European Journal of Geography Volume 6, Number 2:69-84, June 2015
©Association of European Geographers

STRENGTHS AND WEAKNESSES OF THE MICRO-SIMULATION APPROACH TO ANALYSIS OF RESIDENTIAL MOBILITY

Jakub VOREL
Czech Technical University in Prague, Faculty of Architecture, Department of Spatial Planning,
kubavorel@gis.cvut.cz

Abstract

Over the last 20 years, urban micro-simulation models have been the focus of much research. The advent of big data and the solid theoretical base represented by random utility theory, consumer theory and operationalized by discrete choice models seemed to have opened unlimited opportunities for urban micro-simulation. However, initial attempts to replace the traditional aggregated comprehensive urban models with comprehensive micro-simulation models – e.g. ILLUMASS, ILUTE, Oregon and UrbanSim – encountered several methodological obstacles that lowered the overly enthusiastic original expectations.

The aim of this paper is to contribute to the understanding of the strengths and weaknesses of micro-simulation modeling generally, and of residential mobility modeling specifically. The following methodological issues are discussed based on a series of experimental micro-simulation models of residential mobility applied in the catchment area of the medium-sized town of Tábor in the Czech Republic: micro-data availability, methods of data disaggregation, the multicollinearity of environmental factors and the reliability of highly stochastic models.

Keywords: *residential mobility, micro-simulation modeling, discrete choice models, micro-data*

1. INTRODUCTION

Micro-simulation models reproduce processes on the macro-level by modeling many parallel processes on the micro-level (Moeckel, Schürmann, & Wegener, 2002). On the micro-level the micro-simulation modeling analyzes and reproduces the decision making of individual actors.

The micro-simulation models are considered to be an alternative to the aggregated models. Proponents of micro-simulation criticize aggregated models for ignoring the heterogeneity of individual actors and the non-linearity of their interactions. By representing only the collective properties of modeled entities and not their individual ones, the aggregated models ignore the variation of characteristics and choices made by individual actors causing the outputs of models to be biased. The non-linear nature of interactions between individual actors causes the macro-level aggregation of individual processes to accept a number of oversimplified assumptions. Alternatively, the micro-simulation models are based on the premise that macro-level outputs cannot be predicted without simulating the individual actions and interactions of actors (Orcutt, 1957).

There are a number of other arguments supporting the usage of micro-simulation models:

a) the modeling of behavior of individual human actors has a solid theoretical foundation in the microeconomic theories of random utility, consumer behavior and discrete choices;
b) explicit representation of choice processes enables the study of various external (e.g. financial) as well as internal (e.g. cognitive, physical) constraints to the decision making of individuals;
c) building upon the unique individual-centered context allows analysis of the local decision-making factors (Orcutt, 1957).

The first operational micro-simulation model (DYNASIM) is attributed to Guy Orcutt. DYNASIM helped to analyze the impact of government policies on the population based on income distribution (Orcutt, 1957). Since then, micro-simulation models have been utilized in many fields, including the study of spatial diffusion of diseases and innovations (Hägerstrand & Haag, 1973), transportation (Balmer, Axhausen, & Nagel, 2006), residential growth (Donnelly, Chapin, & Weiss, 1964) and residential choice (Kain, 1985). In the 1990s a number of comprehensive urban micro-simulation models for transportation and land-use planning purposes were developed: MASTER and SimDELTA (Feldman et al., 2007), UrbanSim (Waddell, 2002a), ILUMASS (Moeckel, Schwarze, Spiekermann, & Wegener, 2007), ILUTE (Miller, Douglas Hunt, Abraham, & Salvini, 2004), Oregon (Hunt, Abraham, & Weidner, 2010).

The advent of big data and a solid theoretical foundation for modeling human behavior seemed to open unlimited opportunities for the use of micro-simulation techniques. However, initial attempts to replace the traditional aggregated urban models with micro-simulation models encountered several serious obstacles that lowered the original overly enthusiastic expectations. This paper discusses the strengths and weaknesses of micro-simulation models applied on residential mobility. Several experimental applications of residential mobility micro-simulation models in the Czech Republic provide an empirical basis for discussion.

2. MICRO-SIMULATION OF RESIDENTIAL MOBILITY

A set of experimental micro-simulation models was developed with the aim of evaluating the potential and limit of a micro-simulation approach to modeling residential mobility in the specific context of the Czech Republic. The experimental models replicate decisions made by individuals that are related to residential mobility.

According to the stress-resistance theory a decision to change residence is broken down into two steps (Bogue, 1973); an individual first decides to relocate, and then begins to search and choose a new residence. It is assumed that a decision to relocate results from the stress level caused by the discordance between housing needs, aspirations and expectations on one side and actual housing conditions on the other. The choice of residential location is assumed to be affected by the characteristics of the new residence and its surroundings, as well as the personal characteristics of the individual making the choice (Coulombel, 2011; Lee & Waddell, 2010; Pacione, 2009).

The experiments presented here adopt this approach and develop specific micro-simulation models for each step in the decision-making process. Concerning the decision to relocate, it is assumed that the push factors associated with short-distance and long-distance relocations differ (Coulombel, 2011; Pacione, 2009). Therefore, both relocation options – to relocate inside a micro-region or outside a micro-region – are represented in the relocation choice model (RC).

The second decision step – the choice of new residential location – is examined via four experimental models. Two of which were developed to test the influence of the relocation distance on the residential location choices made by individual actors: LC_IN and LC_EX represented the residential choices of individuals relocating inside (internal population) respectively from outside the micro-region (external population).

Other two residential location choice models of internal population (LC_IN) were developed to test the effect of household lifecycle characteristic: one for young families (LC_IN/Y) and another for mature families (LC_IN/M).

The Tábor micro-region was selected for the application of experimental models. It is a typical medium-sized Czech town catchment area with an area of 1,002 km2 and a total population of 80,641 according to the 2011 General Population Census (2011 Population and Housing Census, 2013). The micro-region consists of 79 mostly small municipalities: 56% of the municipalities in the region have less than 200 inhabitants, and together they contain only 6.35% of the total micro-region population; 78.93% of the micro-region population is concentrated in the ten biggest municipalities, each having more than 1,000 inhabitants. The micro-region is dominated by the town of Tábor, the main employment and administrative center of the micro-region and the biggest municipality of the micro-region with a population of 34,430 inhabitants or 43.52% of the total micro-region population (2011 Population and Housing Census, 2013).

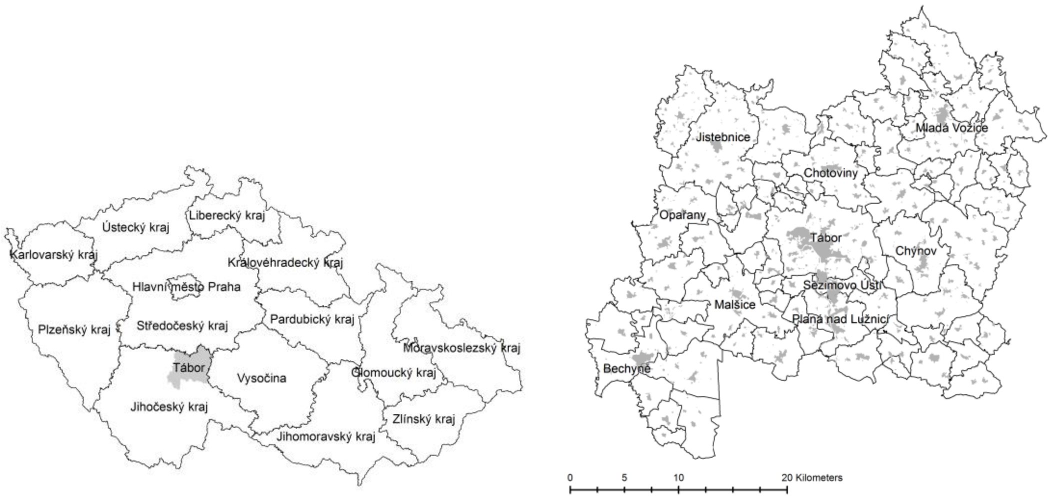

Figure 1. Tábor micro-region (MÚ Tábor, 2012)

The experimental choice models presented here were developed based on two open source software libraries: the Open Platform for Urban Simulation (OPUS) and the Biogeme ("Biogeme", 2013; Waddell, 2013b). The data was stored in the PostGIS database and visualized by QGIS (PostGIS, 2013, QGIS, 2013).

The experimental models provide empirical evidence for the following discussion regarding general strengths and weaknesses of the micro-simulation modeling approach.

3. STRENGTHS

3.1 Simple and comprehensible model definition

The micro-simulation models are intuitively comprehensible, mainly because they operate with individual actors rather than with their abstract aggregates and because they represent the urban processes in a much less abstract way than the aggregated models. Researchers usually have less difficulty imagining the motives, resources, preferences and actions of particular actors compared to the abstract concepts used in the aggregated models.

3.2 Firm foundation in behavioral theories

Discrete choice theory and random utility theory provide a general theoretical framework for the modeling of human decision-making. Discrete choice theory assumes that individual human actors are making choices from a finite number of perceived discrete options and that the options are chosen based on the actors' subjective utility. The random utility theory then explains the actor's choices based on its subjective utility. Based on the theories, the micro-simulation models allow for the analyzing and modeling of the choices made by individual actors – e.g. human actors, households, firms, organizations.

The choices of individual actors are represented by multi-nominal logit (MNL) models (Ben-Akiva & Lerman, 1985; Train, 2009). Logit models link the linear combination of $k \in K$ independent variables x_k and their associated coefficients β_k to the dependent categorical variable J, which represents the set of individual choice options $j \in J$. The logit is equal to the natural log of the odds ratio of the choice probabilities of two alternative choices: examined alternative j and reference alternative j_r.

$$logit(j) = ln\left(\frac{P(j)}{P(j_r)}\right) = \sum_{k=1}^{K} \beta_k x_k \qquad [1]$$

The independent variables x_k represent the characteristics of the choice options j, as well as the personal characteristics of the individual making the choice, and the coefficients β_k represent the effects of the characteristics on their choice probabilities $P(j)$.

The model parameters can be interpreted more directly by the exponentiation of both sides of the equation.

$$odds(j) = \frac{P(j)}{P(j_r)} = \prod_{k=1}^{K} e^{\beta_k x_k} = e^{\sum_{k=1}^{K} \beta_k x_k} \qquad [2]$$

The e^{β_k} (odds ratio) is identical to the odds of the choice probability $odds(j)$ of alternative j, which means that a unit change of independent variable x_k causes the choice probability of alternative j to change e^{β_k} times the choice probability of reference alternative choice j_r (Ben-Akiva & Lerman, 1985; Train, 2009).

The parameters β_k are estimated based on observed choices of individual actors in the past. The estimation process searches for the values of β_k that maximize the model likelihood (Ben-Akiva & Lerman, 1985; Liao, 1994; Train, 2009).

3.3 Providing insight into decision making

Discrete choice models allow the analysis of various factors influencing decision making of an individual actor that are related to the actor's unique context, i.e. personal characteristics of the actor, characteristics of the household, characteristics of the present residential location, similarities between the characteristics of the present and the potential new residential locations and the distance of the potential new residence from the present one.

Based on a review of theoretical as well as empirical research done in the field of residential mobility, a number of potential factors of residential mobility were identified and their potential effects on the decision making of an individual examined (IAURIF, THEMA, 2004, 2005, 2007; Patterson, Kryvobokov, Marchal, & Bierlaire, 2010; Vorel & Franke, 2012, 2012; Waddell & Borning, 2008).

The two tables below present odds ratios e^{β} of three models: relocation choice model RC, location choice model LC_IN for the population living in the micro-region and location choice model LC_EX for the population with a current residence outside the micro-region.

Table 1. The odds ratios e^{β} of the relocation choice RC model for the two options IN and OUT to relocate inside respectively out of micro-region; the statistical significance levels are indicated: * 0.05 (t-value 1.95), **0.001 (t-value 3.29)

Personal characteristics and characteristics of a municipality	**IN** e^{β}	**OUT** e^{β}
Constant	0.1481426**	0.04511999**
Age 0 – 9	7.626736**	4.421725**
Age 10 - 24	5.902288**	5.177954**
Age 25 - 34	9.625348**	8.07094**
Age 35 - 54	2.595435**	1.740428**
Gender (male)	0.4741678**	0.4484016**
Interaction of gender with log of municipality population size	1.063712**	1.068466**
Log of municipality population size	0.8331288**	1.065034**
Log of number of jobs per one economically active inhabitant	1.361854**	1.737465**
Distance from regional urban center - municipality Tábor (km)	0.9757066**	1.023185**

Table 2. The odds ratios e^{β} of the internal LC_IN and external LC_EX residential location choice models with their statistical significance levels indicated: * 0.05 (t-value 1.95), **0.001 (t-value 3.29)

Personal characteristics and characteristics of a municipality	**LC_IN** e^{β}	**LC_EX** e^{β}
Distance from regional urban center - municipality Tábor (km)	0.9581184**	1.030802**
Distance from local urban centers (km)	0.9923988	1.016993**
Log of average distance to train station from municipality (km)	0.4408918**	0.5793915*
Proportion of inhabitants older than 64 years	0.9755486**	0.989207**
Square root of relative increase of number of flats in municipality between 2002 and 2011	2.285352**	1.607787*
Log of municipality population size	2.33224**	2.458236**
Log of land in municipality that is developable (in percentage)	1.048156	1.062803
Log of number of jobs per one economically active inhabitant	1.118945	1.149974
Proportion of municipality area covered by forest	3.011626**	1.906445**
Log of municipality population and the age of individual 0 to 9	1.009594	1.064788**
Log of municipality population and the age of individual 10 to 24	0.9998545	1.094033**
Log of municipality population and the age of individual 25 to 34	1.013541	1.13695**
Log of municipality population and the age of individual 35 to 54	0.9726965	1.084439**
The municipality Tábor	0.4725522**	1.41011**

3.3.1 Characteristics of individuals

The personal characteristics of age, gender, education status, personal income, employment status and, if employed, the economic sector of employment proved to have a significant impact on the residential mobility of the population in the Tábor micro-region. The experimental models applied here confirmed that age significantly influences an individual's decision to relocate, with individuals between 25 and 34 years of age having the highest relocation propensity.

Men have been proven to exhibit a lower propensity to relocate than women, but this gender difference decreases with an increase in municipality population. Relocation models have also proven that individuals between 25 and 34 years of age are most attracted to municipalities with large populations, especially when relocating from a long distance.

Micro-simulation models are usually applied for the prediction of long-term changes, typically from 20 to 25 years, making it necessary to update the characteristics of individuals during the simulation period. Similar to other micro-simulation models – ILLUMASS, ILUTE, UrbanSim or SimDELTA – the experimental choice models presented here applied the exogenous transition probabilities of demographic changes that were based on the following regional demographic statistics: the probabilities of birth based on fertility rates that are specific to the age of women, the probabilities of gender of newly born individual actors and the probabilities of death that are specific to the age of individual actors. Monte Carlo sampling uses the probabilities to change the characteristics of individual actors.

3.3.2 Characteristics of households

Residential location choice is the outcome of collective decisions made by members of a household. Therefore, households are usually considered to be unitary decision-making entities in residential mobility modelling with the following characteristics typically assumed to influence their decision making: number of household members, number of children, number of economically active persons in household, age and ethnicity of head of household, household income (Axhausen, 2005a).

For the application of micro-simulation models presented here, only the data on characteristics of individuals are available. Moreover, available data does not allow aggregation of individuals into households. Therefore, households are not represented as decision-making units and demographic changes of households are not considered.

The age of individuals is the only available characteristic that indicates potential association of individuals to a type of household. It is assumed that decision making of 0-9 and 25-34 year-old individuals is interdependent, as both age groups are associated with young households. A similar interdependency is expected in the case of decision making of 10-24 and 35-54 year-old individuals who are assumed to constitute mature families. To reflect different residential location choice preferences of these two household types, two location choice models were developed: LC_IN/Y for young families and LC_IN/M for mature families.

The experimental models of two household types, presented in Table 3, indicate that the demographic characteristics of households influence their residential choice behavior, especially relocation propensity, preference for new residential development and place attachment.

3.3.3 Characteristics of present and new residential location

The experimental relocation model (RC) indicates that individuals living in more populous municipalities and in municipalities with a high share of jobs per economically active inhabitant have a higher relocation propensity, especially for long-distant relocation out of the Tábor micro-region. Out of the examined characteristics of municipalities the following characteristics increased their residential location choice probability:

- enhanced accessibility to employment, service activities and railway stations;
- increased proportion of the municipality area covered by forest;
- increased supply of new housing units built between 2002 and 2011;

- increased size of the municipality population;
- greater proportion of developable land in the municipality;
- greater number of jobs per economically active inhabitant.

3.3.4 Relocation distance

The comparison of location choice models LC_IN and LC_EX, presented in Table 2, indicates that the location choices made by individuals living inside and outside of the micro-region are based on different location preferences, mainly with regards to the Tábor municipality. Individuals living outside of the micro-region represented by LC_EX, prefer the Tábor municipality more than individuals living in the micro-region represented LC_IN.

3.3.5 Accessibility

The concept of accessibility is central to geography, urban economics and spatial interaction modeling. In the case of the Tábor micro-region, accessibility to two types of urban centers was evaluated: the regional urban center of the Tábor municipality and six local urban centers of Bechyně, Jistebnice, Chotoviny, Mladá Vožice, Sezimovo Ústí, and Chýnov. Location choice models indicate that increasing the driving distance from municipality to urban centers of both types decreased the probability that the municipality would be chosen by an individual living within the micro-region (LC_IN), while it increased that probability for individuals living outside the micro-region (LC_EX). This surprising asymmetry can be explained by the artificiality of the micro-region borders.

Accessibility to railway stations proved to be a significant factor in relocation from inside (LC_IN) as well as outside (LC_EX) the micro-region. Increasing the average distance by 1 km decreased the relative probability of choosing the municipality 0.441 times in LC_IN and 0.579 times in LC_EX.

3.3.6 The effect of place attachment

The residential choice models were used to analyze whether “place attachment” of inhabitants has a significant effect on residential location choice. The place attachment was operationalized in the form of co-location of present and new residence in the predefined municipality clusters. The clusters were composed of neighboring municipalities that were similar with respect to their social and natural characteristics. Seven clusters were created for this purpose, as presented in Figure 2: Bechyňsko, Malšicko-Opařansko, Choustnicko, Mladovožicko, Chýnovsko, Táborsko, Jistebnicko.

Figure 2. Clustering municipalities into seven clusters (MÚ Tábor, 2012)

To study the place attachment, the experimental location choice models had to be extended to include the interaction effects of present and new residential location. The interaction term was defined using odds ratios e^{β} for each cluster s:

$$odds(s) = \frac{P(s)}{P(s_r)} = \quad e^{\beta_s} e^{\beta_s^{in} I_s^{in}} \qquad [3]$$

The $odds(s)$ that express the chance that cluster s will be selected compared to the reference cluster s_r (Táborsko sub-region) is equal to e^{β_s} multiplied by the odds ratio $e^{\beta_s^{in} I_s^{in}}$, where $I_s^{in} = 1$ in the case where the individual is already living in cluster s and $I_s^{in} = 0$ if the individual is living in another cluster.

Table 3 presents odds ratios e^{β} of two location choice models: the LC_IN/Y for young families and LC_IN/M for mature families with place attachment evaluated.

Table 3. The odds ratios e^{β} of the residential location choice model for young (LC_IN/Y) and mature (LC_IN/M) families with the statistical significance levels indicated: * 0.05 (t-value 1.95), **0.001 (t-value 3.29)

Personal characteristics and characteristics of a municipality	**e^{β} for young families LC_IN/Y**	**e^{β} for mature families LC_IN/M**
Distance from regional urban center - municipality Tábor (in km)	0.9486256**	0.9304866**
Square root of relative increase of number of flats in municipality between years 2002 and 2011	1.83572**	1.289052*
Log of land in municipality that is developable (in percentage)	1.158867*	1.289052**
Log of municipality population size	2.008743**	1.865869**
Proportion of municipality area covered by forest	4.585917**	5.165439**
Log of number of jobs per one economically active inhabitant	1.763956**	1.770135**
The proportion of young age population (0-19 years old)	6.845969**	3.540613**
Bechyňsko for non-residents	0.5446973**	0.7323791
Bechyňsko for residents	14.24883**	24.36001**
Choustnicko for non-residents	0.6037126**	0.6481391*
Choustnicko for residents	6.796671**	6.412569**
Chýnovsko for non-residents	1.08377	1.207472*

Chýnovsko for residents	2.885072**	2.513359**
Jistebnicko for non-residents	1.071148	1.481487**
Jistebnicko for residents	1.211388**	2.349163
Malšicko for non-residents	1.219178**	1.404327**
Malšicko for residents	1.399388	4.142991**
Vožicko, for non-residents	0.8513891	0.9986905
Vožicko for residents	9.02857**	11.20724**

Table 3 indicates that the individuals located in clusters Bechyňsko, Choustnicko and Vožicko are characterized by the highest degree of place attachment, which means they have a relatively higher probability of relocating inside of their cluster than relocating out of their cluster. On the contrary, individuals in the other three clusters (Malšicko, Chýnovsko, and Jistebnicko) have a lower degree of place attachment, which means the probability of relocation inside of these clusters is lower for their inhabitants compared to inhabitants of the previous three clusters.

For example, individuals living in the Choustnicko cluster are 6.797 times more likely and individuals living in the Jistebnicko cluster are 1.21 times more likely to relocate within their current clusters than individuals living in the reference Táborsko cluster. This implies that the individuals living in the Choustnicko cluster had a $\frac{6.796}{1.210} = 5.616$ times higher place attachment than the Jistebnicko cluster.

Place attachment can also be illustrated by comparing the relocation preferences of all micro-region populations with regards to the selected cluster. For example, for individuals living in the Choustnicko cluster the choice probability of the same cluster is 6.797 times higher compared to the Táborsko cluster, whereas for an individual living outside of the Choustnicko cluster the choice probability of the Choustnicko cluster is only 0.603 of the Táborsko cluster choice probability. It is possible to conclude that the Choustnicko cluster has a $\frac{6.796}{0.603} = 11.258$ times higher probability of being chosen by an individual already living in this cluster than by an individual living in another cluster.

Figure 3 indicates the low attractiveness of the clusters Malšicko, Chýnovsko, Jistebnicko for their inhabitants and the relatively high attractiveness of these clusters for inhabitants of other clusters.

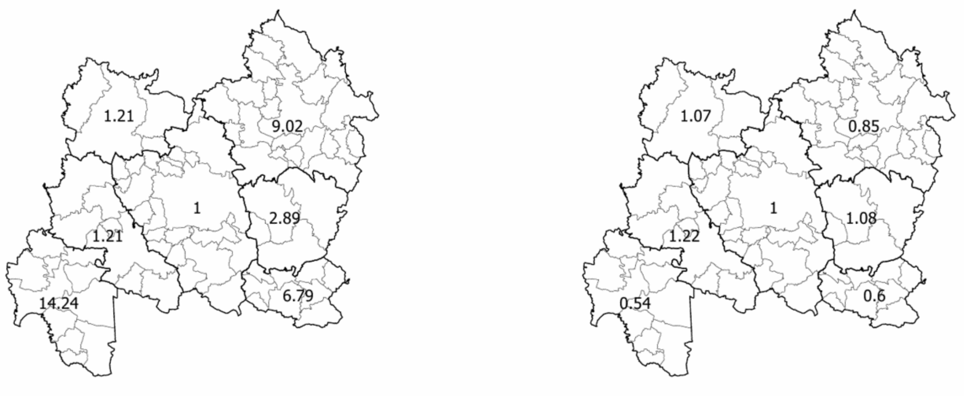

Figure 3. The relative choice probability (odds) of each cluster for its inhabitants (left) and for inhabitants of other clusters (right). Táborsko is the reference cluster for the relative choice probability of other clusters. The relative choice probability depicted is related to choices made by heads of young families.

Building residential choice models for different sub-populations revealed age as a significant factor influencing place attachment. A comparison of young families (0-9 and 25-34 years-old individuals) with mature families (10-24 and 35-54 year-old individuals) indicates a relatively stronger place attachment of mature families in all clusters except Choustnicko and Chýnovsko.

4. WEAKNESSES

4.1 Simple behavior of an individual creates complex outcomes

The non-linear interactions between individual actors lead to complex and unpredictable outcomes on the macro-level, a phenomenon known as the “emergent effect”. This kind of complexity is inherent in all bottom-up, evolutionary models. Macro-level patterns are not a priori stated nor assumed as in the case of aggregated models, but they evolve through spontaneous interaction of individual agents. Consequently, the tractability and controllability of the model behavior is difficult, which constitutes a considerable limitation for building general trust in the micro-simulation models and their acceptance in planning practice.

4.2 High stochasticity

According to random utility theory, the predictions of individual actor choices can, in principle, only be probabilistic. As a result, the number of choices of each municipality is a random variable that can be different for each simulation run. To get the expected value and variance of random variables it is necessary to run the simulation model a number of times. In the case of experimental models presented in this paper, the mean $\bar{f}_m$ and standard deviation sd_m of choices for individual municipalities M were evaluated after 100 runs and the coefficient of variation for each municipality m was derived as $CV_m = sd_m/\bar{f}_m$. The scatterplot below indicates that the stochasticity represented by the coefficient of variance CV_m is inversely proportional to the population size of the municipalities $m \in M$.

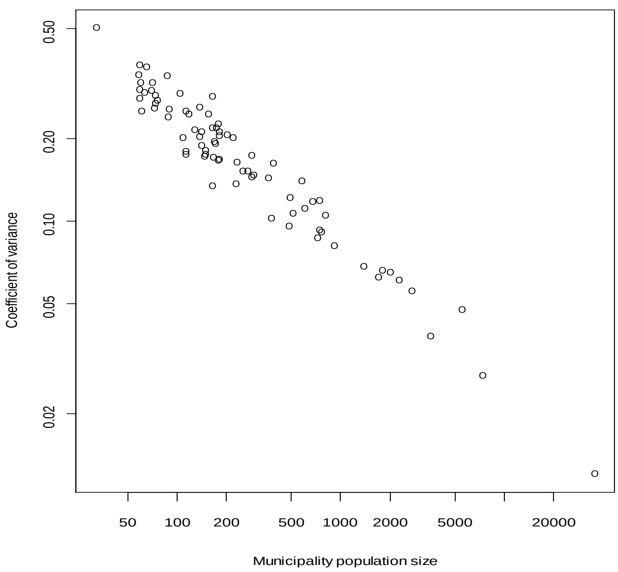

Figure 4. The relation between coefficients of variation CV and the municipality population size f_m for the selected experimental relocation choice model

The relation between stochasticity and the level of disaggregation allows for the decrease in the model stochasticity by increasing the number of choices of individual choice options (here municipality) through the spatial aggregation of choice options (Wegener & Spiekermann, 2011).

The experimental applications presented here indicate that, to attain a coefficient of variation of location choices under 10%, the choice options should have a minimum population size of 1000.

4.3 Limited availability of micro-data

The advantages of conceptual disaggregation of micro-simulation models can be exploited only when it is accompanied by appropriate micro-data. The experimental micro-simulation models presented here are based on three kinds of data: a) data on characteristics of actors making choices, b) data on choices made by actors in history and c) data on choice options.

A residential location choice is generally considered to be the outcome of collective decisions made by members of a household. Current urban simulation models typically consider households to be unitary decision-making entities and their location decisions to be influenced by household characteristics, e.g. number of members, number of children, ethnicity, economic status, occupation of economically active household members, income and number of cars used by household members (Axhausen, 2005b; Simmonds, 2010; Waddell, 2002b).

The Database of Population Relocation (DPR) – the source of data on residential mobility maintained by the Czech Statistical Office (CSO, 2014) – provides information on the residential mobility of individual persons, but does not include any of the characteristics of the household with whom the persons are associated, and therefore, the influence of household characteristics on the residential location choice of individual persons cannot be evaluated.

The reviewed residential choice models that used individual residences as the modelled choices proved to be superior to those using only the residential locations as choice options and indicated that the characteristics of individual residences and their immediate surroundings best explain the relocation choices of residents (Waddell, 2013a). Regrettably, in the Czech Republic, available data includes only the municipalities and not individual residences as residential choices made by individual actors in the past. Consequently, instead of individual residences as choice options the residential choice models use the municipalities with an average area of 12.6 km2 as choice options and all other data on residential characteristics have to be spatially aggregated to that level. This spatial aggregation leads to a significant loss of meaningful information hidden in the intra-municipality spatial variance of residential characteristics. As a consequence, the residential characteristics have insignificant and unstable effects on the observed behavior of individuals. The following residential characteristics aggregated to the municipal level proved not to be predicative of relocation choices made by individual actors: size of housing units in terms of floor area and number of rooms, type of residence (single-family house / apartment building), year of construction and quality of construction.

The lack of micro-data on household mobility and demographic characteristics, as well as demographic changes, is a typical problem faced by a number of reviewed applications of micro-simulation models. Ad-hoc household surveys would be required to complete the missing data; however, performing ad-hoc surveys is too costly and time-consuming. The alternative to the micro-data collection is the use of multi-dimensional statistics for the creation of synthetic micro-data. Such statistics, however, are not available in the Czech Republic at the moment.

4.4 Multicollinearity

Common causes and logical interconnections behind some of the residential mobility factors can manifest themselves through their high correlation. The correlated factors are difficult to analyze by means of regression statistics. The following correlated factors with $r > 0.6$ were therefore excluded from the experimental models: the characteristics of municipality population, number of public services, number of jobs, housing quality, and proportion of housing units in multi-family houses relative to all housing units in the municipality. However, the exclusion led to a reduced number of factors that could be examined.

4.5 Relatively big prediction errors

The relative error RE_m represents the relative difference between the average number of simulated $\bar{S}_m$ and observed O_m choices of individual municipalities m:

$$RE_m = (\bar{S}_m - O_m)/O_m \qquad [4]$$

Figure 5 illustrates the RE_m of location choice model based on 100 simulation runs.

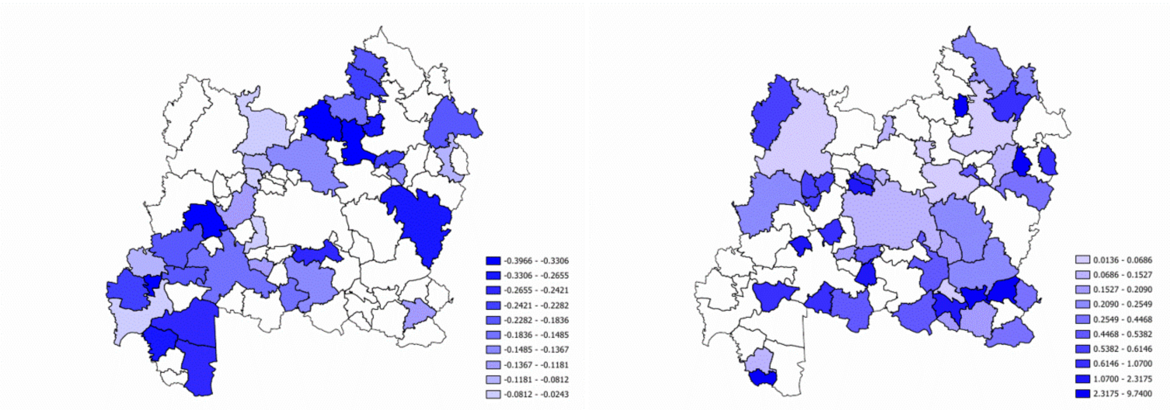

Figure 5. Relative errors of selected residential location choice model; negative errors are on the left; positive on the right.

The scatterplot in Figure 6 indicates that relative errors generally decrease with an increase in the population size of municipalities. Specifically, municipalities with a population of less than 200 are associated with high RE values.

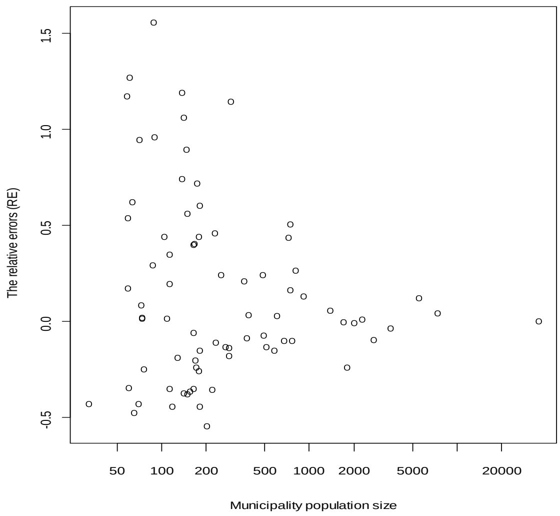

Figure 6. Distribution of relative errors of residential location choice model in relation to the municipality population size.

The high relative errors in the case of municipalities of small population size, are among others, caused by the stochasticity still being partially present in the average model residuals and the under-representation of small municipalities in the model estimation.

The spatial distribution of relative errors presented in Figure 5 indicates the model underestimates the number of location choices in municipalities that are most affected by the suburbanization process. This means that some important factors of residential mobility are still not included in the model; such as the interdependencies of the decision making of household members, the attitudes of municipalities to urban growth management and the real estate market factors.

On the other hand, underestimated relocation choices and location choices in some peripheral municipalities indicate that the catchment area of Tábor, even though corresponding well to the commuting pattern, does not correspond to the relocation pattern of inhabitants.

5. CONCLUSIONS

This paper evaluated the applicability of micro-simulation models on the phenomena of residential mobility in the context of the Czech Republic. Several residential mobility micro-simulation models were assembled for that purpose. The models provided insight into the decision making of individuals related to residential mobility by measuring the effects of examined factors. However, several methodological problems were encountered during the experimentation.

First, the unavailability of data on household location choices and characteristics does not allow for the study of the decision-making of households and instead individual persons have to be considered as autonomous decision makers. The interdependence of the residential location choices among the household members can therefore only partially be represented by assuming the association of individuals in households on the basis of their age characteristics.

Second, the multicollinearity among some of the examined factors strongly limits the evaluation of their real effects on decision making. The multicollinearity is, to a certain extent, caused by the aggregation of residential locations to large zones, which consequently leads to the reduction of inter-zonal variance of characteristics.

Third, the spatial distribution of relative errors indicates that some important factors need to be added to the discrete choice models. Especially, the occupancy of houses, real-estate prices, amount of disposable land for urban development, and land use regulations should be considered as residential mobility factors.

The use of the models for prediction opens another problematic issue. The experimental models presented here confirmed the trade-off between the level of stochasticity measured by the coefficient of variance and the number of simulated choices of individual options. The number of simulated choices is directly proportional to the population size of the municipality. The problem of stochasticity can therefore simply be solved by the ex-ante or ex-post aggregation of choice options, but that means the level of detail generally expected from micro-simulation models can never be attained.

In spite of the above weaknesses, there are good reasons for considering micro-simulation as an alternative to aggregated modeling approaches. First, unlike aggregated models, micro-simulation models establish the relation between the individual agency and the emergence of patterns on the macro-level by means of simulation. This allows the choices and related benefits to be analyzed on an individual level, which makes the micro-simulation models better suited for the assessment of equity issues of urban and regional development.

Second, by reflecting the local variability of personal and environmental characteristics, micro-simulation models offer more reliable predictions on the aggregated level compared to classical aggregated models.

ACKNOWLEDGMENTS

Funding from the Grant Agency of the Czech Republic within the framework of grant P104/12/1948 is gratefully acknowledged.

REFERENCES

2011 Population and Housing Census. 2013. http://www.scitani.cz/

Axhausen, K. W. 2005a. A Dynamic Understanding of Travel Demand: A Sketch. In M. L. Gosselin (Ed.), *Integrated Land-Use Transportation Models: Behavioral Foundations*. Oxford: Elsevier.

Axhausen, K. W. 2005b. A Dynamic Understanding of Travel Demand: A Sketch. In M. L. Gosselin (Ed.), *Integrated Land-Use Transportation Models: Behavioral Foundations*. Oxford: Elsevier.

Balmer, M., Axhausen, K., & Nagel, K. 2006. Agent-Based Demand-Modeling Framework for Large-Scale Microsimulations. *Transportation Research Record*, *1985*(1), 125–134. http://doi.org/10.3141/1985-14

Ben-Akiva, M., & Lerman, S. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand.* The MIT Press.

Biogeme. 2013. http://biogeme.epfl.ch/

Bogue, D. J. 1973. *A Methodological Study of Migration and Labor Mobility in Michigan and Ohio in 1947*. Donald.

Coulombel, N. 2011. Residential choice and household behavior: State of the Art. SustainCity Working Paper, 2.2a, École normale supérieure de Cachan.

CSO. 2014. Czech Statistical Office, Public Database. http://vdb.czso.cz/vdbvo/en/uvod.jsp

Donnelly, T. G., Chapin, F. S., & Weiss, S. F. 1964. *A probabilistic model for residential growth*. Center for Urban and Regional Studies, Institute for Research in Social Science, University of North Carolina. http://books.google.cz/books?id=TWtPAAAAMAAJ

Feldman, O., Simmonds, D., Zachariadis, V., Mackett, R., Bosredon, M., Richmond, E., & Nicoll, J. 2007. SIMDELTA: A microsimulation approach to household location modelling. Berkeley, USA: World Conference on Transport Research Society.

Hägerstrand, T., & Haag, G. 1973. *Innovation Diffusion as a Spatial Process*. University of Chicago Press. http://books.google.cz/books?id=LGxicgAACAAJ

Hunt, J. D., Abraham, J. E., & Weidner, T. J. 2010. Household Behaviour in the Oregon2 Model. In F. Pagliara, J. Preston, & D. Simmonds (Ed.), *Residential Location Choice* (s. 181–208). Springer Berlin Heidelberg. http://www.springerlink.com/content/m329837vl0888660/abstract/

IAURIF, THEMA. 2004. Modèle dynamique de simulation de l'interaction Urbanisation-transports en Région Ile-de-France. Application à la Tangentielle nord. Rapport final de la 1ère année. IAURIF, Université de Cergy Pontoise, THEMA.

IAURIF, THEMA. 2005. Modèle dynamique de simulation de l'interaction Urbanisation-transports en Région Ile-de-France. Application à la Tangentielle nord. Rapport intermédiaire de la deuxième phase année. IAURIF, Université de Cergy Pontoise, THEMA.

IAURIF, THEMA. 2007. Modèle dynamique de simulation de l'interaction Urbanisation-transports en Région Ile-de-France. Application à la Tangentielle nord. Rapport final de la 2ème phase. IAURIF, Université de Cergy Pontoise, THEMA.

Kain, J. F. 1985. *Housing and neighborhood dynamics: a simulation study*. Cambridge, Mass: Harvard University Press.

Lee, B. H. Y., & Waddell, P. 2010. Residential mobility and location choice: a nested logit model with sampling of alternatives. *Transportation*, *37*(4), 587–601. http://doi.org/10.1007/s11116-010-9270-4

Liao, T. F. 1994. *Interpreting probability models: logit, probit, and other generalized linear models*. Thousand Oaks, Calif: Sage.

Miller, E. J., Douglas Hunt, J., Abraham, J. E., & Salvini, P. A. 2004. Microsimulating urban systems. *Computers, Environment and Urban Systems*, *28*(1-2), 9–44. http://doi.org/10.1016/S0198-9715(02)00044-3

Moeckel, R., Schürmann, C., & Wegener, M. 2002. Microsimulation of Urban Land Use (Roč. 42). Prezentováno v 42nd European Congress of the Regional Science Association, Dortmund.

Moeckel, R., Schwarze, B., Spiekermann, K., & Wegener, M. 2007. MICROSIMULATION FOR INTEGRATED URBAN MODELLING. 10th International Conference on Computers in Urban Planning and Urban Management.

MÚ Tábor. 2012. 2. úplná aktualizace územně analytických podkladů pro správní území obce s rozšířenou působností Tábor. Městsky úřad Tábor - odbor územního rozvoje. http://www.taborcz.eu/uap-rok-2012/ds-1570/archiv=0

Orcutt, G. H. 1957. A New Type of Socio-Economic System. *JSTOR: The Review of Economics and Statistics, The MIT Press, 39*(1). http://www.jstor.org/discover/10.2307/1928528?uid=3737856&uid=2&uid=4&sid=21104112937253

Pacione, M. 2009. *Urban geography: a global perspective* (3rd ed). London ; New York: Routledge.

Patterson, Z., Kryvobokov, M., Marchal, F., & Bierlaire, M. 2010. Disaggregate models with aggregate data: Two UrbanSim applications. *Journal of Transport and Land Use*, *3*.

PostGIS. 2013. http://postgis.net/

QGIS. 2013. http://qgis.org/en/site/

Simmonds, D. 2010. The DELTA Residential Location Model. In F. Pagliara, J. Preston, & D. Simmonds (Ed.), *Residential Location Choice: Models and Applications (Advances in Spatial Science)*. Berlin, Heidelberg: Springer.

Train, K. E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Vorel, J., & Franke, D. 2012. *UrbanSIM: návrh metodiky pro založení, správu a aktualizaci vstupních data pro modelovací systém UrbanSim/OPUS*. Praha: Útvar rozvoje hl. m. města Praha.

Waddell, P. 2002a. UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. University of Washington.

Waddell, P. 2002b. UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. University of Washington.

Waddell, P. 2013a. Draft Technical Documentation: San Francisco Bay Area UrbanSim Application. Institute of Urban And Regional Development University of Califormia, Berkeley.

Waddell, P. 2013b. UrbanSim. http://www.urbansim.org/Main/WebHome

Waddell, P., & Borning, A. 2008. Documentation: Puget Sound UrbanSim Application. University of Washington. http://www.psrc.org/assets/2830/CUSPA_UrbanSim_Doc.pdf

Wegener, M., & Spiekermann, K. 2011. From Macro to Micro - How Much Micro is too Much? *Transport Reviews*, *31*(2), 161–177.