Computational Social Sciences

Peter Davis
Roy Lay-Yee

Simulating Societal Change

Counterfactual Modelling for Social and Policy Inquiry

Springer

Computational Social Sciences

Computational Social Sciences

A series of authored and edited monographs that utilize quantitative and computational methods to model, analyze and interpret large-scale social phenomena. Titles within the series contain methods and practices that test and develop theories of complex social processes through bottom-up modeling of social interactions. Of particular interest is the study of the co-evolution of modern communication technology and social behavior and norms, in connection with emerging issues such as trust, risk, security and privacy in novel socio-technical environments.

Computational Social Sciences is explicitly transdisciplinary: quantitative methods from fields such as dynamical systems, artificial intelligence, network theory, agent-based modeling, and statistical mechanics are invoked and combined with state-of-the-art mining and analysis of large data sets to help us understand social agents, their interactions on and offline, and the effect of these interactions at the macro level. Topics include, but are not limited to social networks and media, dynamics of opinions, cultures and conflicts, socio-technical co-evolution and social psychology. Computational Social Sciences will also publish monographs and selected edited contributions from specialized conferences and workshops specifically aimed at communicating new findings to a large transdisciplinary audience. A fundamental goal of the series is to provide a single forum within which commonalities and differences in the workings of this field may be discerned, hence leading to deeper insight and understanding.

Peter Davis • Roy Lay-Yee

Simulating Societal Change

Counterfactual Modelling for Social and Policy Inquiry

Springer

Peter Davis
Department of Statistics
University of Auckland
Auckland, New Zealand

Roy Lay-Yee
COMPASS (Centre of Methods and Policy Application in the Social Sciences)
Research Centre
University of Auckland
Auckland, New Zealand

ISSN 2509-9574 ISSN 2509-9582 (electronic)
Computational Social Sciences
ISBN 978-3-030-04785-6 ISBN 978-3-030-04786-3 (eBook)
https://doi.org/10.1007/978-3-030-04786-3

Library of Congress Control Number: 2018963222

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Statistics New Zealand Disclaimer

Access to the data used in this study was provided by Statistics New Zealand under conditions designed to give effect to the security and confidentiality provisions of the Statistics Act 1975. The results presented in this book are the work of the authors, not Statistics New Zealand.

Acknowledgments

We wish to thank the following for assistance with the production of this book:

The Royal Society of New Zealand, for awarding a 2-year James Cook Fellowship
The Faculty of Arts, especially the Dean, Professor Robert Greenberg, for bridging finance
The Te Pūnaha Matatini Centre of Research Excellence, for salary support
Statistics New Zealand – access to microdata from census and New Zealand Longitudinal Census

COMPASS Research Centre, School of Social Sciences, as host:

Technical team – Kevin Chang, Martin von Randow, Chris Liu
Early software development – Oliver Mannion, Janet Pearson, Jessica McLay
Technical advice and support – Barry Milne, Nichola Shackleton
Use of the remote DataLab facility – located at COMPASS

International adviser – Martin Spielauer
Library services adviser – Mark Hangartner

Finally, thanks to parents as mentors and role models, partners for their forbearance, and families and friends for their encouragement.

Contents

Chapter 1
Introduction

In this book we intend to demonstrate that methodological innovation in the application of quantitative and computational techniques is an important part of the future for a sociology that is a population and policy science able to address some of the big issues facing society. Our book does its work in a particular society – namely, New Zealand – and does so over a defined period of rapid social and economic change leading up to the turn of the millennium. But, just as important, we do our work with the assistance of a uniquely dynamic and representative set of linked data (the longitudinal census), and we do so in a highly innovative and technically accomplished way by building a simulation model that reproduces the principal trajectories of the society and its peoples over this time. This allows us to test hypotheses and create scenarios of wider social and policy interest.

Quantitative and Computational Techniques

Sociology's reliance on quantitative techniques goes back to the earliest days of statistical analysis, at least if we adopt a broad definition of the sociological task. In the first instances of statistical analysis of a public kind, early pioneers sought to derive what we would now regard as social, policy, and public health indicators from laboriously collated data, such as mortality records.

Under the heading "From Political Arithmetic to Social Statistics", Donnelly (1998) provides a historical review of the origins of quantification in the social sciences. On this account the early demand for numerical information was driven by the practical needs of the state and civil society; in other words, statistics was at the time a form of numerical and empirical information about society that might be of interest to the state, hence "State-istics" and statisticians as statists.

The transition from this earlier form of "political arithmetic" to social statistics as we now know it came with the development of new numerical transformations

P. Davis, R. Lay-Yee, *Simulating Societal Change*, Computational Social Sciences, https://doi.org/10.1007/978-3-030-04786-3_1

and analyses that seemed to promise the distillation of empirical social regularities through the collation of individual items of data on a large scale. An early and exemplary case in point was Durkheim's use of population statistics to draw conclusions about suicide as a patterned social phenomenon. His work seemed to indicate that it was possible to extract stable and insightful regularities about suicide – and more broadly about society – from what was otherwise an apparent complexity of a multitude of individual events.

The growing analytical power of quantitative sociology tracked developments in probability theory, survey samples, tabulations, and the emergence and rapid development of multivariate techniques such as regression analysis. It is possible to discern three generations of statistical methods in sociology (Raftery, 2001). In a period of early survey research from the 1940s, sociologists relied on the analysis of cross-tabulations. These represented the quite laborious collection, collation, and aggregation of data in tabular form from the early social surveys. This was the era of techniques for categorical data analysis. The next major development was facilitated by a series of statistical, technical, and computational advances. Thus, from the 1960s quantitative sociologists were able much more readily to access unit-level data from social surveys and carry out advanced statistical techniques on these data. This was the heyday of the general linear model, particularly regression analysis. Finally, by the late 1980s, sociologists were increasingly aware of the potential of new data sources, new statistical methods, and intriguing new analytical challenges, none of which fitted easily into the orthodox regression model. These opportunities have continued to expand, and at a faster pace, with the advent of social media and multiple new sources of data collection beyond the traditional social survey. Thus, we have social networks, spatial data, textual and qualitative data, simulation models, sensor information, complexity analysis, and so on. We draw on these rich traditions in our book, combining access to the administrative data of the census with advanced statistical techniques in its preparation for our work.

By contrast, computational techniques in the social sciences – at least as used in this book – are of rather recent provenance. In their early application, these were procedures required to ease the processing and manipulation of large quantities of social data. An early paper by Anderson and Brent (1991) saw "sociological computing" – as it termed the field – as a potential opportunity missed. In its earlier days, computational sociology was seen as a service function, an applied field, and an area that offered little academic kudos to participants. One consequence of this was that software development along lines suited to the particular needs of sociology was little in evidence.

However, with the revolution accompanying the arrival and expansion of the internet, together with the extraordinary power of contemporary computers, the role of computational social science has become far more ambitious. We are now at a stage not only where the cyberworld plays a crucial part in everyday social interactions – indeed a key part of the symbolic world – but also the power and sophistication of computational techniques give us an opportunity to make faithful representations of the social world "in silico", that is virtual, computational representations of the dynamics of social reality. This is the objective that inspires and

informs our work and that raises it above what might otherwise be a reasonably standard statistical treatment of social change.

This transformation of computational sociology is captured by Macy and Willer (2002) in their review of the field, "From factors to actors". While the focus of this paper is on the analytical potential for sociological advance of one particular computational innovation – agent-based modelling – it does cover the history of "sociological computing" that goes well beyond the traditional service function of facilitating the analytical requirements of both quantitative and qualitative sociology in software development and computing power. In the 1960s the focus was on computational techniques applied at the macro level to gain traction on processes of control and feedback, including at a global level to model demographic changes and assess the ecological limits to growth. From the 1970s investigators introduced the use of individuals as the units of analysis, heralding the advent of microsimulation, which is the technique used in this book. Since the 1980s there has been a third wave in social simulation – agent-based modelling – with the advent of personal computers. This technique is more concerned with theoretical development and explanation than with predictive power, which is the strength of microsimulation and is the focus of this book.

Sociology as a Population and Policy Science

For us sociology is a population science. This has been recently eloquently argued by John Goldthorpe (2015) in his book of that name. We come to this insight by way of our work as social scientists operating in a school of public health (also sometimes interchangeably known as population health). Sociology is necessarily a "social" science, and the sciences of public health address the determinants of health and disease at the level of entire populations.

One important model in this tradition has been that of "social determinants", where structural, cultural, and societal features are seen as crucial in improving health outcomes, alongside the usual suite of behaviour and clinical interventions. In a recent review, House (2016) sees the "social determinants and disparities" model in health as a major influence of sociology in medicine and one that shifts the dominant paradigm from the "supply" side of health policy – better medical care, medicines, and treatments – to the "demand" side (the factors causing ill-health in the first place). On this view a core variable like socio-economic status is a "fundamental cause" that patterns the exposure and experience of entire populations to all health risk factors.

A key discipline in this tradition is that of epidemiology, a form of applied statistical analysis devoted to mapping and understanding disease aetiology. Traditionally epidemiology has focused almost exclusively on the careful mapping of disease events across key population and exposure factors. There is also a dominant emphasis on shaping research design and deploying statistical techniques sufficient to draw strong conclusions. In recent years, however, the social sciences – and more

particularly sociology – have had an influence on the discipline, as reflected in the development of social epidemiology and an engagement with issues of social structure, inequality, and social conflict. Ng and Muntaner (2014) identify this last development as macrosocial epidemiology and see its role in reorienting public health efforts towards social change. Thus, epidemiology draws on the social sciences to take on a stronger population and policy stance.

Another form of applied statistical analysis in the social field is demography – again a sophisticated quantitative framework for understanding and projecting the dynamics of populations. Demography is a science of populations, their structure, and dynamics, and thus its empirical focus is on what might be called the vital statistics of human life – birth, death, marriage, family size and formation, and longevity. Ian Pool (2016) sees obvious synergies with epidemiology and public health, but at core he views demography as a social science having many areas of connection to sociology.

Therefore, sociology "as a population science" has much in common with public health, epidemiology, and demography, and it can be said to have contributed greatly to these disciplines in their conceptual and explanatory frameworks, once they go beyond statistical descriptions and analyses. Sociology differs however in going beyond applied statistical analysis to attempting to identify the mechanisms that help account for the patterning of social outcomes that we document. We see our work as contributing to this tradition in sociology.

We also see our work as essentially a contribution to the developing application of sociology as a policy science. Interpreted in its most straightforward sense, a public policy is an arrangement of policy goals and means in a specified area – such as health or education – that is the result of a government's efforts to change some aspect of its own or social behaviour in order to achieve some public or social purpose (such as improved health of the population). An important means to this end is "policy design", a process by which academics, bureaucrats, and decision-makers fashion a range of instruments that might be considered to assist in the achievement of certain policy goals (Howlett, 2014). Sociology could be contributing to such a concept of policy design by offering information about the state of society and suggesting "instruments" that might be of assistance in achieving social goals.

We are well familiar with this approach when it comes to addressing issues to do with the economy. Various agencies, including the state itself, draw up regulations, settings, and procedures that are designed to achieve certain economic objectives in the areas of productivity, GDP growth, inflation, debt, taxation, and so on. We could equally as well draw up a parallel set of guidelines and objectives in the area of social policy. In this field, sociology should be seen as the contributor of concepts and empirical knowledge to understanding how society works. Social policy on the other hand is an applied discipline seeking to best deploy information and insights from sociology to the achievement of policy goals in the real world (Banton, 2016).

We believe that sociology has a role to play here with providing the conceptual tools, the data sources, and the analytical techniques for addressing society's "big issues". An example of advanced practice in this area that might provide a model of "sociology as a policy science" is the field of medical sociology. This is the largest

specialty within the discipline of sociology; it demonstrates the full range of engagement with the health sector (from the critical to the technocratic) and has contributed substantively to the sector by influencing practice and by changing fundamental concepts of health and medicine (Scambler, 2014). We seek to advance this broad agenda in this book.

New Zealand: A Social Laboratory?

Our focus is necessarily New Zealand, although we believe that our work will be of sufficient generality and broader applicability to be of wider interest. It should also be noted that the size of the country is such that our project is feasible at this scale computationally and conceptually.

One area of considerable significance for this project is the availability for our analysis of data from the New Zealand Census. Not only have these data been made available for external analysis spanning a significant period of time – 1981–2013 – but they are linked for a subset of that data at the level of the individual. This gives our project a special power and owes much to Statistics New Zealand (SNZ), the country's official statistics agency, which has embarked on a process of data linkage that, aside from a few European nations, is almost unprecedented. This has been part of what has come to be called the Census Transformation Programme (Gibb & Shrosbree, 2014), and our project has benefited from this farsighted initiative.

However, there is another reason for our focus on applying simulation techniques to understanding societal change in New Zealand – and that is the country's early reputation for being a pioneer in social policy, such that it was once known as a "social laboratory". Both Australia and New Zealand were new settler societies in which the early colonists felt sufficiently free of the constraints of history and tradition to experiment legislatively, particularly in the founding of key elements of the emerging welfare state (Cox, 2006). This pattern was most evident in the late nineteenth and early twentieth centuries, although such social policy innovation gained renewed momentum in the 1930s, with hallmark legislation such as the 1938 Social Security Act. Ironically, the same political movements initiating these changes to constrain market forces then went on to challenge these arrangements in the 1980s, with a further period of institutional reform and change designed not to constrain but free the market.

This concept of a "social laboratory" provides a link between the country's past as a site of social innovation and our application of simulation techniques to test hypothetical scenarios, scenarios that we can now envisage by way of counterfactual modelling in a simulation environment, but which a century ago could only be approximated by experimentation in a new and emerging society.

This opportunity to test such societal scenarios with computational techniques has only recently become technically feasible. Indeed, a "manifesto" of computational social science was published as recently as 2012, and in a natural science journal (Conte et al., 2012). The emphasis in this manifesto is as much on capturing

the essence of social complexity as it is about making the most of recent advances in computational power and access to "big data".

Our contribution is in attempting to capture the complexity of societal change by deploying advanced computational techniques on census data. In honour of the technical, conceptual, and historical resonances of the idea of a "social laboratory", we have called our simulation model and inquiry tool – SociaLab.

A Period of Rapid Social and Economic Change

New Zealand, like most other developed countries, has changed markedly in the last half century. Shortly after the Second World War the country was still homogeneously British in migrant background, the economy focused on serving the markets of the "mother country", an economy strongly reliant on the rural, primary producing sector, the society, according to Austin Mitchell's designation, "a quarter section paradise" with the nuclear family at its heart, strictly defined gender roles, and the indigenous people barely visible in public affairs. All this has changed, with New Zealand now an ethnically diverse country plying its trade around the world, but particularly in Asia, with non-primary sectors important, diverse household structures, gender roles less rigidly defined, and Māori playing a central role in social, cultural, and political life. In the words of a recent social history of the period, *Changing Times* (Carlyon & Morrow, 2014), this was a trajectory from "dull conformity" to "one of the most ethnically, economically and socially diverse nations on earth".

Changes of this scale over a clear half century during which much of the world was undergoing significant social and economic change is perhaps not to be wondered at. But for New Zealand, a key juncture was the period of the 1980s and 1990s in which the economy was deregulated and the welfare state subject to greater financial and political scrutiny. In a period of about a decade, New Zealand moved from being one of the most regulated social and economic systems to one of the most open and "flexible".

There were two cycles of change over this period, one in the economy, the other in the underpinnings of the welfare state. As Dalziel (2002) states, within a year of the Labour Party coming to power in 1984, regulatory constraints on interest rates, capital movements, and the currency were removed, and agricultural subsidies and tax incentives were being phased out. In a second set of changes introduced by the succeeding National Party government, social policy settings were addressed in what have been seen by commentators as harsh austerity measures and major welfare state retrenchment (Deeming, 2013).

Our model covers this period and we wish to incorporate this phase of rapid institutional change into our analysis. Among other things, we wish to use counterfactual modelling to provide a more nuanced analysis of the impact of those changes. In Dalziel's (2002) analysis, the counterfactual to the longer-term impact of economic policy change in New Zealand is the Australian counterexample over the

same period. For Deeming (2013) a cross-cultural comparison with other OECD countries is the preferred method. We will contribute to this debate by using settings in our model to track and interpret policy changes.

The New Zealand Longitudinal Census

Among the rapid regulatory and structural change experienced by New Zealand has been far-reaching reform of the state sector. Together with other trends in this sector, the country's official statistics agency – Statistics New Zealand – began to engage more with stakeholders both within the traditional bureaucracy and in the wider society. Three related tendencies have become evident in the area of the collection and management of official and administrative data: transforming the census, linking administrative data, and accessing microdata.

A major initiative has been the programme for adapting and potentially transforming the 5-yearly census (Statistics New Zealand, 2012). This has been a major planning and scoping exercise, involving the assessment of a range of options available for scrutiny internationally, including using new technologies, less frequent data collection, adopting continuing surveys, and accessing administrative data for key elements of the traditional census.

A second trend has been the consideration of granting greater access to the vast stores of officially collected data held by the agency for further analysis both by other government agencies and by external researchers. Indeed, this process has gone so far that potentially world-leading developments such as the Integrated Data Infrastructure (IDI) have taken place allowing much great linkage between different administrative data sets (Statistics New Zealand, 2018). This facility can now be accessed by researchers off site via a remote data laboratory service.

Finally, there have been particularly striking developments in access to census and other microdata, previously a highly restricted data source (Statistics New Zealand, 2016). One important initiative of the agency has been to establish a linked census microdata set from 1981 to 2013, providing the opportunity to develop a census-based longitudinal data set for analysis. This is the data set that is the foundation for the work outlined in this book. While there are shortcomings – such as a limited linkage rate – it is almost unprecedented outside the Nordic countries to have population data from a register-type source linked over such a long period of time. Therefore, we seek to make the most of this opportunity in building our model by exploiting many of the analytical advantage this provides.

SociaLab, Our Inquiry Tool

This book outlines the construction and application of an inquiry tool fashioned out of the linked data set derived from the censuses spanning 1981 to 2013. We have called this tool SociaLab, both to signal its historical resonance in New Zealand and

also its ambition to be a simulation model incorporating the structure and dynamics of an entire society over a period of rapid change. The construction of the inquiry tool, as will be evident from the outline in subsequent chapters, is a feat of considerable technical accomplishment that builds on teamwork and years of experience in working with techniques of microsimulation. We are not sure that anything like this has been constructed in New Zealand before, and there are few models of comparable scale and sophistication internationally.

Why an inquiry tool? The object of this modelling exercise is more than demonstrating great technical accomplishment and the potential of adding value to administrative and official data. More than this, we wish to mount a range of analytical ventures and experiments addressing key social and policy questions using our simulation model. In this way we are able to take on issues of societal scale in a rigorous fashion, an ambition that would be almost impossible to achieve by using any other methodological approach. While we may not be able to claim causal efficacy for our ventures and experiments with the inquiry system, we will have applied a strong predictive model with considerable construct and criterion validity at the level of an entire society.

Conclusion

In this book we aim to bring together the major threads contributing to the emergence of a powerful set of tools for the sociological analysis of society. We are fortunate to inhabit a "sweet spot" of impressive developments in quantitative and computational social science, a vibrant debate on sociology as a population and policy science, and an empirical site – New Zealand – where unrivalled data sources provide us with an opportunity to build an inquiry tool capable of addressing some of the great societal issues of our time.

References

Anderson, R. E., & Brent, E. E. (1991). Sociological computing: An opportunity missed? *The American Sociologist, 22*(1), 65–77.

Banton, M. (2016). Reflections on the relation between sociology and social policy. *Sociology, 50*(5), 993–1001.

Carlyon, J., & Morrow, D. (2014). *Changing times: A history of New Zealand since 1945*. Auckland, New Zealand: Auckland University Press.

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., … Sanchez, A. (2012). Manifesto of computational social science. *European Physical Journal-Special Topics, 214*, 325–346.

Cox, L. (2006). The antipodean social laboratory, labour and the transformation of the welfare state. *Journal of Sociology, 42*(2), 107–124.

Dalziel, P. (2002). New Zealand's economic reforms: An assessment. *Review of Political Economy, 14*(1), 31–46.

Deeming, C. (2013). The working class and welfare: Francis G. Castles on the political development of the welfare state in Australia and New Zealand thirty years on. *Social Policy and Administration, 47*(6), 668–691.

Donnelly, M. (1998). From political arithmetic to social statistics: How some nineteenth-century roots of the social sciences were implanted. In J. Heilbron, L. Magnusson, & B. Wittrock (Eds.), *The rise of the social sciences and the formation of modernity*, Sociology of the Sciences Series (Vol. 20, pp. 225–239). Dordrecht, The Netherlands: Springer.

Gibb, S., & Shrosbree, E. (2014). *Evaluating the potential of linked data sources for population estimates: The integrated data infrastructure as an example*. Wellington: Statistics New Zealand. Available at http://archive.stats.govt.nz/methods/research-papers/topss/evaluating-potential-linked-data-sources.aspx

Goldthorpe, J. H. (2015). *Sociology as a population science*. Cambridge, UK: Cambridge University Press.

House, J. S. (2016). Social determinants and disparities in health: Their crucifixion, resurrection, and ultimate triumph (?) in health policy. *Journal of Health Politics, Policy and Law, 41*(4), 599–626.

Howlett, M. (2014). From the 'old' to the 'new' policy design: Design thinking beyond markets and collaborative governance. *Policy Sciences, 47*(3), 187–207.

Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology, 28*(1), 143–166.

Ng, E., & Muntaner, C. (2014). A critical approach to macrosocial determinants of population health: Engaging scientific realism and incorporating social conflict. *Current Epidemiology Reports, 1*(1), 27–37.

Pool, I. (2016). The seminal relationship between demography and sociology. *New Zealand Sociology, 31*(3), 146–167. Available at https://search.informit.com.au/documentSummary;dn=280256964920183;res=IELNZC

Raftery, A. E. (2001). Statistics in sociology, 1950–2000: A selective review. *Sociological Methodology, 31*(1), 1–45.

Scambler, G. (2014). Medical sociology in the twenty-first century: Eight key books. *Contemporary Sociology: A Journal of Reviews, 43*(2), 155–160.

Statistics New Zealand. (2012). *Transforming the New Zealand census of population and dwellings: Issues, options, and strategy*. Wellington: Statistics New Zealand. Available at http://archive.stats.govt.nz/methods/research-papers/topss/transforming-nz-census-pop-dwell.aspx

Statistics New Zealand. (2016). *Microdata output guide*. Fourth Edition. Wellington: Statistics New Zealand. Available at http://archive.stats.govt.nz/~/media/Statistics/services/microdata-access/data-lab/microdata-output-guide-2016.pdf

Statistics New Zealand. (2018). *Integrated data infrastructure*. Retrieved from https://www.stats.govt.nz/integrated-data/integrated-data-infrastructure/

Chapter 2
Conceptual and Analytical Foundations

A research enterprise of the kind outlined in the previous chapter is an ambitious undertaking. A very substantial set of challenges are those associated with constructing the SociaLab model. Even setting up the census data in such a way that it can be analysed in a systematic and rigorous fashion is a major task in itself. Then there is the very demanding statistical work required to provide the basic estimates that will make this model work. However, before we get to the stage of managing and analysing the core data, we need to draw up the fundamental conceptual and analytical insights that inform this project. This is the purpose of this chapter, that is, to draw up the conceptual and analytical foundations that underpin the data management, analysis, and estimation work.

Societal Change as the Starting Point

Our initial focus in this book is to model social and demographic change in New Zealand over the period since 1981. Our starting point is describing societal change and identifying some of its key drivers. Many others have done this before us, often using historical information and standard aggregate measures to get at the bigger picture of an entire society on the move.

One example that comes close to our ambition of scoping societal change in the big picture, though without the underpinning computational modelling, is a piece by Bourne and Rose (2001) on the "changing face of Canada". These authors seek to identify the big social and demographic trends evident across Canada since the Second World War. They perceive four transformations working through the country's social fabric and urban landscape: the demographic transition and changing components of population growth; changes in family structure, domestic relations, and household composition; immigration and increasing social and cultural diversity; and shifts in the linkages between home and work and the changing nature of

P. Davis, R. Lay-Yee, *Simulating Societal Change*, Computational Social Sciences, https://doi.org/10.1007/978-3-030-04786-3_2

state and civil society. These notable social trends are illustrated with descriptive data, sound plausible, and could equally well characterise New Zealand's development over this period.

A second approach is altogether more analytical and seeks less to describe social trends than to unpack the implications of different social changes as modelled using advanced computational techniques. An example of this is the work of Cioffi-Revilla, De Jong, and Bassett (2012) using a combination of evolutionary computation and agent-based modelling to operationalise the social dynamics of a political system evolving through time. The authors use an existing agent-based model of a stylised political system and subject it to the kinds of developments and shocks that its real-world counterpart could experience. For the authors the exercise is a proof of concept that these computational tools can be used more broadly in modelling social change.

The approach adopted in this book falls between these two research strategies; yes, we wish to use computational techniques to model social change along the lines being pioneered in a highly sophisticated way by Cioffi-Revilla et al., but at the same time, we do not wish to be so schematic and stylised as to miss the essential ingredients of social change in the real world that are captured in the much more descriptive review of long-term social and demographic trends in Canada by Bourne and Rose.

A further point of difference is that we will be using the one data collection instrument that takes a snapshot of New Zealand life across the whole population every 5 years, namely, the census. This provides us with an unrivalled data series that gives real meaning to the idea that we are addressing social and demographic change at the level of an entire society over an extended period of time. This can only be done by the census, a remarkable data collection enterprise that regularly and reliably maps the entire New Zealand population. Conventionally the census has been viewed as a relatively passive data source, but a paper drawing on the British context by Killick, Hall, Duff, and Deakin (2016) entitled "The census as an information source in public policy making" highlights a more active role for the census as a data source that can help inform policy discussions. We see our work as being in this tradition.

We are by no means the first people to use the census to track change in New Zealand society. But what is different is that we are achieving this by working not with published, tabulated, or aggregate data – which is the usual approach – but by analysis of the "microdata" generated by people completing the census form. Until recently access to such data was highly restricted. We can now access such information in a much freer way through the data laboratory facility. Thus not only are we modelling social change at the level of the entire society via the regular census data collection, but we are doing this by following the trajectories of individual New Zealanders and their households through time, rather than just relying on a time series of aggregated data.

There are some precedents for this, particularly in the field of family history and social demography, although these approaches have a strong qualitative and historical element; this is in contrast to our approach where we start and finish with census data collections over the last 25 years. For example, Szołtysek and Gruber (2016)

outline a major project designed to recover surviving census records and reconstruct family history across Europe. The authors call the project – Mosaic – "one of the largest infrastructural projects in the history of historical demography and family sociology". In parallel with our work, the investigators sought to harmonise samples of census microdata across time and space. They also envisage that the comparison of such microdata across geographic contexts will allow them to consider the influence of meso-level environmental, socio-economic, and cultural factors. Another initiative has been CEDAR, a project designed to create linked open data with the Dutch census over nearly two centuries, although at the aggregate rather than micro-level (Merono-Penuela Ashkpour, Gueret & Schlobach, 2017). Again, the ethos is one of opening up census data and using it for new purposes.

But Building "from the Bottom Up"

What appears at one level to be an operational breakthrough – that is, being able to access census microdata in a relatively unrestricted way – is also an important conceptual step: we are grounding our analysis of societal change in the reality of decisions made by a myriad of individuals going about their everyday lives (including, as good citizens, the completion of the census form every 5 years!).

A recent example of this approach is that of Billari (2015). Billari makes the distinction between the *discovery* and *explanation* of patterns of human population change. The former is like to occur at the level of aggregate, or macro, data; that is, we are likely to detect large-scale changes using aggregate or macro data, but then we have to try to account for it. And that requires teasing out behaviour at the micro or individual level. Thus macro-level patterns are seen to be the result of actions and interactions at the micro or individual level. Billari's approach has much in common with that of Coleman (1990) and Goldthorpe (2015) where explanations of how population change comes about are seen to be rooted in models of the actions of individuals and families. Billari rejects a simple "rational action theory" approach and draws on the concept of life course to link individual actions into a coherent life trajectory.

Thus, "societal change" is seen to be a reflection at an aggregate level of the patterned outcome of decisions that, fundamentally, are taken by individuals. This is referred to as the principle of methodological individualism: that social phenomena have to be seen to be anchored, and accounted for, ultimately in the actions of individual human beings.

The concept of methodological individualism is not without contention (List & Spiekermann, 2013). At one level, the assertion that good social science explanations should be couched in statements about the actions of individuals seems a reasonable correction to the alternative of holism, where causal and ontological status may be accorded to collective entities such as nations, cultures, or institutions. How can such entities be seen to act and influence? Such an assumption can seem to have metaphysical connotations. At the same time, can all social phenomena be reduced to the actions of individuals in any straightforward way? This approach would seem

to have difficulty in dealing with systems, institutions, and multilevel social complexity more generally. In our work we see individual actions and agency as central but constrained and channelled by contexts, systems, structures, norms, and institutions.

Adding to this notion of building the reality of societal change "from the bottom up" is our use of microsimulation. This is our principal computational device in establishing an inquiry tool that will allow us to interrogate SociaLab so that we can address some of the big questions of social and policy change in New Zealand. As will be described later, this technique functions in our application at the level of identifiable and distinctive individuals who, in aggregate, are representative of the entire New Zealand population.

If our principal computational approach is microsimulation, what is this technique? That is a surprisingly hard question to answer because the practice of microsimulation is a minority pursuit and is frequently used for practical purposes rather than for science-driven investigations and so rarely gains broad academic interest. Spielauer (2011) provides a review of the area. Social science microsimulation is a method for computer modelling the actions and interactions of individuals, particularly over the life course, in order to identify macro outcomes. As computational power has improved, and with growing policy interest, particularly in models that can model and follow the actions of individuals within longitudinal and multilevel perspectives, microsimulation has progressed from a relatively static representation of short-term effects – such as the impact of tax-benefit policy changes – to a much more dynamic representation that can follow individuals over the entire life course and represent system or societal change in the longer term.

Thus, microsimulation can give societal change a "human face" and a grounding in the life trajectories of real people. By estimating these life trajectories in statistical and computational form, we gain traction on the sociological micro-foundations of social change. This, together with the opportunity we have for addressing counterfactuals, gives this approach great analytical power for social and policy inquiry. These two features – micro-foundations for social change and analysis of counterfactuals for social and policy inquiry – bring our work into contact with other social science disciplines. Indeed, Abell (2003) has argued that a search for common ground between economics and sociology in the pursuit of the potential for a unified social science is to be found in a merger or fusion of the rational actor model of the one with the influence of context and constraint of the other. While we would not claim any such ambition, we do see our work as contributing to the search for some common analytical and methodological ground among the social sciences.

Concepts of the Middle Range

Our approach to modelling societal change "builds up" from individual actions and trajectories. These are "real", quantifiable, and tractable by computational means. But this does not mean that only individuals and their actions have meaning in the

social world. Sociology would be greatly impoverished if the reality of other social phenomena were not also recognised. Thus, we attempt to work not only with aggregate outcomes at the macro-level of an entire society nor just with the micro-foundations of individual actions but also with structures, patterns, and determinants in the conceptual and social space in between – that is, at the meso-level (the middle range).

Our approach resonates with a revived interest in sociology in empirical work at the "middle range", work that is neither sparse empiricism nor elevated theorising but that searches for social mechanisms that may help us bridge micro- and macro-levels with meaningful explanations (Edling & Rydgren, 2016). It is argued that this approach encourages us to focus on real and empirical activities that help bring about or generate social phenomena we are seeking to account for. Aside from orienting us towards real-world empirical problems conceptualised at a middle-order range, this approach can also assist in bridging the micro-macro gap which threatens otherwise to be overly dependent on extrapolation from poorly contextualised analytical models of micro-level activity (Raub, Buskens, & Van Assen, 2011).

What are these social entities at the middle range? Neighbourhoods, hierarchies, networks, communities, organisations, systems, norms, and social groups are among the social formations recognised as being in the conceptual space between macroaggregates and micro behaviour. Pawson (2000) has coined the term "middle-range realism", arguing that the actions and intentions of individuals only make sense when seen as being embedded within a social reality that is multilayered, incorporating concepts and social entities of this kind at the meso-level. Such an approach helps advance the cause of a social complexity that goes beyond simple concepts of micro and macro, as well as allying our work with an explanatory paradigm that gives pride of place and substance to social mechanisms of the middle range (Hedstrom & Swedberg, 1996).

Attractive as this ambition might be, the truth is that entities and mechanisms of the middle range are hard to capture in the data that is available to us from the census. One method that we will be adopting to bring these entities into our microsimulation of individual biographies is the concept of a life course – as depicting transitions between significant life stages and important institutional sectors. Ethnicity, region, and socio-economic status are also measures of social context and constraint we can address with our data.

The Life Course as a Key Organising Concept

In keeping with the dynamic thrust of our prime research goal – simulating societal change in New Zealand – we conceptualise the "actors" in our SociaLab model as not only living in context (such as households, neighbourhoods, social groups) but also embarking on journeys that take them into key social role transitions through life. This matches the dynamic requirements of our simulation modelling but also provides a social and institutional texture that might otherwise be missing if we

worked only with a "barebones" approach to individual action. This brings our actors in touch with key social institutions, such as health, education, employment, partnership, household formation, family life, and so on.

Our concept of the life course comes close to that of Levy and Buhlmann (2016) who see it as "an individual movement through social space" that is organised as a series of social fields with distinct cultural and structural characteristics. To an important extent, these life-course biographies or trajectories show a degree of predictability and institutionalisation: hence, there are life phases such as education, employment, and retirement; there are relating institutions such as family and peer group; and the welfare state provides a series of support mechanism at key junctures in life. These elements of predictability and institutionalisation provide a structure and societal pattern to what otherwise is a potential multiplicity of life-course trajectories.

The life-course concept also provides a powerful tool for empirical longitudinal analysis. An example is the US Panel Study of Income Dynamics (PSID), a research platform of half a century's duration (McGonagle, Schoeni, Sastry, & Freedman, 2012). The study is a cornerstone of US social science, generating thousands of papers, sharing data worldwide, hosting millions of visits a year, and more recently being supplemented by enhanced data collections. Such is the age, consistency, and durability of the data that the study is now able to support not only intragenerational but also intergenerational research. Our work does not attempt to match the power of this long-standing and well-funded study, but its eminence highlights the conceptual and analytical power of the life-course concept for core social science research.

For demographers a version of the life course structures the analysis of birth, partnership, family formation, and death (Billari, 2015). Our analysis incorporates these important life events but also tracks the progress of individuals through health, education, entry to the labour market, household and partner formation, employment, household change, and retirement. It is also possible to elaborate this simple and predictable sequence of events to consider the role of other contextual and institutional factors in people's lives.

A recent research example in this genre is the programme of the National Institute for the Study of Ageing and Later Life (NISAL) at Linkoping University, Sweden (Motel-Klingebiel, Hyden, & Cedersund, 2017). Although the focus of the programme is on "ageing" from a gerontological perspective, a key organising concept is that of ageing through the life course and how this is mediated by social structure and social context. The life course is seen as reflective of institutional arrangements in society and the impact of social change. Another contribution that links the life course to wider issues comes from the *Handbook of the Life Course* where O'Rand and Bostic (2016) link the study of the life course to wider macrosocial foundations and social policy imperatives. They see the life course within a larger global context that brings to bear the forces of history and social change. As such, it is a useful analytical tool or lens on the interface between the lives of individuals and wider social forces.

Again, these are ambitious and carefully designed contributions that sketch a grand picture. However, we are limited by the rather sparse nature of data in the

census, our key data source. Our work is therefore necessarily less ambitious, but it does represent a first step in linking a societal-wide source of authoritative data to broader questions of analysis and policy.

But Identifying Social Mechanisms and Social Assets as Well

It is being increasingly argued in the literature that "truly" sociological explanations are those that are able to identify a social mechanism in accounting for a particular outcome (Hedstrom & Swedberg, 1996). Thus, it is no longer sufficient just to identify an empirical regularity of the kind "middle class children do better academically at school". This is not an explanation; rather, it is just a description of an empirical regularity, as it stands. A social mechanism that helps account for the regularity needs to be identified – such as, "middle class children do better academically at school because their parents provide precursors to success, such as cultural capital and encouragement".

One area where this approach has been operationalised to practical effect is in evaluation research and practice (Astbury & Leeuw, 2010). In much evaluation work – say, assessing the effects of a programme – a "black box" approach is adopted whereby the investigator is interested just in evaluating the effects of the intervention rather than going on to identify how such effects might have been produced. This may be a pragmatic approach – does the programme work? – but it does not cast light on potential mechanisms for these effects or assist in building theory and understanding. It is in response to these concerns that practitioners in evaluation research have increasingly sought to unpack the "black box" and identify "underlying entities, processes or structures which operate in particular contexts to generate outcomes of interest".

Ideally, therefore, our microsimulation model of societal change and dynamics should be able to move beyond straightforward empirical regularities of individual behaviour and outcome through the life course to identifying key mechanisms that help account for such regularities. One such set of mechanisms may be the provision of material and non-material assets for success. An example of this can be found in the health area where both structure and agency, as well as material and non-material resources, can be seen as essential building blocks in reducing health inequalities (Abel & Frohlich, 2012).

Again, because of the sparse nature of the data collected in the census, the structure of our simulation modelling may lack some of the complexity that we would ideally like for a full-blown sociological account of social patterns and outcomes. While our simulation modelling will generally be empirically sufficient and validated against predicted outcomes, it may be hard to meet the full standard of sociological explanation that we would otherwise like to set ourselves.

Structure and Agency

The central computational technique we are using in SociaLab is microsimulation. This necessarily operates at the level of the individual, with transitions between states estimated from empirical data. However, these are predictive, behavioural estimates that do not necessarily say anything about the state of the mind of the individual. They cannot necessarily be interpreted as action statements describing conscious decisions taken by individuals faced with competing choices and priorities in real-world settings. So, how much agency can we attribute to the representative "actors" in SociaLab? At present, the behavioural range for such action statements is set by a statistically described distribution of outcome states.

With the availability of the right kind of data, much can be done with the conceptual distinction between structure and agency and their interaction in producing outcomes of interest. Cockerham (2005) develops this area in his review of health lifestyle theory. An excessively individualistic approach risks "blaming the victim" for their own "poor" health behaviour (e.g. smoking, alcohol consumption) when such patterns may be strongly influenced by context. Equally, an orientation that only acknowledges the influence of structural factors, on the other hand, overlooks the role of individual agency in health lifestyle decision-making. In our case, however, given the sparse nature of census data, it may be that we have to be content with identifying and describing behavioural regularities without necessarily attributing any action orientation and decision-making process to them.

Similarly, can we identify the influence of structure and constraint on individual actions in our microsimulation analytical framework? Again, we are dealing with empirical regularities in behaviour as predicted from statistical equations that are estimated from the data. Such regularities are likely to show patterning by key structural factors such as gender, ethnicity, socio-economic status, and location. Therefore, again, SociaLab will rely for its fidelity to the dynamics of real-world social change on probabilistic models working across the major structural features of society. It is doubtful whether these regularities can be interpreted in a richly sociological manner to parse the influence of structure and agency in explaining societal change and other social dynamics. Nevertheless, there are examples where the role of deeply rooted structural conditions can be clearly identified as necessary but not sufficient precursors for certain outcomes. Thus, for example, Bramley et al. (2015) see deep-seated structural preconditions for multiple and severe disadvantages in depressed parts of England, but at the same time, these preconditions are insufficient on their own to account fully for patterns of disadvantage and can be seen to interact with family and individual factors.

Conclusion

The focus of this book is on social change identified at the level of an entire society. New Zealand provides us with that opportunity. It is sufficiently small in scale to allow us to develop a simulation model – SociaLab – that can encapsulate the entire