Computers, Environment and Urban Systems 57 (2016) 93–105

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Agent-based joint model of residential location choice and real estate price for land use and transport model

Chengxiang Zhuge [a], Chunfu Shao [b,*], Jian Gao [c], Chunjiao Dong [d], Hui Zhang [e]

[a] Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, UK
[b] MOE Key Laboratory for Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, 3 Shangyuancun, Xizhimenwai, Beijing 100044, China
[c] School of Science, Beijing Jiaotong University, 3 Shangyuancun, Xizhimenwai, Beijing 100044, China
[d] Center of Transportation Research, The University of Tennessee, 600 Henley Street, Knoxville, TN 37996, USA
[e] School of Traffic and Transportation, Beijing Jiaotong University, 3 Shangyuancun, Xizhimenwai, Beijing 100044, China

A R T I C L E I N F O

Article history:
Received 22 August 2014
Received in revised form 3 February 2016
Accepted 5 February 2016
Available online 19 February 2016

Keywords:
Residential location choice
Real estate price
Agent-based modeling
Land use and transport
MATSim

A B S T R A C T

Residential location choice (RLC) and real estate price (REP) models are traditional and key components of land use and transport model. In this study, an agent-based joint model of RLC and REP (RLC–REP model) was proposed for SelfSim, an agent-based dynamic evolution of land use and transport model. The RLC–REP model is capable of simulating the negotiation between the active household agents (buyers) and owner agents (sellers) using agent-based modeling. In particular, both utility maximization theory and prospect theory were used to develop a utility function to simulate the location choice behavior of active household agents. The utility function incorporates only two variables: house price and accessibility. The latter variable is calculated using MATSim, an activity-based model. The asking price behavior of owner agents is based on three specific rules. The residential location choices of household agents and house prices can be obtained by negotiation. Finally, genetic algorithm was used to estimate the parameters of the RLC–REP model. The calibrated model was tested in Baoding, a medium-sized city in China, and historical validation was performed to assess its performance. The results suggest that the forecasting ability of the RLC–REP model in terms of real estate price is satisfactory.

1. Introduction

The relationship and interaction between land use and transport have been widely recognized. Land use and transport model, such as UrbanSim (Waddell, 2002), has been applied to comprehensively and systematically investigate the transport- and/or land use-related issues. As the core components of the land use and transport model, residential location choice (RLC) model and real estate price (REP) model have received increasing attention. The RLC model is focused on two topics on whether to move and where to move. The REP model is used to predict the real estate price. To date, the majority of land use and transport models have studied the RLC and REP separately (Habib, 2009; Kryvobokov et al., 2013; Moeckel et al., 2007; Waddell et al., 2003), despite a close relationship and strong interaction between them. Briefly, the residential location choice of buyers can affect the real estate price, and vice versa. Therefore, some attempts have been made for a joint study of these two issues with the so-called joint model of RLC and REP (RLC–REP model) (Ettema, 2011; Filatova et al., 2007, 2009; Magliocca et al., 2011, 2014; Parker & Filatova, 2008; Sweet, 2000). However, the RLC–REP models are limited in two aspects. First, the agent-based modeling has been recognized as a promising approach to study the RLC and REP simultaneously at the microscopic level. However, most of the agent-based RLC–REP models were tested in experiments (Ettema, 2011; Filatova et al., 2007, 2009; Magliocca et al., 2011; Parker & Filatova, 2008) and only few of them were used in real-life scenarios (Habib, 2009; Hurtubia et al., 2012), mainly because of the lack of input data, particularly micro-data. Second, the majority of RLC–REP models did not consider the phenomenon of loss aversion in housing market, which is unrealistic and could introduce bias into the simulation of residential location choice. In order to overcome these limitations, an agent-based RLC–REP model was developed with particular attention to the input data and loss aversion. In particular, the proposed RLC–REP model only incorporates two variables: house price and accessibility. The latter variable decreases the demand for the input data and facilitates easy application of the model. In addition, loss aversion is fully considered with the incorporation of prospect theory into the utility function that is used for the decision-making process of house purchase.

* Corresponding author.
E-mail addresses: zgcx615@126.com (C. Zhuge), chunfushao@126.com (C. Shao), gaojian615@126.com (J. Gao), cdong5@utk.edu (C. Dong), zhanghui3052@gmail.com (H. Zhang).

http://dx.doi.org/10.1016/j.compenvurbsys.2016.02.001
0198-9715/

2. Literature review

2.1. RLC and REP models

The decision-making process of residential location can be divided into the following two stages (Lee & Waddell, 2010): (1) The household will make a decision on whether to move, which is also known as residential mobility (Lee & Waddell, 2010) and (2) the households start searching for houses and then decide where to move (Lee & Waddell, 2010). As these stages are continuous, the questions of whether and where to move can be simultaneously studied using, for example, the nested logit model (Lee & Waddell, 2010) and stated preference approach (Kim et al., 2005).

The studies of REP can be categorized into two aspects: (1) how to predict the price and (2) what factors affect the price. The approaches to predict the REP include spatial regression model (Diao & Ferreira, 2010; Haider & Miller, 2000), hedonic model (Dorantes et al., 2011; Iacono & Levinson, 2011; Martinez & Viegas, 2009), multilevel modeling (Habib & Miller, 2008), and geographically weighted regression model (Du & Mulley, 2006). In addition, several factors have been tested to determine the extent to which they affect the REP. These factors include transport network (such as railway and highway) (Dorantes et al., 2011; Habib & Miller, 2008), accessibility (Du & Mulley, 2006; Habib & Miller, 2008; Iacono & Levinson, 2011; Martinez & Viegas, 2009), the built environment (Diao & Ferreira, 2010), and traffic volumes (Kawamura & Mahajan, 2005).

It has been recognized that the RLC and REP are interrelated and interact with each other. Therefore, several attempts have been made for a joint study of these two factors. The bid-rent model was widely used to determine the residential location and house price simultaneously. In a typical bid-rent model, the buyers bid for houses with the objective of maximizing their individual utilities, and the sellers allocate their houses to the buyers who offer the highest prices (Alonso, 1964; Huang et al., 2014). Traditionally, this model is based on equilibrium and it is assumed that every household searching for houses will be allocated and meanwhile, every house in the market will be allocated to the household. In other words, the number of vacant houses in the market is equivalent to the number of households looking for houses (Hurtubia et al., 2012). For instance, the MUSSA is a representative equilibrium-based model (Martínez & Donoso, 2004). Furthermore, the bid-rent model has been extended with the application of agent-based modeling (Ettema, 2011; Filatova et al., 2007, 2009; Magliocca et al., 2011, 2014; Parker & Filatova, 2008; Sweet, 2000). The extended model is capable of dealing with the disequilibrium housing market, as well as simulating the interactions among heterogeneous agents (Huang et al., 2014; Hurtubia et al., 2012), generally including buyers (households) and sellers (landlords).

In the agent-based RLC–REP models, in general, household agents make their purchase decisions based on the utilities of the houses, which were calculated by utility functions (Ettema, 2011; Filatova et al., 2007, 2009; Magliocca et al., 2011; Parker & Filatova, 2008). However, the majority of these utility functions did not consider the difference between the gain and loss utilities. In particular, according to the prospect theory (Kahneman & Tversky, 1979; Kahneman et al., 1991; Mohamed, 2006; Tversky & Kahneman, 1981), the loss of choosing new house decreases utility more than its increase by an equivalent-sized gain. In order to differentiate the gain utility from loss utility, some RLC–REP models have tried to incorporate the prospect theory into the utility functions (Habib, 2009; Magliocca et al., 2014).

To date, several agent-based RLC–REP models have been proposed; however, most of them were tested in experiments (Ettema, 2011; Filatova et al., 2007, 2009; Magliocca et al., 2011; Parker & Filatova, 2008) and only few of them were used in real-life scenarios (Habib, 2009; Hurtubia et al., 2012). Furthermore, there are also some models that have not even been validated (Parker & Filatova, 2008). In general, the agent-based RLC–REP models have a high demand for the micro-input data, which are not available or accessible in most cases. This is probably the main reason for the limited number of cases studies for agent-based RLC–REP models.

2.2. RLC and REP models in the land use and transport model

RLC and REP models are two essential parts of land use and transport models. Currently, the development of land use and transport models is in micro-simulation stage (Iacono et al., 2008). Therefore, this section focuses on the RLC and REP models that are used in typical micro-simulation land use and transport models, including UrbanSim (Waddell, 2002), ILUTE (Roorda et al., 2008), ILUMASS (Moeckel et al., 2007), and PUMA (Ettema et al., 2007).

2.2.1. UrbanSim

As each module in UrbanSim is treated as a plugin, they can be replaced easily according to users' demand. Therefore, various RLC and REP models are available for UrbanSim. Traditionally, the RLC model in UrbanSim was implemented using mobility model and location choice model, which generally used Monte Carlo and logit models, respectively. The REP model in UrbanSim was generally implemented using the hedonic model (Waddell et al., 2003); however, other approaches, such as alternative geographically weighted regression methodology (Kryvobokov et al., 2013), were also applied.

2.2.2. ILUTE

In ILUTE, the RLC model was also made up of mobility and location choice models, which used the discrete-time random parameter model and incorporated reference dependence to establish preference orderings for each active household, respectively. For an REP model, a multilevel model that simultaneously considers both temporal and spatial heterogeneities was developed to predict the real estate price (Habib, 2009).

2.2.3. ILUMASS

Also in ILUMASS, the RLC model composed of the residential mobility and location choice. The former was implemented by Monte Carlo simulation and the latter was a function of supply and demand of the housing market and profitability expectations. The REP model in ILUMASS was implemented by an aggregate function (Moeckel et al., 2007).

2.2.4. PUMA

In PUMA, the RLC and REP models were jointly implemented using the multiagent modeling. The unique feature of the model is that households' decisions are based on perceptions of housing market probabilities (Ettema, 2011). However, the RLC–REP model was only tested in an experiment, rather than a real-life scenario.

2.3. Comments on RLC and REP models

On the basis of the literature review, it can be found that, currently, some attempts have been made to study residential mobility, residential location choice, and real estate price simultaneously at the microscopic level using agent-based modeling. In addition, the agent-based RLC–REP models have also been applied to micro-simulation land use and transport models such as PUMA. Therefore, the proposed RLC–REP model will also become agent-based. However, the agent-based RLC–REP models are limited in the following aspects:

First, most of the agent-based RLC–REP models were tested in experiments, rather than real-life scenarios, which is probably because of the lack of input data, particularly the micro- and disaggregate data.

Second, the majority of the RLC–REP models ignored the phenomenon of loss aversion (Kahneman & Tversky, 1979; Kahneman et al., 1991; Mohamed, 2006; Tversky & Kahneman, 1981). However, it was argued that the loss of choosing new house decreases utility more