Chapter 12
Semi-Artificial Models of Populations: Connecting Demography with Agent-Based Modelling

Eric Silverman, Jakub Bijak, Jason Noble, Viet Cao, and Jason Hilton

Abstract In this paper we present an agent-based model of the dynamics of mortality, fertility, and partnership formation in a closed population. Our goal is to bridge the methodological and conceptual gaps that remain between demography and agent-based social simulation approaches. The model construction incorporates elements of both perspectives, with demography contributing empirical data on population dynamics, subsequently embedded in an agent-based model situated on a 2D grid space. While taking inspiration from previous work applying agent-based simulation methodologies to demography, we extend this basic concept to a complete model of population change, which includes spatial elements as well as additional agent properties. Given the connection to empirical work based on demographic data for the United Kingdom, this model allows us to analyse population dynamics on several levels, from the individual, to the household, and to the whole simulated population. We propose that such an approach bolsters the strength of demographic analysis, adding additional explanatory power.

Keywords Agent-based modelling • Demography • Social simulation

12.1 Motivations: Bringing Together the Statistical and the Simulated Individual

Silverman and Bryden [25] divided current work in agent-based models (ABMs) in social science into two main streams: systems sociology and social simulation. The former focuses on the use of ABMs in an explanatory role, with few predefined

---

E. Silverman (✉) • J. Bijak • J. Noble • V. Cao • J. Hilton
University of Southampton, Southampton SO17 1BJ, UK
e-mail: e.silverman@soton.ac.uk; j.bijak@soton.ac.uk; jn2@soton.ac.uk; vdc1g10@soton.ac.uk; jdh4g10@soton.ac.uk

S.-H. Chen et al. (eds.), *Advances in Computational Social Science: The Fourth World Congress*, Agent-Based Social Systems 11, DOI 10.1007/978-4-431-54847-8_12,

interactions or structures; the models exist to probe elements of social theory, rather than to examine the functionality of specific social systems (e.g., investigations of the behaviour of nation-states, as in [8], or the social effects of different kin systems [23]). Social simulation approaches instead focus on a particular class of behaviour or a particular social system, and may have some link to empirical data (e.g., the reconstruction of the Anasazi population in [1], or the work in [11] on the movement of early Polynesian peoples).

In general, ABMs have become increasingly popular in the social sciences as the methodology has become more established. These models have produced useful insights even when no empirical data was used at all (see, e.g., [24]), and therefore seem to have demonstrated a certain capacity for explanatory power. Meanwhile, some researchers have proposed that ABMs will allow social scientists to more directly integrate their data with simulation approaches [20]. Given the remarkable flexibility the ABMs display, these models can be used for numerous purposes beyond simply explanation and prediction [12].

The methodological challenges facing demography currently (see [26]) have led some demographers to investigate the methodology of agent-based modelling as a potentially valuable addition to the toolkit of the field [4, 5]. In the context of demography, the challenges facing multilevel microsimulation models have alerted some authors to the need to address the issue of the combinatorial explosion of the parameter space in these models [26]. It has also been argued that traditional demographic models cannot fully capture the complexities of micro-level agent behaviour and heterogeneity [26]. ABMs can provide these possibilities [14], as well as a possible platform for understanding social interactions, social networks, and other processes lying at the root of demographic change [5].

Recent work in demography has described the fundamental goals of demographic research—or indeed any field of social science applying statistical methods—as the use of statistical models based on observations (e.g., surveys and censuses) to predict and describe the behaviour of *statistical individuals* [10]. Demography has therefore progressed through a number of different paradigms, ranging from the cross-sectional view, in which events are viewed as separate from individuals, to the event-history perspective, in which demographic events are interdependent on individual life courses. In recent years, the most prominent view has been the *multilevel* paradigm, in which behaviour at the societal level can only be understood by investigating multiple levels of the social world simultaneously [9, 10].

A similar process can be seen within the systems sociology and social simulation communities: in these arenas we focus on *simulated individuals* equipped with some simple behavioural rules, through which we hope to see the emergence of interesting behaviours at the macro level. At the same time, some of these models incorporate feedback loops, social networks, and other related concepts—all of which serve to bring these multiple levels of the social world into play within the simulation.

Thus, we argue that while social science ABMs and demography clearly differ in some respects—for example, in having either theoretical or empirical focus, or using mainly computational or statistical models—the recent history of demographic theory shows synergies developing between the agent-based simulation approach and the multilevel demographic approach. On the technical side, we already see some of these links: some multistate models in demography already assume simple behavioural rules in individuals, while some ABMs of social systems already use substantial empirical information in the model-building process.

Moreover, demographic methods also have significant strengths in the empirical arena, not least the capacity for higher predictive accuracy in many cases, as well as the direct connection to the rich information already embedded in the age structure of populations. Efforts thus far in developing agent-based demography have attempted to incorporate agent-based methods while simultaneously retaining these particular strengths of demography (e.g., [4–6, 16, 29]). Thus, demographic methods can contribute significant additional expertise and allow us to better align models with the real world of empirical observations.

Taking this process further, we propose that directly linking demographic methods with ABM frameworks will allow us to produce models which increase our understanding of population change, while simultaneously helping us to avoid the pitfalls of an *overdependence* on empirical data [26]. ABMs allow us to produce models which have a greater explanatory capacity, while the demographic components allow us to use the inherent flexibility of the ABM approach to generate plausible scenarios within a given parameter space, informed by well-established demographic methods of analysis.

In this framework, such augmented ABMs will allow demographers to examine scenarios over longer time horizons, rather than being heavily limited by data-dependent statistical models. The general consensus is that the predictive horizon of demographic models is about one generation in length [18], but some have argued that scenario-based methods can offer an opportunity to explore a larger space of possible futures (e.g., [3, 30]). Thus, we propose that this new framework can produce a shift within the demographic community: from a focus on prediction and description alone, to an approach which allows for exploration of scenarios of population change, while also providing some insight into the underlying processes.

Hence, we suggest that the social simulation stream of agent-based modelling is of great relevance to demography. The model presented here aims to make progress at this interface, and bring ABMs of population to the fore as a means to both enhance demographic methods, and to tie ABMs more closely to real social systems (an approach we have advocated previously [26]). We refer to this approach as a semi-artificial model of a population to capture the mixture between empirical data and randomly generated agent populations with simple behavioural rules. An example of such a model—inspired by and drawing from the Wedding Ring model seen in [6]—is described in detail in the next section.

12.2 From the Wedding Ring to the Wedding Doughnut

12.2.1 Wedding Ring: Background

In order to illustrate the potential benefits of combining demographic and complexity science approaches, we replicated and expanded upon the model of partnership formation implemented by [6], known as the Wedding Ring model. This model attempts to explain the age-at-marriage patterns seen in modern European states by representing the process of partnership formation as a consequence of social pressure. This pressure arises from contact between married and unmarried individuals within a given social network. This conceptual framework serves to formalise some recent research in social influence and social learning, which has shown that these processes are highly relevant in individuals' decisions to get married (see, e.g., [2, 6, 7]).

Thus, the model represents the spread of marriage through a population as a diffusion process. However, marriage differs from other diffusion processes in that even those experiencing a very high level of social pressure towards marriage cannot get married without finding a suitable unmarried partner [6].

The Wedding Ring is so named because agents live in a one-dimensional ring-shaped world, which changes over time [6]. The agents are thus effectively situated in a cylindrical space, with the circular dimension (arc length) representing space, and the linear dimension time. Each agent's network of relevant others is then defined as a two-dimensional neighbourhood on that cylinder [6].

Within that neighbourhood, the proportion of married agents determines the social pressure felt by an individual agent, which influences that agent's decision to seek out a partner. The overall level of social pressure and the agents' age influence parameter determine the range in which agents search for a suitable partner. As social pressure increases, agents widen their search range, and thus have a greater chance of successfully finding a partner [6]. However, the search is mutual; if one unmarried agent finds another within its acceptable range, marriage may only occur if the suitable partner has the searching agent within its acceptable range as well. Once married, agents may bear children; these children are then placed into the ring-world at a random spot in their parents neighbourhood.

In order to define the network of relevant others in which the agent searches for a partner, each agent is first classified into one of five possible types, according to the age ranges of agents by whom they are most influenced (i.e., similarly by younger and older agents, either mostly or only by older agents, or either mostly or only by younger agents—for further details, see [6]). The size of the spatial interval for the agents' network of relevant others is symmetric around their location, and varies according to the size of the initial population.

12.2.2 Extending the Model: The Wedding Doughnut

In order to better align our model with the demographic processes under study, we altered the Wedding Ring model in several major ways. First, we situated the agents in a toroidal space, as this would allow for a more complete consideration of the impact of spatiality on partnership formation processes. As in the original Wedding Ring, the dimensions of the grid space could be considered not just as spatial distances, but also as social distances [6], or, in our example, a combination of both. The model uses a grid space 72 squares long on each edge, with an initial population of 800 agents, which allowed for a sufficient population density to observe interesting population dynamics.

In our model—hereafter dubbed the Wedding Doughnut, given the new spatial arrangements—several core components of the original model were altered. The Wedding Ring was built on the assumption that each agent had only one spatial coordinate to record, so in the Doughnut we altered the methods used to calculate spatial separation, and substantially altered other parameters in order to allow the agents to properly search this space.

We also allowed the agents to move on the grid. When an agent forms a partnership, they move to a new location, the distance to which is inversely proportional to the size of the network of their relevant others. Any future offspring are placed at this new household. For simplicity, however, we follow the original model's lead and assume that partnerships are permanent, and that agents cannot form partnerships until the age of 16. Future extensions of this model might allow for more detail, such as same-sex partnerships and the possibility of partnership dissolution.

The other details of the model have not changed in this implementation; we use an identical function to the Wedding Ring for age influence (a piecewise linear function which varies with age) and for social pressure (a sigmoid function). For these we used identical parameter values to those in Billari et al. [6] in their base scenario.

12.2.3 Extending the Model: Demographic Elements

The original Wedding Ring made some significant simplifying assumptions in order to make the model run smoothly. Agents died only when reaching age 100, and birth rates were adjusted regularly in order to keep the population constant [6]. These restrictions were removed in the current Doughnut implementation, and ageing, mortality, and fertility were implemented using elements drawn from statistical demography.

We wished to include in this extension a serious consideration of the underlying demographic processes that influence partnership formation. Modern societies are ageing, and older agents have smaller networks of relevant others. Therefore, we added a realistic model of mortality, as otherwise the model risks failing to capture the complexities underlying these trends. Similarly, we added a realistic model of fertility to capture the current shift towards later childbearing and lower birth rates.

The initial population in the model is generated at random, but the distributions of agents by age, sex, and marital status correspond to the observed data from England and Wales in the 1951 Census [21]. As the simulation progresses, fertility and mortality rates are based on empirical data and projections for the UK population. The first 60 years of the simulation use age-specific mortality rates drawn from the Human Mortality Database for 1951–2009 [17].

These rates were then projected forward using the popular method for forecasting mortality developed by Lee and Carter [19]. This method uses the leading vectors of a singular value decomposition of the matrix of centred mortality rates to construct a model for mortality with only one time-varying element. This allows easy forecasting using standard times series methods; more details about procedure and estimation are available in [19]. Projections to 2250 using this method show a continual but slowing increase in life expectancy over the period.

Fertility rates for the simulation were obtained in a similar way. Age-specific fertility rates from 1973 to 2009 for UK women of childbearing age were obtained from the Eurostat database [13], while earlier data for the period 1950–1972 were taken from the Office of National Statistics data for England and Wales [21].

A Lee-Carter model was again fitted to the data to obtain future rates, but, in contrast to the mortality projections, two components of the singular value decomposition of the matrix of fertility rates were needed, as two time indices better captured the trends in fertility. The resultant projections to 2250 see the total fertility rate increase initially before converging at a value just above replacement fertility, and also display a continuation in the empirical trend towards later childbearing in the UK.

12.2.4 Extending the Model: Health Status

In order to make the model more relevant to real-world policy concerns related to the ageing UK population, particularly problems in social care provision, we also added a simplistic model of health status to the Wedding Doughnut. Agents have a probability of transitioning into a state of ill health—here defined as a state of illness or disability requiring significant care—which increases with age. Once agents transition into this state, they remain ill until death. We do not model the impact of illness on mortality, given that an appropriate representation of this would require further analysis on empirical data to determine the effect of unmet care need coupled with a more sophisticated model of illness. As in the real world, males have a slightly higher probability of entering this state than females; parameter values were not based on empirical data, due to the difficulty of obtaining and analysing data on these types of illnesses.

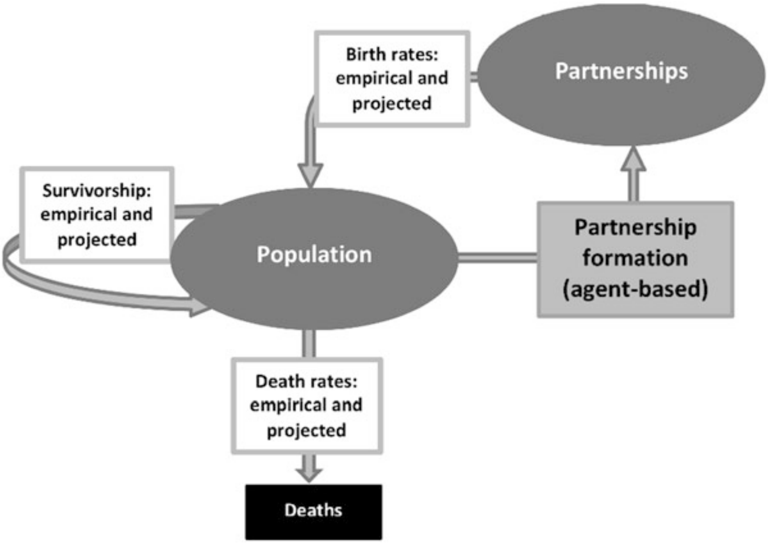

Fig. 12.1 A representation of the flow of the Wedding Doughnut model. Agent-based methods are used to produce a model of partnership formation, which is linked to the other empirical, demographic methods shown here to produce the complete model

While this model of health status is extremely simplistic, we intended this addition to serve as a proof of concept that a simulation built on this semi-artificial framework could produce results which, due to their relationship to empirical data and projections, could bear upon issues relevant to policy-makers. As the results will show, running various scenarios even in this basic model produced interesting results, and future iterations of the model will incorporate a more robust model of health status which would allow for significantly more detail.

12.2.5 The Wedding Doughnut: Flow of the Simulation

The Wedding Doughnut was written in Repast Simphony 2.0, which is a Java-based software package for agent-based modelling. Figure 12.1 illustrates the general flow and structure of the model. The simulation runs for 300 time steps, each time step being 1 year—in contrast to the Wedding Ring, which only ran for 150 years. The first time step corresponds to the calendar year 1951, and hence the simulation extends to 2250.

Each time step proceeds as follows:

1. All agents age 1 year
2. For agents without partners:
    a. For each agent, we find that agent's relevant others
    b. Social pressure is calculated
    c. Potential partners are found
    d. Partnerships form where there is mutual need/pressure
    e. New partners move to a new location
3. For agents with partners:
    a. Check fertility status
    b. Some agents will give birth according to relevant fertility rates
4. For every agent:
    a. Check health status: Some agents will become ill according to relevant rates
    b. Check mortality: Some agents will die according to relevant mortality rates
5. Remove dead agents from the simulation
6. Place newborn agents into the simulation

12.3 Results

Initial runs of the model were designed as a pure replication of the original Wedding Ring: agents were placed on a 1D ring, and none of our modifications were incorporated. Results were consistent with the original paper, which indicated that we could progress to developing our modifications.

Once our modifications based on empirical demographic projections were included, our initial set of 1,800 runs indicated qualitative similarity to the patterns of marriage observed in modern Britain. As shown in Fig. 12.2, the simulation produced populations with proportions of individuals who married at some point during their life course, as we see during 2010–2011 in the UK (the last year for which data is available). The results show greater stochasticity, which is likely due to our small initial population, small simulated world, and the short running time of the simulation by the time it reaches calendar year 2011. The current implementation has the advantage of short running times, and thus allows for a more in-depth analysis, which carries significant benefits when testing a new modelling approach.

Using our initial parameter settings, we were able to broadly replicate the pattern of marriage that we see in empirical data. Sensitivity analysis further showed that significantly altering the basis of the simulation by reducing social pressure or age influence parameters to constants produced marriage hazard rates inconsistent with reality (as shown in Fig. 12.3). The base scenario, in which age influence varies according to the age of the agents and social pressure is calculated using a sigmoid function, more closely resembles hazard rates of marriage that we see in modern European states.

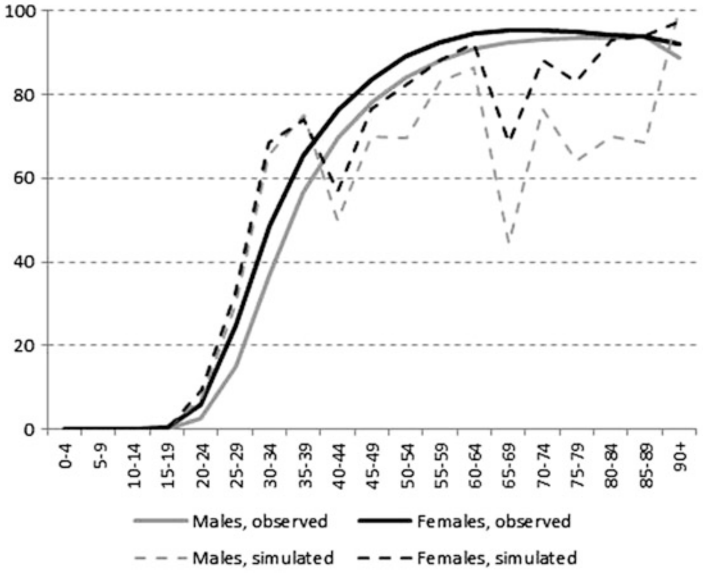

Fig. 12.2 Example of a comparison between observed and simulated populations showing the percentage of the population which has ever married by 2010–2011

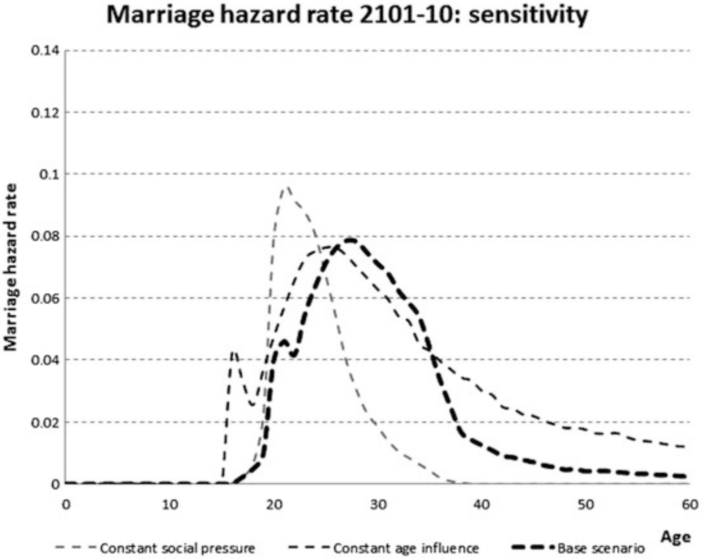

Fig. 12.3 Comparison of marriage hazard rate results in three scenarios: the base scenario, constant social pressure, and constant age influence

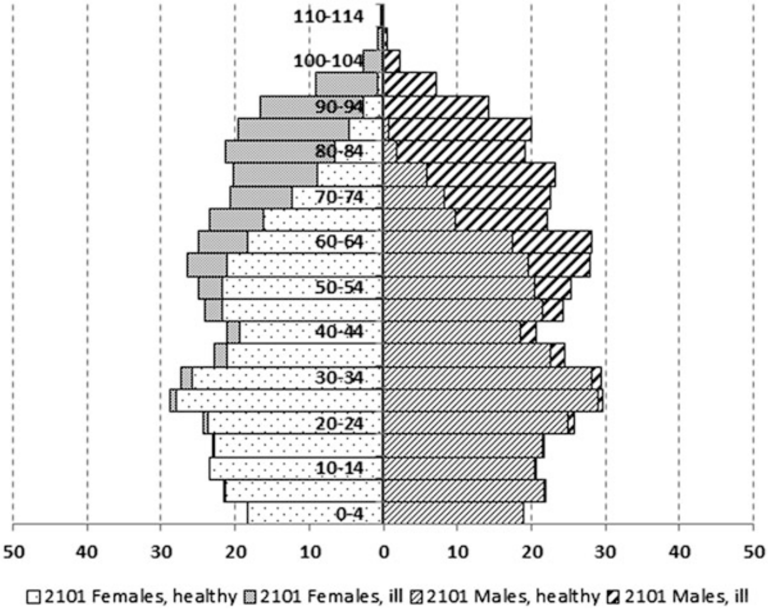

Fig. 12.4 Population pyramid for the year 2101 in our simulated population. Note the increasing incidence of long-term limiting illness as age increases

Figures 12.4 and 12.5 provide a breakdown of a simulated population in the year 2101. While these results obviously cannot be compared with empirical data, we do see a pattern here that intuitively fits our expectations for future population dynamics. As expected in any long-term demographic projection, results at the end of the simulation in year 2250 tend to be more variable, so we chose 2101 for our analysis, as it is more illustrative. In Fig. 12.4, we see that the population becomes increasingly dominated by the older age brackets of society, and further, many of those individuals have become ill and will require care (according to our simplistic model of health status in which only long-term limiting illness is represented, as noted in Sect. 12.2.4). Figure 12.5 shows that as our simulated individuals age, many of those who are ill have not had the opportunity to marry, meaning that they will not have spouses or children available to care for them.

Thus, results show an encouraging parity with empirical data in the early stages of the simulation (which are the only stages for which we can make this comparison). Hazard rates of marriage reflect what we expect in current society, and overall marriage rates, while displaying more stochasticity, are broadly at a level consistent with reality. As the simulation progresses, we see population change that mirrors the expectations of demographers with regard to the dynamics of an ageing population. Agent populations become increasingly dominated by the old and infirm, marriage/fertility rates slowly decline, and as a consequence we see ever-increasing numbers of agents in ill health but who have no family to provide care.

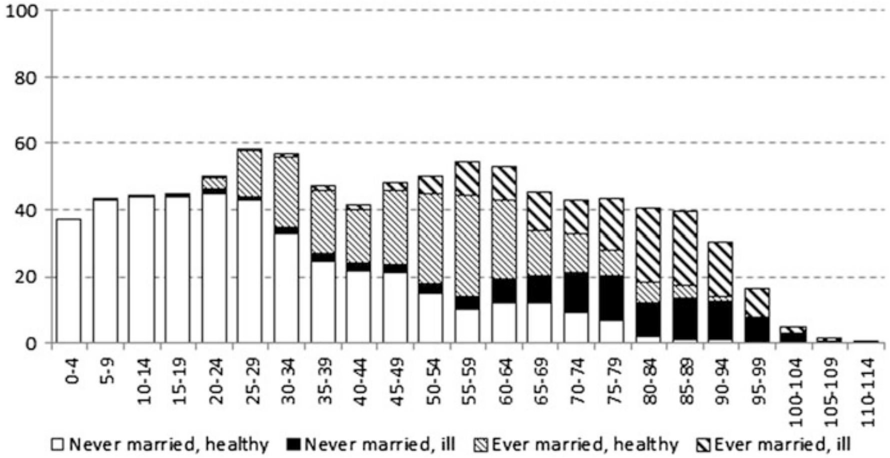

Fig. 12.5 Breakdown of simulated population in the year 2101 by both marriage and health status

12.4 Discussion

The results above demonstrate our original thesis: ABMs of demographic processes, augmented with empirical data on fertility and mortality, can produce results that match our expectations of the demographics of modern European states. Despite the lack of data fully informing our model of health status, we see plausible distributions of healthy and unhealthy agents that illustrate the consequences of Europe's ageing population. This leads us to conclude that the mechanisms used here to drive basic demographic processes have captured the essential elements needed to produce a useful starting point for semi-artificial population models for the UK.

While this is a promising beginning, a number of further extensions are planned. First, our model of health status, while illustrative, is quite simplistic and does not necessarily reflect the complexity of relationships between health and ageing in the modern UK society. As such, this part of the model can be extended by incorporating empirical data on limiting long-term illness in the UK. Further, future expansions of this model will capture the effect of unmet care need on mortality, which is not modelled in the current implementation.

Second, our model of partnership seems to capture the appropriate dynamics, but it does not incorporate the possibility of partnership dissolution, or of same-sex partnerships. Currently, partnerships are defined solely as 'relationships leading to reproduction', but in the UK quite a few children are born in non-cohabiting relationships (10%) or by lone parents (6%) [22]. Future work will capture these subtleties by allowing for greater variety in childbearing partnerships.

Finally, agents on our Wedding Doughnut do not shift position in that space unless they have formed a new partnership. Natural extensions in that respect would require incorporating more sophisticated agent behaviours, particularly in allowing for agent mobility, enabling us to represent the possible causes and effects of within-country migration.

Nevertheless, in its current form the Wedding Doughnut provides an illustrative example of the power of the combined ABM-demography approach.[1] We propose that combining ABMs and statistical demography can help both disciplines in combination to improve their explanatory relevance, deepen our understanding of demographic processes, and increase our appreciation of the links between macro-level effects and micro-level behaviour. This may move us closer to the ideal expressed by the British demographer John Hajnal, who aspired to the construction of models which "involve [. . .] more cognition than has generally been applied" [15, p. 321].

Acknowledgements The authors would like to thank our anonymous reviewers for their comments, which helped improve our earlier draft. We would also like to thank attendees of the 4th World Congress on Social Simulation in Taipei for their valuable feedback. 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

1. Axtell RL, Epstein JM, Dean JS, Gumerman GJ, Swedlund AC, Harberger J, Chakravarty S, Hammond R, Parker J, Parker M (2002) Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. Proc Natl Acad Sci USA 99(3):7275–7279
2. Bernardi L (2003) Channels of social influence on reproduction. Popul Res Policy Rev 22(56):527–555
3. Bijak J (2010) Forecasting international migration in Europe: a Bayesian view. Springer series on demographic methods and population analysis, vol 24. Springer, Dordrecht
4. Billari FC (2006) Bridging the gap between micro-demography and macro-demography. In: Graziella G, Vallin J, Wunsch G (eds) Demography: analysis and synthesis, vol 4. Academic Press, New York, pp 695–707
5. Billari F, Prskawetz A (eds) (2003) Agent-based computational demography: using simulation to improve our understanding of demographic behaviour. Physica Verlag, Heidelberg
6. Billari F, Aparicio Diaz B, Fent T, Prskawetz A (2007) The wedding-ring: an agent-based marriage model based on social interaction. Demogr Res 17(3):59–82. http://www.demographicresearch.org/Volumes/Vol17/3/17-3.pdf
7. Bühler Ch, Fratczak E (2007) Learning from others and receiving support: the impact of personal networks on fertility intentions in Poland. Eur Soc 9(3):359–382
8. Cederman LE (1997) Emergent actors in world politics: how states and nations develop and dissolve. Princeton University Press, Princeton
9. Courgeau D (2007) Multilevel synthesis: from the group to the individual. Springer series on demographic methods and population analysis, vol 18. Springer, Dordrecht
10. Courgeau D (2012) Probability and Social Science: Methodological relationships between the two approaches. Methodos series, vol 10. Springer, Dordrecht
11. Di Piazza A, Pearthree E (1999) The spread of the 'Lapita people': a demographic simulation. J Artif Soc Soc Simul 2(3). http://jasss.soc.surrey.ac.uk/2/3/4.html

---

[1] See [27] for a further explication of this approach, and [28] for an additional example using a different modelling platform.

12. Epstein JM (2008) Why model? J Artif Soc Soc Simul 11(4):12. http://jasss.soc.surrey.ac.uk/11/4/12.html
13. Eurostat (2011) Eurostat statistics database: domain population and social conditions. http://epp.eurostat.ec.europa.eu. Accessed 27 Oct 2011
14. Gilbert N, Troitzsch KG (2005) Simulation for the social scientist, 2nd edn. Open University Press, Maidenhead
15. Hajnal J (1955) The prospects for population forecasts. J Am Stat Assoc 50(270):309–322
16. Heiland F (2003) The collapse of the Berlin wall: simulating state-level East to West German migration patterns. In: Billari F, Prskawetz A (eds) Agent-based computational demography, Heidelberg: Physica-Verlag, pp 73–96
17. Human Mortality Database (2011) Human mortality database. http://www.mortality.org/. Accessed 26 July 2011
18. Keyfitz N (1981) The limits of population forecasting. Popul Dev Rev 7(4):579–593
19. Lee RD, Carter LR (1992) Modeling and forecasting U.S. mortality. J Am Stat Assoc 87(419):659–671
20. Moss S, Edmonds B (2005) Towards good social science. J Artif Soc Soc Simul 8(4):13. http://jasss.soc.surrey.ac.uk/8/4/13.html
21. Office for National Statistics (1998) Birth statistics series FM1 (27). Office for National Statistics, London
22. Office for National Statistics (2011) Live births in England and Wales by characteristic of mother 1-2010. Office for National Statistics, Titchfield. Retrieved from www.ons.gov.uk on 10 June 2013
23. Read DW (1998) Kinship based demographic simulation of societal processes. J Artif Soc Soc Simul 1(1). http://jasss.soc.surrey.ac.uk/1/1/1.html
24. Schelling T (1978) Micromotives and macrobehavior. Norton, New York
25. Silverman E, Bryden J (2007) From artificial societies to new social science theory. In: Almeida e Costa F, Rocha LM, Costa E, Harvey I, Coutinho A (eds) Advances in artificial life, 9th European conference, ECAL 2007 proceedings. Springer, Berlin, pp 645–654
26. Silverman E, Bijak J, Noble J (2011) Feeding the beast: can computational demographic models free us from the tyranny of data? In: Lenaerts T, Giacobini M, Bersini H, Bourgine P, Dorigo M, Doursat R (eds) Advances in artificial life, ECAL 2011. MIT Press, Cambridge, pp 747–754
27. Silverman E, Bijak J, Hilton J, Cao VD, Noble J (2013) When demography met social simulation: a tale of two modelling approaches. J Artif Soc Soc Simul 16(4):9. http://jasss.soc.surrey.ac.uk/16/4/9.html
28. Silverman E, Hilton J, Noble J, Bijak J (2013) Simulating the cost of social care in an ageing population. In: Rekdalsbakken W, Bye RT, Zhang H (eds) Proceedings of the 27th European conference on modelling and simulation. Digitaldruck Pirrot, Dudweiler, Germany, pp 689–695
29. Todd PM, Billari FC, Simão J (2005) Aggregate age-at-marriage patterns from individual mate-search heuristics. Demography 42(3):559–574
30. Wright G, Goodwin P (2009) Decision making and planning under low levels of predictability: enhancing the scenario method. Int J Forecasting 25(4):813–825