Brigham Young University
BYU ScholarsArchive

International Congress on Environmental Modelling and Software

7th International Congress on Environmental Modelling and Software - San Diego, California, USA - June 2014

Jun 16th, 3:40 PM - 5:20 PM

Residential mobility and model complexity - an agent-based modelling experiment on a small shrinking town in Eastern Germany

Carsten M. Buchmann
Helmholtz Centre for Environmental Research, carsten.buchmann@ufz.de

Nina Schwarz
Helmholtz Centre for Environmental Research, nina.schwarz@ufz.de

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Part of the Civil Engineering Commons, Data Storage Systems Commons, Environmental Engineering Commons, Hydraulic Engineering Commons, and the Other Civil and Environmental Engineering Commons

---

Buchmann, Carsten M. and Schwarz, Nina, "Residential mobility and model complexity - an agent-based modelling experiment on a small shrinking town in Eastern Germany" (2014). *International Congress on Environmental Modelling and Software*. 13.
https://scholarsarchive.byu.edu/iemssconference/2014/Stream-C/13

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

International Environmental Modelling and Software Society (iEMSs)
7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,
Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2014-proceedings

Residential mobility and model complexity – an agent-based modelling experiment on a small shrinking town in Eastern Germany

Carsten M. Buchmann[a], **Nina Schwarz**[a]

a: Department of Computational Landscape Ecology, Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, D-04318 Leipzig, Germany,
carsten.buchmann@ufz.de; nina.schwarz@ufz.de.

Abstract: Residential mobility is one of the major drivers of urban land use change. For urban planning it is therefore essential to understand patterns of population distribution and residential mobility in cities. Agent-based modelling has been found to be the most suitable method to investigate the complex interplay of processes, including individual location decisions. While numerous of such models have been developed in recent years, the challenge of adequate model complexity is still a hot topic. In this talk we present an agent-based model on residential mobility in the shrinking town Delitzsch in eastern Germany. The model is parameterized by a detailed data base resulting from an extensive questionnaire survey. Using this data-driven approach, it is possible to systematically explore the role of model complexity for describing relocation patterns, for example by setting agents' decision parameters either equal for all households or different for various groups, defined by household properties (like income, composition, lifestyle). In the most detailed model version, agents' decision characteristics are drawn from empirically observed distributions for the given combination of up to six household properties. Furthermore, the complexity of the demography module can be varied. Preliminary results indicate that at larger spatial scales (e.g. moves within the city) detailed parameterization of the decision process does not provide much additional information, whereas model complexity can potentially help to better understand patterns at the scale of neighbourhoods. Moreover, experiments varying structural complexity can provide guidance for further structural refinements in model design und hence help improving model realism.

Keywords*:* residential mobility; shrinking city; agent-based model; model complexity; agent heterogeneity.

1. Introduction

Understanding urban land use change is a crucial prerequisite for effective urban planning. This holds for growing and expanding cities, a well-studied process (e.g. Dietzel et al., 2005; Arsanjani et al., 2013), but also for a shrinking context, where infrastructure becoming ineffective, and empty spaces and high vacancies that bear high uncertainty for any investment, pose a particularly hard challenge for planners (Schwarz et al., 2010). Modelling can assist and guide decision-makers, and different types of models are being used for this purpose, like structural equation models or cellular automata models (see summary in Schwarz et al., 2010). Particularly in the situation of urban shrinkage, agent-based modelling has been assessed to be the only method that offers the possibility to reflect all processes that characterize the specific situation in enough detail (Schwarz et al., 2010). Going along with this advantage of describing detailed processes, one inherent characteristic of these models is their high level of complexity compared to most other modelling methods (Parker et al., 2003; An et al., 2012). Even though agent-based modelling of social systems can look back on more than two decades of experience (Gilbert, 2004), finding the right level of complexity is still a major challenge for modellers (e.g. Parker et al., 2003; Grimm et al., 2005; Gilbert, 2004; Huang et al., 2013; Voinov et al., 2014). In this paper, we apply an agent-based model that is based on an extensive data set on residential choice and mobility of individual households and explore the role of different aspects of model complexity on patterns of residential mobility. We use the city of Delitzsch, a small shrinking town in Eastern Germany, as a case study.

The aim of the project in which the presented modelling study is embedded is to develop management strategies for the shrinking town Delitzsch to become more energy-efficient. A basic prerequisite for effective investments into new energy-efficient technology and infrastructure is to get a better understanding of current and future distribution of the population within the city. For this purpose a detailed questionnaire survey on different aspects of residential mobility was conducted in order to provide data to feed into an agent-based model that explains and simulates scenarios of future patterns of population distribution and residential behaviour. Thus, the model shall primarily explain the specific situation in Delitzsch. For this reason we developed the model in a fashion that allows considering as many aspects as possible that have been found relevant for explaining residential mobility in the survey data set. This complex model is well suited to explore the role of complexity in describing individual households' residential decisions, specifically whether to move and where to move within town, since one can systematically reduce complexity while observing model outcomes (compare Gilbert, 2004; Grimm et al., 2005). Complexity in agent-based models can be categorized as (1) structural complexity (the number and the detail of aspects and processes considered, including agents and the environment) and (2) agent heterogeneity (to what extent different agents are different). While the first strongly relates to the number of model parameters the latter mostly limits the number of values certain agent-related parameters can take (the special case of different/heterogeneous behavioural rules affects both kinds of complexity). Varying agent heterogeneity can be seen as equivalent to the approach described and suggested by Cox et al. (2006) for general mechanistic models. In this paper, the authors demonstrate the reduction of complexity by setting parameters constant instead of eliminating them as a promising way to reduce model complexity without missing structural realism of the modelling system. Here we explore the role of complexity on residential mobility by varying heterogeneity of the household agents with respect to the reasons to change housing (demographic events and other stimuli) and with respect to the decision where to move including the procedure of search and evaluation of potential candidates. To test for structural model complexity, the sensitivity of model output is evaluated in dependence of two potentially important model parameters on the decision to move (incorporation of costs related to relocation and specific preferences for building types and quarters). Refinement and testing of the model is, however, still work in progress. Therefore we do not focus on absolute values, but rather on general trends and patterns in dependence of parameters and complexity settings.

2. Methods

2.1. Data on households, living units and relocation behaviour of households

The questionnaire survey comprised 60 questions on household characteristics, household living situation and both past and envisioned relocation behaviour, as well as preferences related to the choice of residential location. 1015 respondents answered the questionnaire, while the effective sample size for analysis was between 300 and 600 when considering missing values. Examples for such variables are questions like: "how many flats did you visit before your last relocation?", "what was the reason for your last move?" or "what would be the maximum acceptable rent?", as well as a list of possible pull-factors concerning the last relocation, such as "was it an important factor to have a balcony when deciding for your current home?". Analysis of these variables revealed that the following household properties can play an important role in describing different aspects of residential behaviour: Age, household type (composition: single, single + kid(s), couple, couple + kid(s)), income, lifestyle (traditional, moderate and progressive based upon Sinus® milieus), and tenure (owner or tenant). Analysis method chosen to determine the importance of household properties for specific aspects were random forest (R statistical software, package randomForest), since it is a non-parametric method that allows analysis of response variables following different distributions and delivers a measure of variable importance which is straightforward to interpret. Results were, however, mostly comparable to classical statistical modelling techniques (multiple regression, GLMs or multinomial regression, dependent on the characteristics of the respective variable of interest).

To account for the variety of relationships between household properties and relocation aspects, we developed an approach to group households differently for every relevant aspect (for example the number of visited flats or whether a balcony is an important aspect) with respect to the most important properties (important meaning statistically explaining a high share of variance of that aspect; random forest analysis). For every relevant aspect, households were grouped according to their properties beginning with the most important property until at least 20 cases were left in the survey data set, which could be used to derive the respective value for model parameterization.

The model uses lists of all living units (single family homes and flats in larger apartment buildings) and households in Delitzsch as an input. House and living unit properties with respect to sanitation state as well as the availability of balconies and gardens which can affect relocation decisions of households, are still preliminary estimates (per districts), since not all relevant information has been provided by project partners yet. Data sources were the city administration and the questionnaire survey.

2.2. The model

The model is agent-based, with individual households being the autonomous agents that have their own history (time since last move). Households are characterized by their members (all characterized by age) defining the household type, income, lifestyle and tenure. The most important properties define the household group for any relevant model aspect (see above). The environment is spatially explicit with a spatial resolution of 10x10 meters and contains infrastructure, other points of interest (e.g. green spaces, schools, shopping areas) as well as living units. The latter are characterized by building type (7 combinations of construction period and either single family home or apartment building), their number of rooms, the costs, properties like elevator, balcony, terrace/garden, etc., and their location.

A typical simulation run starts with the initialization during which households are initially distributed to living units in a simple procedure. Then every simulated year consists of (1) a call of a population model, in which households get older, are possibly subject to some event (e.g. birth) and triggers for relocations are allocated to households (section 2.2.1), and (2) a call of a relocation model in which households with a trigger try to move to a more adequate living unit (2.2.2). Interactions between household can be indirect (competition for living space) or direct (one preference potentially considered is to have “likeable” households in the neighbourhood, implemented by a similar lifestyle). Since the model shall be used for a stakeholder dialog it is implemented in C++ language, including Qt-routines for the creation of a graphical user interface.

2.2.1. Population model

The population model is executed in the beginning of each yearly time step. All persons age by one year and the model assigns triggers for relocations to households. Triggers can be based on a demographic event (e.g. birth leading to one more person and hence to the requirement of a larger living unit), on a model rule (when kids reach the age of 6 years, the household might consider moving closer to a school), or one of five other reasons often stated in the survey (e.g. “I wanted to become owner” or “I could afford a better flat”). Events and according triggers are illustrated in Fig. 1. Separation and the moving out of the parents’ home when becoming an adult initially lead to single persons without living unit (collected in a “pool” of single persons). These persons can either form a new single household, form a 2-adults household with another singe from the pool, join a single-adult household (the latter two options if the other person has similar age and lifestyle) or can emigrate (probabilities 30%, 30%, 30% and 10%, respectively).

The assignment of each event/trigger is continued until certain rates are fulfilled: Demographic rates are used to calculate number of births, deaths, separations, emigrations and immigrations. Rule based triggers follow from persons reaching certain ages. For the other triggers, the rates are based on the share of households that stated these as the reasons for their last relocation (survey), scaled to rates by a factor in order to relate these reasons with demographic rates. Since the values from the survey only indicate the number of moves that had occurred due to these reasons, they miss the cases in which these reasons (triggers) have not led to a successful move. Therefore, we scale the rates for the other triggers by a factor determined by the number of moves in the last year and the number of triggers set last year, indicating the general rate of success in releasing triggers (in the first year the factor is set to 5).

The probability of being subject to receiving a trigger depends on the time since the last move (as a correlation between the time since the last move and the propensity to move) and the stated probability to move in the next five years (both obtained from the survey data), as well as the share of households in the respective group having stated this trigger being important. Every household can only be assigned one trigger every year, and unresolved triggers are deleted every year.

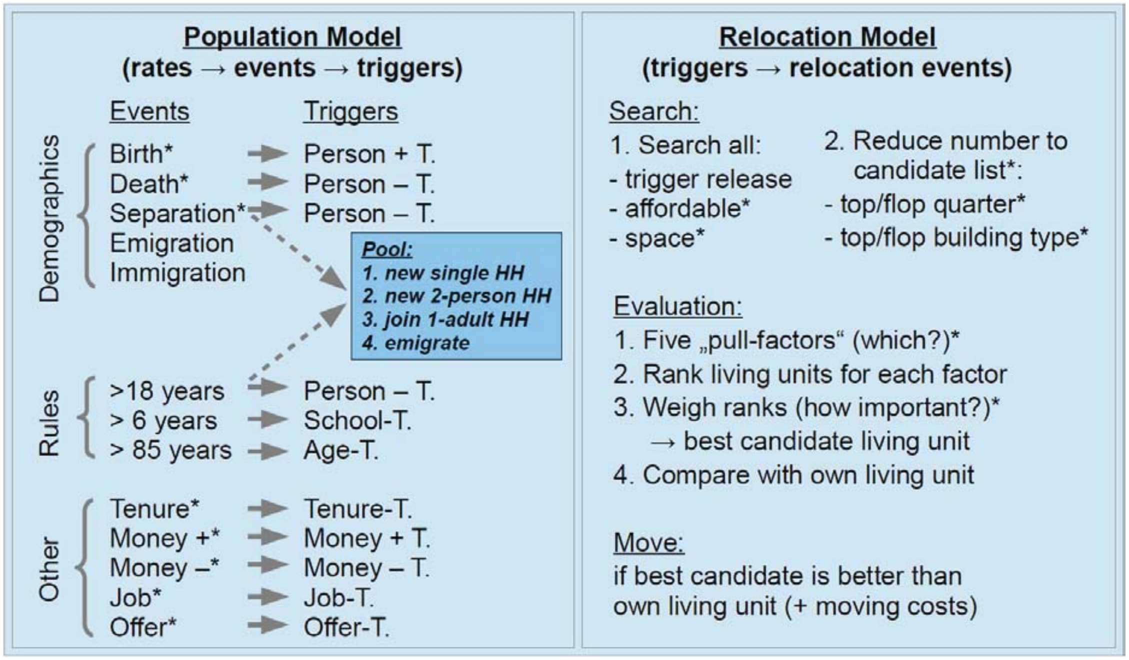

Figure 1. Schematic overview of the two main sub-models and their important components and concepts. Asterisk (*) indicates a variable that can be either homogeneous or heterogeneous among agents; "HH" stands for households, "top" and "flop" refer to the most and least favoured quarters and building types.

2.2.2. Relocation model

For all households that have received a trigger, all flats that are affordable, provide enough space (thresholds determined from the questionnaire), and which release the trigger (for example have one more room, are close to a school), are stored in a list. Only living units for newly created households from the pool as well as households triggered with "there was a good offer" do not have to fulfil the trigger constraint. In the next step, this list is reduced to the "number of flats visited", drawn from the empirical Poisson distribution (survey) accounting for the three most favoured and three least favoured quarters as well as the three most and least favoured building types. For determining the preference for certain quarters, the property of the quarter itself (do households generally like to stay there or leave?) is combined with the household-specific preference for certain quarters. The probabilities at which the affectivity towards quarters and building types are considered or ignored and are free model parameters.

All remaining living units in the list are first evaluated against each other by ranking them with respect to the five most important pull factors, weighing the ranks by the relative importance of these pull-factors and summing the weighted ranks in order to determine the best candidate. Weights are equivalent to the percentage of households in the respective group that have stated the aspect to be important. This living unit, the best of the candidate list, is then compared with the currently inhabited living unit by the same procedure and the household moves to the new location if the score of the alternative is higher than the score of the actual living unit plus X%. X can be interpreted as the moving costs and is a free model parameter.

2.2.3. Simulation scheme

For simplicity, simulations shown here stem from a scenario with a population number of 26,000 being constant over time due to equal birth and death rates as well as immigration and emigration rates, respectively (values in the range of statistical data for Delitzsch, 2012). Simulation time is 25 years with the first five years being considered as 'burn in' allowing the system to stabilize to some extent.

The role of model complexity is addressed on the one hand by testing sensitivity of parameters and the role of decision aspects in controlling model results, specifically the moving costs and the probabilities of considering house and quarter preferences. On the other hand, applying the grouping approach, we can describe household agents in a very complex empirically driven heterogeneous way. This affects the population model (the probability of getting certain triggers and the general propensity to move plus its dependence on the time since the last move) as well as various aspects of the relocation model (length of candidate list, thresholds for the cost and space ratios, the list of pull-factors accounted in the evaluation, the importance of these factors for weighting them). We systematically investigate the role of this heterogeneity by comparing results of the complex model with simulations of simple homogeneous agents only in one, or in both, the population and the relocation model.

3. Results and discussion

The number of moves simulated by the model is generally lower than data for Delitzsch from the last five years indicate (see Fig. 2, values for Delitzsch in figure caption). This is probably due to the fact that we do not consider all possible triggers that may play a role, but only those we could identify from the survey, and the fact that the environment in our simulations is static, whereas in reality changing infrastructure, service providers etc. also trigger relocations. Moreover, the number of moves decreases over time, indicating that households manage better and better to meet their specific preferences. Interestingly, the trend is consistent between models of different complexities at the scale of the whole city (parallel lines in the left panel of Fig. 2), with the complex population model (heterogeneous agents) yielding lower numbers than the simple population model, because heterogeneous agents have lower specific probabilities for certain triggers, while homogeneous agents all have a probability of 1.

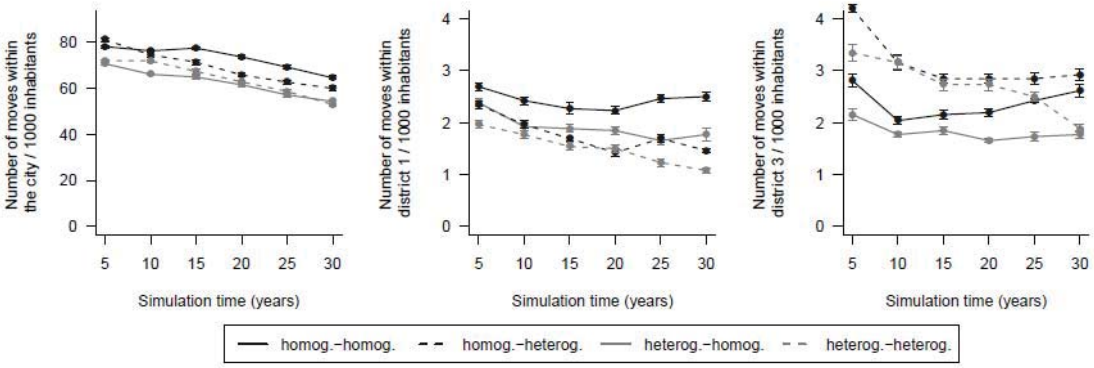

Figure 2. Simulation results on the number of moves within the city and the two largest districts in the city per 1,000 inhabitants. Analysis of data on moves in Delitzsch in 2005, 2009, 2010, 2011 and 2012 and from the questionnaire survey yields comparable values (+/- SE) of 73.3. (+/- 2.8), 8.4 (+/- 0.8) and 12.3 (+/- 1.3), respectively. For line style see legend (for example "homog.-homog." stands for homogeneous agents in the population model as well as in the relocation model). Error bars indicate the standard error of the mean of 20 replicates.

Results of simulations with heterogeneous agents in the relocation model are more complex, which we can better understand when looking at the number of moves within two selected districts (the largest in terms of population, middle panels in Fig. 2). The complex (heterogeneous) relocation model is able to better explain a clear difference between both districts, which indicates that group specific household preferences can be necessary to explain spatial patterns at finer scale, while this effect might be blurred at larger scales. However, results on the district scale still show too small numbers for moves (empirical mean values of 5 years: 8.4 and 12.3 moves per 1,000 inhabitants).

The model reveals different sensitivity towards different model parameters: Figure 3 shows that an increase in movement costs drastically reduces the number of residential relocations, variation in costs between 0 and 0.1 has nearly no effect on the number of moves. In contrast, the relevance of the aversion to or preference for certain house types and quarters shows no effect on the number of moves in the two largest districts. The described sensitivity patterns are independent of the level of heterogeneity applied for the household agents.

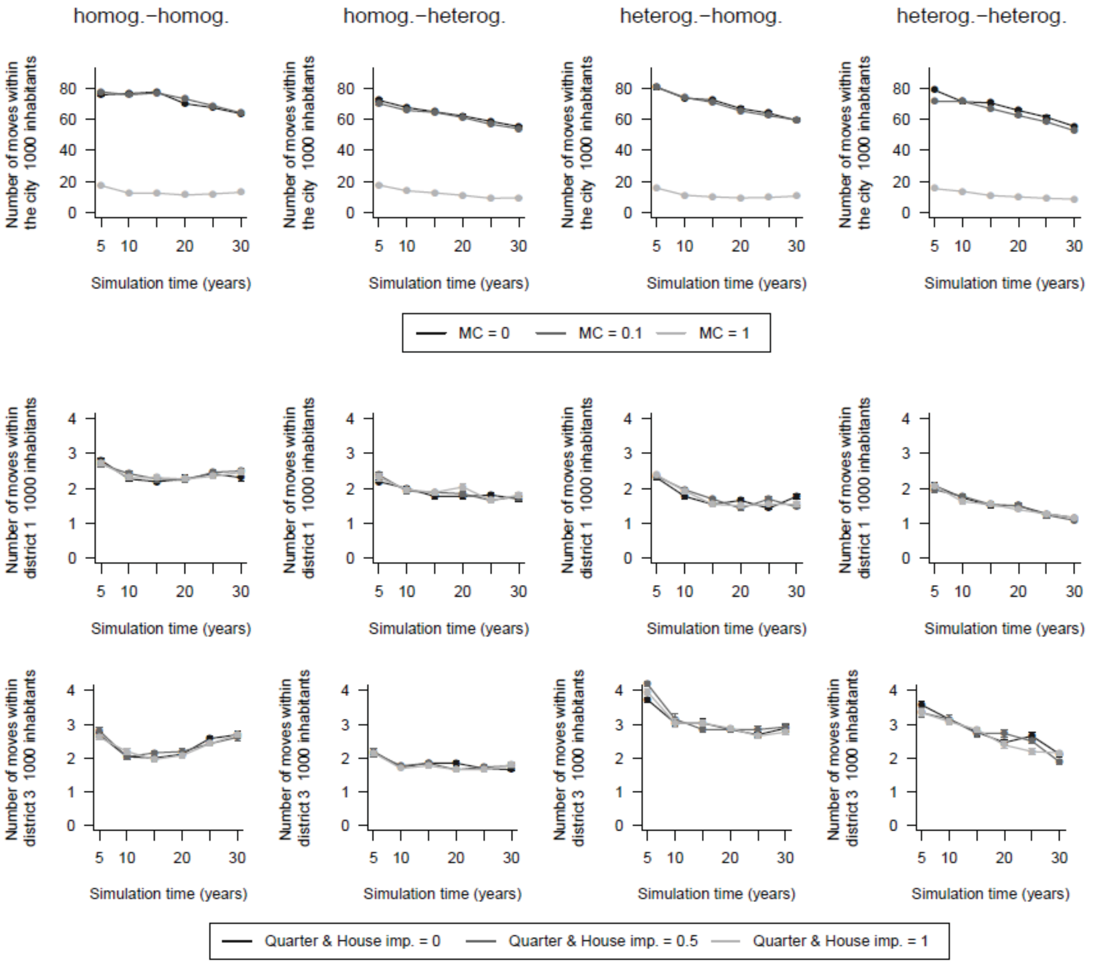

Figure 3. Simulation results on the number of moves within the city of model runs in which the relative movement costs (MC) were varied (upper row), and on the number of moves within the two largest districts in dependence on the probability of preferences for house type and quarter being considered for the decision, i.e. the importance of these aspects for the decision (rows 2 and 3). Tested aspects are illustrated as different colours (see legend); different columns are for models with different agent heterogeneity.

4. Conclusions

Being embedded into a large trans-disciplinary project providing a good data base on household behaviour offers us the opportunity to compare an agent-based model on residential mobility characterized by a comparatively high level of agent heterogeneity with simplified model versions. While previous studies working with mostly structurally simpler models have found agent heterogeneity to be a crucial feature also for aggregated patterns (e.g. Brown and Robinson, 2006, see also references in Huang et al. 2014), preliminary results of our model suggest that heterogeneity might be particularly important at finer spatial scales. Differences between complex and simple model versions (in terms of agent heterogeneity, but also for structural complexity as the consideration of house and quarter preferences) are, however, surprisingly low – also for patterns of income distribution or vacancy rates (results not shown). Besides the explanation that these aspects of complexity provide little added value, such a finding can also indicate either that other components of the model might outplay the expected effects or that the current implementation might not be a good representation of the crucial aspects of the real process. For example, in our modelling experiment, the small effect of the preference for a specific house type or quarter possibly could be due to the actual implementation which is to only consider the three most and least favoured options. Since these do not extraordinarily differ between household groups, considering small differences in the level of acceptance for the least favoured options seems to be another implementation worth to be tested before stating that this issue is not important for residential mobility in Delitzsch. Little effect of agent heterogeneity might in our case also be caused by a strong and overriding effect of the initial distribution of households, for which group specific preferences are not considered.

The principle of parsimony suggests the ‘simplest possible model’ as adequate, and most scientists agree upon the fact that simpler models tend to give greater insights of general validity (Gilbert, 2004; Grimm et al., 2005; Parker et al., 2003; Voinov et al., 2014). Besides that, simple models require significantly less computational resources (in our example simulation time of the most complex model version is approx. 8 times the simulation time of the most simple version). Nonetheless, simulation models, particularly agent-based models, still become more and more complex (e.g. Edmonds and Moss, 2005; Gaube and Remesch, 2013). Moreover, incorporation of more detailed data and adding more detail in the model structure (both aspects increase complexity) are often stated as possible or promising solutions if simulation results fail to meet expectations. These facts demonstrate the ongoing controversy and debate on the right level of model complexity (compare An et al., 2012). With our experiment, we could so far not answer the question on the most adequate level of complexity. However, our preliminary results indicate that agent heterogeneity can have different effects on patterns at different scales. Moreover, (missing) sensitivity of results to different levels of structural complexity can give helpful guidance for rethinking model design aspects and thus improving model structure. We see potential for the chosen approach of testing the effect of model complexity on results of different scales, by reducing complexity in terms of agent heterogeneity and model structure, also for other scientific fields in which agent-based modelling is applied. This is particularly the case in complex systems like (non-urban) land-use change where different agent properties can be expected to affect different decisions or aspects of decisions differently, and hence agent heterogeneity potentially is a crucial model feature.

REFERENCES

An, L., 2012. Modeling human decisions in coupled human and natural systems: Review of agent-based models. Ecol. Modell. 229, 25–36.
Arsanjani, J.J., Helbich, M., Vaz, E.D.N., 2013. Spatiotemporal simulation of urban growth patterns using agent-based modeling: The case of Tehran. Cities 32, 33–42.
Brown, D.G., Robinson, D.T., 2006. Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. Ecol. Soc. 11.
Cox, G.M., Gibbons, J.M., Wood, A.T.A., Craigon, J., Ramsden, S.J., Crout, N.M.J., 2006. Towards the systematic simplification of mechanistic models. Ecol. Modell. 198, 240–246.
Dietzel, C., Herold, M., Hemphill, J.J., Clarke, K.C., 2005. Spatio-temporal dynamics in California’s Central Valley: Empirical links to urban theory. Int. J. Geogr. Inf. Sci. 19, 175–195.

Edmonds, B., Moss, S., 2005. From KISS to KIDS - An 'anti-simplistic' modelling approach, in: Davidsson, P and Logan, B and Takadama, K (Eds.), Multi-agent and multi-agent-based simulation. Springer, Berlin, Germany. pp. 130–144.
Gaube, V., Remesch, A., 2013. Impact of urban planning on household's residential decisions: An agent-based simulation model for Vienna. Environ. Model. Softw. 45, 92–103.
Gilbert, N., 2004. Agent-based social simulation : dealing with complexity. Complex Syst. Netw. Excell. 9, 1–14.
Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310, 987–91.
Huang, Q., Parker, D.C., Filatova, T., Sun, S., 2014. A review of urban residential choice models using agent-based modeling. Environ. Plan. B Plan. Des. 40, 1–29.
Parker, D.C., Manson, S.M., Janssen, M. a., Hoffmann, M.J., Deadman, P., 2003. Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review. Ann. Assoc. Am. Geogr. 93, 314–337.
Schwarz, N., Haase, D., Seppelt, R., 2010. Omnipresent sprawl? A review of urban simulation models with respect to urban shrinkage. Environ. Plan. B Plan. Des. 37, 265–283.
Voinov, A., Seppelt, R., Reis, S., Nabel, J.E.M.S., Shokravi, S., 2014. Values in socio-environmental modelling: Persuasion for action or excuse for inaction. Environ. Model. Softw. 53, 207–212.