Micro interaction modeling with the recursive structure for land transaction and activity chain in urban networks

INU021001

Risa KOBAYASHI
kobayashi@bin.t.u-tokyo.ac.jp

Eiji Hato
hato@bin.t.u-tokyo.ac.jp

Abstract

The microscopic land and transportation interaction model framework and the estimation method for it are proposed. The proposed model consists of a land transaction by landowner based on random utility theory and an activity chain based on a dynamic discrete choice model. As an estimation method, we proposed a new recursive algorithm in which the results of each sub-model become explanatory variables. The proposed model is based on economic theory as well as previous LUTI models. They have only dealt with the interaction between land and transportation on a zonal basis so far. In contrast, the proposed model is formulated on a link level, and the submodels for both land and transportation are disaggregated. The empirical model estimation results show that the proposed method can remove the parameter bias in a reasonable time. We obtained this result by comparing algorithms that assume multiple behavioral equilibria with input land transaction data and activity-chain obtained from a survey of visitors in a local Japanese city. This paper explicitly shows how to a micro land-transport interaction model from three perspectives: formulation, data used, and empirical estimation methods. The proposed recursive model framework and algorithm have demonstrated their practical applicability to street-level policymaking.

INTRODUCTION

Land-transport interactions occur not only at the urban scale but also at the scale of pedestrian activity. This trend is highlighted by the recent situation where owners are changing their use due to reduced pedestrian flow and stay by the lockdown.

The land-transport interactions have been studied for a long time as a land-use-transportation interaction (LUTI) model for urban and transportation planning at the urban scale (*1*). To understand the interaction of link scales and provide evidence for planning, a model framework that allows for theoretical analysis, such as LUTI, is required. However, the spatial resolution of typical LUTI models, which are applied from urban areas to local governments, is too coarse to be used to pedestrian-scale. This research aims to propose a microscopic joint model framework between a land transaction by landowners and link-based behavior.

The LUTI model differs depending on what kind of market is assumed and what stage of interaction is focused on. This interaction model has also been evolving into a model framework consisting of more disaggregated sub-models. Even though how to use the lands almost requires decisions by landowners practically, there are fewer LUTI models with sub-models based on landowners than on households and firms.

At the same time, the LUTI model, which is based on microeconomics, assumes a behavioral equilibrium or defines market-clearing between land component and transportation component. Therefore, since the interaction tends to be complex, the method of parameterization has been proposed (*2*).

To solve this complex problem, recursive methods are an approach to characterize the interaction between land and transportation and provide empirical findings. Recursive methods are a powerful approach for analyzing dynamic decisions. One of these theoretical methods (*3*), the model using the Bellman equation, has recently been applied to transportation planning theoretically (*4–6*) and empirically (*7–9*). However, the applicability of the recursive method to LUTI has not been clarified. At the same time, it is necessary to examine the computational cost, which is the key challenge in demonstrating the recursive method. In addition, as far as we know, there are no models in the literature on how microscopic land and transportation interact.

This paper empirically proposes a framework and estimation method for a microscopic land-transport interaction model, drawing on elements of LUTI, which aims at econometric analysis and prediction. The proposed model framework consists of a utility-maximizing discrete choice model on the land component and a dynamic discrete choice model on the transportation component. The probability of selling or buying land, including the choice probability of other entities, and the expected utility to the destination are expressed by recursive equations, respectively. Thus, the challenge is to solve the immovable point problem, similar to that of the previous aggregate equilibrium land-transportation model. Furthermore, we propose an estimation method for a multi-agent decision model and evaluate it in an empirical context. The results of the proposed model suggest the micro-interactions between land and transportation and the optimal estimation method.

Section 2 provides a review of previous LUTI models. The review will focus on developing the theory underlying the land component and the challenges in demonstrating it. Section3 proposes the microscopic interaction model and its estimation method. Section4 outlines the data used in the empirical model estimation and its background. Section 5 describes the sampling method and model specification for the estimation and then presents the results. Finally, section 6 discusses what the proposed model has achieved and the new challenges it has identified.

LITERATURE REVIEW

The theoretical model of land-transportation interaction has been studied extensively from the perspective of microeconomics. As a classic example, Alonso's (1964) (*10*) theory of urban economics assumes that the optimal locations of households and firms are determined via bid-rents (bid-rent theory). This model theoretically shows the frame in which transportation affects location by including the transportation costs of firm products in this bid-rent. In addition, Rosen (1974) (*11*) defined the hedonic price, which assumes that the attributes or characteristics of goods, especially housing, potentially affect the price (hedonic theory). In this way, the hedonic theory can take the individuality of land into account more than in urban economic theory, simplifying the urban structure. Furthermore, Mcfadden (1978) (*12*) proposed the multinominal logit (MNL) model, which is a random utility model in the context of housing location choice (utility-theory). The RUT includes observed attributes in the deterministic term of the utility function and assumes unobserved attributes and preferences in the error term. This character allows modifying the hedonic price function.

In the context of land or housing markets, these three methods have been improved to be integrated. Ellikson's (1981) bid-auction approach (*13*) is a residential choice model based on Mcfadden's (1978) MNL-based willingness to pay using hedonic theory. This bid-auction approach was also proved to be consistent with the choice process at equilibrium (*14*). This proof defines bid-rent theory as an application of hedonic theory and also enables parameterization by maximum likelihood estimation as in the logit model. At about the same time, the location choice of supply-side firms was also constructed using a logit model based on profit maximization (*15, 16*). The Random Bidding and Supply Model (RB&SM) (*17*), which assumes a logit model for both the demand and supply sides of land, has a recursive structure in the behavioral functions of the demand and supply sides. And it states that the equilibrium point can be represented by the solution vector of the immovable point problem. Since it is a system of non-linear equations, they present a solution algorithm and conclude that it can be applied by parameterizing the solution vector using maximum likelihood estimation. RB&SM integrates the externality resulting from the allocation of traffic networks by Markov Traffic Equilibrium (*18*) in the form of a bid function, and an interactive bi-level model of land and transportation is proposed (*19, 20*). Although the equilibrium algorithm has been put to practical use, only numerical simulations have been conducted at the stage when the framework was proposed.

On the other hand, in the transportation component, RUT is the theoretical basis of travel demand prediction and management of travel behavior.

In recent years, based on the bid-auction approach model, Hong Lo and his co-authors have proposed several land-transportation interaction models mainly for TOD policy evaluation, focusings on firms that simultaneously engage in railroad and housing development projects (*21–24*). These models describe the interaction between land use and transportation as a two-stage optimization problem. The optimal number of housing units to be supplied based on the objective function is the upper-level problem. In contrast, the equilibrium problem of location behavior and traffic behavior for the number of housing units to be supplied is the lower-level problem. In addition, this model is unique in that it does not use price as a direct signal. The solution is presented by equilibrium constrained mathematical programming. However, as with the general problem of land-use and transportation interaction modeling, papers leave an empirical application.

The travel demand model handled by the Transportation component of LUTI is the four-step demand model, which used to be the main approach, but has evolved into the activity-based approach in recent years (*25*). The four-step demand model is aggregated and assumes that trip purpose does not affect demand. At the same time, the activity-based approach is disaggregated and focuses on the interdependent choice of activity-travel patterns. Therefore, the activity-based approach is disaggregated and focuses on the interdependent choice of activity-travel patterns. However, the activity-based approach has a problem that the large choice set of the activity-travel pattern is too large.

A growing body of literature has proposed several models to address this issue. For example, in recent years, the activity-travel model of the Markov decision process (MDP) has been proposed, which assumes that travel patterns are generated by sequential choices, instead of sampling choices as combinations of places, purposes, and means of activities, which can lead to a large number of choices. The model of Västberg et al.(2020) (*9*) is formulated as a dynamic discrete choice model (DDCM) that can be estimated and simulated and has a theory in common with the recursive logit model (RL) (*4*) for the path choice problem and the Bellman equation that describes the future utility. Oyama (2017) (*7*) extended these dynamic discrete choice models, especially the RL model, by adding a discount factor.

As has been pointed out, the LUTI model is challenging to measure empirically, and many of the proposed models have been limited to numerical simulations (*1, 25*). The main reasons for this are the availability of data and the complexity of the calculations (*26*).

In the transportation component, the activity-based model is disaggregated and does not require zone aggregation. In contrast, all land component models assume the zone level, which means that the actual estimation and simulation will need aggregate data. The activity-based model is disaggregated and does not require zone aggregation. In contrast, both land component models assume the zone level and rely on aggregate data in empirical estimation and simulation (*2*). The subdivision of the zone also points out the heterogeneity of the data variance (*26*).

In addition, the dynamic discrete choice model for travel demand and behavior with the Bellman equation has been developed as a transportation model alone. Still, it has not been applied as a transportation component in LUTI. It is not used as a transportation component in LUTI. The common problem is that the computational load of the value function is high, and the algorithm needs to be devised. The development of an algorithm for solving the immovable point problem of these nested equations is similar to the model proposed by Martínez and Henríquez (2007) (*17*).

In this paper, we empirically integrate land transactions and transportation behavior at the microscale in the context of utility theory. By modeling based on RUT in the land component, the utility function can be expressed in land attributes. Thus, it can be treated in a disaggregated manner. In addition, by assuming land ownership and land transaction data in the zone as input data for estimating the land component model and proposing a disaggregated model in which the land owned by the user is used as a choice, the size of choice set is reduced. We propose a new recursive estimation algorithm to enable parameterization in the demonstration, which has been a challenge in the previous estimation. And then, we show the estimation results and characterize the algorithm and model. In this way, we contribute to the analysis using a microscopic land-transportation interaction model that shares the basis with previous models.

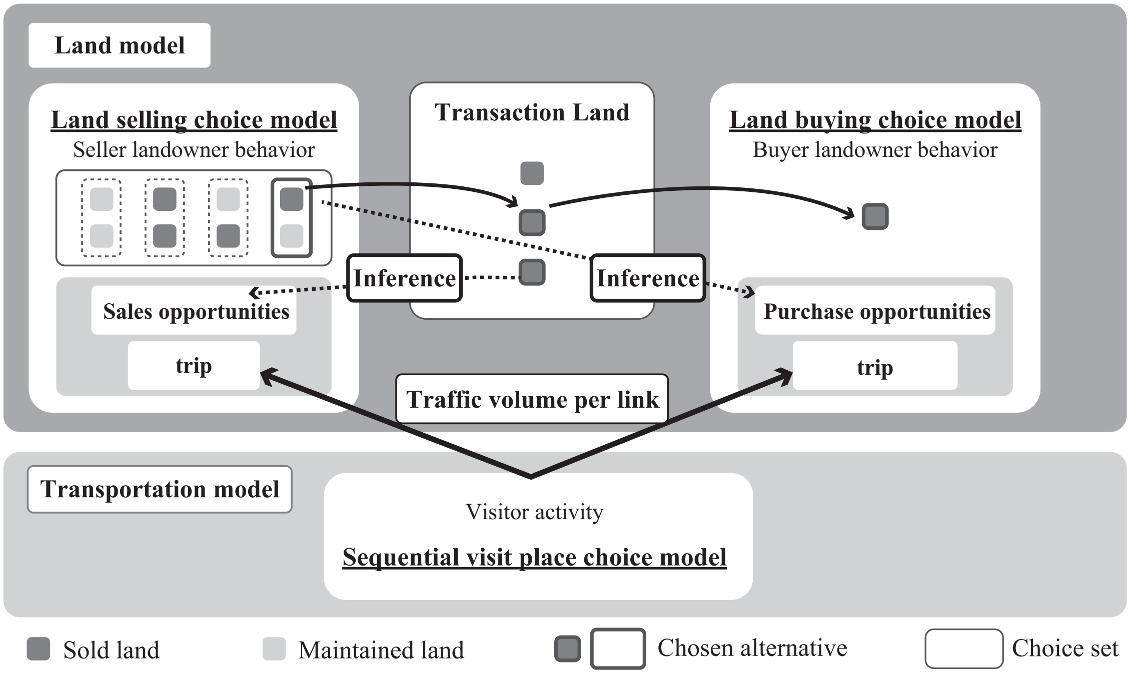

FIGURE 1 Concept of the model framework

MODEL

Land selling choice model

The landowner chooses a combination $i = \{\{i\},\{i\}^-\}$ of land to sell$\{i\}$ and land to keep $\{i\}^-$ from the set of land he owns I^s. A multinominal logit model represents the choice probability.

$$P^s_{\{\{i\},\{i\}^-\}} = \frac{e^{V_{\{\{i\},\{i\}^-\}}}}{\sum\limits_{\{\{i\},\{i\}^-\}\in I^s} e^{V_{\{\{i\},\{i\}^-\}}}} \tag{1}$$

Using these assumptions, we set the deterministic component of the utility as follow:

$$V_{\{\{i\},\{i\}^-\}} = \boldsymbol{\theta}^{\mathrm{T}}_{\{i\}^-}\boldsymbol{X}_{\{i\}^-} + \alpha_{\{i\}^-}\sum_{l,i_r\in I^s} NQ_l + \gamma_{\{i\}}\sum_{l,i_l\in I} \delta^{buy}_l(\hat{\boldsymbol{\theta}}_{buy}) \tag{2}$$

$\boldsymbol{X}$ is the vector of explanatory variables, which are exogenous variables such as land attributes. N is the exogenous traffic demand volume, and Q_l defines the choice probability of link l. Therefore, by adding the 2nd term, the model explicitly incorporates the traffic volume for each link l based on the choice probability derived by the sequential visit place choice model described below. The third term represents the estimated purchase volume on link l to which land i belongs and explicitly deals with the seller is inferring the buyer's behavior. The parameter to be estimated in this model is:

$$LL(\boldsymbol{\theta}_{sell}) = \sum_s \log P^s_{\{\{i\},\{i\}^-\}}(\boldsymbol{\theta}_{sell}) \tag{3}$$

$$\max_{\boldsymbol{\theta}_{sell}} LL(\boldsymbol{\theta}_{sell}) \tag{4}$$

$$\text{subject to} \quad \sum_{l,i_l\in I} \delta^{buy}_l(\hat{\boldsymbol{\theta}}_{buy}) \tag{5}$$

Let us consider the likelihood of the seller model. When estimating the parameters by likelihood maximization, we can obtain the parameters to solve an optimization problem with the estimated purchase volume (third term) as the constraint.

Land buying choice model

$$P_{j^b} = \frac{e^{V_j^b}}{\sum_{j \in J} e^{V_j^b}} \tag{6}$$

Using the assumptions for the buyer, we set the deterministic component of the utility as follow:

$$V_j^b = \boldsymbol{\theta}_j^{\mathrm{T}} \boldsymbol{X}_j + \alpha_j \sum_{r,j \in J} NQ_l + \gamma_j \sum_{l,i_l \in I} \delta_l^{sell}(\hat{\boldsymbol{\theta}}_{sell}) \tag{7}$$

The third term is the estimated volume of sale of the link l to which the land j belongs, inferred from the estimated parameters defined in the seller model.

Let us consider the likelihood function as well as the seller model. Then, estimates satisfying Eq. (8), expressed for the log-likelihood function, are given solutions to a constrained optimization problem conditional on the estimated sale amount.

$$LL(\boldsymbol{\theta}_{buy}) = \sum_b \log P_{j^b}(\boldsymbol{\theta}_{buy}) \tag{8}$$

$$\max_{\boldsymbol{\theta}_{buy}} LL(\boldsymbol{\theta}_{buy}) \tag{9}$$

$$\text{subject to} \quad \sum_{l,i_l \in I} \delta_l^{sell}(\hat{\boldsymbol{\theta}}_{sell}) \tag{10}$$

Sequential visit place choice model

It is based on the RL model (*4*) and the DRL model (*7*).

For agent n, we assume a Markov process that sequentially choices a place to visit and reaches the final destination. In this case, if we define the vector of places to visit based on sequential choice as a tour σ, the probability of choosing a tour is as follows:

$$P_n(\sigma_n = [s_1, s_2, \cdots, s_T]) = \prod_{\tau=1}^{T-1} P^d(s_{\tau+1} \mid s_\tau) \tag{11}$$

d is the final destination, and $P^d(s_{\tau+1} \mid s_\tau)$ is the conditional choice probability.

Let us consider a directed graph $\mathscr{G} = (\mathscr{E}, \mathscr{S})$. $\mathscr{E}$ and $\mathscr{S}$ are the set of edges and visit place, respectively.

The agent chooses from the set of possible transitional visit places the alternative that maximizes the sum of instantaneous utility and the maximum expected utility discounted by a discount factor β.

The expected utility can replace the value function of the Bellman equation.

$$V^d(s_\tau) = \mathbb{E}\left[\max_{s_{\tau+1} \in \mathscr{S}(s_\tau)} \{ v(s_{\tau+1} \mid s_\tau) + \beta V^d(s_{\tau+1}) + \mu_{s_\tau} \varepsilon(s_{\tau+1}) \} \right] \tag{12}$$

$v(\cdot)$ is the deterministic utility component characterized by the unknown parameter. The random component $\varepsilon(s_{\tau+1})$ is assumed to be an i.i.d generalized extreme value distribution with non-negative scale parameter μ. The probability of the visit place is Eq. (13).

$$P^d(s_{\tau+1} \mid s_\tau) = \frac{e^{\frac{1}{\mu}\{v(s_{\tau+1}|s_\tau)+\beta V^d(s_{\tau+1})\}}}{\sum_{s'_{\tau+1} \in \mathscr{S}(s_\tau)} e^{\frac{1}{\mu}\{v(s'_{\tau+1}|s_\tau)+\beta V^d(s'_{\tau+1})\}}} \tag{13}$$

Since we were assuming a generalized extreme value distribution, Eq. (12) can be expressed in log-sum form, and by taking exponents on both sides, we obtain Eq. (13) and (14).

$$V^d(s_\tau) = \begin{cases} \mu \ln \sum_{s_{\tau+1} \in \mathscr{S}(s_\tau)} \delta \cdot e^{\frac{1}{\mu}\{v(s_{\tau+1}|s_\tau)+\beta V^d(s_{\tau+1})\}}, & s_\tau \neq d \\ 0, & s_\tau = d \end{cases} \tag{14}$$

$$e^{\frac{1}{\mu}V^d(s_\tau)} = \begin{cases} \sum_{s_{\tau+1} \in \mathscr{S}(s_\tau)} \delta \cdot e^{\frac{1}{\mu}\{v(s_{\tau+1}|s_\tau)+\beta V^d(s_{\tau+1})\}}, & s_\tau \neq d \\ 1, & s_\tau = d. \end{cases} \tag{15}$$

To calculate the choice probability in Eq.(12), it needs to solve the Bellman equation. In this study, we describe the algorithm for solving the equation according to the RL model.

$|S| \times |S|$ matrix $\mathbf{M}^d$ and $|S| \times 1$ vector $\mathbf{z}^d$ is defined as follows:

$$z_{s_\tau} = \begin{cases} \sum_{s_{\tau+1} \in \mathscr{S}(s_\tau)} M_{s_\tau, s_{\tau+1}} (z_{s_{\tau+1}})^\beta, & s_\tau \neq d \\ 1, & s_\tau = d, \end{cases} \tag{16}$$

$$M_{s_\tau, s_{\tau+1}} = \delta(s_{\tau+1} \mid s_\tau) e^{v(s_{\tau+1}|s_\tau)} \tag{17}$$

Using $\mathbf{M}$ and $\mathbf{z}$, the value function expressed in Equation 14 can be obtained as a solution to a linear equation. Following Oyama 2017, Eq. (17) is solved by iterating the calculation until the fixed point is converged.

$$\mathbf{z} = \mathbf{M} \odot \mathbf{X}(\mathbf{z}) + \mathbf{b} \tag{18}$$

$$\mathbf{X}(\mathbf{z}) = \mathbf{z}^\beta$$

Note that the proposed land and transportation interaction model framework is consist three sub-models. And then, the interaction between the selling land choice and the buying land choice is represented by the recursive inclusion of the estimated sale or purchase volume for each link in the utility function.

Estimation algorithm

In this subsection, we present the estimation algorithms to be considered in estimating the parameters of the proposed model. First, we review the estimation algorithms for the dynamic discrete choice models in the transportation component. Then we describe the computational algorithms for the entire land and transportation interaction model framework proposed in this study.

Algorithms have been proposed to solve the parameters of dynamic discrete choice problems under the assumption of the infinite horizons, such as the nested fixed-point algorithm(NFXP: (*27*)), the nested pseudo-likelihood algorithm(NPL: (*28*)) and the mathematical programming with equilibrium constraints(MPEC: (*29*)). NFXP is an algorithm that consists of an outer-loop to search for structural parameters and an inner-loop to solve the fixed-point problem of the value function. In contrast, MPEC is an algorithm that simultaneously calculates parameters and value functions as an optimization problem with the expected value function as a constraint and the likelihood function as the objective function. Su and Judd (2012) (*29*) state that MPEC can speed up the computation because it only needs to evaluate the expected value function once, while NFXP has a computational complexity of (iterations of outer-loop)*(iterations of inner-loop). However, Iskhakov et al. (*30*) state that there is almost no difference in computational complexity between NFXP and MPEC. On the other hand, NPL is an algorithm that iteratively calculates the update of the value function and the search for structural parameters.

Incidentally, the dynamic discrete choice model assuming a finite horizon can calculate the value function by backward induction, and the fixed-point problem does not arise. When the value function is a system of linear equations, the value function can be obtained by computing the inverse matrix in the estimation of the discrete choice model of the Bellman equation under the assumption of the Markov process, which formulates the path choice problem (*4*). On the other hand, when the value function is a system of non-linear equations, NFXP is improved, and a two-step estimation is performed. (*5, 7*)

The figure 2 shows the entire estimation procedure, including the land and transportation components and how to interact.

(1) In the first step, we define the initial value and then execute the inner loop to calculate the value function according to the algorithm described above in the transport component. When the inner loop converges, the outer loop is executed to estimate the parameters of the visit place choice model using the maximum likelihood estimation method. (2) The inner and outer loops are repeated until convergence is achieved. The parameter vector at the convergence point is updated as the estimated parameters of the visit place choice model (left half of the figure). After the parameters are estimated, we estimate the volume of visits for each link.

(3) After these steps(1)-(2) of the transportation component have been carried out, the parameters of the land selling choice model are estimated using the maximum likelihood estimation method with the estimated volume of visits and purchase volume for each link as inputs. The estimated parameters are then used to estimate the land sale volume for each link. (4) Once the land selling choice model is estimated, we estimate the parameters of the land buying choice model using the estimated land sale volume per link and the volume of visits as inputs, as in step (3), and estimate the estimated purchase volume. (5) The estimation of the land selling choice model and the land buying choice model is looped in the land component until convergence is achieved. The parameter vector at the convergence point is updated as the estimated parameters of the land component. Note that when the iteration count m in the land component is 1, the observed transactions are an exogenous explanatory variable instead of the estimated purchase volume. Thus, the parameters of the land selling choice model are estimated.

We define "One-way" as a calculation procedure that can be terminated only by this step. The term "One-way" indicates that the transaction volume is given exogenously, without considering the estimated sale and purchase volume for the agent who makes the visiting choice. In contrast, the land transaction considers the volume of visits.

We show the procedure for the "Integrated" case. After step (5), we calculate the transaction coefficient TR for each link from the estimated sale and purchase probabilities. The transaction

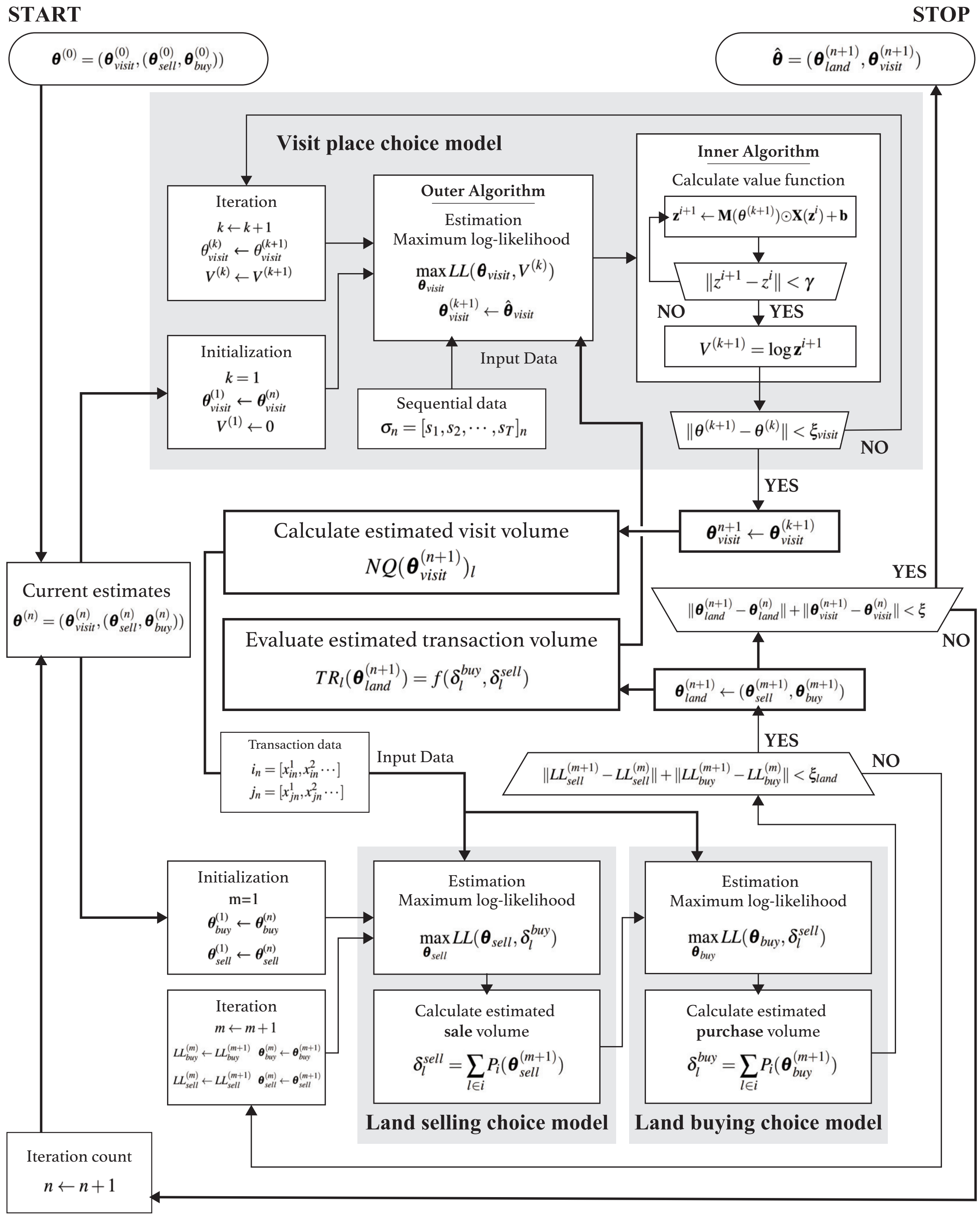

FIGURE 2 Model framework in detail as "integrated" estimation algorithm

coefficient TR is expressed by the following equation.

$$TR(\boldsymbol{\theta}_{land}^{(n+1)}) = \begin{cases} -\frac{\delta_l^{buy}+1}{\delta_l^{sell}}, & \delta_l^{buy} = 0,\ \delta_l^{buy} < \delta_l^{sell} \\ -\frac{\delta_l^{buy}}{\delta_l^{sell}}, & \delta_l^{buy} \neq 0,\ \delta_l^{buy} < \delta_l^{sell} \\ 1 & \delta_l^{buy} = \delta_l^{sell} \\ \delta_l^{buy} - \delta_l^{sell} & \delta_l^{sell} = 0, \delta_l^{buy} > \delta_l^{sell} \\ \frac{\delta_l^{buy}}{\delta_l^{sell}} & \delta_l^{sell} \neq 0, \delta_l^{buy} > \delta_l^{sell} \end{cases} \tag{19}$$

When the coefficient TR is positive, it means that the demand exceeds the supply, and when it is negative, it means that the supply exceeds the demand. The larger the absolute value of the coefficient, the more active the transactions are. Note that when the estimated sale volume is equal to the estimated purchase volume, the coefficient is set to 1. The convergence of "Integrated" is determined by using the estimated parameters of the transportation component and the estimated parameters of the land component. If it does not show a constant value, update the iteration count n and return to step (1).

DATA

Study Area

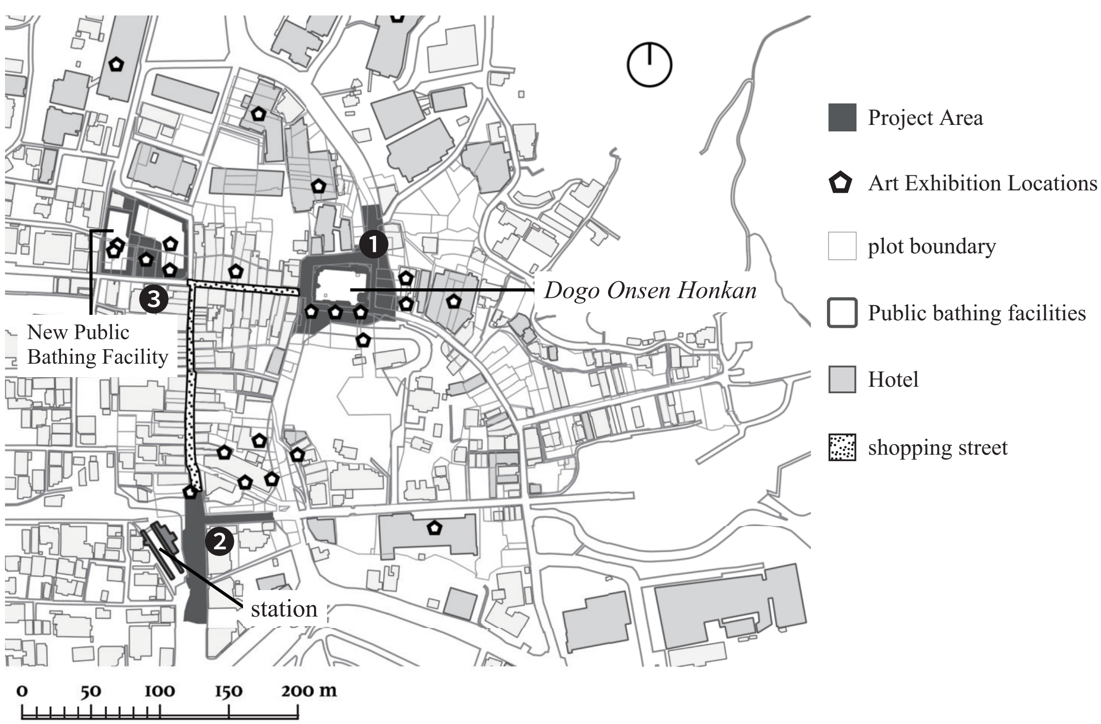

FIGURE 3 Study area with plot boundary, location of central facilities, and urban renewal project zone, 2017

The proposed model is experimented with for Dogo as a Japanese sightseeing area in Matsuyama city. The figure 3 shows the study area in this paper. This study area includes 593 landowners and 605 plots where the land survey was conducted. The major facilities in the area are also shown in the figure. The bathing facility(*Dogo Onsen Honkan*) is the center of tourism with cultural heritage value. The station is the entry point for public transportation. And the shopping street connects the station to the bathing facility and the hotel district.

The following is an overview of the urban planning project that led to the land survey and the visitor survey. From 2004 to 2009, this area underwent an urban renewal project to improve convenience and add value for tourists. Specifically, as shown in the figure, the local city of Matsuyama constructed pedestrian paths around the bathing facilities (*Dogo Onsen Honkan*) in 2007 (the area 1 in fig3). Later, it implemented the same project in front of the station in 2009(the area 2 in fig3).

In parallel with these improvements, it planned to renovate the bathing facilities (*Dogo Onsen Honkan*). First, the plan was to build a new bathing facility for tourists as an alternative facility announced in 2015. Then, the new bathing facility opened in the fall of 2017(area 3 in fig3). The renovation of the bathing facility(*Dogo Onsen Honkan*) has begun in January 2019 and is scheduled to be completed in December 2024.

Considering the above public works, we defined the period as follows. (1-1) the 1st improvement term is from April 1, 2004, to March 31, 2009. (1-2) the 1st term is from April 1, 2009, to January 30, 2013. (2-1) the 2nd improvement term is the period from January 31, 2013, to September 27. (2-2) the 2nd term is the period from September 28, 2017, to January 31, 2021.

About Data

We conducted the tourist behavior survey in two phases: 2009 (the 1st term) and 2017 (the 2nd term). In both surveys, we distributed questionnaires to tourists staying in the area and later collected them by mail. In addition, in the 2017 survey, questionnaires were also distributed to tourists when they checked into their local hotels. The items interviewed were all the places they visited from home to home, the departure and arrival times of the places they stayed, and travel mode. In the 2009 survey, we received 734 responses, and in the 2017 survey, we received 548 responses.

To model sequential visit place choice behavior, for each response, we extracted respondents who (1) had an activity chain with a hotel as origin-destination and (2) visited four or more places, including a hotel. As a result, we obtained a sample of 88 cases from the 2009 survey and 118 cases from the 2017 survey.

The data of land transaction trajectories in the Dogo area was obtained from the *Whole Real Estate Registration* provided by *Real Estate Registration Information Service* operated by the Japanese government. *Whole Real Estate Registration* includes the owner and area of each plot and the reason and date of the transaction.

We opted for transactions in which the reason for the transaction was "acquisition" or "sale" because the model assumes the sale and purchase of land. Also, we converted a list of land-owning for each landowner based on *Whole Real Estate Registration* for each year.

In Addition, we converted the "property attached map" provided *Real Estate Registration Information Service* to a polygon map to identify the plots' location and shape. By combining the *Real Estate Registration* with the plot map, we obtained the geographical and property-based information for each plot.

ESTIMATION

We present estimation results to show the differences in the performance of the estimation algorithms for the visit place choice model. After that, the difference between assuming interaction between sub-models, not assuming interaction, and assuming only one-way behavior is presented.

In the first step of this section, we show the sampling method of the set of alternatives for the land and transportation sub-models. Next, we describe the model specification.

The model was estimated according to the following procedure. Firstly, we compare the NFXP algorithm (*27*) and NPL algorithm (*28*) proposed for dynamic discrete choice problems in terms of speed and accuracy of the calculation for the proposed visit place choice model. Then, based on the calculation results, we estimate the transportation component using the algorithm with superior speed and accuracy of the calculation.

Finally, we compare the differences between the three methods in terms of convergence, accuracy, and calculation speed: the one-way method, in which only the transportation component affects the land component, and the integrated method, in which the land component also affects the transportation component. The latter is, in other words, the simultaneous estimation of the land transaction model and the visit place choice model. We also estimate the model not to consider all interactions to account for parameter bias.

Sampling methods of the choice set

In the case of the land component, we sampled a set of alternatives because the number of alternative plots is likely to be huge.

The sampling was conducted by year and by the landowner. First, in the land selling choice model, we sampled four combinations of sold plots and those that the landowner did not sell for each set of owned land, including the actual choice. The reason for sampling four alternatives was that the percentage of landowners who owned two or fewer pieces of land was 86.1% in the first period and 83.4% in the 2nd period.

While, in the land buying choice model, the sampling was done for each number of purchases. We sampled four alternatives for each choice, including the actual choice of one land. The alternatives are randomly sampled among the lands whose ownership was transferred for all reasons in the year of sale and purchase.

Specification of the utility function

A detailed description of the attribute used in each sub-model is given in Table 1.

In the land transaction model, the 1st in the deterministic utility of land selling and land buying choice model (Eq.(2) and (7)) is expressed as follows:

$$\boldsymbol{\theta}_{\{i\}^-}^{\mathrm{T}}\boldsymbol{X}_{\{i\}^-} = \theta_{\{i\}^-,CCdist.}x_{CCdist.}(\{i\}^-) + \theta_{\{i\}^-,FL}x_{FL}(\{i\}^-) \tag{20}$$

$$\boldsymbol{\theta}_j^{\mathrm{T}}\boldsymbol{X}_j = \theta_{j,CCdist.}x_{CCdist.}(j) + \theta_{j,FL}x_{FL}(j) \tag{21}$$

$x_{CCdist.}$ represents the aggregation of the land $\{i\}^-$or$j$. x_{FL} denote the length of the frontage. All the attributes are related to the usability of the land and can be measured from the map.

As a result of trials with various combinations of parameters, most land components, including AS, did not converge. In this study, we adopted as model specification the combination in which the model converges, and the utility function is composed of the largest number of variables.

In the transportation model, to compare the calculation speed of the value function con-

TABLE 1 Description of attribute variables.

Selling and buying land choice		
Parameter	Attribute	Description
CC dist.	Cluster centroid distance	Average distance from the centroid of land sale or purchase to the centroid of clusters calculated from Ward's method clustering of previously owned land(/10m)
AS	Area size	Total area indicated in the real estate registration cadastre of the land to be sold or purchase(normalized, m^2)
FL	Length of frontage	The length of the side facing the road, or 0 if the parcel does not face the road (/10m)
Est visit vol.	The volume of visits	the value obtained by allocating the estimated parameters of the visit place choice model for each year according to (*4*) and aggregating the allocation results for each link (/10)
Est land vol.	Estimated volume of buys or sales	the number of lands purchase and sale per link, calculated respectively from the land buying and selling model estimation parameters

Visit place choice		
Parameter	Attribute	Description
S	Bathing facilities	Dummy variable with 1 being Bathing facilities.
R1	Store & Shop	Dummy variable with 1 being the store & shop for visitor
R2	Historical site	Dummy variable with 1 being the historical site.
Trans dum	Dummy variable of land transactions	Dummy variable becomes 1 when a land transaction for each link to which the location belongs (for each period). For the "non-interaction" and "One-way" method
Trans land	The transaction coefficient	The land transaction coefficient defined Eq(19). Based on the estimated volume.
Dist.	Distance	Distance between the location of visit place(/100m)
PJ1	Project in front of the PB	Dummy variable with 1 being the location to visit within the project area (1) 50m buffer shown in Fig.3.
PJ2	Project in front of the station	Dummy variable with 1 being the location to visit within the project area (2) 50m buffer shown in Fig.3.

cerning the number of parameters, the deterministic utility $v(s_{\tau+1} \mid s_\tau)$ is specified as:

$$v(s_{\tau+1} \mid s_\tau; \boldsymbol{\theta}) = \theta_S x_S + \theta_{R1} x_{R1} + \theta_{R2} x_{R2} + \theta_{R3} x_{R3} + \theta_{Dist.} x_{Dist.} + \theta_{PJ1} x_{PJ1} + \theta_{Est.landvol} TR_{l,Est.landvol} \quad (22)$$

$$v(s_{\tau+1} \mid s_\tau; \boldsymbol{\theta}) = \theta_S x_S + \theta_{R1} x_{R1} + \theta_{R2} x_{R2} + \theta_{R3} x_{R3} + \theta_{Dist.} x_{Dist.} + \theta_{Est.landvol} TR_{l,Est.landvol} \quad (23)$$

$$v(s_{\tau+1} \mid s_\tau; \boldsymbol{\theta}) = \theta_S x_S + \theta_{R1} x_{R1} + \theta_{R2} x_{R2} + \theta_{Dist.} x_{Dist.} + \theta_{Est.landvol} TR_{l,Est.landvol} \quad (24)$$

$$v(s_{\tau+1} \mid s_\tau; \boldsymbol{\theta}) = \theta_S x_S + \theta_{R1} x_{R1} + \theta_{Dist.} x_{Dist.} + \theta_{Est.landvol} TR_{l,Est.landvol} \quad (25)$$

$$v(s_{\tau+1} \mid s_\tau; \boldsymbol{\theta}) = \theta_S x_S + \theta_{Dist.} x_{Dist.} + \theta_{Est.landvol} TR_{l,Est.landvol} \quad (26)$$

x_S, x_{R1}, x_{R2} and x_{R3} represent the uses of land. $x_{Dist.}$ the shortest network distance between places. x_{PJ1} is a dummy variable representing the location around the site of the urban renewal project shown in Figure 2. $TR_{l,Est.landvol}$ is the estimated transaction volume estimated from the land transaction model.

Estimation results

NFXP v.s. NPL in transportation component

Firstly, we compare NFXP and NPL algorithms for calculating the value function of the specified visit place choice model. Then, we define the likelihood as a measure of model accuracy and compare the computational speed for each number of parameters. We use data from two-time points to verify that the computational algorithm is robust to the data. There is no difference in the number of states between the number of parameters at each time point. The proposed model was estimated using R and a single core on Intel Core i9-9900K. In the past, there have been studies that simultaneously estimated the discount factor (*7*). However, since our study does not strongly focus on the effect of the discount factor, we defined $\beta = 0.1$.

For all parameter numbers, the NFXP algorithm was faster than NPL with almost the same

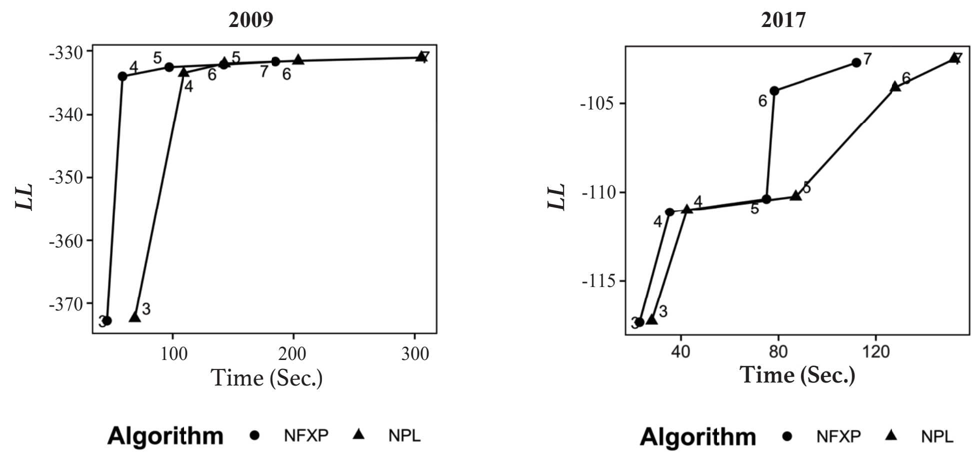

FIGURE 4 Log-likelihood of the visit place choice model over NFXP and NPL algorithm. The numbers in the figure represent the number of parameters. The number of states in 2009: 74, the number of states in 2017: 44

accuracy (Fig. 4). As the number of parameters increases, the computation time itself and the difference in computation time between algorithms increase. In general, the likelihood increases as the number of parameters increases, as the 2017 results show. Still, the 2009 results suggest that a small number of variables may explain the results. One of the limitations of the present calculation is the combination of parameters. Not all parameter combinations are estimated; however, the NFXP algorithm is more effective in computational speed and accuracy.

"One-way" v.s. "Integrated"

Next, we estimate the proposed land-transportation interaction model in three ways: one-way and integrated (Fig. 2)and non-interaction. As in Fig.4, data from two time points are used. In addition, the transportation component uses the model specified by the seven explanatory variables (Eq.22) from the results shown in Fig.4. The optimization algorithm for the land component side is the BFGS method through the standard package of R for both the land selling choice model and the land buying choice model.

We show the estimation results in Table5. These results showed that there were parameters that converged in both estimation algorithms. The computation time for Integrated was four times longer than that for One-way for the 1st and seven times longer for the 2nd term. However, when we reduced the total number of parameters of the land component to 7, the estimation time of the 2nd term was reduced to 3 times. In general, the estimation time tends to increase as the number of parameters increases, which is supported by this result. For each sub-model, the 2nd term land selling choice model and the visit place choice model show higher likelihood by the integrated method. Thus, the results show that in the proposed framework, the interaction in the direction of the response of visit place choice to land transactions improves the likelihood of each sub-model only at 2nd term. However, we can also conclude that the one-way method can also be used in a shorter time to obtain almost the same accuracy as the integrated method.

The remarkable result to emerge from the interaction term in each component is the parameter biases. In the case of non-interaction, which does not take mutuality into account, even

TABLE 2 Estimation results.

1st. Term												
Non-interaction					One-way			Integreted				
Param.	Selling	Buying	Param.	Visit	Selling	Buying	Visit	Param.	Selling	Buying	Param.	Visit
CC dist.	-2.29	-1.30	S	1.89	-2.53	-1.21	1.84	CC dist.	-2.47	-1.25	S	1.58
t-value	-6.85	-3.24	t-value	5.76	-5.71	-3.23	5.67	t-value	-5.60	-3.29	t-value	4.49
FL	2.62	0.67	R1	1.62	1.50	0.38	1.60	FL	1.59	0.51	R1	1.42
t-value	22.28	1.65	t-value	6.14	10.63	0.85	6.10	t-value	11.49	1.12	t-value	5.27
Est visit vol.	-	-	R2	1.98	2.43	-1.86	1.96	Est visit vol.	1.11	-2.11	R2	1.88
t-value	-	-	t-value	7.53	2.76	-0.55	7.45	t-value	3.91	-0.89	t-value	7.49
Est land vol.	-	-	Trans dum	0.40	-7.17	-1.03	0.55	Est land vol.	-6.32	-0.53	Trans land	0.17
t-value	-	-	t-value	2.28	-13.61	-1.41	1.87	t-value	-12.66	-0.80	t-value	1.68
			Dist.	-0.03			-0.02				Dist.	-0.02
			t-value	-0.42			-0.37				t-value	-0.34
			PJ1	0.76			0.80				PJ1	1.07
			t-value	2.72			2.86				t-value	3.26
			PJ2	1.32			1.34				PJ2	1.45
			t-value	6.11			6.15				t-value	6.16
$L(0)$	-1541.35	-79.02		-399.09	-1541.35	-79.02	-399.14		-1541.35	-79.02		-399.14
LL	-928.86	-37.93		-315.05	-722.24	-36.29	-315.73		-727.53	-37.14		-315.77
ρ^2	0.40	0.51		0.19	0.53	0.53	0.19		0.53	0.52		0.19
Sample	1938	57		88	1938	57	88		1938	57		88
Time(sec.)		41.56		162.11		1620.48	174.08					7398.94

2nd. Term												
Non-interaction					One-way			Integreted				
Param.	Selling	Buying	Param.	Visit	Selling	Buying	Visit	Param.	Selling	Buying	Param.	Visit
CC dist.	-2.80	-2.59	S	0.34	-2.52	-2.00	0.36	CC dist.	-2.55	-2.02	S	-0.05
t-value	-4.53	-1.94	t-value	0.74	-3.86	-1.40	0.77	t-value	-3.84	-1.41	t-value	-0.11
FL	2.40	0.40	R1	1.50	1.82	0.23	1.41	FL	1.84	0.22	R1	1.82
t-value	17.68	1.26	t-value	2.32	11.91	0.69	2.25	t-value	11.95	0.67	t-value	2.85
Est visit vol.	-	-	R2	0.77	0.86	-1.53	0.76	Est visit vol.	0.90	-7.73	R2	0.91
t-value	-	-	t-value	1.57	2.07	-0.40	1.55	t-value	3.80	-0.40	t-value	1.81
Est land vol.	-	-	Trans dum	0.07	-4.03	-1.33	0.55	Est land vol.	-4.02	-1.16	Trans land	-2.94
t-value	-	-	t-value	0.35	-7.80	-1.61	0.75	t-value	-7.78	-1.44	t-value	-2.53
			Dist.	0.20			0.19				Dist.	0.22
			t-value	3.87			3.83				t-value	4.09
			PJ1	1.02			0.99				PJ1	-1.36
			t-value	2.50			2.43				t-value	-1.26
			PJ2	-0.09			-0.04				PJ2	-0.21
			t-value	-0.14			-0.07				t-value	-0.33
$L(0)$	-957.45	-58.22		-122.77	-957.45	-58.22	-122.74		-957.45	-58.22		-122.77
LL	-601.65	-31.13		-107.08	-549.50	-28.77	-106.84		-538.96	-28.82		-103.57
ρ^2	0.37	0.45		0.07	0.42	0.49	0.07		0.43	0.49		0.10
Sample	1180	42		118	1180	42	118		1180	42		118
Time(sec.)		21.62		87.77		1032.85	151.473					8157.43

stronger biases emerge from the parameter ratios than in the other cases.

First, the estimated volume of visits parameter, $\alpha_{\{i\}^-,j}$, in the land component was significant in the land selling choice model for both terms and estimation algorithms: 2.43, 1.11, 0.86, 0.90.

Next, we discuss the parameters of Est land vol.$\gamma_{\{i\},j}$ in the land component, i.e., estimated purchase volume in the case of the land selling choice model and estimated sale volume in the case of the land buying choice model. Contrary to expectations, only the land selling choice model has significant negative parameters in both terms: -7.17, -6.32, -4.03, -4.02. This indicates that land with a low bought probability is sold, indicating an oversupply.

To compare the parameter bias among the estimation algorithms, the parameter ratios of Est visit vol and Est land vol were calculated. The parameter ratio for the one-way method in the 1st is -2.95, and that for the integrated method is -5.70. Thus, the integrated method was less than twice as sensitive as the one-way method. This tendency is stronger for the 2nd term. This suggests that the integrated method, with its recursive structure, contributes to removing parameter bias in the land component.

Finally, the dummy variable of the land transaction parameter(Trans dum) in the transportation component represents the visitor's attitude toward the concluded land transaction. On the other hand, the transaction coefficient parameter represents the attitude toward the estimated supply and demand. The Trans land parameter is 10% significant in the 1st term and 5% significant in the 2nd term, while the Trans dum parameter is not significant in the 2nd term. The results suggest that the recursive algorithm and the supply and demand index setting found visitors' positive or negative responses to land supply and demand activity.

In summary, we found that the proposed integrated method could reduce the parameter biases compared to the one-way method, but the computational cost was more expensive. However, even though the cost is expensive, it can be calculated within a reasonable time.

CONCLUSION

This paper has proposed the microscopic joint model framework between a land transaction by landowners and link-based behavior and the estimation method for this model.

In the general LUTI model, the land use and transportation assignment model are combined to represent the equilibrium of interaction, land use, and transportation patterns based on the land supply and demand functions. These studies have treated land transactions and transportation choice as aggregate quantities for each spatial unit of the model, a zone of about 1 km in size. In recent years, these models have evolved into activity-based models of transportation choice. However, in this paper, we proposed and demonstrated a land and transportation interaction model in which both components are at the disaggregated level.

Our model defines the plot as the unit of each land transaction and tries to relate land transactions directly to pedestrian behavior. While past studies were based on the results of zone-based questionnaires, long-term, systematic data accumulation is underway for land ownership data. Thus, we believe that this type of modeling will undoubtedly become more promising as pedestrian behavior analysis using location data becomes possible.

Essentially, the choice of streets by pedestrians and the transaction of land by landowners in response to the increase or decrease in the walking flow should mediate a dynamic interaction to balance each other's decision-making. In addition, of course, tourist pedestrians' choice of activities in urban networks is also based on dynamic programming within a limited time of stay, so there is a recursive structure between the sequences of choices by agents.

This recursive structure of land use and activity choice in urban networks required us to explicitly describe each other's micro-level interactions in the model and propose a unique algorithm to estimate the solution. If we do not describe the interactions included initially in each other's behavior, the parameter vector will contain biases.

We have been able to estimate model parameters in a reasonable time using real data. In addition, the comparison results of the value function calculation algorithm for the discounted RL model-based transportation component indicate that the NFXP algorithm is superior in terms of final accuracy and speed. For the two estimation methods, "One-way" and "Integrated," the integrated method requires more computation cost than the one-way method but can remove the bias of the parameters. In other words, by endogenizing the interaction and dynamics of land and transportation behaviors using a recursive computational approach, we succeeded in parameterization with behavioral equilibrium constraints.

These results also support that the microscopic joint model framework of land and transportation interaction is empirically valid under certain conditions. And also, We have demonstrated that the previous stochastic and random utility theory-based land and transportation interaction model can have a common theory.

The most notable limitation is that the land transaction model has not converged depending on the combination of parameters, i.e., the solution is not uniquely determined. Especially, we found that this result tends to occur strongly when the model includes the area of size parameter. Another limitation is the estimation of price. Our research has succeeded in stochastically calculating the estimated land transaction volume and estimated visitation volume for each link. However, when existing planning valuations require price estimates, we will need to define a function that reflects those probabilities as monetary values.

Compared to the previous land and transportation interaction model, which consists of an aggregate land model, our research presents a method that empirically enables the modeling of landowners' disaggregate behavior. However, the microeconomic evidence is insufficient for the previous studies.

In summary, our research provides the first basis and specific issues for a new method to quantitatively and economically analyze microscopic land and transportation interaction. Defining an activity-based model for the transportation component and demonstrating it in a large-scale census is expected to be applied to cities with other characteristics.

To solve the new issue through our works, we plan to apply learning theory to the estimation method because our research is also a parameterization of multiple stochastic models. Future works should focus on the alignment with economic theory.

REFERENCES

1. Wegener, M., Overview of land use transport models. In *Handbook of transport geography and spatial systems*, Emerald Group Publishing Limited, 2004.
2. Hurtubia, R., F. J. Martinez, and M. Bierlaire, A quasi-equilibrium approach for market clearing in land use microsimulations. *Environment and Planning B: Urban Analytics and City Science*, Vol. 46, No. 3, 2019, pp. 445–468.
3. Ljungqvist, L. and T. J. Sargent, *Recursive macroeconomic theory*. MIT press, 2018.
4. Fosgerau, M., D. McFadden, and M. Bierlaire, Choice probability generating functions. *Journal of Choice Modelling*, Vol. 8, 2013, pp. 1–18.
5. Mai, T., M. Fosgerau, and E. Frejinger, A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological*, Vol. 75, 2015, pp. 100–112.
6. Mai, T., A method of integrating correlation structures for a generalized recursive route choice model. *Transportation Research Part B: Methodological*, Vol. 93, 2016, pp. 146–161.
7. Oyama, Y. and E. Hato, A discounted recursive logit model for dynamic gridlock network analysis. *Transportation Research Part C: Emerging Technologies*, Vol. 85, 2017, pp. 509–527.
8. Zimmermann, M., T. Mai, and E. Frejinger, Bike route choice modeling using GPS data without choice sets of paths. *Transportation research part C: emerging technologies*, Vol. 75, 2017, pp. 183–196.
9. Västberg, O. B., A. Karlström, D. Jonsson, and M. Sundberg, A dynamic discrete choice activity-based travel demand model. *Transportation science*, Vol. 54, No. 1, 2020, pp. 21–41.
10. Alonso, W., A theory of the urban land market. *Papers in Regional Science*, Vol. 6, No. 1, 1960, pp. 149–157.
11. Rosen, S., Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, Vol. 82, No. 1, 1974, pp. 34–55.
12. McFadden, D., Modeling the choice of residential location. *Transportation Research Record*, , No. 673, 1978.
13. Ellickson, B., An alternative test of the hedonic theory of housing markets. *Journal of Urban Economics*, Vol. 9, No. 1, 1981, pp. 56–79.
14. Martinez, F. J., The bid—choice land-use model: an integrated economic framework. *Environment and Planning A*, Vol. 24, No. 6, 1992, pp. 871–885.
15. Anas, A., *Residential location markets and urban transportation : economic theory, econometrics, and policy analysis with discrete choice models*. Studies in urban economics / under the editorship of Edwin S. Mills, Academic Press, 1982.
16. Anas, A. and C. Chu, Discrete choice models and the housing price and travel to work elasticities of location demand. *Journal of Urban Economics*, Vol. 15, No. 1, 1984, pp. 107–123.
17. Martínez, F. J. and R. Henríquez, A random bidding and supply land use equilibrium model. *Transportation Research Part B: Methodological*, Vol. 41, No. 6, 2007, pp. 632–651.
18. Baillon, J.-B. and R. Cominetti, Markovian traffic equilibrium. *Mathematical Programming*, Vol. 111, No. 1, 2008, pp. 33–56.

19. Briceño, L., R. Cominetti, C. E. Cortés, and F. Martínez, An integrated behavioral model of land use and transport system: a hyper-network equilibrium approach. *Networks and Spatial Economics*, Vol. 8, No. 2, 2008, pp. 201–224.
20. Bravo, M., L. Briceño, R. Cominetti, C. E. Cortés, and F. Martínez, An integrated behavioral model of the land-use and transport systems with network congestion and location externalities. *Transportation Research Part B: Methodological*, Vol. 44, No. 4, 2010, pp. 584–596.
21. Ma, X. and H. K. Lo, Modeling transport management and land use over time. *Transportation Research Part B: Methodological*, Vol. 46, No. 6, 2012, pp. 687–709.
22. Ma, X. and H. K. Lo, On joint railway and housing development strategy. *Procedia-Social and Behavioral Sciences*, Vol. 80, 2013, pp. 7–24.
23. Ng, K. F. and H. K. Lo, Optimal housing supply in a bid-rent equilibrium framework. *Transportation Research Part B: Methodological*, Vol. 74, 2015, pp. 62–78.
24. Ng, K. F. and H. K. Lo, On joint railway and housing development: Housing-led versus railway-led schemes. *Transportation Research Part B: Methodological*, Vol. 106, 2017, pp. 464–488.
25. Acheampong, R. A. and E. A. Silva, Land use–transport interaction modeling: A review of the literature and future research directions. *Journal of Transport and Land use*, Vol. 8, No. 3, 2015, pp. 11–38.
26. Wegener, M., From macro to micro—how much micro is too much? *Transport Reviews*, Vol. 31, No. 2, 2011, pp. 161–177.
27. Rust, J., Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica: Journal of the Econometric Society*, 1987, pp. 999–1033.
28. Aguirregabiria, V. and P. Mira, Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models. *Econometrica*, Vol. 70, No. 4, 2002, pp. 1519–1543.
29. Su, C.-L. and K. L. Judd, Constrained optimization approaches to estimation of structural models. *Econometrica*, Vol. 80, No. 5, 2012, pp. 2213–2230.
30. Iskhakov, F., J. Lee, J. Rust, B. Schjerning, and K. Seo, Comment on "constrained optimization approaches to estimation of structural models". *Econometrica*, Vol. 84, No. 1, 2016, pp. 365–370.