Choosing the Choice: Reflections on Modelling Decisions and Behaviour in Demographic Agent-Based Models

March 29, 2017

1 Introduction

As demonstrated throughout this Special Issue, agent-based models have become important tools with which demographers can study mechanisms of social reality, enhancing the analysis of population processes with multi-level influences, feedback effects, the presence of networks and interactions. Still, the related methodology remains under development: in particular, as evident in the nascent demographic agent-based literature, there is hardly any discussion on operationalising the actual decision making processes, through which the agents could acquire some more agency beyond reflexive reactions to simulated stimuli.

One of the most pervasive features of making decisions in the social realm is their uncertainty: the actors involved often do not – and cannot – know the full consequences of their decisions, nor even the full circumstances under which the decisions are made. However, in agent-based models, designed to reflect the complex features of social reality, uncertainty is typically being underplayed. As a result, the simulated agents lack psychological realism with respect to how they deal with ambiguity. Similarly, an analysis of additional layers of uncertainty, such as those related to model choice, model parameters, stochastic nature of the computer code, and so on, are often missing in model-building endavours. On the other hand, there is an acute need for bridging the gap between agent-based modelling and statistical inference, which can provide an appropriate language for dealing with uncertainty (see e.g. Heard, Dent, Schifeling, and Banks 2015).

In this paper, we discuss selected challenges of choosing an appropriate model of choice in the context of decision making in agent-based demography. The argument is illustrated by examples related to demographic agent-based models, with focus on migration. In particular, in Section 2 we elaborate on the role of the context of the decision processes. In Section 3, we discuss some specific aspects of the mechanics of choice: time, uncertainty, as well as the heterogeneity of agents. Finally, in Section 4, we conclude with a list of tentative recommendations for further work in the area, focusing on the mechanisms of decision making, multi-disciplinary approaches, choice between the choice models as such, a multi-model framework, and modularity of agent-based modelling.

2 Context Matters

This section stresses the axiomatic status of the current models of choice in agent-based simulation, and discusses the numerous interfaces between choice and context at which modelling decisions must be made.

That the context in which a decision is made plays a role in the choice is inarguable: context can strongly influence the decision process itself (Ben-Akiva et al., 2012). The challenge arises in determining what the context is, and having done so, deciding what of it to consider, and how to operationalise it. The former is the easier task, since the context is the universe up to the moment of decision. The latter are more problematic, and the 'state of the art' is for now, just that - an art.

The problems of inclusion and operationalisation of context are best tackled through an iterative process of development and evaluation against empirical data, in agreement with proposals from Cioffi-Revilla (2010) and from the ecological literature on Agent-Based Models (Grimm et al., 2005; Schmolke, Thorbek, DeAngelis, & Grimm, 2010). The Pattern-Oriented Modelling approach of Grimm et al. (2005) describes a process of matching empirical patterns at multiple levels in order to determine the suitability of the model under study. Cioffi-Revilla (2010) recommends a process of movement from simple to complex models, where at each stage an additional component of the phenomenon under study can be added, until a satisfactory model is found.

This iterative approach places considerable demands on the modeller in terms of developing the model, but also in exercising their judgement as to what should be included, and identifying where a component should be iterated out. This is however an expensive exercise, both because augmenting a model is time consuming, and because evaluating complex models poses considerable challenges. The challenge of evaluation can be somewhat ameliorated by the use of techniques like sensitivity analysis (see Thiele, Kurth, and Grimm (2014) for a review of several techniques, and Oakley and O'Hagan (2004) for an alternative approach), which can support the modeller in identifying how the individual parameters of a model contribute to overall outputs. A key benefit to this is that parameters which do not significantly impact the results of a model can be removed. This too has limitations, because as we discuss later in relation to heterogeneity in agents (section 3.3), context can take the form of both process, and parameter, and sensitivity analysis is directly informative only about the latter. As a result, the onus remains on the modeller to use their best judgement. The application of such judgement can be documented, however: Schmolke et al. (2010) suggest that developmental cycles are recorded in a standardised format, the TRACE framework, so the process of rejection of alternatives through which the final model came about can be reconstructed.

As an example, we might consider social context, and how it is treated in modelling one demographic process. The social context of a decision is a broad sphere touching as it does on social norms, the role of the decision maker's social network, and so forth. Social context, and indeed any context, may interact with decision making by being represented as a parameter to the decision model, or through impacting on the outcome of enacting a decision. Naturally, where the context is an input to the decision model, the modeller must consider whether process it arises from is within the scope of the model, or should be externalised. Here, we will follow the lead of Klabunde and Willekens (2016), and examine how social context is operationalised when modelling migration.

The prototypical approach to modelling choice is to employ some form of random utility model (Baltas and Doyle (2001) have reviewed the key types, in the context of marketing research), where all decision makers consider the same factors in their decision making, and choose between the same options. Considering migration, social networks have a clear role (G. S. Epstein & Gang, 2006), both through social influence on decision making, and in the process by impacting how far a migration decision succeeds. Naturally, the two are not indivisible, since success or failure carries implications for how that actor influences others.

If we limit ourselves further, and consider only how social influence might be modelled, then there are two key questions: what is the unit of influence; and how does influence interact with decision making. It will be useful to address these one at a time, although the two are of course intertwined, since influence is interpreted information that affects behaviour. This is also demonstrated by Mason, Conrey, and Smith (2007) in their review of models of social influence, where many models treat these aspects in combination.

The flow of influence and the flow of information are interdependent processes, the latter of which has attracted considerable attention in the last few years, with an increasing appreciation of the role of social networks in social influence. Here, rather than the structure over which influence flows, we will focus on what flows over it. Where information is transacted, it may be about attitudes, or actions, or outcomes, and can correspond strongly or weakly to the truth. The mechanics of flow can also vary considerably, for example transmission might occur only between direct links, or be passed along many, and may pass in only a single direction.

The distinction between acts, and attitudes is an interesting one largely unexplored in this arena. Typically, we observe the actions of others, rather than getting accurate insight into their reasoning. Where actions are what agents observe, their function is usually socially normative, for example in Silveira, Espíndola, and Penna's (2006) model of rural-urban migration, an agent's utility is partially a function of how similar their behaviour is to that of their neighbours. The contrasting approach is to communicate a representation of the attitudes of the agent, used by Espindola, Silveira, and Penna (2006) in an earlier simulation of rural-urban migration, where agents compare 'satisfaction'.

Neither approach is a complete account. While mimicry seems to be fundamental to many aspects of human behaviour (Chartrand & Bargh, 1999), we are not simply impersonators and instead interpret the actions of others. Absent telepathy however, we can only draw inferences based on observation of their actions[1]. A potential approach to this might be one similar to that of Jern, Lucas, and Kemp (2011) who propose an inverted generative model[2] to explain how people might translate observed behaviour into preferences, although this presages a utility type mechanism for decision making. This also has implications for more high level models of the interaction between context and decision making process, for example that of Ben-Akiva et al. (2012), where there are explicit models of the interaction between context and choice. An inverted model which does not sufficiently capture the context in which an agent's decision making took place, i.e. the observing agent is underinformed about inputs to the other's decision process, or constraints on their actions, may well lead to faulty inferences about what drove their behaviour. Although this is realistic, in that our own inferences about the motivations of others are not always accurate, it may not be desirable in all models.

The other usual interpretation of social information flow is as transmission of information about the state of the world, for example information about expected income, as in models by Filho, de Lima Neto, and Fusco (2011), and Klabunde (2014). Relating this back to our previous mention of faulty inference, this raises the issue of decision making in an uncertain world, since there is no guarantee that information is accurate. In both models, agents take the average incomes of others in their social network as their expected income were they to relocate. Although some distortion may be introduced by the limited sample, the reported income of other agents is perfectly accurate. A potential extension might be to consider inaccuracy arising from cognitive bias; for example, Mather, Shafir, and Johnson (2000) found that people distorted their recollections in favour of the choices they had made, suggesting that agents might exaggerate their success. Alternatively, McKenzie, Gibson, and Stillman (2013) found that potential migrants had unrealistically negative expectations about earnings and employment. They suggest this may partly arise from attempting to reduce demands for remittances, i.e. agents underreport their incomes. In either case, this suggests that information transmission is not always simply a passive process. Influence may also not be exclusively limited to information, and could for example take the form of the provision of help, emotional assistance, or material support - transactions where information is present, but secondary.

McKenzie et al. also suggest that overweighting of negative experiences plays a role, which demonstrates the importance of considering how information is incorporated into the decision process. In this case the suggestion is that another cognitive bias, loss aversion, is a salient feature. This could be reflected by using a decision model explicitly incorporates loss aversion, for example the Prospect Theory (Kahneman & Tversky, 1979), or Cumulative Prospect Theory (Tversky & Kahneman, 1992). Both variants of Prospect Theory are intended to mimic human decision making, by distorting the perception of high and low probabili-

[1]Which may well include communicating about their internal state.

[2]A generative model produces decisions, given preferences over outcomes. The inversion provides estimates of preferences, given observed decisions.

ties towards certainty, and shifting the value of gains and losses such that large gains are progressively less significant, and losses more so. The wider implication is that social information flow may be distorted at the source, or through interpretation, and that agent-based models are uniquely suited to capture this. We have focused here on descriptive theories of decision making, precisely because the aim is to model human behaviours as as they are manifested, rather than how they should be. Other approaches to decision models are available, for example heuristic methods (Gigerenzer & Goldstein, 1996), which emphasise the bounded rationality of human decision making, and Bayesian approaches, which address learning as a component of the decision process (Robbins, 1964). In an agent-based modelling context, Gray, Bijak, and Bullock (2016) have contrasted Bayesian, heuristic, and Cumulative Prospect Theory approaches in a simulation of alcohol misuse disclosure among pregnant women.

As we have intimated, we believe that introducing such additional complications to a simulation necessitates an iterative approach (Cioffi-Revilla, 2010). We also believe that the implemented model of decision making should be treated as a feature to iterate on, and that different models should be contrasted against one another (Rossiter, Noble, & Bell, 2014). This suggests that modularity should be a key consideration in model design, to support the modeller in evaluating the relative merits of different decision models, given the purpose of the overall research.

The diversity of approaches to this single issue is reflective of the open ended nature of agent-based modelling. A recurring feature however, is that an appreciation is needed of the multiple roles of contexts in decision making. Interpersonal context, for example, is a critical feature of information transmission, but also plays a role in the subjective value of outcomes. This is a challenge, since it is not clear that the two are readily separable, and our understanding of interpersonal context is incomplete. However this presents us with great opportunity, in using simulation to expand our understanding both of the role of social processes, and their mechanics; and in following the example of Tversky and Kahneman by forging inter-disciplinary collaborations with experts in the micro-domain of behaviour.

3 Auxiliary factors in modelling decisions.

This section focuses on two auxiliary model elements that are often understandably overlooked in implementations of agent decision models, but are also key assumptions. The first is the treatment of time in models of decision making, while the second is the extent to which individual agents differ in their approaches to making decisions in the same context.

3.1 Time

Time is explicitly represented in almost all agent-based models, either in discrete or (pseudo) continuous form, in order to allow the dynamic unfolding of the modelled processes by 'stepping' from one time-step to the next (or between time-ordered events in continuous-time discrete-event simulation). However, time often does not enter into the decision making process of the simulated individuals. This results in the exclusion of a number of factors that may potentially be significant in eventual model outcomes. Firstly, individuals do not always make decisions instantaneously; as Willekens, Klabunde, and ??? (2017) describe elsewhere in this special issue, the Theory of Planned Behaviour (Ajzen, 1991) can be used to describe the gradual process of intention forming, influence-absorption, information gathering and planning that terminate in a migration event. This drawn-out decision making process allows the possibility that agents may change their minds. Secondly, past experiences affect decisions in the present. This brings up questions about how the passage of time affects memory and attitudes. Finally, agents differ in their preferences between rewards realised immediately and equivalent rewards realised at some point in the future.

There are many good reasons why ignoring time in models of decision making might be valid or even desirable (and, as modellers, we have often done so). As was noted above, the context and research question must of course inform modelling decisions. It may be that the decisions in question have only short-term consequences, and so a detailed consideration of how agents treat time in their decision making be completely superfluous.

That said, as demographers, we are generally concerned with decisions which affect the life-course. These necessarily have consequences which extend far beyond the present (Willekens, 2001). For example, the decision to have a child may affect not only a parent's immediate satisfaction with life, labour market participation, income streams, consumption and so forth, but will continue to do so for the rest of their life. Future 'opportunity costs' must also be borne in mind (Butz & Ward, 1979): taking such a decision now may widen or narrow future options in various other areas of life. Similar considerations are salient for migration and partnership decisions, and indeed for many other demographic and social phenomena, for which the questions of time and sequencing of events tend to be too often ignored (see Abbott (2001) for an excellent overview).

This assumed sensitivity of demographic decision-making to the horizons over which the resultant costs and benefits are realised suggests that we should at least acknowledge what assumptions we make about how people value present and future rewards. The flexibility of agent-based models suggests that are is also scope to examine the effect of decision models that include time preference. Considering migration models, agents may learn about the potential rewards form migration from their social environment, and such information will influence a prediction about their future conditional on their migration decision. The influence of this prediction on the final decision may then depend on the time horizon over which rewards are realised. A brief summary of some current thinking on time-dependence in decision making follows; including such models in demographic models will certainly be challenging, and in many cases impossible, but being aware of the potential problems with making simplifying assumptions about treatment of time in decision models is valuable, and at least helps a modeller anticipate when such assumptions may fall down and why.

Many economic models do of course include time preference in decision making through the use of discounting; in migration modelling, this element is already present in the neoclassical framework, whereby the potential future gains from migration are discounted (see Massey et al. (1993) for an overview). This approach assumes that rewards received in the present are valued more highly by decision makers than those that will only be realised later. The most common functional form for such discounting is the exponential (O'Donoghue & Rabin, 2000):

$$U^t(u_t, u_{t+1}, \ldots, u_T) = \sum_{\tau=t}^{T} u_{t+\tau}\delta^{\tau}$$

Here, U^t, denoting present utility at time t, is a function of the stream of individual rewards u_t at each time point, weighted according to δ^{τ}, with δ being the discounting factor. This form has nice properties in that it is time-consistent; agents who prefer one option over another today will do so until the reward has been realised. As an example, consider a utility maximising individual who saves money to fund a migration attempt, and must choose between two destinations - one which is closer and thus cheaper to travel to, while the alternative is more distant and thus requires a longer saving period, but is also more attractive. If such a migrant discounts future rewards exponentially, once they have made a choice, such a decision will remain optimal until the trip is made, assuming the material circumstances do not change.

However, experimental evidence for both humans and animals suggests that this is not how time discounting is actually practised (Boyer, 2008). While we do prefer rewards now to those later, we do not do so in a consistent manner; preference orderings may change as we move forward in time. The functional form suggested by this sort of behaviour is hyperbolic (Benhabib, Bisin, & Schotter, 2010):

$$U^t(u_t, u_{t+1}, \ldots, u_T) = \sum_{\tau=t}^{T} \frac{u_{t+\tau}}{1+\delta\tau}$$

Such behaviour can be approximated by the use of quasi-hyperbolic or present-weighted exponential discounting, which, as the names suggests, accounts for time-inconsistency by including a weight on present rewards in the function, either fixed or variable (Benhabib et al., 2010). To return to the example of the migrant above, a hyperbolically or quasi-hyperbolically discounting agent may begin to save with the more distant, more attractive destination in mind, but actually move to the closer host country once they had saved enough for it to be an immediate option, even without receiving any new information about either.

Unfortunately, this is not the limit of potential complications around time. A large number of other contextual factors make a difference to the extent of discounting. For instance, it has been found that smaller rewards are more heavily discounted than larger rewards; that people treat a reductions in time to rewards differently to equivalent delays; and future losses treated as more significant than equivalent gains (Read & Loewenstein, 2000).

Another factor to consider is human reflexivity; often, we know that we are weak-willed, and we therefore take steps to restrain the compulsiveness of our future selves (O'Donoghue & Rabin, 2000). Such 'pre-commitment' may take the form of imposing extra costs on oneself should the planned course of action be deviated from. For instance, making a promise or public commitment to a colleague or friend introduces the risk of social costs should the actor renege.

To give a demographic example of how these models of time-preference could be important, the work of Wrede (2011) is instructive. He investigated the effect of hyperbolic discounting on fertility in a simple three period model where mothers could have both immediate rewards from motherhood, and delayed rewards in the form of being looked after in old age. He found that the introduction of hyperbolic discounting reduces numbers of births if this second, investment, motive is dominant, but that the ability to offset the immediate costs of motherhood through the financial tools might mitigate these effects. The idea of pre-commitment is also mentioned (although not modelled); in this context, an individual might limit their future freedom of action through sterilisation. Agent based models are flexible enough to embed such time-preference models in their wider social context (discussed in Section 2).

A crucial point is how much of these empirical and theoretical considerations with respect to time should be included in simulation in a particular case. A model of migration may include exponential or hyperbolical discounting of future income streams associated with migration, as is common in many analytic models (Carrington, Detragiache, & Vishwanath, 1996). This may be crucial to understanding behaviour, as perhaps an adjustment period of obtaining country-specific human captial leads to initially poor earnings (Dustmann, Fadlon, & Weiss, 2011), and so the pay-off to migration is disproptionately located at a longer time horizon. Alternatively, as a simplification, agents may be constructed to only consider a choice between current earnings and earnings that could be achieved on arrival. It may be that this type of simplification with respect to time-dependent decisions is sufficient to capture the essence of the process at hand. However, in general we would suggest that given that many demographic decisions have long term consequences, an awareness of the effect of the choices one is making about time and decision making is, we suggest, invaluable, and testing of different models of time-preference should be considered by modellers.

In order to allow for discounting of future reward streams, we must allow for agents to predict their future earnings and/or ability to save conditional on migration. Of course, in practice, predictions of such quantities are rarely perfect and can easily turn out incorrectly (Dustmann, 1997). Sozou (1998) describes how the hyperbolic discounting scheme can be derived from an exponential scheme where the reward may be lost in the intervening time period between intial consideration of the choice and realisation of the outcome; as the agent becomes more certain about their eventual receipt of any given reward, the attractiveness of the option increases. The possibility for error in prediction is just one of the many sources of uncertainty surrounding the agent's decision-making process and the attempt to model it.

3.2 Uncertainty

Uncertainty is a pervasive feature of any social, and thus also demographic reality. This uncertainty stems both from a lack of knowledge about processes and phenomena (*epistemic uncertainty*), as well as the inherent randomness of the world (*aleatory uncertainty* – see e.g. O'Hagan 2004). This means that social actors make decisions under conditions of imperfect knowledge, but also that the reality is random, unpredictable and can yield many surprises. In agent-based modelling, these features need to be reflected in the simluated representations of the world, both with respect to agents themselves, as well as features of their environment.

As discussed in Section 2, human decisions are made under the conditions of bounded rationality and loss aversion. In the context of uncertainty, risk aversion is equally important. The underpinning theory, dating back at least to Pratt (1964) and Arrow (1965), looks at the departures from the expected utility decision framework, and stipulates that risk-averse decision makers are willing to choose options with lower expected utility as long as they involve less, or no risk. Both aleatory and epistemic uncertainty are important here: through learning, agents can try reducing the epistemic uncertainty, but especially in the context of time-dependent agent based models, some of the uncertainty is irreducible, especially with respect to the decisions about future. This constitutes one of the main arguments for the importance of time, as argued in Section 3.1. Besides, varying attitudes to risk are an important source of heterogeneity of the agents, discussed in more detail in Section 3.3.

In the context of agent-based modelling of migration – itself one of the most uncertain demographic processes – there are some aspects where uncertainty just asks for being included in the models. For example, the uncertain nature of future benefits from migration, an important driver of population flows, has already been reflected in the neoclassical theories of migration, whereby earnings are weighted by the probability of employment, and future income streams are discounted to reflect the inter-temporal nature of decision making (Massey et al., 1993). In the New Economic Theory of migration, uncertainty comes to the fore of the decision-making process at the family or household level, and migration becomes one of the important tools of managing labour market risks (Stark & Blum, 1985). Uncertainty is at its most acute when we consider migration caused by extreme events, such as armed conflict or environmental catastrophes – under such circumstances, decisions need to be made rapidly, sometimes with a very limited insight into available options and the potential trade-offs between costs, benefits and risks.

At the level of model, there are different layers of uncertainty: in the parameters, model specification, computer code, observations, not to mention the inherent residual variability (Kennedy & O'Hagan, 2001). There are different ways to treat these manifestations of uncertainty in agent-based models. The use of statistical emulators or meta-models (Kleijnen & Sargent, 2000) – statistical models approximating the key uncertain relationships between the parameters of the underlying complex computational models – has been suggested by Kennedy and O'Hagan (2001) and Oakley and O'Hagan (2004), and more recently, specifically in the context of agent-based modelling, by Heard et al. (2015) and Hilton and Bijak (2016). Alternatives include the use of approximate Bayesian computation, as advocated by (Grazzini, Richiardi, & Tsionas, 2017), or hybrid approaches, whereby emulators are enhanced by direct samples from the output of the underlying agent-based models (Kamiński, 2015).

All these different approaches aim to describe uncertainty in complex models in a coherent way, and to allow meaningful (and ideally computationally efficient) inference from observed data. The methodological developments in this area are likely to continue: Heard et al. (2015) noted a need to further enhance the statistical methods specifically designed for dealing with agent-based models. However, even now the available tools enable a statistically rigorous – if at times approximate, as in the case of emulators – analysis of uncertainty in selected aspects of the models at hand. Coupled with the explicit acknowledgement of uncertainty in human decision making, an honest account of the aleatory and epistemic limitations of agent-based models seems indispensable for their further methodological advancement and uptake, in demography and beyond.

3.3 Heterogeneity in Decision- Making

In the context of the study of expert decision making, psychologist James Shanteau has criticised psychology for a focus on the study of the "Generalised Normal Adult Human Mind", while ignoring differences between people and contexts (Shanteau & Edwards, 2015). Agent-based models provide an opportunity to study two potential ways in which individuals may differ in methods of decision-making. The first - somewhat easier to operationalise and analyse - can be termed *parametric* differences, in which individuals are thought of as having the same underlying model of decision, but are different only in the way that these are parametrised. Many agent-based models include such differentiation. As an example, in the J. M. Epstein (2002) model of civil disobedience, individuals differ in the extent of their aversion to risk, but for a given level of risk aversion, agents will behave in the same way (holding all else constant). Notably, this heterogeneity in risk aversion is crucial to the behaviour of the model. Without it, the 'revolts' against authority which were a central feature of the model would never be triggered.

Secondly, individuals may differ in the *methods* they use to come to decisions in the same situations. Sociological models regarding frames and scripts provides some justification for thinking that this may be the case (Kronenberg, 2014)[3]. These suggest that individuals first attempt to select a frame for the

[3]Thanks to Sebastian Pink for drawing our attention to this strand of the literature.

situation they encounter, where are frames are mental model which answer the question: 'What kind of situation is this' . Conditional on the selected frame individuals, will then look to choose a relevant script - in answer to the hypothetical question: "What am I expected to do in this situation?" (ibid; 99). Thus, individuals may differ widely in their behaviour should they frame the situation differently, or if their cultural experiences suggest a different script. This second type of heterogeneity is potentially more difficult to operationalise, yet is more fundamental to the way the decision processes are described.

To put such discussion in a migration context, Bijwaard (2008) shows the importance of considering both temporary and permanent migration within a single migration stream; we can interpret this as migrants adopting different decision-making strategies, or equivalently adopting different frames and scripts. The use of agent-based models allows a range of potential scripts to be explored, with fewer constraints on what these scripts must look like.

As well as differences between agents, theories of frame and script selection also suggests difference within single agents over time; difference frames and scripts may be selected in response to different contexts, developing experiences, or influences from peers. This leads to discussion of another form of difference in decision making which may be important to demographers, that of differences over the life course. Migrants who were once intending to save for a period and return home to their family may change their mind and settle permanently in response to a new interpretation of their identity and circumstances (Constant & Massey, 2002).

As with representations of time, the relevance of heterogeneity in decision making is likely to be very context dependent. If heterogeneity exists, however, different mixes of agents could produce different macro-level results where interactions between agents feature heavily in the simulation, because of the possibility of non-linear interactions and emergence. Thus, researchers may wish to consider whether decision-making methods are likely to differ between individuals in their case.

4 The way forward: Multi-model approaches and modularity

In beginning to answer the question of which model of choice behaviour is the right one, we would do well to borrow an insight from computational neuroscience, and recognise that a productive understanding of decision making in a demographic context requires understanding at multiple levels (Marr, 1982; Marr & Poggio, 1976). High level models like the Theory of Planned Behaviour (Ajzen, 1991) can assist in understanding what decisions are being made, and why, and on the sequencing of the underpinning processes. Models of process, and the integration of context like that of Ben-Akiva et al. (2012), can inform about how decision making takes place, and what the processes involved are. Finally, these must be coupled with an appropriate implementation.

All of these problems are far from trivial, particularly as there is little consensus about which models of decision are best in which contexts. How then should we proceed with modelling when the fundamental questions about how individuals make decisions, how they treat time, and how they differ between themselves in these factors are still very uncertain? One way forward is to use the same simulation set-up to analyse multiple models of behaviour (J. M. Epstein, 2013; Grimm et al., 2005; Rossiter et al., 2014). This multi-model approach allows the modeller to attempt to distinguish between more and less plausible models of behaviour in the particular research context. Conditional on the simulation set-up being appropriate, more plausible models may be better able to reproduce multiple empirical patterns at varying scales (Bianchi, Cirillo, Gallegati, & Vagliasindi, 2008; Werker & Brenner, 2004).

This commitment to analysing multiple models of behaviour within a single simulation project is facilitated by a modular approach to simulation design, so that different models of behaviour can be swapped in and out without changing the underlying conditions of the simulation (J. M. Epstein, 2013), as can be seen in Gray et al. (2016). Such an approach reflects standard object-oriented design principles, but is not always reflected in an academic context (see also Rossiter (2015) for discussion of the application of software engineering principles to social simulation). Modularity also allows for the possibility of easy extension by other researchers with access to the code, allowing for further behavioural models to be examined within the same simulation framework. On this theme, Bell, Robinson, Malik, and Dewal (2015) highlight the atomistic nature of the agent-based modelling discipline, and advocate the introduction of Agent-based Modelling Primitives (AMPs) to enable general components of agent-based models to be packaged and re-used.

So, in the light of the above arguments, how to choose the appropriate representation of the human decisions and behaviour to be used in the agent-based modelling of demographic processes? First, we need to remember that agent-based models are a very convenient tool, with which we can integrate behavioural theory with social theory and data. This implies that as suggested in the Manifesto of computational social science (Conte et al., 2012) such models would ideally need to collect additional information on human decision making through bespoke cognitive experiments, instead of merely relying on hypotheses or assumptions (Courgeau, Bijak, Franck, & Silverman, 2016). In this way, the choice of a particular choice model will become an empirical issue, but specific to a given context and the decision problem of interest. Following this approach would allow for including empirical micro-foundations in computational demographic models in an open and transparent way and provide a natural, bottom-up and empirical way if evaluating the suitability of particular choice theories for the problem at hand.

Second, the existing methods of the statistical design of experiments (e.g. Chaloner and Verdinelli 1995, Kleijnen and Sargent 2000) would allow for adopting a systematic approach to model design and data collection. In this way, following the suggestions of Courgeau et al. (2016), the simulations would be built iteratively, by applying the experimental design principles to identifying

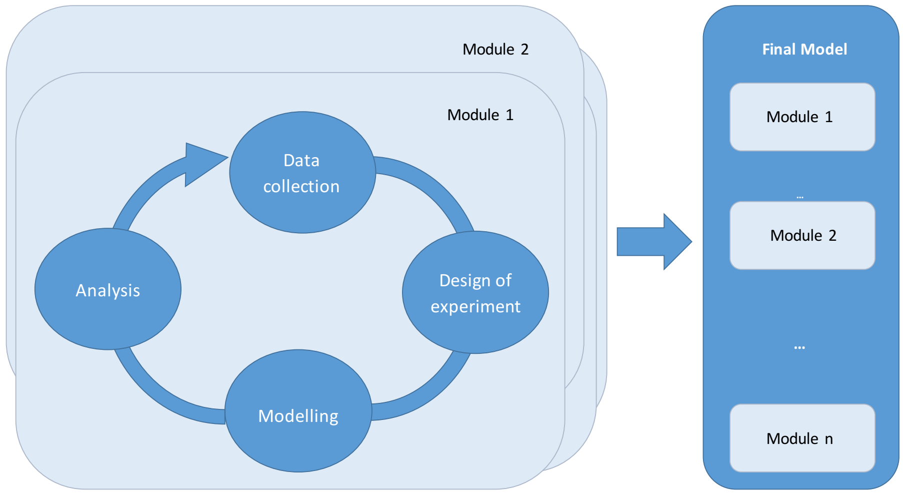

Figure 1: Choosing the choice: An outline of the approach

the elements of the model which require additional empirical information, and then by collecting this information. In this way, the important aspects of the decision process discussed above context, time and heterogeneity can be explicitly built into the model, through its specification or parameters. Then, they can be learnt about through repeated collection of information including through experiments. The modularity of the model would help this process, as it would allow for fine-tuning the different aspects of the model independently, before they can be assembled together. An outline of the process is illustrated in Figure 1.

How might this work in practice? We recommend to begin the model building process by identifying the key features and the context of the phenomena being modelled. In the example of migration, the key aspects of the model would include not only the different types of heterogeneous agents – migrants, non-migrants, institutions, intermediaries, policy makers, and so on – but also the available contextual information: the known and relevant features of the geographic, social, and economic environments in which the agents operate. The selection of these features would be informed by the theoretical literature on migration (see Massey et al. 1993 for examples), and to the greatest degree possible made realistic by including the empirical information – quantitative as well as qualitative – on what is known about migrant decision processes. The staged nature of decisions, as stipulated by the Theory of Planned Behaviour, could also be incorporated at this stage.

The second step of the model-buidling process would involve constructing a prototype, where the features that cannot be benchmarked to real-world data would be parameterised. Specifically in the context of decision making, such parameters could describe preferences related to risk or utility, aspects of decisions related to uncertainty or imperfect knowledge, time discounting. The choice of a model of decision making could be also parameterised: there may be many candidate frameworks (such as various heuristics, the Cumulative Prospect Theory or models of Bayesian reasoning – see Gray et al. 2016), each with their idiosyncratic features, that can apply to different situations or types of agents. Crucially, for the purpose of constructing the agent-based models, all the unknown parameters would need to be described in probabilistic terms, and embedded within a framework of statistical design of experiments Chaloner and Verdinelli (1995). In this way, the prototype agent-based model could be executed for different combinations of parameters, yielding different outputs.

The third step of the process would consist of calibrating the outputs of the models against the selected aspects of social reality, ideally in a probabilistic manner (Hilton & Bijak, 2016). Once the model is calibrated, its statistical properties can be analysed. In particular, sensitivity of the outputs to the different parameters can be assessed (Oakley & O'Hagan, 2004) – and the parameters that do not influence the model behaviour can be removed for the sake of parsimony. Conversely, we may want to find the empirical basis for those parameters, which appear to be highly influential for the chosen outputs, for example by collecting new data on the specific aspects of the agent-based model.

In particular, by calibrating the model, we can learn the choice of an appropriate representation of the decision processes of different agents. If the results of calibration do not point out to choosing a specific set of decision rules and models for various agents, this aspect of modelling can be enhanced by conducting cognitive experiments under controlled conditions. The sensitivity analysis would help identify, which aspects of the decision making process are the most important from the point of view of generating the observed outcomes – this could be the treatment of time, uncertainty, or other parameters. In the fourth step of the process, the decision problems faced by the agents can be reproduced under laboratory conditions in a series of tasks to be solved by human volunteers participating in such experiments. The additional insights from this exercise would enable fine-tuning and re-parameterising of the agent-based model in the light of experimental findings, as proposed in the Manifesto of computational social science (Conte et al., 2012). The second, third and fourth steps above can then be iterated until no more improvements can be reasonably made – what will be left is the residual uncertainty that is an irreducible feature of all models, and especially such complex ones.

Several examples of agent-based models presented in this Special Issue already contain some of the building blocks of the proposed approach. The multistage nature of the decision making processes in the context of migration has been explored by Willekens et al. (2017) and Kley (2017), both drawing from the Theory of Planned Behaviour (Ajzen, 1991) and grounded in available empirical information. In addition, Warnke, Reinhardt, Klabunde, Willekens, and Uhrmacher (2017) have provided a stochastic description of the underlying decision making processes, and have proposed a corresponding programming language designed to facilitate the implementation and execution of agent-based models. The natural next step would be to use these building blocks, and others – such as the experimental design, cognitive psychological experiments on decision making, or statistical analysis of uncertainty – for constructing more robust and empirically grounded models of human decisions, which could then be embedded within agent-based models of human populations.

As argued throughout the Special Issue, all these features, and more, make agent-based modelling a very attractive analytical proposition for demographers. The potential gains in understanding of the underlying population processes cannot be overstated, and further links connecting agent-based and statistical demography are yet to be explored (see for example Bijak and Bryant (2016)). That at present there is no methodological consensus on modelling agency and human decisions is not an obstacle, but rather a challenge to guide the further work in this area. Given that over the last 15 years agent-based models have slowly begun to enter the demographic mainstream (Billari & Prskawetz, 2003; Van Bavel & Grow, 2016), the gaps in knowledge in the existing approaches have become clearly visible. In our view one of the important gaps is related to the agency and decision making processes of the simulated agents, and the related aspects, some of which context, time, and a variety of forms of decision making are discussed above. Now is the time to initiate the discussion about choosing the choice.

References

Abbott, A. (2001). *Time matters. On theory and method.* Chicago: University of Chicago Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Arrow, K. J. (1965). *Aspects of The Theory of Risk-Bearing (Yrjö Jahnsson Lectures).* Helsinki: Yrjö Jahnssonin Säätiö.

Baltas, G., & Doyle, P. (2001). Random utility models in marketing research: a survey. *Journal of Business Research*, *51*(2), 115–125.

Bell, A. R., Robinson, D. T., Malik, A., & Dewal, S. (2015). Modular ABM development for improved dissemination and training. *Environmental Modelling & Software*, *73*, 189–200.

Ben-Akiva, M., de Palma, A., McFadden, D., Abou-Zeid, M., Chiappori, P. A., de Lapparent, M., ... Walker, J. (2012). Process and context in choice models. *Marketing Letters*, *23*(2), 439–456.

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, *69*(2), 205–223.

Bianchi, C., Cirillo, P., Gallegati, M., & Vagliasindi, P. A. (2008). Validation in agent-based models: An investigation on the CATS model. *Journal of Economic Behavior & Organization*, *67*(3-4), 947–964.

Bijak, J., & Bryant, J. (2016). Bayesian Demography 250 Years after Bayes. *Population Studies*, *70*(1). (Forthcoming)

Bijwaard, G. E. (2008, dec). Immigrant migration dynamics model for The Netherlands. *Journal of Population Economics*, *23*(4), 1213–1247.

Billari, F., & Prskawetz, A. (Eds.). (2003). *Agent-Based Computational Demography: Using Simulation to Improve Our Understanding of Demographic Behaviour*. Dordrecht: Plenum.

Boyer, P. (2008). Evolutionary economics of mental time travel? *Trends in Cognitive Sciences*, *12*(6), 219–224.

Butz, W. P., & Ward, M. P. (1979). The Emergence of Countercyclical U . S . Fertility. *American Economic Review*, *69*(3), 318–328.

Carrington, W., Detragiache, E., & Vishwanath, T. (1996). Migration with endogenous moving costs. *The American Economic Review*, *86*(4), 909–930.

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, *10*(3), 273-304.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

Cioffi-Revilla, C. (2010). A methodology for complex social simulations. *Journal of Artificial Societies and Social Simulation*, *13*(1), 7.

Constant, A., & Massey, D. S. (2002, sep). Return Migration by German Guestworkers: Neoclassical versus New Economic Theories. *International Migration*, *40*(4), 5–38. doi: 10.1111/1468-2435.00204

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., . . . Helbing, D. (2012). Manifesto of computation social science. *European Journal Special Topics*(214), 325–346.

Courgeau, D., Bijak, J., Franck, R., & Silverman, E. (2016). Model-Based Demography: Towards a Research Agenda. In J. Van Bavel & A. Grow (Eds.), *Agent-based modeling in population studies: Concepts, methods, and applications.* Dordrecht: Springer. (Forthcoming)

Dustmann, C. (1997). Return migration, uncertainty and precautionary savings. *Journal of development economics*, *52*(2), 295–316.

Dustmann, C., Fadlon, I., & Weiss, Y. (2011). Return migration, human capital accumulation and the brain drain. *Journal of Development Economics*, *95*(1), 58–67.

Epstein, G. S., & Gang, I. N. (2006). The influence of others on migration plans. *Review of Development Economics*, *10*(4), 652–665.

Epstein, J. M. (2002). Modeling civil violence: an agent-based computational approach. *Proceedings of the National Academy of Sciences of the United States of America*, *99 Suppl 3*(2), 7243–50.

Epstein, J. M. (2013). *Agent_Zero: Towards Neurocognitive Foundations for Generative Social Science.* Princeton: Princeton Univeristy Press.

Espindola, A. L., Silveira, J. J., & Penna, T. J. P. (2006). A Harris-Todaro Agent-Based Model to Rural-Urban Migration. *Brazilian Journal of Physics*, *36*(3A), 603–609.

Filho, H. S. B., de Lima Neto, F. B., & Fusco, W. (2011). Migration and social networks - An explanatory multi-evolutionary agent-based model. In *2011*

IEEE symposium on intelligent agent (IA) (pp. 1–7). IEEE.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, *103*(4), 650–69.

Gray, J., Bijak, J., & Bullock, S. (2016). Deciding to Disclose: A Decision Theoretic Agent Model of Pregnancy and Alcohol Misuse. In J. Van Bavel & A. Grow (Eds.), *Agent-based modeling in population studies: Concepts, methods, and applications.* Dordrecht: Springer. (Forthcoming)

Grazzini, J., Richiardi, M. G., & Tsionas, M. (2017). Bayesian estimation of agent-based models. *Journal of Economic Dynamics & Control*, *77*(1), 26–47.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., ... DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science*, *310*(5750), 987–991.

Heard, D., Dent, G., Schifeling, T., & Banks, D. (2015). Agent-Based Models and Microsimulation. *Annual Review of Statistics and Its Application*, *2*(April 2015), 259–272.

Hilton, J., & Bijak, J. (2016). Deciding to Disclose: A Decision Theoretic Agent Model of Pregnancy and Alcohol Misuse. In J. Van Bavel & A. Grow (Eds.), *Agent-based modeling in population studies: Concepts, methods, and applications* (pp. 211–235). Dordrecht: Springer.

Jern, A., Lucas, C. G., & Kemp, C. (2011). Evaluating the inverse decision-making approach to preference learning. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 24* (pp. 2276–2284). Curran Associates, Inc.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, *47*(2), 263–291.

Kamiński, B. (2015). Bayesian estimation of agent-based models. *European Journal of Operational Research*, *247*(3), 859–866.

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(3), 425–464.

Klabunde, A. (2014). Computational Economic Modeling of Migration. *Ruhr Economic Papers 471*.

Klabunde, A., & Willekens, F. (2016). Decision-making in agent-based models of migration: state of the art and challenges. *European Journal of Population*, *32*(1), 73–97.

Kleijnen, J. P., & Sargent, R. G. (2000). A methodology for fitting and validating metamodels in simulation. *The American Economic Review*, *120*(1), 14–29.

Kley, S. (2017). Facilitators and constraints at each stage of the migration process. *Population Studies*, *71*(This issue).

Kronenberg, C. (2014). Frames, scripts and variable rationality: An integrative theory of action. In G. Manzo (Ed.), *Analytical sociology: Actions and networks* (pp. 97–123). Chichester: Wiley.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Rep-*

resentation and Processing of Visual Information. San Francisco: W. H. Freeman and Company.

Marr, D., & Poggio, T. (1976). From understanding computation to understanding neural circuitry. *Neurosciences Research Progress Bulletin*, *15*, 470–488.

Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating Social Influence Processes: Dynamic, Multidirectional Flows of Influence Within Social Networks. *Personality and Social Psychology Review*, *11*(3), 279–300.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & E, T. J. (1993). Theories of International Migration: Review and Appraisal. *Population and Development Review*, *19*(3), 431–466.

Mather, M., Shafir, E., & Johnson, M. K. (2000). Misremembrance of Options Past: Source Monitoring and Choice. *Psychological Science*, *11*(2), 132–138.

McKenzie, D., Gibson, J., & Stillman, S. (2013). A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad? *Journal of Development Economics*, *102*(2788), 116–127.

Oakley, J. E., & O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *66*(3), 751–769.

O'Donoghue, T., & Rabin, M. (2000). The economics of immediate gratification. *Journal of Behavioral Decision Making*, *13*(2), 233–250.

O'Hagan, T. (2004). Dicing with the unknown. *Significance*, *1*(3), 132–133.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, *32*(1–2), 122–136.

Read, D., & Loewenstein, G. (2000). Time and decision: introduction to the special issue. *Journal of Behavioral Decision Making*, *13*(2), 141–144.

Robbins, H. (1964). The empirical bayes approach to statistical decision problems. *The Annals of Mathematical Statistics*, *35*(1), 1–20.

Rossiter, S. (2015). Simulation Design: Trans-Paradigm Best-Practice from Software Engineering. *Journal of Artificial Societies and Social Simulation*, *18*(3), 9.

Rossiter, S., Noble, J., & Bell, K. R. W. (2014, sep). Social simulation comparison in arbitrary problem domains: first steps towards a more principled approach. In S.-H. Chen, T. Terano, R. Yamamoto, & C.-C. Tai (Eds.), *Advances in computation social science: The fourth world congress* (pp. 329–345). Taipei: Springer.

Schmolke, A., Thorbek, P., DeAngelis, D. L., & Grimm, V. (2010). Ecological models supporting environmental decision making: A strategy for the future. *Trends in Ecology and Evolution*, *25*(8), 479–486.

Shanteau, J., & Edwards, W. (2015). Decision Making by Experts: Influence of Five Key Psychologists. In E. A. Wilhelms & V. F. Reyna (Eds.), *Neuroeconomics, judgement, and decision making* (pp. 3–26). New York: Psychology Press.

Silveira, J. J., Espíndola, A. L., & Penna, T. (2006). Agent-based model to rural-urban migration analysis. *Physica A: Statistical Mechanics and its Applications*, *364*, 445–456.

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society B: Biological Sciences*, *265*(1409), 2015–2020.

Stark, O., & Blum, D. E. (1985). The New Economics of Labor Migration. *The American Economic Review*, *75*(2), 173–178.

Thiele, J. C., Kurth, W., & Grimm, V. (2014). Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and R. *Journal of Artificial Societies and Simulation*, *17*(3), 11.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Van Bavel, J., & Grow, A. (Eds.). (2016). *Agent-based Modelling in Population Studies: Concepts, Methods, and Applications.* Dordrecht: Springer. (Forthcoming)

Warnke, T., Reinhardt, O., Klabunde, A., Willekens, F., & Uhrmacher, A. M. (2017). Modelling and Simulating Decision Processes of Linked Lives – An Approach Based on Concurrent Processes and Stochastic Race. *Population Studies*, *71*(This issue).

Werker, C., & Brenner, T. (2004). *Empirical Calibration of Simulation Models.* Max Planck Institute of Economics, Jena.

Willekens, F. (2001). Theoretical and technical orientations toward longitudinal research in the social sciences. *Canadian Studies in Population*, *28*(2), 189–217.

Willekens, F., Klabunde, A., & ??? (2017). ??? *Population Studies*, *??*(??), ?

Wrede, M. (2011). Hyperbolic discounting and fertility. *Journal of Population Economics*, *24*(3), 1053–1070.