Feeding the Beast: Can Computational Demographic Models Free Us from the Tyranny of Data?

Eric Silverman, Jakub Bijak, and Jason Noble

Institute for Complex Systems Simulation
Centre for Population Change
School of Social Sciences
University of Southampton
e.silverman@soton.ac.uk

Abstract

Since its inception, ALife has moved from producing large numbers of highly-idealised, theoretical models towards greater integration with empirically collected data. In contrast, demography — the interdisciplinary study of human populations — has been largely following the principles of logical empiricism, with models driven mainly by data, and insufficient attention being paid to theoretical investigation. Such an approach reduces the ability to produce micro-level explanations of population processes, which would be coherent with the phenomena observed at the macro level, without having to rely on ever-increasing data demands of complex demographic models. In this paper we argue that by bringing ALife-inspired, agent-based methods into demographic research, we can both develop a greater understanding of the processes underlying demographic change, and avoid a limiting over-dependence on potentially immense sets of data.

– But you are paying a lot of money for the dragon!
– And what, should we just give it to the citizens instead? [...] I see you know nothing about the principles of economics! Export credit warms up the economy and increases the global turnover.
– But it also increases the dragon as such – I stopped him. – The more intensely you feed him, the bigger he gets; and the bigger he gets, the higher his appetite. What kind of a calculation is it? He will finally devour you all!
Stanisław Lem, *Pożytek ze smoka [The Use of a Dragon]* (1983/2008: 186)

Introduction

After attending the very first Artificial Life conference in 1987, the evolutionary biologist John Maynard Smith famously quipped that ALife appeared to be "fact-free science". His comment was made in response to early ALife work (see, e.g., Langton, 1989) that tended to be abstract and conceptual, not to mention ontologically ambitious, making no connection to empirical data in the conventional sense.

Over time, the early enthusiasm for highly abstract models in ALife has lessened somewhat, as it has become increasingly clear that making such models empirically relevant involves a highly contentious theoretical commitment to artificial life as an instantiation of biological life (Silverman and Bullock, 2004). Instead, abstract and conceptual ALife models have come to be viewed as tools for theoretical enquiry (Di Paolo et al., 2000), i.e., ways of explaining the qualitative dynamics of complex systems. At the same time, some modellers under the ALife banner have moved toward a greater connection with empirical data (e.g., Toquenaga et al., 1995; Smith V., 2008). ALife has experienced greater scientific respectability, we maintain, due to the collective recognition that modelling and simulation stand alongside theory generation and data collection in the scientific cycle — or, as Rossiter et al. (2010) put it, models are "first class citizens of science".

Thus, ALife has been in a somewhat unique position: starting from methods almost completely disconnected from empiricism, the field has gradually moved toward a greater integration with empirical data, while retaining a focus on using simulation as a tool for theoretical investigation. In this paper we consider a discipline which appears to be following the opposite trajectory. Demography — the interdisciplinary study of the development of human populations — has long been a field devoted to predictive statistical modelling based on vast storehouses of data, while theory-building has mostly taken a back seat.

Demography's intense devotion to data has served the field well when making projections of future demographic change in human populations. Nevertheless, traditional demographic methods struggle to develop well-founded *explanations* of these changes, going beyond simple generalisation of the observables (Burch, 2003). One of the motivations driving ALife's shift toward greater connection to empirical data has been the recognition that neither theory nor data alone are enough to provide coherent explanations of phenomena. ALife has addressed this dilemma by bringing more data into a largely theory-focused modelling enterprise, and we propose that demography, in order to develop beyond its current epistemological limits, must also make a move towards the centre by incorporating conceptual and theoretical investigation into its heavily data-focused framework.

The scientific benefit to such an approach would be the enrichment of the theoretical foundations of demography. In this paper however we will also discuss another, perhaps more pragmatic, advantage to ALife-inspired demographic models: as a means for escaping some of the burdens of the time-consuming and combinatorially expensive data collection required to continue in the traditional fashion.

We begin our discussion in the next section with a summary of demography's struggles with its data-collection demands. We then move on to suggesting some potential applications of agent-based models for demographic research, describing the relevant strengths and weaknesses of the approach. Next, a detailed analysis of several demographic simulation models allows us to develop a more nuanced understanding of how agent-based models may provide new utility and insight. Finally, we offer our conclusions, and suggest some directions for future work in this area.

Motivation: Meet The Beast

In the context of large, policy-focused projects in social science, modelling and simulation in some form has become ever more important as a means of providing useful information to stakeholders. Models provide a means of producing predictions or characterisations of complex systems which can give the stakeholder what they need: a target number, a summary of current numbers, or numbers to be wary of. However, many such modelling projects can become quite large and unwieldy. We often find that we require extensive amounts of data in order to feed into a large-scale model (hereafter, for illustrative purposes, referred to as 'the beast'), and the process of collecting that data is inevitably expensive and time-consuming. Plus, as our models get increasingly complex, the beast becomes ever hungrier.

Demography offers a unique predictive potential given the information embodied in the age structure of populations. However, for reasons we will discuss later, these predictions still remain largely uncertain. In an effort to alleviate some of the epistemological limitations, recent work in demography has attempted to bridge the gap between micro- and macro-level analysis (Courgeau 2007 and the MicMac project — see Willekens 2005 and Zinn et al. 2009). Advances in event-history analysis and microsimulations linked with multilevel statistical analysis have been offered as potential solutions to the micro-macro divide. However, such methods still have one major weakness: potentially enormous requirements for data due to the 'combinatorial explosion' of the parameter space. So, even as extended modelling frameworks, such as MicMac, try to bridge the micro-macro gap by producing linked simulations at both levels of analysis, we still find ourselves hamstrung by the need for large amounts of data.

Thus, we see demography reaching for more sophisticated modelling paradigms and for ways to produce more micro-level explanations of factors that drive population change. Unfortunately, current modelling methods require us to continue 'feeding the beast': pumping models full of ever-increasing amounts of data, each dataset requiring vast amounts of resources (time and money) to collect. This has further impacts on the overall modelling enterprise: turn-around time for producing models grows out of control; stakeholders find themselves confronted with nigh-incomprehensible models and endless reams of data; and the primacy of post-hoc statistical analyses inevitably leads us toward certain types of models which seem to fit the data well.

In the other part of the methodological spectrum, the use of agent-based models à la ALife in recent years has become increasingly popular in certain areas within the social sciences. Starting from Schelling's (1978) famous residential segregation model, and moving on to Axelrod's *Complexity of Cooperation* (1984), Cederman's (1997) work on international relations, and the current wide spectrum of agent-based models in social science (cf. Epstein, 2008 or Gilbert and Troitzsch, 2005), the prospect of using agents to examine properties of human societies which are difficult or impossible to measure empirically has become increasingly attractive. Understandably, many social scientists are excited by the possibility of examining fundamental properties of social phenomena without being forced to devote excessive resources to primary data collection.

To date, much of the work extolling the virtues of agent-based models for the social sciences have focused on the potential explanatory benefits (Epstein, 2008). After all, by examining the processes occurring between agents, perhaps we can gain a greater understanding of how macro-level societal effects happen (although even this is debatable; see Sawyer 2005). However, we feel that another, perhaps more immediate benefit of agent-based modelling has been largely ignored in the literature: the prospect of escaping the expensive, time-consuming process of continual data collection.

Thus, in this paper we propose that agent-based models, informed by work in artificial life and social simulation, can provide a way forward for demographers who seek to escape the 'hungry beast' of highly data-driven research. This approach allows us to create models which develop a new understanding of population change without undue dependence on excessive empirical data, and create an environment in which models can be continually tweaked and worked on as new information comes to light, rather than simply sitting in stasis until the next wave of surveys comes back.

In the next section we will briefly outline the current state-of-the-art in demography, with focus on the contemporary limits of demographic knowledge.

Where the Beast Lies: Demographic Knowledge and its Limits

Demography is currently facing major epistemological challenges. In particular, demographers' knowledge seems to have reached its limits with respect to the predictability of future population developments, as well as the ability to combine micro- and macro-level information and to find a compromise between the complexity and simplicity of analytical tools. This section discusses these issues in more detail.

The first problem with the limits of demographic knowledge is the issue of predictability. Amongst social science disciplines, demography has a unique predictive potential. Unlike in economics or sociology, very important information on the future development of populations is already embodied in their own age structures. The main mechanism of demographic dynamics is known, too: human populations change through births, deaths and, if considered at sub-global levels, migrations. However, when considered on their own, these three components of population change remain largely uncertain (Hajnal, 1955; Orrell, 2007). They also differ with respect to their degree of predictability: mortality is considered to be the best-predictable component; migration — the worst; fertility being usually located in the middle (National Research Council, 2000).

In the context of the uncertainty of forecasts, predictability limits have been extensively discussed elsewhere (Keyfitz 1981; Keilman's contribution to Willekens 1990; de Beer 2000; Bijak 2010), with two main methodological conclusions. Firstly, it is argued that demography should embrace uncertainty more closely (Alho and Spencer, 2005), in particular by moving from traditional deterministic projections to probabilistic forecasts. Secondly, there is an agreement that with longer horizons — beyond 10 to 20 years — uncertainty anyway becomes too large to be usefully described in probabilistic terms, and hence there is a need to turn to scenario-based approaches (see also Orrell and McSharry, 2009; Wright and Goodwin, 2009). An open question is, which elements should be included in such scenarios and how should they be constructed?

The second limitation of demographic knowledge stems from the problem of aggregation. Populations are composed of individuals and, as argued by Courgeau (2007), focusing exclusively on macro or micro-level analysis can generate problems with either ecological or atomistic fallacy. Whilst demography until the 1980s was almost entirely preoccupied with the macro level, and since then increasingly more with the individual level (mainly in a form of the event-history analysis allowing for microsimulations), attempts to bridge both levels are much more recent (Willekens, 2005; Courgeau, 2007; Zinn et al., 2009). The multi-level models are usually also multi-state, states being for example age groups, educational classes, or states of health. In such models, individuals move between the states according to some transition probabilities, usually estimated on the basis of large-scale representative surveys, population registers or census data.

The main challenge with the multi-level approaches lies with their potentially enormous data requirements owing to the combinatorial explosion of the parameter space at different levels. That is exactly where the beast lies: Burch (2003) identified it to be the realm of logical empiricism, on which demography was — and still is — over-reliant. This philosophy focuses on observable phenomena and attempts to create generalisations solely on an empirical basis. As a result, in contemporary research problems driven by real-life questions concerning more complex phenomena, the beast can quickly become insatiable.

The third epistemological dilemma of contemporary demography stems directly from the previous two. At its core there is a question, whether complex models are more useful to aid prediction and decision making than their simpler counterparts. In terms of predictive performance, there is no evidence that complex models perform better (Ahlburg, 1995; Smith S.K., 1997). If that is the case, there might be a temptation to follow the Occam's razor principle (or the KISS principle in complexity science), disregard the additional subtleties involved in the modelling process and opt for simplicity instead (Bijak, 2010). However, such approaches may not increase our understanding of the underlying mechanisms, and are largely limited to shorter time horizons of decision making. To move beyond that, a different approach to modelling would be required.

From this perspective, the following section discusses the applicability of agent-based models in demography, with focus on how they could address the three challenges mentioned above.

Agent-Based Demography: Avoiding the Beast

In their seminal book, Billari et al. (2003) present a compelling argument for the use of agent-based models in demography, or what they refer to as 'agent-based computational demography' ('ABCD'). Their enthusiasm for this approach stems from the potential for agent-based models to build theories regarding social processes that underlie demographic change. They describe a new ethos for simulation in demography, in which "the simulation is used first of all to develop and explore theories rather than to evaluate empirically the consequences of given rates/probabilities" (Billari et al., 2003, p. 11).

The suitability of agent-based models for exploring theories is certainly attractive for social scientists, as such models are well-positioned to examine the link between individual behaviour and higher-level organisation (Silverman and Bryden, 2007). In demography, agent-based models provide a potential platform in which the dynamic relationship between the micro- and macro-levels of a simulated population can be more fully represented. While in recent years multi-level microsimulation models, such as the ones mentioned above, have become increasingly popular, these modelling platforms still fail to capture the influence of micro-level behaviour and agent heterogeneity on macro-level entities, and indeed the feedback of those entities on agent behaviour. Nor do they capture social interactions, formation of social networks, or other elements which may contribute to the social processes underlying demographic change — here, agent-based models are more suitable (Gilbert and Troitzsch, 2005).

Beyond these theoretical benefits, we propose that in the specific context of demography, agent-based modelling offers a possible means to escape some limitations to knowledge imposed by the currently dominant data-based methodological paradigm. The first limitation — the one of predictability — points us toward the potential for using agent-based models for generating scenarios, which would produce useful insights about demographic change over a longer time horizon. A great advantage of agent-based models lies in their suitability for exploring a set of scenarios based upon varied parameter settings. Modellers can develop such scenarios based on variations within a parameter space, which allow them to examine how these parameters affect agent behaviour (and, in appropriately designed models, how those behaviours affect macro-level entities). In the development process, boundaries to the scenarios are limited only by the modellers' imagination rather than by data availability alone.

The second challenge for demographers — the aggregation problem — again points toward agent-based models as a possible way forward. After all, some ambitious social simulations not only include individual agents, but may also include macro-level components and thus allow for feedbacks between individuals, as well as between micro- and macro-level (Billari et al., 2003; Murphy, 2003; Silverman and Bryden, 2007). This would allow the modeller to neatly side-step the problem of focusing exclusively on either the micro- or macro-level. Of course, this second challenge also allows the beast to begin rearing its ugly head. As mentioned in the previous section, the prevalence of the logical empiricist approach in demography places a certain primacy on deriving sensible results *only* from empirical observation (Burch, 2003). This naturally leads demographers to seek out ever larger and more comprehensive data sets, each more expensive and time-consuming to collect than the last.

We then find ourselves sat facing the third challenge — that of simplicity. The beast gets hungrier for more data, and the sets of numbers which need crunching continue to grow in response. Agent-based models, however, necessitate a different approach: *data* is given less primacy than *parameters*. Rather than extrapolating from a given dataset about a population, social simulations will attempt to *generate* a society using the given parameters. The latter can certainly be informed by real-world data whenever they are available.

Thus, a type of modelling used quite often to represent complex systems might require less numerical data input than traditional methods. In certain contexts, the social scientists may not find any data necessary at all — as in Schelling (1978), which demonstrated a possible mechanism for residential segregation based on individual behaviour without requiring any data, and only using a single parameter. As an additional benefit, agent-based models can more sensibly be informed by qualitative data than traditional methods, given that such data often explicitly attempts to "elicit agent models directly rather than inferring them from behavior" (Chattoe, 2003, p. 52).

So, agent-based models can present demographers with a way to avoid the beast and get away from ravenous traditional models which require regular feedings of painstakingly collected data. However, using such models in demography may require a certain shift in focus: agent-based models are better-suited for exploring theories and scenarios than for making firm predictions (Epstein, 2008). Therefore, perhaps we may take inspiration from John Hajnal — himself one of the most prominent demographers of the 20th century — and focus on building models which "involve less computation and more cognition than has generally been applied" (1955, p. 321). As we shall see in the following section, attempts to bolster the power of traditional data-driven models have not always been successful — and agent-based models have already been proven useful in some areas of demography.

Analysis: Case Studies of Demographic Models

In demography, there are notable examples of models that fell short of their proclaimed aims due to the presence of the data-hungry beast. With respect to approaches spanning the micro and macro levels, an interesting attempt to apply methods from the system dynamics tradition to a demographic problem — migration — was the one by Weidlich and Haag (1988). However, solutions proposed by the authors, based on systems of differential equations, despite their mathematical sophistication and elegance, did not become a part of demographers' toolkit. There were several reasons for this. Some reviews of Weidlich's and Haag's book stressed that their method did not take into account heterogeneity of migration with respect to age, sex and past migration history[1]. Other points of criticism were that the approach did not model agents at all, thus not exploring the underlying social complexity in full, and did not provide many examples of empirical applications, mainly due to very large data requirements[2]. Finally, the quasi-deterministic nature

[1] Daniel Courgeau's review of Weidlich and Haag (1988). *Population* 46(5), 1991: 1298–1299.

[2] J. Barkley Rosser's review of W. Weidlich's (2002) book "Sociodynamics: A Systematic Approach to Mathematical Modelling

of the models made them overly reliant on analytical solutions to the system of differential equations describing the dynamics of the migration system in question.

More recently, the MicMac project, as previously mentioned, aimed to develop a new methodology for dynamic microsimulation in demography (Willekens, 2005; Zinn et al., 2009). The final MicMac model consists of a macro-level part, which examines demographic change at the population level (known as *Mac*), together with a micro-level model that examines demographic events at the micro level (known as *Mic*). The model aims to bridge the micro-macro gap, providing a comprehensive modelling package which can pinpoint the influences of micro-level demographic events on macro-level demographic change (see also Billari et al., 2006).

In practice, however, the beast once again rears its head, and data requirements in this case are substantial. The microsimulation portion of the model (Mic) requires a significant amount of detailed micro-level data to implement, especially on transition rates between all possible demographic states for individuals[3]. The macro-level model (Mac) also requires extensive data about transition rates in order to run. Given that Mic includes 12 variables for each individual, very large amounts of input data are required to produce age- and time-specific transition rates between all possible states.

In turn, from the opposite — agent-based — end of the modelling spectrum, one example of an agent-based model producing some historical demographic insight is the model of the Kayenta Anasazi civilisation (Axtell et al., 2002). The model attempts to explain the rapid decline of the Kayenta Anasazi tribe in Long House Valley in northeastern Arizona, United States. The Anasazi tradition began in the area around 1800 B.C., when maize was introduced as a major agricultural crop. Around 1300 A.D., the population declined rapidly, and eventually there was a mass exodus from the valley.

The model of Axtell et al. (2002) consists of a digital reconstruction of the Long House Valley landscape, constructed using existing knowledge of the environmental conditions at that period in history. The agents themselves represent households, individual people being more difficult to identify with any reliability using the existing archaeological data. Each household has certain rules of behaviour which specify how it will select its dwelling and planting locations during each calendar year based on how successful it has been at satisfying its nutritional needs.

The model seemed to produce a simulated population which closely followed the ebbs and flows of the real Anasazi population in Long House Valley. Interestingly, however, the model shows that some small sustainable population could have remained in the northern part of the valley, even as the environmental conditions started to degrade toward 1300 A.D.; this contrasts with the real population, in which the remaining people joined the mass exodus leaving the valley.

This model thus demonstrates that the demographic changes which affected the Anasazi population in this area can be explained at least in part by an agent-based model with simple behavioural rules. As the environment degrades over time, and the agents must continue to look for fertile ground in which to plant their fields, the simulated population shifts northward, just as the real Anasazi had done. In contrast, a demographic model which did not capture these rules of individual behaviour may have been able to accurately portray the changes occurring at an aggregate level, but would not be able to explain *why* those changes occurred.

Of course one might ask, how did this model keep the beast from getting out of control? The model clearly incorporated many pieces of information from a variety of disciplines. However, it is interesting to note in this case *how* the beast was fed. The archaelogical data used here had been previously gathered by another project for archaeological purposes, and there is a distinct absence of resource-consuming data-collection exercises directly linked to the model. Interestingly from a demographic point of view, next to available archaelogical information, the model was able to incorporate *qualitative* data in the form of ethnographical research: the agents' behavioural rules were formulated by distilling ethnographic knowledge about the Anasazi civilisation into simple rules driving their migration and agricultural activities.

Another example of an agent-based model producing demographic insight is a recent study of marriage offered by Billari et al. (2007). Their model was constructed as an attempt to bridge the gap between two different perspectives which predominate in the study of the timing of marriage in populations: macro-level statistical modelling used by demographers, and micro-level studies performed by psychologists and economists examining the partner (mate) search process. In this context, an agent-based model is seen as a possible way to "account for macro-level marriage patterns while starting from plausible micro-level assumptions" (Billari et al., 2007, p. 60).

The resulting model assumes that the formation of marriage partnerships is the result of social interaction between heterogeneous agents. The model attempts to demonstrate the link between these interactions and marriage patterns by simulating the impact of the availability of mates and the desirability of marriage, which is affected by the influence of relevant others in an agent's social network. The results show that the model can reproduce the hazard functions of marriage observed at the population level in the real world.

in the Social Sciences." *Discrete Dynamics in Nature and Society*, 3, 2005: 331–335.

[3] See Deliverable D9 of MicMac: "Report on Data Requirements of MIC" by F Willekens, J de Beer and N van der Gaag: http://www.nidi.knaw.nl/Content/NIDI/output/micmac/micmac-d9.pdf

The performed sensitivity analysis suggests that the results are robust to changes in the relevant simulation parameters.

The findings of Billari et al. (2007) have important implications for demographers wishing to avoid the beast. As the authors note, the model uses substantial simplifying assumptions: placing the agents in a one-dimensional, circular space; leaving out additional social complexities such as courtship or divorce; and focusing only on age and location as agent attributes, ignoring kinship, education, occupation, socio-economic status, or other similar factors. In fact, the simulation almost entirely ignores any empirical data, with the exception of the initial population which is generated with an age distribution reminiscent of 1950s America.

Despite the paucity of data, however, the simulation-based demographic models seem to produce at least plausible micro-level explanations of macro-level phenomena. In the work of Billari et al. (2007) this concerns the influence of social pressure to get married within a social network, and the variation of the size of that network by age is a determinant of the desirability of marriage. In contrast, a macro-level statistical model of marriage timing would not be able to provide this sort of micro-level explanation — and would require significantly larger investments into data collection in order to function. In turn, the study of Axtell et al. (2002) captured the main factors behind the expansion and twilight of the Anasazi population. One could imagine painstaking efforts to reconstruct birth, death and migration rates based on fragmented pieces of historical information, but the ensuing results of traditional demographic predictions would be too uncertain to be meaningful. The beast might be fed and sated — but our understanding of the underlying processes would be no greater.

Conclusions

Our discussion and analysis have demonstrated that traditional demographic methods, while highly accomplished in producing data-driven population projections, face some major epistemological and pragmatic challenges. The overall focus on data over theoretical investigation has hampered demography's ability to provide explanations of demographic change, while the hunger of the beast of logical empiricism traps demographers in continuous cycles of expensive and time-consuming data-collection.

As we have seen, the application of agent-based methods inspired by contemporary ALife work to demography provides a means to lessen some of these burdens on population researchers. The resultant increased focus on explanation over producing projections from empirical data could allow demographers to develop more coherent micro-level explanations of macro-level demographic change. More pragmatically, the concomitant reduction in data dependence would reduce the hunger of the beast, allowing demographers more freedom to produce varied and ambitious models while also removing the restrictive timetables imposed by lengthy and expensive data-collection processes.

So far, all applications of agent-based models to population change, such as the ones mentioned earlier in our analysis, have been performed separately, abstracting away from the main mechanism and inertia of population dynamics. The challenge ahead is to build models which would combine various features of demographic processes and yield artificial populations equipped with real-world characteristics. In that respect, agent-based demography is not only interesting as a research field, but also as a promising venue for answering questions relevant to policy makers. Moreover, it provides the users of the final research output with more possibilities for interacting with the researchers, by engaging in the experimentation with the artificial worlds created. For both parties involved in the process — researchers, as well as the end-users of research — this can bring about a better understanding of the underlying population processes, which itself can be a very important gain from the whole modelling exercise.

From these points of view, agent-based demography seems to be an innovative way of moving the whole research field in a new direction, towards the middle ground on the theory-data spectrum. For the ALife community, this 'dialectic' position would open up a whole new, fascinating field of research with direct applications to real-world problems. However, building agent-based models to population questions would require that demographers use more imagination than in pure data-driven modelling, in line with Hajnal's *credo*. Agent-based modelling can offer a solution, which has to be based on cognition and thinking about mechanisms (Burch, 2003; Chattoe, 2003), while taking into account these pieces of information (data) that are already available. In that respect, the rule of the thumb for agent-based demographers who would like to strike the delicate balance between empiricism and explanation is: we should feed the beast where feasible — but not more.

Acknowledgements

This work was supported by the EPSRC grant EP/H021698/1 "Care Life Cycle", funded within the "Complexity Science in the Real World" theme.

References

Ahlburg, D.A. (1995). Simple versus complex models: Evaluation, accuracy and combining. *Mathematical Population Studies*, 5(3): 281–290.

Alho, J. M. and Spencer, B. D. (2005). *Statistical Demography and Forecasting*. Berlin–Heidelberg: Springer.

Axelrod, R. (1984). *The Evolution of Cooperation*, New York: Basic Books.

Axtell, R.L., Epstein, J.M., Dean, J.S., Gumerman, G.J., Swedlund, A.C., Harberger, J., Chakravarty, S., Hammond, R., Parker, J., and Parker, M. (2002). Population Growth and Collapse in a Multi–Agent Model of the Kayenta Anasazi in Long House

Valley. *Proceedings of the National Academy of Sciences*, Colloquium 99(3): 7275–79.

de Beer, J. (2000). Dealing with uncertainty in population forecasting. *Working Paper of the Department of Population Statistics*. Voorburg: Central Bureau of Statistics.

Bijak, J. (2010). Forecasting International Migration in Europe: A Bayesian View. *Springer Series on Demographic Methods and Population Analysis*, vol. 24. Dordrecht: Springer.

Billari F., Aparicio Diaz B., Fent T. and Prskawetz A. (2007). The "Wedding–Ring". An agent–based marriage model based on social interaction. *Demographic Research*, 17(3): 59–82. http://www.demographic–research.org/Volumes/Vol17/3/17–3.pdf

Billari F., Fent T., Prskawetz A. and Scheffran J. (eds.) (2006). *Agent–Based Computational Modelling: Applications in Demography, Social, Economic and Environmental Sciences*. Heidelberg: Physica Verlag.

Billari F.C., Ongaro F. and Prskawetz A. (2003). Introduction: Agent–Based Computational Demography. In: Billari F and Prskawetz A (eds.), *Agent–Based Computational Demography: Using Simulation to Improve Our Understanding of Demographic Behaviour*. Heidelberg: Physica Verlag (pp. 1–17).

Billari, F.C. (2006). Bridging the gap between micro–demography and macro–demography. In: G. Graziella, J. Vallin and G. Wunsch eds. *Demography: analysis and synthesis*, Vol 4. New York: Academic Press (Elsevier) (pp. 695707).

Burch, T.K. (2003). Data, Models, Theory and Reality: The Structure of Demographic Knowledge. In: Billari F and Prskawetz A (eds.), op. cit. (pp. 19–40).

Cederman, L.–E. (1997). *Emergent Actors in World Politics: How States and Nations Develop and Dissolve*. Princeton, NJ: Princeton University Press.

Chattoe, E. (2003). The Role of Agent–Based Models in Demographic Explanation. In: Billari F and Prskawetz A (eds.), op. cit. (pp. 41–54).

Courgeau, D. (2007). Multilevel Synthesis: From the Group to the Individual. *Springer Series on Demographic Methods and Population Analysis*, vol. 18. Dordrecht: Springer.

Di Paolo, E. A., Noble, J., and Bullock, S. (2000). Simulation models as opaque thought experiments. In M. A. Bedau, J. S. McCaskill, N. H. Packard, and S. Rasmussen (Eds.), *Artificial Life VII: The Seventh International Conference on the Simulation and Synthesis of Living Systems* (pp. 497–506). Cambridge, MA: MIT Press/Bradford Books.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, 11(4)–12.

Gilbert, N. and Troitzsch, K.G. (2005). *Simulation for the Social Scientist*. 2nd ed. Maidenhead: Open University Press.

Hajnal, J. (1955). The prospects for population forecasts. *Journal of the American Statistical Association*, 50(270): 309–322.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review*, 7(4): 579–593.

Langton, C.G. (1989). Artificial life. In: Langton C G (ed.), *Proceedings of the Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems (ALIFE '87)*, Redwood City, CA: Addison–Wesley (pp. 1–48).

Lem, S. (1983). Pożytek ze smoka. Reprinted in: (2008) *Kongres futurologiczny*. Opowiadania Ijona Tichego. Warsaw: Agora (pp. 179–187). [Own translation – JB]

Murphy, M. (2003). Bringing Behavior Back into Micro–Simulation: Feedback Mechanisms in Demographic Models. In: Billari F and Prskawetz A (eds.), op. cit. (pp. 159–174).

NRC [National Research Council], (2000). *Beyond six billion: Forecasting the world's population*. Washington, DC: National Academies Press.

Orrell, D. (2007). *The Future of Everything: The Science of Prediction*. New York: Thunders Mouth Press.

Orrell, D. and McSharry, P. (2009). System economics: Overcoming the pitfalls of forecasting models via a multidisciplinary approach. *International Journal of Forecasting*, 25(4): 734–743.

Rossiter S., Noble J. and Bell K.R.W.(2010). Social simulations: Improving interdisciplinary understanding of scientific positioning and validity. *Journal of Artificial Societies and Social Simulation*, 13(1)–10.

Sawyer, R.K.(2005). *Social emergence: Societies as complex systems*. New York: Cambridge University Press.

Schelling, T. (1978). *Micromotives and Macrobehavior*. Norton, New York, USA.

Silverman, E. and Bullock, S. (2004). Empiricism in Artificial Life. In *Artificial Life IX: Proceedings of the Ninth International Conference*, MIT Press: Cambridge, MA, USA (pp. 534–539).

Silverman, E. and Bryden, J. (2007). From artificial societies to new social science theory. In Almeida e Costa F, Rocha L M, Costa E, Harvey I and Coutinho A (eds.) *Advances in Artificial Life, 9th European Conference, ECAL 2007 Proceedings*. Berlin–Heidelberg: Springer (pp. 645–654).

Smith, S.K. (1997). Further thoughts on simplicity and complexity in population projection models. *International Journal of Forecasting*, 13(4): 557–565.

Smith, V. (2008). Evolving an agent–based model to probe behavioral rules in flocks of cowbirds. In Bullock S, Noble J, Watson R and Bedau M A (eds.), *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*. Cambridge, MA: MIT Press (pp. 561–568).

Toquenaga, Y., Kajitani, I., and Hoshino, T. (1995). Egrets of a feather flock together. *Artificial Life*, 1(4): 391–411.

Weidlich, W. and Haag, G. (eds.) (1988). *Interregional migration: Dynamic theory and comparative analysis*. Berlin–Heidelberg: Springer.

Willekens, F. (1990). Demographic forecasting; state–of–the–art and research needs. In CA Hazeu and GAB Frinking (eds.), *Emerging issues in demographic research*. Amsterdam: Elsevier (pp. 9–66).

Willekens, F. (2005). Biographic forecasting: bridging the micro–macro gap in population forecasting. *New Zealand Population Review*, 31(1): 77–124.

Wright, G. and Goodwin, P. (2009). Decision making and planning under low levels of predictability: Enhancing the scenario method. *International Journal of Forecasting*, 25(4): 813–825.

Zinn, S., Gampe, J., Himmelspach, J. and Uhrmacher, A. M. (2009). MIC–CORE: a tool for microsimulation. In Rossetti M D, Hill R R, Johansson B, Dunkin A and Ingalls R G (eds.) *Proceedings of the 2009 Winter Simulation Conference*. IEEE (pp. 992–1002). http://ieeexplore.ieee.org