Chapter 1
Introduction

Jakub Bijak

Population processes, including migration, are complex and uncertain. We begin this book by providing a rationale for building Bayesian agent-based models for population phenomena, specifically in the context of migration, which is one of the most uncertain and complex demographic processes. The main objectives of the book are to pursue methodological advancement in demography and migration studies through combining agent-based modelling with empirical data, Bayesian statistical inference, appropriate computational techniques, and psychological experiments in a streamlined modelling process, with the overarching aim to contribute to furthering the model-based research agenda in demography and broader social sciences. In this introductory chapter, we also offer an overview of the structure of this book, and present various ways in which different audiences can approach the contents, depending on their background and needs.

1.1 Why Bayesian Model-Based Approaches for Studying Migration?

Migration processes are characterised by large complexity and uncertainty, being some of the most uncertain drivers of population change (NRC, 2000). At the same time, migration is one of the most politically sensitive demographic phenomena in contemporary Europe (Castles et al., 2014). In a nutshell, migration is an increasingly more powerful driver of overall population dynamics across developed countries (Bijak et al., 2007; Castles et al., 2014), is socially and politically contentious, as well as being a top-priority, high-impact policy area (e.g. European Commission, 2015, 2020; UN, 2016). The so-called Syrian asylum crisis of 2015–16, and its impact on Europe and European policy and politics are prime examples of the urgent need for sound and robust scientific advice in this domain.

J. Bijak, *Towards Bayesian Model-Based Demography*, Methodos Series 17,
https://doi.org/10.1007/978-3-030-83039-7_1

Unfortunately, theoretical foundations of migration remain weak and fragmented (Arango, 2000; McAuliffe & Koser, 2017), which is also to some extent true for other areas of demography (Burch, 2018). In the case of migration, tensions and trade-offs between high-level structural forces shaping the population flows and the agency of individual migrants are explicitly recognised as defining aspects of population mobility (de Haas, 2010; Carling & Schewel, 2018). Complex interrelations between various types of migration drivers operating at different levels – from individuals, to groups, to societies and states – call for more sophisticated methods of analysis than has been the case so far (Van Hear et al., 2018).

For all these reasons, among the different areas of population studies, there is a strong need to increase our understanding of migration processes. Addressing the challenges of the future requires the ability to comprehend and explain migration much better and more deeply than ever before. Currently, there is a gap between the demand for knowledge about migration, and the state of the art in this area.

From the point of view of quantitative population studies, especially those focused of human mobility, there is an acute need to fill a crucial void in formal modelling by offering new insights into the explanation of the underlying processes. Only in that way can social science help address important societal and population challenges: how the demographic processes, such as migration, can be better understood, predicted and managed. Previous efforts in that domain were largely constrained to simple approaches, with the explanatory endeavours lagging behind (for a review of formal modelling approaches from a predictive angle, see Bijak, 2010).

This book offers to fill this methodological void by presenting an innovative process for building simulation models of social processes, illustrated by an example of asylum migration, which aims to integrate behavioural and social theory with formal methods of analysis. Its key contribution is to combine in one book, novel methods and approaches of migration modelling, embedded in a joint analytical framework, while addressing some of the well-recognised philosophical challenges of model-based approaches. In particular, our main innovations include insights into human decisions and applying the formal rigour of statistical analysis to evaluate the modelling results. This combination offers novel and unique insights into some of the most challenging areas of demography and social sciences more broadly. It also bears a promise of influencing not only academics, but also practitioners and decision makers – in the area of migration and beyond – by offering methodological advice for policy-relevant simulations, and by providing a framework for decision support on their basis.

1.2 Aims and Scope of the Book

This book presents and reflects on the process of developing a simulation model of international migration route formation, with a population of intelligent, cognitive agents, their social networks, and policy-making institutions, all interacting with one another. The overarching aim of this work is to bring new insights into the

theoretical and methodological foundations of demographic and migration studies, by proposing a blueprint for an interdisciplinary modelling process. In substantive terms, we aim at answering the following general question: how to introduce theoretical micro-foundations to demographic simulation studies, in particular, those of migration flows?

To that end, the book proposes a process for developing such micro-foundations for migration and other population studies through interdisciplinary efforts centred around agent-based modelling. The design of the modelling approach advocated in this volume follows recent developments in demography, computational modelling, statistics, cognitive psychology and computer science. In addition, we also offer a practical discussion on application of the proposed model-based approach by discussing a range of programming languages and environments.

In terms of the application area, the book sets out to address one of the most uncertain, complex and highest-impact population processes – international migration – which is situated at the intersection between demography and other social sciences. To address the challenges, we build on the existing literature from across a range of disciplines, incorporating in practice some of the ideas that have been proposed in terms of furthering the philosophical, theoretical and methodological perspectives involving computational social modelling.

Throughout this book, the methodological challenges of studying migration are thus addressed by bringing together interdisciplinary expertise from demography, statistics, cognitive psychology, as well as computer and complexity science. Combining them in a common analytical framework has a potential to move beyond the current state of affairs, which is largely developing in silos delineated by disciplinary boundaries (Arango, 2000). The proposed solutions can offer broader and generic methodological suggestions for analysing migration – a contemporary topic of global significance.

In particular, we offer a template for including in computational demographic models psychologically realistic micro-foundations, with an empirical basis – an aspect that is contemporarily lacking not only in migration research, but also in population studies more broadly. At the same time, the approach advocated here enables us to acknowledge and describe the fundamental epistemological limits of migration models in a formal way. To that end, some of the broader objectives of this programme of work include: identifying the inherently uncertain aspects of migration modelling, formally describing their uncertainty, providing policy recommendations under different levels of predictability of various processes, and finally offering guidance for further data collection.

In terms of the scope, the book discusses in detail the different stages and building blocks for constructing an empirically grounded simulation model of migration, and for embedding the modelling process within a wider framework of Bayesian experimental design. We use statistical principles to devise innovative computer-based simulation experiments, and to learn about the simulated processes as well as individual agents and the way they make decisions. The identified knowledge gaps are filled with information from dedicated psychological experiments on cognitive aspects of human decision making under uncertainty. In this way, the models are

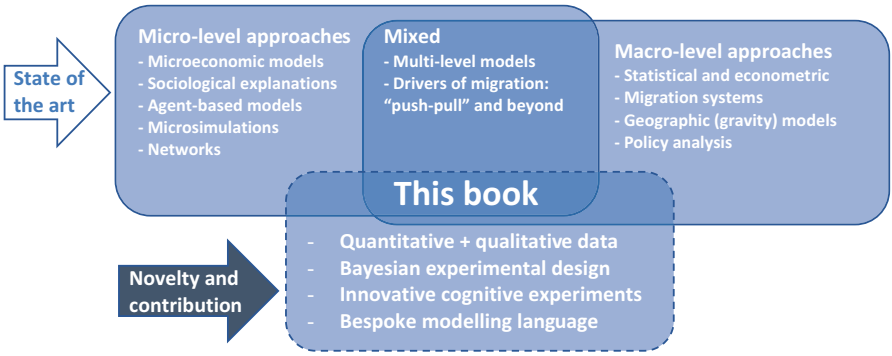

Fig. 1.1 Position of the proposed approach among formal migration modelling methods. (Source: own elaboration, based on Bijak (2010: 48))

built inductively, from the bottom up, addressing important epistemological limitations of population sciences.

The book builds upon the foundations laid out in the existing body of work, at the same time aiming to address the methodological and practical challenges identified in the recent population and migration modelling literature. Starting from a previous review of formal models of migration (Bijak, 2010), our proposed approach is specifically based on the five elements that have not been combined in modelling before. In particular, the existing micro-level approaches to migration studies, including microeconomic and sociological explanations, as well as inspirations from existing agent-based and microsimulation models, are combined here with macro-level statistical analysis of migration processes and outcomes, with the ultimate aim of informing decisions and policy analysis (see Fig. 1.1).

The novel elements included in this book additionally include combining qualitative and quantitative data in the formal modelling process (Polhill et al., 2010), learning about social mechanisms through Bayesian methods of experimental design, as well as including experimental information on human decision making and behaviour. Additionally, we develop further a dedicated programming language, ML3, to facilitate modelling migration, extending the earlier work in that area (Warnke et al., 2017). These different themes draw from the existing state of the art in migration modelling, and enhance it by adding new elements, as summarised in Fig. 1.1.

From the scientific angle, we aim to advance both the philosophical and practical aspects of modelling. This is done, first, by applying the concepts and ideas suggested in the contemporary literature to develop a model of migration routes in an iterative, multi-stage process. Second, these parallel aims are addressed by offering practical solutions for implementing and furthering the model-based research programme in demography (van Bavel & Grow, 2016; Courgeau et al., 2016; Silverman, 2018; Burch, 2018), and in social sciences more broadly (Hedström & Swedberg, 1998; Franck, 2002; Hedström, 2005).

The book draws inspiration from a wide literature. From a philosophical perspective, key ideas that underpin the theoretical discussions in this book can be found in Franck (2002), Courgeau (2012), Courgeau et al. (2016), Silverman (2018) and Burch (2018). The practical aspects of the many desired features of modelling involved, including the need for modular nature of model construction, were called for by Gray et al. (2017) and Richiardi (2017), while the need for additional, non-traditional sources of information, including qualitative and experimental data, was advocated by Polhill et al. (2010) and Conte et al. (2012), respectively.

At the same time, methods for a statistical analysis of computational experiments have also been discussed in many important reference works, for example in Santner et al. (2003). Specific applications of the existing statistical methods of analysing agent-based models can be found in Ševčíková et al. (2007), Bijak et al. (2013), Pope and Gimblett (2015) or Grazzini et al. (2017). The use of such methods – mainly Bayesian – have also been suggested elsewhere in the demographic literature, for example by Willekens et al. (2017). To that end, we propose a coherent methodology for embedding the model development process into a wider framework of Bayesian statistics and experimental design, offering a blueprint for an iterative process of construction and statistical analysis of computational models for social realms.

1.3 Structure of the Book

We have divided this book into three parts, devoted to: Preliminaries (Part I), Elements of the modelling process (Part II), and Model results, applications, and reflections (Part III). This structure enables different readers to focus on specific areas, depending on interest, without necessarily having to read the more technical details referring to individual aspects of the modelling process.

Part I lays down the foundations for the presented work. Chapter 2 focuses on the rationale and philosophical underpinnings of the Bayesian model-based approach. The discussion starts with general remarks on uncertainty and complexity in demography and migration studies. The uncertainty of migration processes is briefly reviewed, with focus on the ambiguities in the concepts, definitions and imprecise measurement; simplifications and pitfalls of the attempts at explanation; and on inherently uncertain predictions. A risk-management typology of international migration flows is revisited, focusing on asylum migration as the most uncertain and highest-impact form of mobility. In this context, we discuss the rationale for using computational models for asylum migration. To address the challenges posed by such complex and uncertain processes as migration, we seek inspiration in different philosophical foundations of demographic epistemology: inductive, deductive and abductive (inference to the best explanation). Against this background, we introduce a research programme of model-based demography, and evaluate its practical applicability to studying migration.

Part II presents five elements of the proposed modelling process – the building blocks of Bayesian model-based description and analysis of the emergence of migration routes. It begins in Chap. 3 with a high-level discussion of the process of developing agent-based models, starting from general principles, and then moving focus to the specific example of migration. We review and evaluate existing examples of agent-based migration models in the light of a discussion of the role of formal modelling in (social) sciences. Next, we discuss the different parts of migration models, including their spatial dimension, treatment of various sources of uncertainty, human decisions, social interactions and the role of information. The discussion is illustrated by presenting a prototype, theoretical model of migrant route formation and the role of information exchange, called Routes and Rumours, which is further developed in subsequent parts of the book, and used as a running example to illustrate different aspects of the model-building process. The chapter concludes by identifying the main knowledge gaps in the existing models of migration. This chapter is accompanied by Appendix A, where the architecture of the Routes and Rumours model is described in more detail.

Chapter 4 introduces the motivating example for the application of the Routes and Rumours model – asylum migration from Syria to Europe, linked to the so-called European asylum crisis of 2015–16. In this chapter, we present the process of constructing a dedicated knowledge base. The starting point is a discussion of various types of quantitative and qualitative data that can be used in formal modelling, including information on migration concepts, theories, factors, drivers and mechanisms. We also briefly present the case study of Syrian asylum migration. Subsequently, the data related to the case study are catalogued and formally assessed by using a common quality framework. We conclude by proposing a blueprint for including different data types in modelling. The chapter is supplemented by detailed meta-inventory and quality assessment of data, provided in Appendix B and available online, on the website of the research project Bayesian Agent-based Population Studies, underpinning the work presented throughout this book (www.baps-project.eu).

Chapter 5 is dedicated to presenting the general framework for analysing the results of computational models of migration. First, we offer a description of the statistical aspects of the model construction process, starting from a brief tutorial on uncertainty quantification in complex computational models. The tutorial includes Bayesian methods of uncertainty quantification; an introduction to experimental design; the theory of meta-modelling and emulators; methods for uncertainty and sensitivity analysis, as well as calibration. The general setup for designing and running computer experiments with agent-based migration models is illustrated by a running example based on the Routes and Rumours model introduced in Chap. 3. The accompanying Appendix C contains selected results of the illustrative uncertainty and sensitivity analysis presented in this chapter, as well as a brief overview of software packages for carrying out the experimental design and model analysis.

The cognitive psychological experiments are discussed in Chap. 6, following the rationale for making agent-based models more realistic and empirically grounded. Building on the psychological literature on decision making under uncertainty, the

chapter starts with an overview of the design of cognitive experiments. This is followed by a presentation of three such experiments, focusing on discrete choice under uncertainty, elicitation of subjective probabilities and risk, and choice between leading migration drivers. We conclude the chapter by providing reflections on including the results of experiments in agent-based models, and the potential of using immersive interactive experiments in this context. Supplementary material included in Appendix D contains information on the study protocol and selected ethical aspects of experimental research and data collection.

Chapter 7, concluding the second part of the book, presents the computational aspects of the modelling work. We discuss the key features of domain-specific and general-purpose programming languages, by using an example of languages recently developed for demographic applications. In particular, the discussion focuses on modelling, model execution, and running simulation experiments in different languages. The key contributions of this chapter are to present a bespoke domain-specific language, aimed at combining agent-based modelling with simulation experiments, and formally describing the logical structure of models by using a concept of provenance modelling. Appendix E includes further information about the provenance description of the migration simulation models developed throughout this book, based on the Routes and Rumours template.

Part III offers a reflection on the selected outcomes of the modelling process and their potential scientific and policy implications. In particular, Chap. 8 is devoted to discussing the results of applying the model-based analytical template, combining all the building blocks listed above, and aimed at answering specific substantive research questions. We therefore follow the model development process, from the purely theoretical version to a more realistic one, called Risk and Rumours, subsequently including additional empirical and experimental data, in the version called Risk and Rumours with Reality. At the core of this chapter are the results of experiments with different models, and the analysis of their sensitivity and uncertainty. Subsequently, we reflect on the model-building process and computational implementation of the models, as well as their key limitations. The chapter concludes by exploring the remaining (residual) uncertainty in the models, and highlighting areas for future data collection. The underlying model architecture is an extension of the Routes and Rumours one, presented in Chap. 3 and Appendix A.

Subsequently, in Chap. 9, we outline the scientific and policy implications of modelling and its results. First, we discuss perspectives for furthering the model-based research agenda in social sciences, reflecting on the scientific risk-benefit trade-offs of the proposed approach. The usefulness of modelling for policy is then explored through a variety of possible uses, from scenario analysis, to foresight studies, stress testing and calibration of early warnings. To that end, we also present several migration scenarios, based on two models introduced in Chap. 8 (Risk and Rumours, and Risk and Rumours with Reality), aiming to simulate the impacts of actual policy decisions using an example of a risk-related information campaign. The chapter concludes with a discussion of the key limitations and practical recommendations for the users of the model-based approach.

The discussion in Chap. 10 focuses on the key role of transparency and replicability in modelling. Starting from a summary of the recent 'replicability crisis' in psychology, and lessons learned from this experience, we offer additional arguments for strengthening the formal documentation of the models constructed, including through the use of formal provenance modelling. The general implications for modelling and modellers, as well as for the users of models, are presented next.

Finally, the simulation results serve as a starting point for a broader reflection on the potential contribution of simulation-based approaches to migration research and social sciences generally. In that spirit, Chap. 11 concludes the book by summarising the theoretical, methodological and practical outcomes of the approach presented in the book in the light of recent developments in population and migration studies. We present further potential and limitations of Bayesian model-based approaches, alongside the lessons learned from implementing the modelling process proposed in the book. Key practical implications for migration policy are also summarised. As concluding thoughts, we discuss ways forward for developing statistically embedded model-based computational approaches, including an assessment of the viability of the whole model-based research programme.

1.4 Intended Audience and Different Paths Through the Book

The book is written by an interdisciplinary team with combined expertise in demography and migration studies, agent-based simulation modelling, statistical analysis and uncertainty quantification, experimental psychology and meta-cognition, as well as computer programming and simulations. We hope to demonstrate how adopting such a broad multidisciplinary approach within a common, rigorous and formal research framework opens up further exciting research possibilities in social sciences, and can help offer methodological recommendations for policy-relevant simulations. Practical applications are aided by intuitive programming advice for implementing and documenting the Bayesian model-based approach to answer real-life scientific and policy questions.

This book is primarily intended for academic and policy audiences, and aspires to stimulate new research opportunities. We hope that the presented work will be of interest to two types of academic readers. First, for demographers, sociologists, human geographers and migration scholars, it provides new methodological and philosophical insights into the possibilities offered by applying statistical rigour and empirical grounding of model-based approaches. In this way, we hope that computational demography – and demography and social sciences more generally – will benefit from engagement with new statistical, cognitive and computer science perspectives through formal, interdisciplinary modelling endeavours, which are offered throughout the whole book.

Second, for statisticians, complexity and computer scientists, as well as experimental psychologists, the book presents a case study of how the methods and approaches developed in their respective disciplines can be used elsewhere, under a common analytical umbrella. Demography can offer here a fascinating and contemporary area for the application of such research methods in a truly multi-disciplinary manner, opening up the scope for further methodological advancements. For such readers, the respective Chaps. 3, 4, 5, 6, and 7 are likely to be of interest, alongside Part III.

For non-academic readers from the areas of policy, government and civil service, working on migration, asylum, and in related domains, such as border protection, humanitarian aid, service provision, or human rights, the relevant outcomes are summarised primarily in Part III, tailored for practical applications. The focus of that part is on illustrating the possible uses of simulations by policy makers to test different scenarios concerning migration and related processes. Here, and particularly in Chap. 9, we present several ways to evaluate the efficacy of migration management measures through simulations and experimentation on a computer (*in silico*), under controlled, yet realistic conditions. More generally, such results can be of interest for policy think-tanks, government and parliamentary researchers, advisors, and independent experts as well.

Finally, the book can be used as supplementary reading for postgraduate courses, doctoral studies, and dedicated professional development training programmes, especially in the areas of formal and statistical demography, complexity science, or formal sociology. Here, we assume the prior knowledge of basic tenets of modelling and Bayesian statistics, and where relevant refer the readers to some of the key reference works and textbooks. Selected excerpts from the book, especially from Part I, can be also suited for final-year undergraduate courses in demography and complexity science, especially on methods-oriented programmes.