{\psi_j}}\right) \quad (9)$$

where Φ represents a cumulative distribution function and σ_{ψ_j} is the scale parameter. A standard normal distribution for the error term ψ_j is assumed in this study.

Two nests of competing choices can end a property listing: sold and termination, among which only one would happen. The termination nest includes the choices of termination & withdrawal, termination & lease, and termination & resale. These nests of two choices are viewed as competing events because they are two different processes that have different underlying logics. Once a property is listed on the market, a sale may happen at any point in time, based on factors such as market conditions and whether an acceptable bid has been made, which makes it a hazard-type process. As time passes, the seller may choose to terminate the listing or a sale can happen anytime if a buyer places an acceptable bid. Thus, the listing remains active on the market until either a sale or termination occurs. Therefore, the two underlying processes for sale and termination compete, and eventually one would happen.

The competing hazard model formulation is inspired by several studies including Rashidi & Ghasri (2019) and Henley (1998). Rashidi & Ghasri (2019) utilized a competing accelerated failure model to jointly model the residential relocation reason and timing decisions. They considered three reasons for relocation, and their competing

hazard model accounted for the possibility that more than one reason can contribute to a relocation decision. Therefore, they developed a competing survival formulation to account for the interdependencies between the possible outcomes. However, this study differs from Rashidi & Ghasri (2019) in that the competing outcomes (i.e., sale and termination) cannot happen at the same time. The interpretation of the relationship between the competing choices (i.e., sale and termination) in the proposed model is similar to that in Henley (1998), which applied a competing risk hazard to a discrete competing hazard to model relocation timing and tenure decisions. Henley (1998) assumed that the latent durations of possible tenure states (i.e., transition to owner-occupation, public rental or private rental) are independent random variables, and the final event is dependent on which of these durations is the shortest. The model proposed in this study also assumes that the TOM of a listing is determined by the latent duration of which choice (sale or termination) is the shortest.

Since sale and termination cannot occur simultaneously, the outcome hazard can be expressed as the summation of the probability of each event occurring (Mohammadian & Rashidi, 2007). The hazard, survival, and likelihood functions are presented below:

$$h(t) = h_{sale}(t) + h_{term}(t) \tag{10}$$

$$S(t) = S_{sale}(t) \times S_{term}(t) \tag{11}$$

$$L = \prod_{i=1}^I [h_{sale}(t)^{R_{si}} \times S_{sale}(t) \times h_{term}(t)^{R_{ti}} \times S_{term}(t)] \tag{12}$$

where R_{si} and R_{ti} are binary variables indicating whether the property listing is sold or terminated. Unique sets of parameters are estimated separately for the competing events in the hazard duration model.

4.3 Copula-based joint model

This study applies a copula-based joint model to investigate the correlation between a property listing's outcome and its TOM duration, the dependent variables of the discrete choice, and hazard duration models. A copula function C_θ , as shown in Equation 12, can join multiple marginal probability distributions (i.e., $F_1(y_1)$, $F_2(y_2)$) into a multivariate probability distribution $F(y_1, y_2)$ (Sklar, 1996).

$$F(y_1, y_2) = C_\theta(u_1 = F_1(y_1), u_1 = F_2(y_2)) \tag{13}$$

where θ is the copula parameter. Suppose the copula parameter is found to be statistically significant, in that case, it implies that there is a certain dependency between the error terms (or unobserved factors) of $F_1(y_1)$ and $F_2(y_2)$. The final model specifications presented in this study are from a Gumbel copula. The equation for the Gumbel copula is provided below:

$$C_\theta(u_1, u_2) = \exp\left(-\left[(-\ln(u_1))^\theta + (-\ln(u_2))^\theta\right]^{1/\theta}\right) \tag{14}$$

where, $u_1 = F_1(\beta_j x_j)$, and $u_2 = F_2\left(\frac{\ln(t_j) - \alpha_j W_j}{\sigma_{\psi_j}}\right)$. Maximum likelihood estimation is used to estimate the parameters. The likelihood function of the copula-based joint model is presented in Equation 14 (Spissu et al., 2009).

$$L = \prod_{i=1}^I \left[\left\{ \prod_{j=1}^J \frac{1}{\sigma_{\psi_j}} \times \frac{\partial C_{\theta}(\mu_{i1}^j, \mu_{i2}^j)}{\partial \mu_{i2}^j} \times f_{\psi_j} \left(\frac{\ln(t_{ji}) - \alpha_j W_{ji}}{\sigma_{\psi_j}} \right) \right\}^{R_{ji}} \right] \quad (15)$$

where $\frac{\partial C_{\theta}(\mu_{i1}^j, \mu_{i2}^j)}{\partial \mu_{i2}^j} = u_2^{-1} (-\ln u_2)^{\theta-1} C_{\theta}(u_1, u_2) [(-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}]^{\frac{1}{\theta}-1}$ is the partial derive concerning C_{θ} with respect to μ_{i2}^j (Bhat & Eluru, 2009; Spissu et al., 2009), f_{ψ_j} is the probability density function, and R_{ji} is a binary indicator on whether alternative j is the outcome of property listing i . All the models are coded and estimated using the statistical software Gauss (Gauss, Aptech Inc., 2023).

5 Experimental results

5.1 Model specifications

MNL and NL models are examined and compared using goodness-of-fit measures to estimate the probability of listing outcomes. The NL model results in lower Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, which indicates a better fit to the dataset. Therefore, only the estimation results of the NL model are presented in this paper, and the NL model is applied in the copula-based joint approach. For the hazard duration model component, three sets of parameters are estimated separately for the different listing outcomes: 1) sold listings, 2) terminate & withdraw and terminate & lease, and 3) terminate & resale. This is to better capture the distinctive effects of the variables on the different listing outcomes. It is noted that two of the termination outcomes, terminate & withdraw and terminated & lease, are estimated using the same set of parameters because the number of listings which are terminated & leased is relatively small compared to the others. These observations are not sufficient to estimate a unique set of parameters for this listing outcome.

Various copula functions can be used to examine the correlation structure, such as the Gaussian copula, the Farlie-Gumbel-Morgenstern (FGM) copula, and the Archimedean class of copula, etc. Several copula functions were explored in this study, and the Gumbel copula is found to be the one with the best statistical fit and interpretable results. Therefore, only the estimation results of the Gumbel copula model are presented in the paper. The estimation results of the NL model component and the hazard model component are summarized in Table 1 and Table 2, respectively. The Gumbel copula model has a loglikelihood of -71057 and a ρ^2 value of 0.1 against a null model with equal probability across choice alternatives and constant hazard rates. As a comparison, an independent model was also estimated by setting the correlation parameter to 1. The estimation results of the independent model are included in Appendix A. The Gumbel copula model has a slightly higher log-likelihood than the independent model. The higher log-likelihood of the Gumbel copula model also suggests that the data is a factor of the error correlation model structure.

The Gumbel copula cannot account for negative dependency. Its correlation parameter is always greater than or equal to 1, with a parameter of 1 representing independence. It is believed that dependency between the different termination choices and the TOM would be different. Therefore, to better capture the potential differences in dependency, three unique copula parameters θ are estimated between the TOM of terminated listings and the three termination choices. The three copula parameters are all statistically significant and greater than 1. This indicates the existence of unobserved factors that can influence both the sellers' choices of when to terminate the property listings and what actions to take afterward. The choice of terminate & resale has the largest copula parameter in magnitude. This shows that it has the strongest correlation with the TOM, among all three termination options.

The Gumbel copula is suitable for joint distributions that have a strong correlation at the higher values but a weak correlation at the lower values (Bhat & Eluru, 2009). In this case, it is reasonable that the correlation between the probability of listing termination and its TOM duration is stronger for those with longer TOM durations. For all three termination choices, a positive correlation is found between its error term ε_j and the error term ψ_j of the TOM hazard model. Spissu et al. (2009) explained the interpretation of the copula correlation parameter. A positive correlation parameter implies the unobserved factors that can increase the likelihood of choosing to terminate a property listing would probably decrease its TOM until the termination.

Table 1. Estimation results of the copula-based joint model – NL model component

NL model component	Termination Options					
	Terminate & withdraw		Terminate & lease		Terminate & resale	
Variable	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
Copula parameter	1.16	6.19	1.14	6.53	1.72	14.14
Kendall's τ	0.14	-	0.13	-	0.42	-
Constant	3.67	2.98	3.09	1.63	4.41	3.77
Dwelling type – detached	-0.07	-1.35	-0.24	-2.69	-0.08	-1.57
Region - Peel	0.15	3.62	-	-	0.17	3.78
Number of parking spots	-0.14	-8.35	-0.21	-4.57	-0.14	-7.76
Natural logarithm of dwelling area	-0.24	-3.29	-0.16	-1.72	-0.12	-1.80
Natural logarithm of lot size	-0.10	-11.33	-0.12	-8.97	-0.11	-11.94
Natural logarithm of asking price	1.49	20.12	1.50	13.55	1.33	20.70
Number of subway stations within a 500-meter radius	-0.35	-2.57	-0.25	-1.63	-0.29	-2.32
Number of BRT stations within a 500-meter radius	-0.14	-2.38	-0.10	-1.27	-0.13	-2.39
Number of streetcar stops within a 500-meter radius	-0.22	-6.85	-0.18	-8.32	-0.17	-9.62
Natural logarithm of average residents' age of the DA	-2.12	-13.97	-2.62	-9.33	-2.12	-14.37
Natural logarithm of median household income of the DA	-1.16	-13.45	-1.05	-7.88	-1.10	-13.32
Sale-to-list ratio of the DA	-1.12	-25.03	-1.20	-15.78	-1.10	-25.43
The scale of the nest	4.61	2.47	-	-	-	-

Table 2. Estimation results of the copula-based joint model – hazard model component

Hazard model component	Sold	Termination Options				
		Parameter	t-stat	Terminate & withdraw		Terminate & resale
Variable	Parameter			t-stat	Parameter	t-stat
Constant	-8.95	-19.18	-8.79	-7.56	-8.95	-15.43
Dwelling type - condo	1.14	11.71	1.17	5.47	1.40	10.93
Dwelling type - townhouse	0.24	6.59	0.35	5.37	0.37	10.16
Region - Durham	0.15	4.87	-	-	-	-
Region - York	0.16	6.57	-	-	0.13	4.96
Region - Toronto	0.16	3.24	-	-	-	-
Number of washrooms	-0.06	-4.72	-0.08	-4.01	-0.02	-1.54
Number of parking spots	-	-	-	-	-0.03	-1.98
Have an attached parking garage	-0.10	-4.34	-	-	-0.09	-3.17
Have a detached parking garage	-	-	0.22	2.71	0.11	2.14
No parking garage	0.22	5.56	0.17	2.31	0.13	2.70
Natural logarithm of dwelling area	-0.16	-3.58	-	-	-0.15	-3.11
Natural logarithm of lot size	0.06	5.81	0.06	2.31	0.10	7.26
Natural logarithm of asking price	0.78	21.18	0.74	10.16	0.75	18.36
Average residents' age of the DA	0.22	2.74	0.39	2.33	0.52	5.07
Percentage of movers in the DA	-	-	-	-	0.61	3.13
Sale-to-list ratio of the DA	-	-	-	-	0.13	3.24
Number of streetcar stops within a 500-meter radius	-0.03	-3.76	-	-	-	-
Scale parameter	0.91	127.92	0.99*	24.16	0.97	41.49
			0.80**	18.00		

5.2 Model interpretations

Due to the nature of the dataset, only generic variables are available for the NL model components. Therefore, alternative-specific parameters are estimated using the sold alternative as the reference. The other three alternative outcomes are bundled under the “termination” nest. The scale parameter of the nest is statistically different from 1 at the 95% confidence level, indicating a valid nesting structure. Most of the parameters are statistically significant at the 95% confidence level. Only a few parameters that are statistically significant at the 90% confidence level are included in the model for comparison purposes.

As expected, dwelling attributes have a significant influence on the listing outcomes. It is observed that being a detached house and having more parking spots can negatively affect the probability of terminating the listing.

* is the scale parameter for terminate & withdrawn, ** is the scale parameter for terminate & lease

This implies that residential properties that are houses and/or have more parking spots are more likely to stay on the market until they are sold. In addition, the dwelling area and lot size of a property are shown to have negative effects on the probabilities of termination choices, whereas the asking price has positive effects. Intuitively, properties with higher asking prices, smaller dwelling areas, and/or smaller lots are less likely to be sold.

There are three parameters measuring the accessibility to public transit: 1) the number of subway stations within a 500-meter radius of the property, 2) the number of streetcar stops within a 500-meter radius of the property, and 3) the number of bus rapid transit (BRT) stations within a 500-meter radius of the property. As shown in Figure 4, the subway and streetcar systems are only available in the City of Toronto, whereas the BRT system is only available in York Region. All three parameters have negative coefficients for the three termination choices. This indicates that if the number of such transit stations increases in the vicinity of the listed property, its listing is less likely to be terminated and it is more likely to be sold. In general, people are more willing to live in areas with easy access to public transport. Therefore, it is reasonable to observe that properties with better transit access are more likely to be sold.

The model also includes two socio-economic variables generated from the Census data that negatively affect the probability of termination outcomes: average residents' age and the natural logarithm of median household income of the DA where the property is located. This implies that the property neighborhoods with more elderly people and/or higher median household income are less likely to be taken off the market before selling. This is probably because such neighborhoods may have the qualities that generally attract residential property buyers, such as greenness, quietness, cleanliness, safety, etc. Moreover, the sale-to-list ratio of the DA where the property is located is also included in the model to account for the market influence on the listing's probability of sale. This variable shows the ratio between the sold and listed properties in the month before the listing is sold or terminated. It appears that if the sales-to-listing ratio increases, the likelihood of terminating the listing would decrease. Intuitively, a larger sales-to-listing ratio would indicate a hot market in which the property is more likely to be sold.

From the results of the hazard model component, it is observed that many of the variables have similar effects on the TOM of both sold and terminated listings. For example, dwelling attributes such as dwelling types, lot size, number of washrooms, and asking price of the listing have statistically significant coefficients for all of the listing outcomes. It is found that condo units, townhouses, properties without garages, properties with larger lot sizes, and/or properties with higher asking prices will likely stay longer in the market regardless of the listing outcomes. In contrast, having more washrooms shortens a property listing's TOM, consistent with the observation from Allen et al. (2018). Moreover, some dwelling attributes only have effects on the TOM of certain listing outcomes. Having an attached garage and larger dwelling areas may shorten a listing's TOM until a successful sale or a terminate & resale decision, whereas having a detached garage may increase a listing's TOM until all types of termination decisions. Meanwhile, having more parking spots would decrease a listing's TOM until a terminate & resale decision. Such results suggest that having an attached garage, larger dwelling areas, and/or more parking spots may be more desirable dwelling attributes than having a detached garage, hence that can help decrease the TOM until a sale or terminate & resale decision.

In terms of accessibility measurements, the number of streetcar stops within a 500-meter radius of the property is the only statistically significant variable on the hazard model. It can negatively affect a listing's TOM until a successful sale. This means that the more streetcar stops within the vicinity of a listing property, the faster it is likely to be sold. Similar to the NL model component, the hazard model component also contains the

sales-to-listing ratio as a market indicator variable. This ratio can positively impact a listing's TOM until a terminate & resale decision. It is possible that in a hot market, even if a property cannot be sold within an expected timeframe, the sellers would be more inclined to stay in the market longer and look for an opportunity to sell, as opposed to terminate and relist the property sooner. Lastly, the scale parameters estimated for the variance of the error terms are statistically significant, indicating that important unobserved factors influence the TOM durations of property listings.

5.3 Elasticity analysis

The estimated model parameters for the hazard duration models can intuitively provide the effects of variables on TOM durations in terms of direction and magnitude because they are in a log-linear regression formulation. However, for the discrete choice models, the estimated parameters can only reveal the directional effects of the variables, and an elasticity analysis needs to be conducted to quantify how changes in the variables can impact the listing outcomes (Loa & Habib, 2023). Therefore, the sample average direct elasticities of some key variables in the NL component of the copula-based joint model are calculated and summarized in Table 3.

Table 3. Direct elasticity of continuous variables in NL component

Variable	Terminate & withdraw	Terminate & lease	Terminate & resale
Number of parking spots	-0.49	-1.04	-0.64
Natural logarithm of dwelling area	-4.13	-3.72	-2.63
Natural logarithm of lot size	-1.77	-2.99	-2.55
Natural logarithm of asking price	47.15	63.99	53.51
Number of subway stations within a 500-meter radius	-0.02	-0.01	-0.01
Number of BRT stations within a 500-meter radius	-0.01	-0.01	-0.01
Number of streetcar stops within a 500-meter radius	-0.08	-0.05	-0.05
Natural logarithm of average residents' age of the DA	-17.81	-29.66	-22.60
Natural logarithm of median household income of the DA	-30.81	-37.47	-37.04
Sale-to-list ratio of the DA	-0.83	-0.99	-1.00

In terms of dwelling attributes, the dwelling area of a property seems to have the greatest impact on choosing to terminate the listing. For example, a 1% increase in the natural logarithm of the dwelling area can reduce the probability of withdrawing the property from the market by 4%. The elasticities of the asking price on the termination probabilities appear to be the highest among all continuous variables. Particularly, its effect on converting the listing to a lease or re-listing the property as another sale is larger than withdrawing it from the market. It is shown that a 1% increase in the natural logarithm of the asking price would lead to a 54% increase in the probability of re-listing it as another sale. In contrast, it would result in a 47% increase in the probability of withdrawal. The average residents' age and median household income of the DA also have moderate influences on the termination probabilities. A 1% increase in the natural logarithm of median household income can lead to a 37% decrease in the probability of terminating and reselling the property. The transit accessibility measures appear to have the smallest influence on the termination probabilities. All of their elasticities are smaller

than one. In comparison, the effect of the sale-to-list ratio is slightly larger than that of the transit accessibility measures. Specifically, it has a greater impact on the probability of resale than the other two termination probabilities. A 1% increase in the sale-to-list ratio can cause a 1% decrease in the probability of terminating and reselling the property. This shows that in a hot market or population residential areas, sellers are less likely to terminate the existing listings and they are more likely to stay in the market until the properties are sold.

5.4 Model predictions

A randomly selected 15% of the original dataset is used as a holdout sample to investigate the performance of the copula model on predicting the TOM of property listings. This holdout sample was not included in the training dataset when estimating the model specifications. Both the copula model and the independent model were tested using the holdout sample to predict the TOM durations. Figure 6 presents a density plot showing the distribution of the observed TOM and the predicted TOM using both models. The copula model has an R2 value of 0.24, whereas the independent model has an R2 value of 0.19.

It appears that the independent model generally underpredicts the TOM durations. Compared to the observed data, the independent model would result in a lot more listings with TOM in less than 15 days. Moreover, the independent model fails to predict the listings with TOM longer than 50 days. On the contrary, the density distribution of the copula model is comparable to the observed data between 0 to 15 days, and between 50 to 70 days. However, the copula model tends to overestimate the number of listings with TOM between 15 to 45 days. It is possible that when correlating the termination choices with the listings' TOM until termination, the model tends to overestimate their TOM durations.

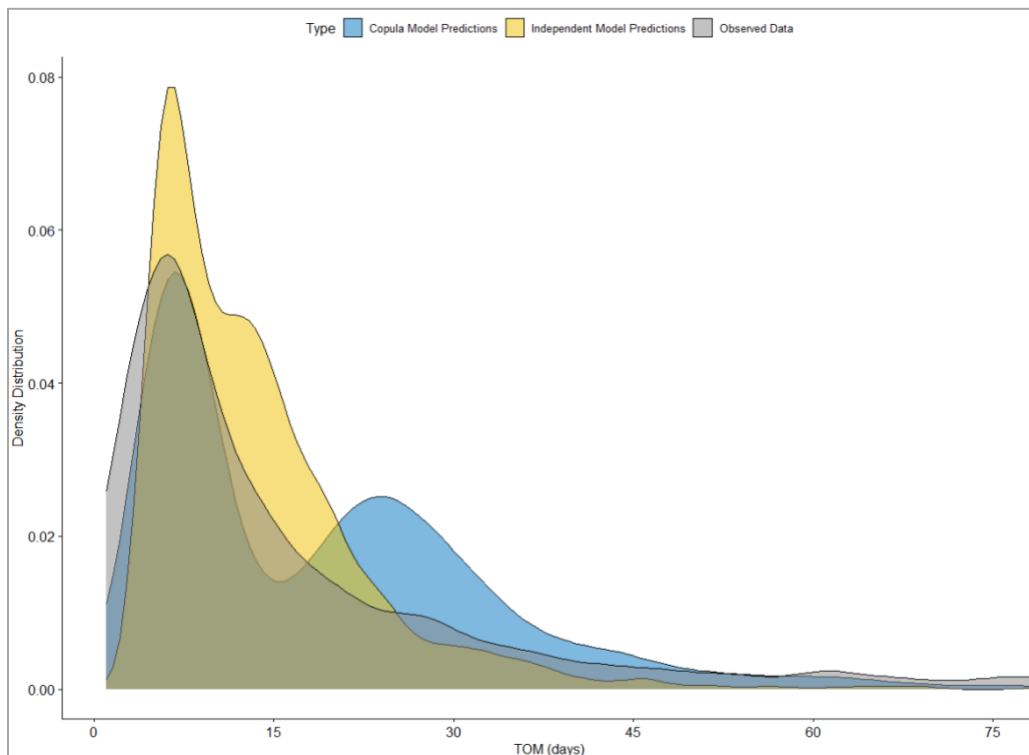


Figure 6. Density plot of observed and predicted TOM distributions

6 Implications of empirical findings

The results of this study offer insights into the determinants of sellers' choices on terminating property listings and their TOM durations, along with the correlations between them. The variables utilized in the model are commonly used variables for microsimulation LUTI models. The data on dwelling attributes may be obtained from private agencies or open data platforms available in the study area. The transit accessibility measures can be computed from General Transit Feed Specification (GTFS) data made available by local transit agencies. The aggregated socio-economic variables of the DA can be retrieved from the census data.

Some factors appear to significantly affect both TOM durations and termination outcomes, such as dwelling area, asking price, lot size, number of parking spots, and the sale-to-list ratio of the DA, etc. Specifically, the asking price of a listing turns out to be a significant and influential factor. This is expected because price may be the primary consideration of a residential property buyer, and few people have the budget for expensive properties.

Some factors are influential to TOM durations but are insignificant for the listing termination choices, such as the number of washrooms and types of garages. It is reasonable that such dwelling attributes would affect a listing's TOM because certain features may be more popular on the market than others. The model results prove that there are correlations between the termination choices of a listing and its TOM duration, and the model estimations may be biased if the correlations are ignored. In addition, the competing hazard approach utilized in this study attempts to examine the differences between the determinants of the TOM durations of sold and terminated listing, and the result suggests that the differences are minimal.

6.1 Proposed integration with a LUTI framework

The model estimated in this study can provide empirical evidence for several key assumptions often included in the housing market-clearing process of LUTI models. For example, LUTI microsimulation frameworks such as ILUTE (Farooq & Miller, 2012; Rosenfield et al., 2013) and SimMobility (Zhu et al., 2018) have assumed parameters for a seller's decision to exit the housing market. SimMobility assumes a maximum TOM of 210 days for unsold properties (Zhu et al., 2018). In contrast, ILUTE simulates a seller's market exit decision based on the number of failed attempts to sell the property (Farooq & Miller, 2012). However, the development of these assumptions remains unclear, and their validity has yet to be examined.

In comparison with the findings of this study, the 210-day maximum TOM assumption seems to be an oversimplification of the market behavior. This study finds that most unsold property listings are only available for a month or two before termination. In addition, there are no explicit statements in SimMobility and other microsimulation models about what happens to the properties that are taken off the market after the maximum TOM duration. This study observes that around 70% of the terminated listings are re-listed within the same year, attempting to sell again. The existing LUTI models cannot capture such behavior and its impact on market dynamics. This study can fill this gap using the proposed joint model to estimate the maximum TOM durations for each simulated residential property and the likelihood of exiting the market for each simulated seller in a LUTI microsimulation model. Figure 4 presents a conceptual flow diagram illustrating how the proposed model can be integrated into the housing market component of a LUTI framework.

The proposed model primarily assists with the sellers' market exit decisions, which is lacking in most of the LUTI models. When a seller enters the market for the first time,

his/her property for sale would become available in the active dwellings' pool. Meanwhile, the households that wish to relocate are in the active buyers' pool. The market-clearing mechanism employed in the LUTI model would attempt to match the active dwellings with the potential buyers, usually through a bid-auction process. If the property for sale can be successfully matched with a buyer, then it would proceed to the transaction process. However, if the property fails to be matched with a buyer, then the proposed model would come into play to determine whether it would exit the market or stay in the market for another simulation cycle. In this case, the probabilities of the three termination choices would be estimated to determine whether the property would be withdrawn from the market, re-listed for sale in the next cycle, or converted to a lease. Meanwhile, the TOM of the property would also be predicted to decide when the termination decision would be made. If the time until the next cycle exceeds the property's TOM and its termination decision is to be withdrawn from the market, then the property would be removed from the active dwellings' pool and would participate in the next market-clearing cycle. Otherwise, the property would remain in the pool and continue to be matched with potential buyers until it is sold or reaches its TOM for a termination decision.

The market-clearing frequency varies between different LUTI models. Some models assume a simulation period of one year, such as RELU-TRAN (Anas & Arnott, 1991, 1993) and UrbanSim (Waddell, 2010), while others choose to use smaller time steps. For example, the ILUTE housing market is cleared every month (Rosenfield et al., 2013), whereas SimMobility attempts to simulate daily transactions in its housing market sub-model (Zhu et al., 2018). Therefore, the integration between a LUTI model and the proposed model is subject to the structure of the LUTI model. There may inevitably be discrepancies between the LUTI simulation steps and the timeframe of the proposed model. The proposed model analyzes the sellers' behaviors using real-world housing market data, which operates in the timestep of days. However, for LUTI models, the feasibility and necessity of simulating a daily housing market clearance is subject to the project scope and resource constraints. Nevertheless, the results of the proposed model can still provide useful insights for LUTI simulation regarding the sellers' market exit decisions. They may affect the results of the LUTI model on whether a dwelling for sale should remain active in the simulated market for more than one simulation cycle, which would also affect the choice sets of the housing buyers in subsequent simulation cycles.

The dataset used in this study reveals that most of the property listings would stay active for one or two months before it is withdrawn from the market. If the LUTI model clears its housing market in a shorter time frame, such as by month or by day, then the proposed model can be integrated into each simulation period to determine the sellers' market exit decisions and timestamps. For LUTI models with smaller simulation steps, the proposed model can help pinpoint the precise estimated timestamp and probability of either terminating, withdrawing, or re-listing a property given that it has not been sold within an expected time. Even if the LUTI model has a simulation step of one year, the proposed models can still help understand whether the unsold properties should remain active in the market for the next stimulation cycle. It is not uncommon for sellers to withdraw a property listing if it cannot be sold for a long time. Therefore, in a one-year LUTI simulation, if a property is not sold in the current cycle and the proposed model estimates that it has a higher probability of being withdrawn from the market, then it should not appear in the active dwelling pool for the next cycle. However, if the proposed model reveals that the property is more likely to be re-listed on the market, then it might enter the active dwelling pool for the next cycle for another attempt to sell.

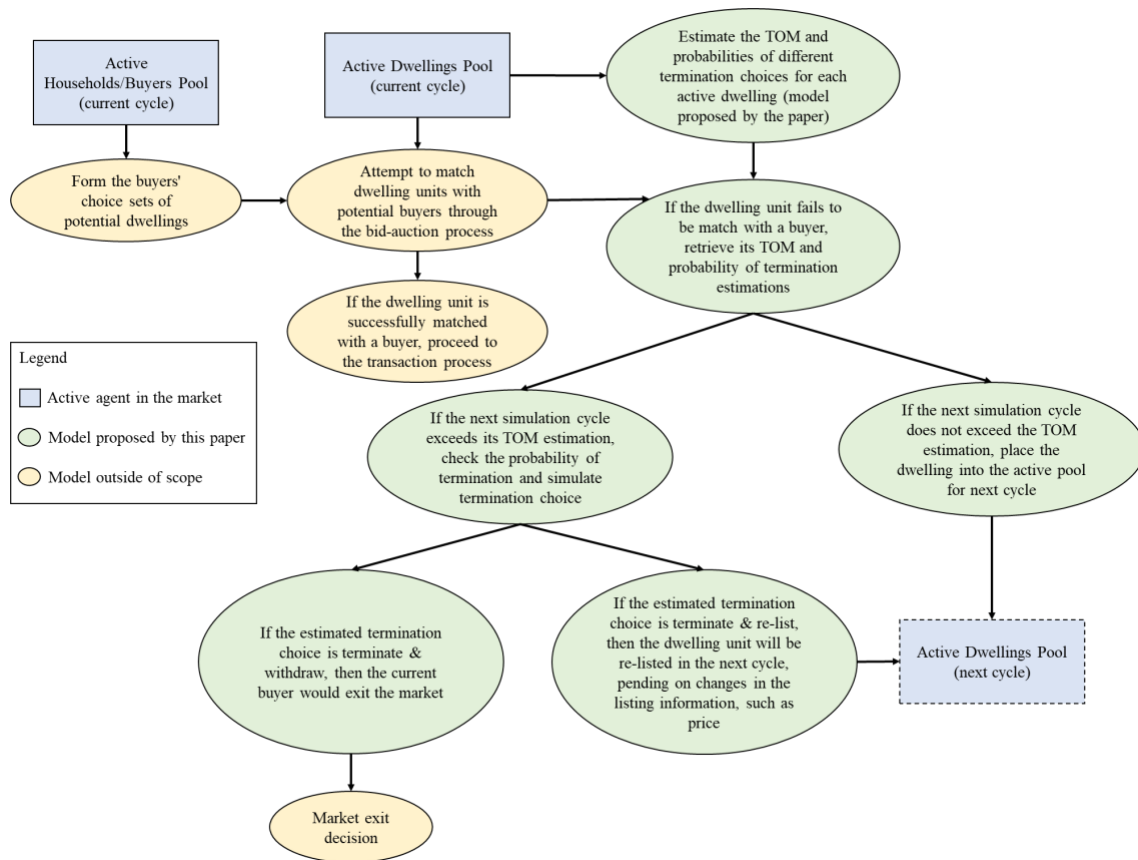


Figure 7. Conceptual diagram integrating the proposed model to an LUTI housing market-clearing process

7 Conclusion and future work

This study investigates the determinants of residential property listings' time-on-market (TOM) durations and the listing outcomes and correlations between them. Utilizing real estate listing data collected from the Toronto Regional Estate Board and Information Technology Systems Ontario, a competing hazard duration model for estimating TOM and a nested logit model for estimating listing outcomes are joined through a Gumbel copula structure. Four listing outcomes are considered in the model: 1) sold, 2) terminate & withdraw, 3) terminate & lease, and 4) terminated & resale. The sold alternative is used as a reference in the model, while the other three alternatives are bundled in a "termination" nest.

The model results suggest statistically significant and positive correlations between the three termination choices and the TOM duration of a listing. This indicates that the unobserved factors that may increase a seller's probability of terminating a list would also increase its TOM duration until the termination. The variables examined for the model include dwelling attributes, geographical variables, transit accessibility measurements, and socioeconomic variables derived from the Census data. Elasticities of some key variables of the nested logit model are computed to quantify their effects on the listing outcomes.

It is found that the asking price has large and positive effects on both the TOM duration. In addition, an increase in the number of parking spots, dwelling area and lot size would decrease the likelihood of a listing being terminated, and having more

washrooms and larger dwelling areas would reduce a listing's TOM duration. In terms of transit accessibility measurements, it is observed that an increase in the number of subway stations, BRT stations, and streetcar stops can decrease a listing's probability of being terminated. Moreover, for the hazard models, statistically significant scale parameters are found for the variance of the error terms, implying that important unobserved factors influence the property listings' time until sold or termination.

The model estimated in this study can be useful for inferring key assumptions in the housing market-clearing process of LUTI microsimulation models. The existing market clearing mechanisms in microsimulation models either do not allow the sellers to exit the market before finding buyers or set a maximum number of days or sale attempts in the market. The empirical evidence for such assumptions is unclear and yet to be examined. This study can fill this gap by providing an empirical model that can estimate the maximum TOM duration and the probability of sale and termination for each simulated residential property in a LUTI microsimulation model, introducing more realistic market dynamics for the housing market-clearing process.

Like any other research, this study has some limitations. The models only include variables about the property listings, such as dwelling attributes; however, no information is available about the sellers. The models are expected to be improved if there are socio-economic variables about the sellers. However, due to the nature of the data sources and privacy concerns, it is difficult, if not impossible, to collect such information. Future studies are encouraged to explore new techniques, such as data fusion methods, to gather socio-economic variables for the proposed model. Although the model is an improvement of the existing applications in LUTI models, it is not the perfect representation of reality. Future studies can investigate different model formulations including the assumption that the competing outcomes are conditional. Comparisons can be made between the performances of the assuming non-competing outcomes, unconditional competing outcomes, and conditional competing outcomes, like Leszczyc and Timmermans (2002). In addition, future studies are recommended to test the implementation of the proposed model in LUTI microsimulation frameworks. Comprehensive empirical experiments should be conducted to verify the simulation results with the proposed model against historical housing market data. It should be noted that 2021 was, perhaps, an "unusual" year in that it was in the middle of the COVID-19 pandemic and the Toronto housing market was (perhaps surprisingly) exceptionally active during this period. Data from "more normal" time periods would be useful to verify the TOM behavior observed in this study.

References

- Allen, M. T., Cadena, A., Rutherford, J., & Rutherford, R. C. (2015). Effects of real estate brokers' marketing strategies: Public open houses, broker open houses, MLS virtual tours, and MLS photographs. *Journal of Real Estate Research*, 37(3), 343–369. <https://doi.org/10.1080/10835547.2015.12091422>
- Allen, M. T., Dare, W. H., & Li, L. (2018). MLS information sharing intensity and housing market outcomes. *The Journal of Real Estate Finance and Economics*, 57(2), 297–313. <https://doi.org/10.1007/s11146-017-9612-5>
- Anas, A., & Arnott, R. J. (1991). Dynamic housing market equilibrium with taste heterogeneity, idiosyncratic perfect foresight, and stock conversions. *Journal of Housing Economics*, 1(1), 2–32. [https://doi.org/10.1016/S1051-1377\(05\)80023-3](https://doi.org/10.1016/S1051-1377(05)80023-3)
- Anas, A., & Arnott, R. J. (1993). Development and testing of the Chicago prototype housing market model. *Journal of Housing Research*, 4(1), 73–129.
- Benefield, J., Cain, C., & Johnson, K. (2014). A review of literature utilizing simultaneous modeling techniques for property price and time-on-market. *Journal of Real Estate Literature*, 22(2), 149–175. <https://doi.org/10.1080/10835547.2014.12090387>
- Benefield, J. D., & Hardin, W. G. (2015). Does time-on-market measurement matter? *The Journal of Real Estate Finance and Economics*, 50(1), 52–73. <https://doi.org/10.1007/s11146-013-9450-z>
- Bhat, C. R., & Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transportation Research Part B: Methodological*, 43(7), 749–765. <https://doi.org/10.1016/j.trb.2009.02.001>
- Bian, X., Waller, B. D., & Wentland, S. A. (2016). The role of transaction costs in impeding market exchange in real estate. *Journal of Housing Research*, 25(2), 115–135. <https://doi.org/10.1080/10835547.2016.12092115>
- Caudill, S. B., Bogosavljevic, K., Johnson, K. H., & Mixon, F. G. (2022). Time on the market and probability of sale using a generalized geometric hazard model. *Journal of Econometric Methods*, 11(1), 1–16. <https://doi.org/10.1515/jem-2020-0017>
- Das, A. (2007). Testing a hazard model for the housing market in New Orleans. *American Journal of Economics and Sociology*, 66(2), 443–456.
- Farooq, B., & Miller, E. J. (2012). Towards integrated land use and transportation: A dynamic disequilibrium based microsimulation framework for built space markets. *Transportation Research Part A: Policy and Practice*, 46(7), 1030–1053. <https://doi.org/10.1016/j.tra.2012.04.005>
- Gauss, Aptech Inc. (2023). Aptech. Retrieved from <https://www.aptech.com/>
- Golshani, N., Shabanpour, R., Mahmoudifard, S. M., Derrible, S., & Mohammadian, A. (2018). Modeling travel mode and timing decisions: Comparison of artificial neural networks and copula-based joint model. *Travel Behavior and Society*, 10, 21–32. <https://doi.org/10.1016/j.tbs.2017.09.003>
- Habib, K. M. N., & Miller, E. J. (2006). Modeling individuals' frequency and time allocation behavior for shopping activities considering household-level random effects. *Transportation Research Record*, 1985(1), 78–87. <https://doi.org/10.1177/0361198106198500109>
- Habib, M. A. (2009). Microsimulating residential mobility and location choice processes within an integrated land use and transportation modelling system (Ph.D. thesis), Department of Civil Engineering, University of Toronto, Toronto, Ontario.
- Henley, A. (1998). Residential mobility, housing equity and the labor market. *The Economic Journal*, 108(447), 414–427.

- Kiefer, N. M. (1988). Economic duration data and hazard functions. *Journal of Economic Literature*, 26(2), 646–679.
- Leszczyc, P. T. L. P., & Timmermans, H. (2002). Unconditional and conditional competing risk models of activity duration and activity sequencing decisions: An empirical comparison. *Journal of Geographical Systems*, 4(2), 157–170. <https://doi.org/10.1007/s101090200083>
- Liu, Y., Miller, E. J., & Habib, K. N. (2022). Inferring trip destination purposes for trip records collected through smartphone apps. *Journal of Transportation Engineering, Part A: Systems*, 149(2), 04022145. <https://doi.org/10.1061/JTEPBS.TEENG-7350>
- Liu, Y., Miller, E. J., & Habib, K. N. (2023). A review of the housing market-clearing process in integrated land-use and transport models. *Journal of Transport and Land Use*, 16(1), 335–359. <https://doi.org/10.5198/jtlu.2023.2268>
- Loa, P., & Habib, K. N. (2023). Identifying the determinants of anticipated post-pandemic mode choices in the Greater Toronto Area: A stated preference study. *Transportation Research Record*, 2677(6), 199–217. <https://doi.org/10.1177/03611981221145133>
- Miller, E. J. (2018). The case for microsimulation frameworks for integrated urban models. *Journal of Transport and Land Use*, 11(1), 1025–1037. <https://doi.org/10.5198/jtlu.2018.1257>
- Mohammadian, A., & Rashidi, T. H. (2007). Modeling household vehicle transaction behavior: Competing risk duration approach. *Transportation Research Record*, 2014(1), 9–16. <https://doi.org/10.3141/2014-02>
- Ozonder, G., & Miller, E. J. (2021). Longitudinal investigation of skeletal activity episode timing decisions – A copula approach. *Journal of Choice Modelling*, 40, 100306. <https://doi.org/10.1016/j.jocm.2021.100306>
- Rashidi, T. H., & Ghasri, M. (2019). A competing survival analysis for housing relocation behavior and risk aversion in a resilient housing market. *Environment and Planning B: Urban Analytics and City Science*, 46(1), 122–142. <https://doi.org/10.1177/2399808317703381>
- Rosenfield, A., Chingcuanco, F., & Miller, E. J. (2013). Agent-based housing market microsimulation for integrated land use, transportation, environment model system. *Procedia Computer Science*, 19, 841–846. <https://doi.org/10.1016/j.procs.2013.06.112>
- Sklar, A. (1996). Random variables, distribution functions, and copulas: A personal look backward and forward. *Lecture Notes-Monograph Series*, 28, 1–14.
- Spissu, E., Pinjari, A. R., Pendyala, R. M., & Bhat, C. R. (2009). A copula-based joint multinomial discrete–continuous model of vehicle type choice and miles of travel. *Transportation*, 36(4), 403–422. <https://doi.org/10.1007/s11116-009-9208-x>
- Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9780511805271>
- Waddell, P. (2010). Modeling residential location in UrbanSim. In F. Pagliara, J. Preston, & D. Simmonds (Eds.), *Residential location choice* (pp. 165–180). Berlin: Springer. https://doi.org/10.1007/978-3-642-12788-5_8
- Zhu, Y., Diao, M., Ferreira, J., & Zengras, C. (2018). An integrated microsimulation approach to land-use and mobility modeling. *Journal of Transport and Land Use*, 11(1), 633–659. <https://doi.org/10.5198/jtlu.2018.1186>