1 Introduction

Lavinia Mitton, Holly Sutherland and Melvyn Weeks

Microsimulation models use micro-data on persons (or households, or firms or other micro-units) and simulate the effect of changes in policy (or other changes) on each of these units. Differences before and after the change can be analysed at the micro-level or aggregated to show the overall effect of the change. It is the dependence on individual information from the micro-data at every stage of the analysis that distinguishes microsimulation models from other sorts of economic, statistical or descriptive models.

Modern policy problems require analysts to capture the interactions between policy and the complexities of economic and social life, as well as between policies of different types. Microsimulation is increasingly a technique that is employed to analyse these problems. At the same time, developments in computing power and analytical techniques allow a greater sophistication in the view of the world that microsimulation models can attempt to portray, and hence in the range of questions that they may address. This book brings together examples of microsimulation modelling that are at the frontiers of developments in the field, either because they are extending the range of techniques available to modellers, or because they demonstrate new applications for established methods.

The problem of determining the impact of a proposed change in policy has been succinctly summarised by Heckman and Smith (1995, p. 87) as being:

> the fundamental evaluation problem that arises from the impossibility of observing what would happen to a given person in both the state where he or she receives a treatment (or participates in a program) and the state where he or she does not. If a person could be observed in both states, the impact of the treatment on that person could be calculated by comparing his or her outcomes in the two states, and the evaluation problem would be solved.

In principle, the effect of a policy change can be isolated by the comparison of 'before' and 'after' observations, or comparison of two groups that are considered identical except that one group has received the treatment. However, despite the use of a range of statistical techniques, the difficulty of creating *ceteris paribus* conditions may result in selection bias, distorting the true impact of the treatment. Often we do not observe all the significant characteristics that need to be controlled for. Furthermore, we may wish to use models to inform the design of new types of policy or to predict their impact in changed social or economic conditions. In these situations, microsimulation models can provide a consistent and structured framework in which to explore a range of 'what if' questions about the outcomes of policy reforms.

Traditionally, microsimulation models are divided into two types: *static* and *dynamic*. Microsimulation is essentially a set of methods for the generation of missing information, and the distinction between static and dynamic depends on the particular method that is used. Most critical is the method for *ageing* the micro-units (Harding, 1996; Merz, 1991). Static models typically use a combination of re-weighting of micro-units and indexation of money amounts to update cross-sectional micro-data to the required point in time. Some static models may use no ageing at all, and may operate in terms of the time at which the underlying data were collected. Dynamic ageing, on the other hand, changes the characteristics of the micro-units in response to accumulated experience or the passage of time. At the most basic level, units are older by a year in each year of the updating. As the unit gets older, combinations of stochastic and deterministic methods are used to predict changes in status. In models of persons, the changes in status typically include labour force participation, co-habitation and parenthood. At each stage, incomes are estimated, based on current status and circumstances and past history. Dynamic models generate long-term or lifetime data describing each micro-unit. They may operate in a time warp, abstracting from real changes such as economic growth, or they may either predict the future or fill in missing information about the past (see Harding, 1990). Dynamic ageing is sensitive to assumptions about macro conditions (such as unemployment) and dynamic models may be used to explore the effects of incorporating alternative assumptions about the future, or alternative representations of the past.

In principle, static or dynamic models may be augmented by introducing behavioural response, which allows the calculation of second-order effects due to changes in, for example, labour supply or fertility, following a policy change. For dynamic models, incorporating behavioural response means altering the nature of the transition probabilities that are used to age the micro-units. In practice, this is rarely done.

The way in which behaviour change is estimated is clearly an important issue. The principal drawback of behavioural models derived from cross-sectional data is one of interpretation. Given that cross-sectional models lack an explicit time dimension, it is not possible to examine the time profile of the impact of a particular policy change. Further, it is not possible to differentiate between the two major problems in predicting individual behaviour – namely heterogeneity and time dependence (see Heckman, 1981). The increasing availability of panel data provides us with prospects for improvement on two fronts: first, the opportunity to estimate econometric models of behavioural response that incorporate 'dynamics'; and second, better-founded estimation of transition probabilities that are used to age the sample in a dynamic microsimulation model.[1]

The choice of whether to use a static or dynamic microsimulation model, and whether to include behavioural modelling, depends in principle on the policy question to be addressed and also on the quality and suitability of available data. (In practice, it also depends on the institutional context and the speed with which an answer to the question is necessary. These issues are considered below.)

Static microsimulation is generally used when only cross-sectional information is needed to answer a policy question. For example, in this volume, Kaplanoglou models the distributional effect of a change in tax rates. Dynamic methods are used when a set of repeated cross-sections is needed, as when modelling the effects of pension reform (see Bonnet and Mahieu in this volume), or when long-range future prediction is required for a current sample, such as in Hancock's paper exploring the financing of long-term care. In either case, whether behavioural responses to the policy changes are modelled depends on a number of factors. In some circumstances we may want to know the first-round effects (for example, if we are concerned about the perception of the fairness of the reform) or we may want to know both the first-round effects and the full effects because we believe behaviour will take some time to change. In addition, whether we model changes in behaviour depends on the availability of suitable data for the estimation of behaviour, and whether we believe that our estimation is sufficiently reliable to make the exercise worthwhile. However, in cases where the proposed policy is designed to alter behaviour – such as the provision of subsidised child-care, considered by Duncan and Weeks in this volume – then an estimate of the revenue cost using only a deterministic set of rules will, in general, generate a misleading estimate of the overall impact.

On the other hand, the introduction of a behavioural component will also introduce an additional element of uncertainty, generated by the parameter estimates. All microsimulation model results are subject to some degree of error. The great strength of microsimulation based on micro-data is that the full range of variation in circumstances is captured.

However, many of the processes that are part of building any microsimulation model – updating and uprating, imputing missing values and reweighting, as well as estimation – involve some degree of aggregation and approximation. The effect of sampling error on the reliability of a range of static model outputs has been calculated by Pudney and Sutherland (1994). The same authors have estimated the additional error due to a labour supply response model (Pudney and Sutherland, 1996).[2] It is often a question of judgement as to whether the bias introduced by failing to capture important effects (such as behaviour change, or differences between small groups) is outweighed by additional uncertainty due to imperfect estimation procedures or underlying data. The calculation of confidence intervals around model results that account for *all* sources of error remains a major challenge.

In dynamic microsimulation modelling, the simulation of a realistic ageing of the population depends on the existence of reliable transition probabilities. Since these probabilities are estimates, they naturally introduce an additional element of uncertainty. Given the nature of these models, any error in one component is likely to be compounded with repeated use during the ageing process. Devising methods to assess the extent of error in dynamic model outputs, particularly those that cannot be subjected to direct or 'collateral' comparison with independent information, is currently 'an art rather than a science' (see Caldwell and Morrison in this volume).

The categorisations 'static', 'dynamic' and 'behavioural' are useful from the model-builder's perspective because the nature of the work involved, and the physical resources required to carry out these three types of modelling are quite different. For these reasons the sections in this book follow traditional divisions: the first section includes papers that use static methods innovatively; the second focuses on developments in dynamic model building and the third on modelling behaviour change. This introduction considers the papers in each section in turn and concludes with some reflections on the process of model building.

1.1 New directions for microsimulation

Part One of this volume contains chapters on what would traditionally be termed static models. However, it is increasingly the case that microsimulation models do not fall conveniently into the established categories of static and dynamic models, either with or without behavioural change. Indeed, some of the contributors to this volume would challenge the taxonomy that we have set out above. Thus, many of the chapters in Part One incorporate an innovative dimension that makes the choice of ageing method less relevant as a classification criterion. For example, the pensions model described by Rake could be defined as static in that it employs a single cross-section for its data source and uses static ageing. But it also contains elements in common with dynamic models in relation to the methods used to simulate lifetime earnings and pension contributions of the sample individuals that were not recorded in the original survey data.

The aim of Rake's chapter is to illustrate how microsimulation can be used when conducting cross-country comparative social policy research. In the field of social policy, cross-national comparisons are frequently limited to the descriptive: the researcher can become overwhelmed by the variation between national policies. Microsimulation can strip away some of the underlying national variations in order to improve our understanding of the link between policy and outcome. Thus, Rake is able to show the effect of the British, French and German pensions systems in 'exaggerating, replicating or mitigating the earnings differential between women and men'.

Another innovation in microsimulation is the linking of micro-models to macro-models of various kinds. The paper by Cameron and Ezzeddin describes a preliminary exercise to link two established models together in order to incorporate some macroeconomic second order effects. The models are a static microsimulation model and a regional input–output model for Canada. The linked models simulate both the direct and indirect effects on micro-level economic well-being of various tax/transfer and social policy alternatives in the Canadian Provinces. They can also be used to assess the distributional effects on households of changes to the economic climate and the industrial base. Although its preliminary nature means that some of the links are crude (the adjustments to micro-level earnings are proportional, for example), this paper demonstrates the potential for integrating personal and industrial sector models.

The models described in Part One are not used simply to measure the effects of actual or prospective policy reforms – the original motivation for the development of static microsimulation models. Policy simulations are also used as analytical devices to understand the operation of existing systems (as in Rake's paper) and the sensitivity of results to conventional assumptions. For example, Kaplanoglou uses microsimulation methods to discover that the apparent mild progressivity of the Greek indirect tax system is due solely to taxes on private transport: if car taxes are treated as road use charges rather than pure taxes, the regressive parts of the system are left to dominate the overall pattern. Some of the apparent progressivity arises from the lower rates of car ownership among less well off groups. Based on her results from simulating a uniform VAT as an alternative to the immensely complex Greek indirect tax structure, Kaplanoglou is able to make a strong argument for simplifying Greek indirect taxes.

Decoster and Van Camp also use microsimulation to explore the sensitivity of results to conventional assumptions, at the same time as providing a distributional analysis of the 1988 Belgian tax reform. They consider the issue of the unit of analysis and the effect of the choice between fiscal unit and household on conclusions about the distributional characteristics of the tax system. Using a statistical matching procedure, a link is made between an administrative fiscal data set and the household budget survey for Belgium. This allows personal tax liabilities to be imputed for each fiscal unit within each household in the survey, and hence for the implications of either unit of analysis to be explored. One finding is that the redistributive power of the tax system was reduced by the reform, and that this conclusion is *not* sensitive to the choice of unit. However, for both the pre-reform and post-reform tax systems, the choice of the unit of analysis *does* affect conclusions about the redistributive effect of the systems themselves. In both cases, use of the fiscal unit leads to *higher* parameters of the tax system than if the household is used.

Taking the issue of sensitivity of results in another direction, the paper by O'Donoghue *et al.,* focuses on the comparability of model results across countries. Using a prototype of the EUROMOD model, a static tax–benefit model for the whole European Union, it addresses the question of the sensitivity of European model results to the way in which incomes are measured across household types and across countries. It shows that the country composition of quantiles of the European income distribution can be sensitive to the choice of equivalence scale, adjustments for apparent differences in the quality of micro-data, and exchange rates. It implies that the evaluation of policy at the European level requires careful interpretation in the light of the assumptions that have been chosen.

Finally, the paper by Walker *et al.,* extends the use of static microsimulation into a new policy area: the growth in expenditure on the Australian Pharmaceutical Benefits Scheme (PBS) which subsidises the cost of medicines. The model uses static ageing techniques and detailed information on a range of types of prescribed medicines. The paper illustrates the potential of the model by analysing the likely outcomes for the present scheme in 2020 under three scenarios: an ageing population; a continued upward trend in medicine costs; and a general improvement in Australians' health leading to lower usage of prescribed medicines. The analysis finds that increases in drug prices are likely to have the greatest impact on the cost of the PBS, and population ageing the least impact, but suggest that improvements in Australians' health have the potential to limit cost increases significantly.

1.2 Dynamic modelling

The development of microsimulation models can challenge the results achieved through other methods. Two of the papers in Part Two – by Caldwell and Morrison, and by Bonnet and Mahieu – confront other types of model that are traditionally used to analyse policy issues related to pensions. Bonnet and Mahieu contrast the use of dynamic microsimulation with that of Computable General Equilibrium 'overlapping generations' models. Their dynamic model is able to study the transfers within generations as well as between them. Caldwell and Morrison confront the dynamic model, DYNACAN, with results from the Canadian actuarial model of pensions in a more direct way: it is expected to produce results that are consistent with the more traditional actuarial approach.

Population ageing and slower rates of economic growth raise many questions about the future of intergenerational public transfers in countries such as France with generous pay-as-you-go public pension systems. Bonnet and Mahieu describe their dynamic microsimulation model, *Destinie*, and use it to explore the implications of six alternative economic, demographic and policy scenarios. Since they are concerned to compare the microsimulation approach with overlapping generations models, their focus is mainly on the differential effects on successive generations. However, they also examine effects by gender and income level within cohorts.

As Caldwell and Morrison point out in their paper, validation is a vital part of integrating the use of microsimulation models into the policy development process. They present a range of types of validation and reconciliation for two dynamic models sharing a common basic structure (DYNACAN for Canada and CORSIM for the US). They note the lack of literature or theory on which to base a validation exercise, but are nonetheless optimistic, maintaining that validation is 'not a problem to be avoided, but an asset to be exploited'. It can be seen as an opportunity for improving understanding of the modelling process itself.

Hancock uses dynamic microsimulation to simulate the contributions that older people will make towards the cost of care in a residential home, should they need it. She simulates what older people in a relatively high-risk group might pay towards care costs, both now and in 15 years' time. She explores a range of charging options including the use of housing wealth. Of particular interest to prospective builders of dynamic microsimulation models is Hancock's 'progress through small steps' approach. Most dynamic model construction projects involve teams of people, a long-term and large-scale resource commitment (by social science standards) and the associated costs of management and co-ordination. In contrast, Hancock single-handedly focuses on the new and particular issues raised by her research questions, leaving aside for later development aspects of the model that are not central to them.

At a practical level, building microsimulation models is all about detail, on the one hand, and finding ways of representing complex processes in a tractable form, on the other. Chénard's paper describes the solution found to one particular problem faced by the Canadian dynamic model, DYNACAN. It does so in a way that graphically illustrates to the non-practitioner the process of constructing a dynamic microsimulation model, at the same time as documenting, for the fellow-modeller, a neat solution to a difficult problem. The problem is migration. While migration itself affects a whole family, keeping the model consistent with external totals ('alignment') must be done on an individual basis. The technique developed in the paper is based on the *pageant* principle ('many are called but few get chosen'). It allows alignment on an individual basis at the same time as transition on a family basis, and is in principle applicable to problems other than migration.

1.3 Modelling behavioural response

Three papers include some behavioural response modelling. The chapter by Swan examines the distortionary effects on migratory behaviour of unemployment benefit in Sweden. The model *Sverige* is in the early stages of development as a dynamic microsimulation model. The design is fairly standard, but with two important innovations. The first is that the underlying data include information on the location in Sweden of every person in the sample to within 100 square metres. Thus there is enormous potential to explore the spatial aspects of policy. In this chapter, a logistic regression approach is used to estimate separately the effects on migration of changes in unemployment and changes in unemployment benefits. The results show that there is a significant – but small – effect of benefit levels on migration. However, the effect is purely monetary in the sense that unemployment itself does not appear to have a significant effect.

The second innovative feature that is planned is the ability to choose between alternative labour market functions in the model. The standard treatment is to use 'natural rate of unemployment' theory to achieve alignment during simulation. Using other theories would give rise to different results for changes in regional and national unemployment rates following migration. In the case considered by Swan, the small predicted migration effect would give rise to – at most – small changes in unemployment under any theoretical assumption. However, the general prospects for users of being able to choose their preferred theoretical framework, is a significant step. Indeed, the laying bare of underlying theoretical assumptions is an innovation in this context.

The two remaining chapters both focus on the problems of modelling labour supply responses. Unlike the majority of models of labour supply, the study by Aaberge *et al.*, considers the *joint* labour supply of household members. A second innovation is the incorporation of demand constraints. Otherwise unconstrained choices are adjusted by the likelihood of obtaining jobs with given hours and wages combinations. Previous studies (see van Soest, 1994) have noted the empirical tendency of labour supply models to overestimate predicted part-time employment. This is due, in part, to the focus on the supply-side characteristics of individuals, thereby ignoring the influence of the fixed costs of employment on the availability of part-time employment.

Using Italian survey data, the authors simulate the impact of a number of tax reforms including the introduction of a flatter profile of tax rates and a negative income tax regime. The reforms involve incentives for some people to work less and others to work more, such that the more productive decide to work longer hours. However, if the quantity constraints on hours choice are removed, an increase in the participation rates for individuals in the poorest income deciles is predicted. This result provides further evidence that models of labour supply that ignore demand-side factors will on average over-predict participation rates for lower income groups whose opportunity set consists mainly of home production and difficult-to-find part-time jobs.

The problem of the tendency to over-predict part-time working in discrete choice models of labour supply is also addressed in the chapter by Duncan and Weeks. They recognise that a discrepancy between observed and predicted choice is not a problem if the analyst is simply interested in predicting aggregate frequencies. However, if the estimation model is linked to a microsimulation model in order to predict the costs of, for example, the introduction of subsidised child-care, then the within-sample forecasts of the underlying choice model need to be accurate.

The chapter assesses the performance of a number of transition estimators, including the maximum probability rule estimator and estimators based upon calibration. Using both a Monte Carlo study and labour market data from a household survey, substantial differences in performance between the alternative estimators are found. Significantly better properties are exhibited by those based on calibration, where the baseline model is adjusted to remove discrepancies between observed and predicted outcomes.

1.4 Investing in microsimulation models

Several of the chapters in this volume refer to their models as being prototypes, 'preliminary' or work in progress. In practice, most microsimulation models are – as far as their builders are concerned – major enterprises requiring many person-years of expertise, attention to detail and stamina. As investment in research capacity, they can all be considered to be work in progress in some sense or another. They may develop one step at a time, as resources permit (in this volume, see Hancock). Or they may be set up as large-scale enterprises with multiple goals and a relatively long time horizon (in this volume, see Caldwell and Morrison, and Chénard (DYNACAN); Swan (*Sverige*); O'Donoghue *et al.,* (EUROMOD)).[3] It is also the case that some models may never be distinct objects with identifiable histories. They may be a collection of procedures and techniques that are assembled, re-assembled or discarded in the search for methodological improvement (in this volume see Aaberge *et al.*; Duncan and Weeks). Some model builders consider themselves to be answerable to a set of 'users' or 'clients' (see Caldwell and Morrison in this volume, and also Immervoll *et al.*, 1999). This can constrain the model development process. It can also provide a valuable focus and discipline. For others, there are no distinctions or separations between model builder, model user or user of the model's output.

These differences are related to the range of types of institutions in which microsimulation model construction and analysis takes place (see Sutherland, 1998). This volume combines chapters reporting on modelling efforts in government departments with modelling developments that are carried out in academic environments. The institutional differences can be important in understanding the motivation for the project and the constraints under which it operates. For example, modelling by government departments is often initiated by policy-makers' needs to find answers to specific questions. These may be regular or multiple needs, encouraging investment within government in the development of durable and flexible models. The incentives in the academic world for this investment are less obvious, although the EUROMOD project shows that it is possible. It is also quite clear that innovation in technique is not the preserve of academic modellers: in some countries, many of the most exciting projects are carried out by government analysts. At the same time, academics do have the freedom to look beyond the analytical needs of current policy agendas, both in the direction of technical virtuosity and in order to model independent, alternative or dissident policy ideas.

Thus there is not only a role for microsimulation modelling in each type of institution, but also a set of good reasons to encourage and maintain a