The role of microsimulation in the development of public policy

Roy Lay-Yee and Gerry Cotterell
COMPASS Research Centre
The University of Auckland

Contents

Abstract

This paper provides an introduction to the method of microsimulation, which underpins the Modelling the Early Life Course project being undertaken by the Centre of Methods and Policy Application in the Social Sciences (COMPASS) at The University of Auckland. The project is funded by the Ministry of Science and Innovation, now part of the Ministry of Business, Innovation and Employment. Since the inception of microsimulation in the 1950s, its use for policy purposes has extended from the economic to other domains as data availability and technological developments have permitted. Microsimulation focuses on modelling individual units and the micro processes that affect their development, be they lives or other trajectories. It comes in various types, for example arithmetical or behavioural, and static or dynamic. It has its own distinctive model-building process which relies on data and derived parameters. The utility of microsimulation for policy development lies in its ability to combine multiple sources of information to answer "what if" questions on a complex phenomenon of interest.

Introduction

This paper provides an introduction to the method of microsimulation and its role in public policy. The Centre of Methods and Policy Application in the Social Sciences at The University of Auckland has been funded by the Ministry of Science and Innovation – formerly the Foundation for Research Science and Technology, and now part of the Ministry of Business, Innovation and Employment, to produce a microsimulation-based policy tool to model aspects of the early life course, particularly focusing on family factors that influence child outcomes. The paper firstly provides a brief history of microsimulation, which tracks the development of this method since its inception in the 1950s. This is followed by an explanation of microsimulation itself, the differences between its various types, and the model building procedure. We end with assessing the utility of microsimulation for policy development, and setting out its strengths and weaknesses as a method for this purpose. In an appendix, we provide selected examples of a number of existing microsimulation models in use.

A brief history

The academic literature generally acknowledges the pioneering work of Guy Orcutt in his 1957 paper, "A new type of socio-economic system", as providing the foundation for the field of microsimulation. In his paper, Orcutt posited that microsimulation models consisting of "various sorts of interacting units which receive inputs and generate outputs" (Orcutt, 1957, p. 117) could be used to investigate "what would happen given specified external conditions and governmental actions" (Orcutt, 1957, p.122). However, the potential of this new approach was slow to materialise because of limitations in computing power and the lack of suitable data.

As these respective limitations were gradually overcome with the rise of technology and data collection, the use of microsimulation increased. In the 1970s, microsimulation models were being used in the USA to assist the development of social policy (Citro & Hanusek, 1991), and by 1990 "microsimulation had become widespread enough in the domain of tax and transfer analysis" (Anderson & Hicks, 2011, p. 1). Merz identified six major microsimulation projects in the USA and Europe in the 1980s, growing to 18 in the 1970s and then to 33 in the 1980s (Merz, 1991). Further rapid advances in computing power, along with information sharing, served to increase the use of microsimulation modelling, and in 2005, the International Microsimulation Association was formed (Anderson & Hicks, 2011, p. 1) (http://www.microsimulation.org/). Microsimulation models have historically been employed for tax and transfer policy purposes (Harding, 1996; Gupta & Kapur, 2000; Harding & Gupta, 2007), but their use has extended to other social domains (Gupta & Harding, 2007). Even today, despite the advances in technology and data availability microsimulation projects typically require a large investment of time and resources.

What is microsimulation?

As a form of social simulation, the microsimulation variant can be seen to be located between and contrasted with system dynamics – with a macro focus and low complexity – on one side, and agent based modelling – with a behavioural focus often relying on notional rules – on the other (Gilbert & Troitzsch, 2005; Spielauer, 2011).

The core of microsimulation has been defined as “a means of modelling real life events by simulating the actions of the individual units that make up the system where the events occur” (Brown and Harding, 2002), and as “computer-simulation of a society in which the population is represented by a large sample of its individual members and their behaviours” (Spielauer, 2011). This has been broadened to encompass its role in policy so that “microsimulation models are computer programs that simulate aggregate and distributional effects of a policy, by implementing the provisions of the policy on a representative sample of individuals and families, and then summing up the results across individual units using population weights” (Martini & Trivellato, 1997, p. 84). Microsimulation models are based on individual-level data, or microdata relating to the characteristics and behaviours of individuals. The method endeavours to approximate an experiment whereby the potential range of outcomes is “simulated or imputed on the basis of a set of assumptions about the behavioural reactions of individuals following changes ... brought about by the introduction of a policy” (Zucchelli, Jones & Rice, 2010, p. 3).

The microsimulation method can account for the complex nature of social relationships and the heterogeneity of social groups (Zaidi, Harding & Williamson, 2009; Zucchelli, Jones & Rice, 2010). It relies on data from the real world to create an artificial one that mimics the original, but upon which virtual experiments can be carried out (Gilbert & Troitzsch, 2005). Microsimulation operates at the level of individual units, for example children, each possessing a set of associated attributes as a starting point. A set of rules, typically derived from statistical analyses, is then applied in a stochastic manner to each and every individual to simulate changes in state or behaviour. Such a model (Rutter, Zaslavsky & Feuer, 2011; Spielauer, 2007) can essentially generate a set of diverse synthetic biographies for the base sample of individuals. Modifications of influential factors can then be carried out to test hypothetical “what if” scenarios on a key downstream outcome of policy interest. Simulation at the micro level enables the assessment of the impact of potential policy changes for subgroups as well as aggregates of the population.

Last but not least, the technology requirements for implementing microsimulation are demanding, especially in relation to data sources (for inputs), and computer hardware and software (for implementation and outputs).

Types of microsimulation

There exists a range of types of microsimulation for policy purposes, which can be categorised on various dimensions. Zucchelli, et al. propose a taxonomy that comprises "arithmetical versus behavioural models and static versus dynamic models" (Zucchelli, Jones & Rice, 2010, p. 6). The primary features of each of these types are outlined below.

Arithmetical microsimulation models are typically run to estimate distributional and budgetary change occurring as a response to changes in taxes, benefits and wages. These models ignore the behavioural responses by individuals to the policy change being examined. For example, in the case of a model simulating tax and benefit changes, the simulation mimics adjustments to real disposable income following a policy change, and assumes that the behaviour of the individual does not change (Bourguignon & Spadro, 2006). Such models are useful in highlighting the impact of reforms, and in ascertaining those who benefit and those who lose from the modelled reforms, along with their characteristics (Zucchelli, Jones & Rice, 2010).

In contrast, behavioural microsimulation takes into account changes in the behaviour of the individuals in the simulation in response to the policy parameters modified. For example, changes to tax-benefit policy introduced into a model are considered to impact on the options favouring the individual, and may lead to a change in the amount of labour they choose to supply (Zucchelli, Jones & Rice, 2010, p. 8).

Static simulation models restrict their analysis to a single point in time or a set of points in time, without modelling the processes which drive the changes over time (Spielauer, 2011, p. 2). Thus static models evaluate the putative state of each individual under a changed set of policy rules (Brown & Harding 2002). Static models are typically used to model changes in taxes and social security benefits. A data-based static microsimulation model consists of two parts: (1) a baseline database – containing information on individual or family/household units, particularly sociodemographic characteristics and economic information that bears a relationship with a set of policies; and (2) a set of accounting rules – computer language instructions that produce for each unit the results of, for example, alternative tax or transfer policies and procedures (Martini & Trivellato, 1997).

"Dynamically ageing microsimulation models, on the other hand, involve updating each attribute for each micro-unit for each time interval" (Brown & Harding, 2002, p. 3.). Dynamic microsimulation models are constructed using either stochastic or deterministic algorithms. If the former, then the model is "based on conditional probabilities that certain economic or social conditions or processes will exist or occur" (Brown & Harding, 2002, p. 3.). If the latter, it is rule-based so that a condition will trigger a state or event. Spielauer (2011) argues that behaviours studied in dynamic microsimulation are of two types: "behaviour that produces events that take place over time such as demographic events, for example, marriage, divorce, death and economic events such as leaving the labour force", and "behaviour producing feedback reactions of individuals and/or families to changes in external circumstances, notably to changes in public policies" (Spielauer, 2011, p. 4).

Distinctions can be made among dynamic microsimulation models themselves depending on the nature of the base sample of individuals and the process by which it is aged (Spielauer, 2011):

1. actual *or* synthetic base sample, i.e. the base sample comes from an actual survey, *or* it is synthesised from various data sources;
2. open *or* closed population , i.e. other individuals can be introduced to the base sample as the simulation progresses, *or* membership of the sample is fixed;
3. cohort *or* population, i.e. a cross-sectional sample that is aged and eventually dies out, or a sample of the population whose representativeness is maintained over time;
4. discrete time *or* continuous time, i.e. individuals' characteristics are simulated at fixed time intervals, for example yearly, *or* their characteristics are assigned and can change at any time;
5. case-based *or* time-based, i.e. each individual's life course is simulated independently from beginning to end, *or* all individuals have their characteristics updated at the same time points.

The process of microsimulation

The process of building a microsimulation model follows several linked stages and components (Caro, et al., 2012; Cassells, Harding & Kelly, 2006; Zaidi & Rake, 2001):

1. Conceptualisation – considerations:
    a) setting clear objectives delineating the focus and scope of the model;
    b) deciding on the level of simplicity or complexity of the model;
    c) ensuring that the framework accounts for core processes of the phenomenon being modelled;
    d) recognising that proper configuring of model structure and pathways is critical;
    e) balancing the dual goals of explanation and prediction.
2. Computing platform – considerations (Percival, 2007; Scott 2003):
    a) following good IT practice;
    b) procuring adequate and sufficient resources;
    c) building a developer and user platform that is fit for purpose, reliable, and well supported.
3. Data integration – considerations:
    a) identifying and accessing various data sources, i.e. quantitative, qualitative, findings from other studies, or guesstimates;
    b) generating data to form the base file (initial conditions);
    c) generating parameters (from data analyses) to inform simulation rules;
    d) combining both data and parameters into a cohesive model.
4. Implementation – steps:
    a) initialisation, pre-simulation – setting up the base file;
    b) simulation – invokes a stochastic process in which a (transition) probability is compared to a number from a random draw;
    c) validation – comparing against both internal and external benchmarks;
    d) calibration (before simulation) or alignment (after simulation) to benchmarks.
5. Application:
    a) projection of the current state into the future;
    b) simulating the impact of potential policy change or intervention over time.

Is microsimulation useful for policy development?

Large sums of money are spent on public policy sometimes without a clear understanding of how they will impact on or benefit different groups in the population. The complexity of public policy issues calls for innovative methods to inform decision-making. Anchored in empirical data, microsimulation modelling is a prime candidate able to represent systems and processes with the potential to capture and test policy intricacies (Brown & Harding 2002). By attempting to mimic an experiment, it is a relatively inexpensive means of testing policy options to ensure ineffective polices are not implemented (Mitton, Sutherland & Weeks, 2000). Microsimulation has been primarily used for investigating the effects of tax and benefit policies but its use is spreading rapidly to other domains such as health policy (Glied & Tilipman, 2010; Ringel, et al., 2010; Rutter, Zaslavsky & Feuer, 2011).

Microsimulation techniques bring a range of benefits to social policy modelling, including the ability to change a greater variety of parameters independently and the capacity to provide considerably more accurate estimates and detailed projections of the distributional effects of changes. Two key advantages of microsimulation models are that: 1) they can replicate the complexity of the policy structures, transfers, and settings; and 2) they can be used to forecast the outcomes of policy changes and "what if" scenarios (i.e. the counterfactual where the results describe what, under specified conditions, may happen to particular individuals and groups) (Brown & Harding 2002).

Microsimulation models are important tools for the development of policy for other reasons. Policy evaluation is typically undertaken "by ex-post techniques which by definition are used to evaluate the impact of interventions and programmes following their implementation" (Zucchelli, Jones & Rice, 2010, p. 2). Use of microsimulation removes the need to collect information after the implementation of the policy – allows the testing of a wide range of scenarios – with the range only being limited by the data that are available. Microsimulation also allows ex-ante examination of the impact of policies before their implementation under different conditions. In this role the models can be used to assist in recognising ineffectual policies prior to their implementation (Zucchelli, Jones & Rice, 2010).

Strengths and weaknesses

"To simulate a society realistically requires detailed data, complicated models, fast computers, and extensive testing" (Spielauer, 2010, p. 1).

The primary strength of microsimulation techniques is their use of actual individual-level data which allows them to reproduce social reality and the intricacy of policy structures. The model can then be used to estimate the outcomes of "what if" scenarios (Brown & Harding, 2002, p. 4).

Spielauer (2010) notes that microsimulation is certainly the preferred modelling choice in three situations: (1) if population heterogeneity matters and if there are too many possible combinations of considered characteristics to split the population into a manageable number of groups; (2) if behaviours are complex at the macro level but better understood at the micro level; and (3) if individual histories matter, that is, when processes possess memory (Spielauer, 2010, pp. 6-8).

The strengths of microsimulation for policy purposes can be summed up as follows:

a) provides another piece of evidence to the jigsaw;
b) can tackle problems perhaps intractable by conventional means;
c) models processes or pathways that may be amenable to policy influence;
d) can account for social complexity, heterogeneity, and change;
e) can combine relevant information from disparate sources into a cohesive whole;
f) can be used to carry out virtual experiments, i.e. testing policy scenarios with various underlying assumptions;
g) produces aggregate and distributional results, and allows subgroup analysis perhaps for targeting of policy interventions.

There are typical weaknesses of the microsimulation method that are not necessarily inherent given more sophisticated modelling:

a) macro social outcomes tend to be merely an aggregation of micro ones;
b) no emergence of macro social phenomena from micro ones;
c) no interaction between individuals;
d) may be no feedback loops or reciprocal effects;
e) models tend to be predictive rather than explanatory;
f) pathways tend to be simplified;
g) true behaviour, e.g. motivation or intention, tends to be neglected;
h) tends to be no influence of exogenous macro factors (e.g. economic climate);
i) high demand for data and computer technology.

Conclusion

Microsimulation employs a system approach which can take into account social complexity, heterogeneity, and change. Microsimulation relies on the availability and quality of individual-level data that are representative of the target population. These data can come from various sources, which microsimulation is able to combine into a cohesive whole. Microsimulation also relies on the incorporation of parameters for the core processes underlying the system of interest. Thus a virtual (and realistic) world can be created, on which experiments can be undertaken and scenarios tested. Such a model is a simplification of reality but is nevertheless a powerful source of indicative information that can be used alongside other evidence for policy.

References

Anderson RE, Hicks C. (2011). Highlights of contemporary microsimulation. Social Science Computer Review, 29(1), 3-8.

Bourguignon F, Spadaro A. (2006). Microsimulation as a tool for evaluating redistribution policies. Journal of Economic Inequality, 4, 77-106.

Brown, L, Harding A. (2002). Social modeling and public policy: Application of microsimulation modeling in Australia. Journal of Artificial Societies and Social Simulation 5(4)6.

Caro JJ, Briggs AH, Siebert U, Kuntz KM. (2012). Modelling good research practices – overview: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Medical Decision Making, 32, 667-677.

Cassells R, Harding A, Kelly S. (2006). Problems and prospects for dynamic microsimulation: A review and lessons for APPSIM. Discussion Paper no. 63. National Centre for Social and Economic Modeling, University of Canberra.

Citro CF, Hanusek EA (eds), 1991. The uses of microsimulation modeling. Volume 1. Review and recommendations. Washington DC: National Academy Press.

Gilbert N, Troitzsch K. (2005). Simulation for the social scientist. Maidenhead: Open University Press.

Glied S, Tilipman N. (2010). Simulation modeling of health care policy. Annual Review of Public Health, 31, 439-455.

Gupta A, Harding A (eds.), (2007). Modelling our future: Population ageing, health and aged care. Amsterdam: Elsevier.

Gupta A, Kapur V. (eds.), (2000). Microsimulation in government policy and forecasting. Amsterdam: North-Holland.

Harding, A. (ed.), (1996) Microsimulation and public policy, Contributions to Economic Analysis Series. Amsterdam: North Holland.

Harding A, Gupta A (eds.), (2007). Modelling our future: Population ageing, social security and taxation. Amsterdam: Elsevier.

Li J, O'Donoghue C. (2012). A methodological survey of dynamic microsimulation models. UNU-Merit Working Paper #2012-002. Maastricht University.

Martini A, Trivellato U. (1997). The role of survey data in microsimulation models for social policy analysis. Labour, 11(1), 83-112.

Merz J. (1991). Microsimulation — A survey of principles, developments and applications. International Journal of Forecasting, 7(1), 77-104.

Mitton L, Sutherland H, Weeks M. (2000). Microsimulation modelling for policy analysis: Challenges and innovations. Cambridge: Cambridge University Press.

O'Donoghue C. (2001). Dynamic microsimulation: A survey. Brazilian Electronic Journal of Economics.

Orcutt G. (1957). A new type of socio-economic system. Review of Economics & Statistics, 39(2):116-23.

Percival R. (2007). APPSIM - Software selection and data structures. Working Paper. National Centre for Social and Economic Modeling, University of Canberra.

Ringel JS, Eibner C, Girosi F, Cordova A, McGlynn EA. (2010). Modeling health care policy alternatives. Health Services Research, 45, 1541-1558.

Rutter CM, Zaslavsky AM, Feuer E J. (2011). Dynamic microsimulation models for health outcomes. Medical Decision Making, 31(1), 10-18.

Scott A. 2003. A computing strategy for SAGE: 2. Programming considerations. Technical Note no. 3. London: Citeseer.

Spielauer M. (2007). Dynamic microsimulation of health care demand, health care finance and the economic impact of health behaviours: Survey and review. International Journal of Microsimulation 1(1): 35-53.

Spielauer M. (2011). What is social science microsimulation? Social Science Computer Review, 29(1), 9-20.

Zaidi A, Rake K. 2001. Dynamic microsimulation models: A review and some lessons for SAGE. Discussion Paper no. 2. London: Citeseer.

Zaidi, A, Harding A, Williamson P (eds.), (2009). New frontiers in microsimulation modeling. Public policy and social welfare, vol. 36. England: Ashgate Publishing.

Zucchelli E., Jones AM, Rice N. (2010). The evaluation of health policies through microsimulation methods. Working Paper 10/03. University of York.

Appendix

Examples of usage for policy

This section highlights a number of existing microsimulation models used for policy purposes. The list is not intended to be exhaustive, merely to provide an indication of the broad sweep of models in use across various domains and countries. There are useful reviews and compendia of existing microsimulation models that describe their features in more detail (for example: Li & O'Donoghue, 2012; O'Donoghue 2001).

APPSIM (Australian Population and Policy Simulation Model)

APPSIM, developed by the National Centre for Social and Economic Modelling (NATSEM), is a dynamic model which simulates the life cycle of 200,000 individuals (1% sample of census) from 2001 to 2050. It shows how the Australian population develops over time under various scenarios, and allows the social and fiscal impacts of policy changes over time to be simulated.
http://www.natsem.canberra.edu.au/models/appsim/

EUROMOD (Europe)

EUROMOD, based at the Institute for Social & Economic Research, University of Essex, is a static tax-benefit model for the European Union (2000s). It enables researchers and policy analysts to calculate, in a comparable manner, the effects of taxes and benefits on household incomes and work incentives for the population of each country and for the European Union as a whole.
https://www.iser.essex.ac.uk/euromod/

FEM (Future Elderly model – USA)

FEM, developed by the RAND Roybal Center for Health Policy Simulation, is a demographic and economic simulation model designed to predict the future costs and health status of the elderly and explore the implications for policy. It uses a representative sample of 100,000 Medicare beneficiaries aged 65 and over drawn from the Medicare Current Beneficiary Surveys. Each beneficiary in the sample is linked to Medicare claims records to track actual medical care use and costs over time.
http://www.rand.org/labor/roybalhp/projects/health_status/fem.html

LIFEPATHS (Canada)

LIFEPATHS, developed by Statistics Canada, is a dynamic model of individuals and families from an 1872 birth cohort, to today. It creates synthetic life histories from birth to death that are representative of the history of Canada's population. It can be used to evaluate government programs, or to analyse societal issues of a longitudinal nature, e.g. intergenerational equity.
http://www.statcan.gc.ca/microsimulation/lifepaths/lifepaths-eng.htm

MIDAS (Microsimulation for the Development of Adequacy and Sustainability – Belgium, Germany, Italy)

MIDAS is a dynamic population model. Starting from a cross-sectional dataset representing a population of all ages at a certain point in time (early 2000s), the life spans of individuals are simulated over time. So new individuals are born, go through school, marry or cohabit, enter the labour market, divorce, retire and, finally, die. During their active years, they build up pension rights, which result in a pension benefit when they retire.
http://www.plan.be/publications/Publication_det.php?lang=en&TM=30&KeyPub=781

POHEM (Population Health Model – Canada)
POHEM is a dynamic microsimulation model intended to represent the lifecycle dynamics of the population. The simulation creates and ages a large sample until death. The life trajectory of each simulated person unfolds by exposure to different health events. The model combines data from a wide range of sources, including cross-sectional and longitudinal surveys, cancer registries, hospitalisation databases, vital statistics, census data, and treatment cost data as well as parameters in the published literature.
http://www.statcan.gc.ca/microsimulation/health-sante/health-sante-eng.htm#a2

SADNAP (Social Affairs Department of the Netherlands Ageing and Pensions)
SADNAP is a dynamic microsimulation model for estimating the financial and economic implications of an ageing population and evaluating the redistributive effects of policy options. It simulates the life paths of a sample of the Dutch population using transition probabilities on demographic events. The model uses administrative data sets on pension entitlements and payments.
http://ima.natsem.canberra.edu.au/IJM/V4_1/Volume%204%20Issue%201/5_IJM2011_van_Sonsbeek_CORRECTED_GD_JMS.pdf

STINMOD (Static Incomes Model – Australia)
STINMOD is a static microsimulation model of Australia's income tax and transfer system, developed by the National Centre for Social and Economic Modelling (NATSEM). The model is mostly used to analyse the distributional and individual impacts of income tax and income support policies and to estimate the fiscal and distributional impacts of policy reform.
http://www.natsem.canberra.edu.au/models/stinmod/