# Microsimulations of Public Policy

J. Theodore Anagnoson
Department of Political Science
California State University, Los Angeles
Los Angeles, CA 90032-8226
anag999@silcom.com
(323) 343-2245 (office)
(323) 343-6452 (fax)

## Abstract

The purpose of this chapter is to introduce microsimulations of public policy to public managers. Microsimulations are large models that start with a large-scale survey of the public. To the survey are added several kinds of information: data from other surveys and databases, imputations and statistical matches, program rules, and behavioral assumptions. Microsimulations are different from the usual surveys that are taught in graduate programs of public administration in that they can be used to make estimates of the number of participants in different eligibility categories and the potential costs of programs and potential policy changes to those programs. The past decade has seen the continued development of some existing microsimulations and the new development of several simulations in different countries. Issues that public managers may face when they come into contact with microsimulations are discussed.

## I. Introduction

Consider these questions:

- What is the effect of a change in the income tax rate upon different types of families ?
- What would it cost to add a prescription drug or a particular kind of cancer screening benefit to the Medicare program?
- If Congress changes the Aid to Families with Dependent Children (AFDC) program to require a maximum five year lifetime limit on the receipt of funds from AFDC, what proportion of the welfare population will face a time limit?
- If we wanted to restore benefits to certain categories of immigrants in the food stamp program, what would be the cost, and how many people would participate?

Each of these questions could be answered in part in only a limited way with a survey, like the Current Population Survey (CPS) or the Survey of Income and Program Participation (SIPP), because crucial features are not present in the average survey. We don't know, for example, from surveys, what the respondents' health insurance premiums are, or the exact amount or components of their taxes, or whether they are eligible for welfare, Medicaid, or other programs.

When we ask survey respondents whether they received aid from public welfare or health programs, we frequently find far fewer respondents admitting to having received such aid than administrative data indicates. For Medicaid, for example, the average survey will result in half or fewer of the number of respondents that the Health Care Financing Administration knows from administrative data are on the program. In order to simulate

the operations of various programs and aspects of the tax code, we have to add information to the average survey. Adding that information is what a microsimulation of public policy is all about.

A microsimulation of public policy begins with a large survey of the public or the relevant group, to which we add information about the rules and operations of the programs we are interested in and how the members of the population affected by that program behave. We may trend these data into the future, weighting the results to replicate future projections (static aging), or even allowing, with a certain probability, people to live, die, retire, enroll in Medicare or Medicaid, or enroll in other programs (dynamic aging). We can, with this information, answer questions about costs and program enrollment in the future, deal with the interrelationships among programs in a systematic way and more thoroughly than could be done by hand or on "the back of an envelope," and estimate how the participants in these programs would be affected by the proposed changes.

**II. What is microsimulation?**

Trippe and Stavrianos (1998, 2) define a microsimulation model as one that "simulates how a change to a government transfer program would affect the costs and caseload of that program." Citro and Hanushek state that microsimulations are "large, complex models that produce estimates of the effects on program costs and who would gain and who would lose from proposed changes in government policies ranging from health care to welfare to taxes" (National Research Council, 1991). Mathematica Policy Research, the firm that maintains the Micro Analysis of Transfers to Households (MATH) model, defines microsimulation as "a type of computer program that simulates how a welfare program would operate under proposed changes and how participants would be affected" (Mathematica Policy Research, http://www.mathematica-mpr.com/math-2.htm.

As such, simulations of public policy models begin with a large scale survey of the public, often the Current Population Survey in the U.S. A microsimulation is much more than just the survey, however. To the survey is added considerable information from other surveys, the program rules for the programs being simulated, interactions among those programs, and how the people in the survey behave. Static or dynamic aging techniques will then be used to simulate the population during the year being investigated if that year is different from the simulations input data year.

In short, microsimulations of public policy are much more than just surveys; they are entire worlds unto themselves. Therein lies both their advantages and possibilities, and some central problems.

**III. Social Science Simulations During the 1990s**

We should distinguish microsimulations of public policy from other kinds of simulations, particularly the agent-based simulations that have become very popular during the 1990s. The history of simulation is basically that microsimulations of public policy of the type described immediately above were first developed and used in the 1960s and 1970s. During the 1980s, the U.S. federal government cut back its funding, although some, if not all, of the models continued at least to maintain the status quo. Later in the 1980s and certainly during the 1990s, there has been a resurgence of interest in these kinds of public policy models both in the United States and abroad, and at the same time several other kinds of simulations have been developed and have attracted considerable interest (for a more detailed history of social science simulation, see Gilbert and Troitzsch, 1999).

Gilbert and Troitzsch (1999) divide the social science simulation enterprise in the late 1990s into the following categories (see also Gilbert, 1995):

- Systems dynamics and world models are macro level difference or differential equation models.
- Microanalytical simulation models are the type we call "microsimulation models of public policy" in this chapter.
- Queuing models or discrete event models focus on a system and its events, activities and delays.
- Multilevel simulation models attempt to model the interactions between the macro level and the micro levels for a given population or group.
- Cellular automata "model social dynamics where the focus is on the emergence of properties from local interactions" (Gilbert and Troitzsch, 1999).

- Multi-agent models model somewhat more complex situations from cellular automata in which the 'agents' control their own actions, typically on a probability basis, based on their perceptions of the environment (Gilbert and Troitzsch, 1999). Agent-based models have become very popular in the 1990s, and a growing literature can be found in the newly founded (volume 1 is 1998) *Journal of Artificial Societies and Social Simulation*, available freely on the Internet (http://www.soc.surrey.ac.uk/JASSS/). A commonly used computer program to build agent-based models is SWARM, recently developed by the Santa Fe Institute (Santa Fe Institute, 1998; Terna, 1998).
- Learning and evolutionary models incorporate learning over time in response to environmental constraints or changes.

This chapter considers on the second category, microanalytical simulation models or 'microsimulations of public policy'.

A common feature of many kinds of simulations, including both agent-based models and microsimulations of public policy, is the use of probabilities of interacting or performing certain events over time. With the agent based simulations, those are often probabilities of interacting. In microsimulations of public policy, some models use dynamic models of aging in which each person in the survey has a certain probability of dying, marrying, retiring, etc. as each year advances, as mentioned above.

**IV. Major microsimulation efforts**

Microsimulation efforts are underway in a number of countries:

- Australia. The National Centre for Social and Economic Modelling (NATSEM), established in 1993 as part of the Faculty of Management at the University of Canberra, maintains a number of microsimulation databases for retirement, social security, and research on the structure of Australian society. For more information, see: http://www.natsem.canberra.edu.au/.

- Canada. A variety of microsimulation models are used in several Canadian federal departments; Statistics Canada has developed and provides several such microsimulations to interested users free. For more information, see: http://www.statcan.ca/english/spsd/index.htm. Also see the web site of Burkby Consulting, which has done some of the developmental work on the simulations: http://www.comnet.ca/~burpee/index.htm.
- Britain. Extensive microsimulation efforts, including a microsimulation Unit based in the Department of Applied Economics at the University of Cambridge (http://www.econ.cam. ac.uk/dae/mu/microsim.htm) and another center, CRESS (the Center for Research on Simulation in the Social Sciences) (http://alife.ccp14.ac.uk/cress/research/simsoc/cress.html). The European Commission is funding an integrated European benefit-tax model at the Cambridge University center (http://www.econ.cam.ac.uk/dae/mu/emod.htm). The Institute for Fiscal Studies, an Economic and Social Council (ESRC) research center, has a page on "Microsimulation and Economics," including descriptions of the Institute's tax and benefit model and its construction of a tax-benefit model for the Czech republic (http://www1.ifs.org.uk/research/basicResearch.HTM).

In the United States, the following microsimulations are being used for the analysis and estimation of public policy changes:

- The ASPEN Microsimulation Model of the US Economy is being developed at the Sandia National Labs using "evolutionary learning and parallel computing" to model the entire U.S. economy. This model is agent-based (search for Aspen at http://www.sandia.gov).
- CORSIM, the Cornell Dynamic Microsimulation Model, maintained at Cornell University, is an agent-based microsimulation model of individual and family behavior through time and has been used for projects relating to the U.S. Social Security program, household wealth accumulation, private pensions, and others (http://misic.soc.cornell.edu). The model includes a wide variety of information including demographic, employment, and earnings information. The Cornell organization is developing projects in both Canada and Sweden.

- Mathematica Policy Research (MPR) maintains the family of MATH (Micro Analysis of Transfers to Households) models, centered on the analysis of changes to the Temporary Aid to Needy Families (TANF) and Food Stamps programs. One of the clearest explanations of microsimulations of public policies is on their web site (http://www.mathematica-mpr.com/MICROSIM.HTM). The model, in a personal computer version, may be purchased from MPR for less than $100 U.S.
- The Social Security Administration has two microsimulation models. The MINT (Modeling Income in the Near Term) model projects retirement incomes of persons in the Survey of Income and Program Participation (SIPP) until 2020; the Historical Cohort model simulates Social Security proposals on a sample of just-retired persons in the March, 1994 Current Population Survey. For more information, see Bailey, Cohen and Iams (1998).
- The Center for Policy Research at Syracuse University has linked data from several different waves of the National Long-Term Care Survey and Medicare to produce the Dynamic Microsimulation of Elders' Health and Well-Being (http://www-cpr.maxwell.syr.edu/demogctr /microsim.htm
- Faculty at the University of California's campuses at Davis and Irvine have developed a Program for Improved Vehicle Demand Forecasting Models, based on extensive surveys of California households, and designed to predict vehicle ownership, miles of travel, and fuel demand (http://128.200.36.2/its/research/fuel.html).
- The Urban Institute has developed and maintains the Transfer Income Model (TRIM2), one of the earliest and most extensive models, and presently being developed for a world wide web interface. The Institute web site has a number of papers using TRIM2 data and results, but no explicit description of the model. The best description is found in Giannarelli (1992) and Giannarelli and Moore (1997). The Washington, D.C. based organizations Center on Budget and Policy Priorities and Citizens for Tax Justice also use the model, and the MATH model began as an earlier version of TRIM. Lewis and Michel (1990) contains a history of the development of the TRIM model. See http://www.urban.org.
- U.S. Los Alamos National Labs have developed "The Transportation Analysis Simulation System, aimed at predicting transportation and land use issues based on survey data from specific geographic areas. A remote user interface, allowing access to data sets residing at the lab, is under development (http://www-xxx.tsasa.lanl.gov/transims_intro.html).

As can be seen, many of the microsimulations used by U.S. government agencies have been developed by consulting organizations and academic departments, with the federal agency providing the funding to develop and maintain the model.

World wide web pages with links to lists of either books or other microsimulation sites include the Center for Research on Simulation in the Social Sciences (http://alife.ccp14.ac.uk/cress/research/simsoc/microsim.html and Dr. Troitzsch's web sites at the University of Koblenz-Landau, http://www.uni-koblenz.de/~kgt/Learn/Textbook/node155.html for microsimulation units, and http://www.uni-koblenz.de/~kgt/Books.html for books.

V. Review of the Literature

At this point, the literature on microsimulation can be divided into three categories:

- Book length reviews, of which the most accessible in both price and comprehensibility are National Research Council, 1991, Lewis and Michel (1990), and Giannarelli (1992). More recent reviews include Harding (1996), a collection of papers presented at the International Association for Research into Income and Wealth's special conference on microsimulation in 1993, and Conte, Hegselmann, and Terna (1997), another collection representing the range of interests into microsimulation in the early to mid-1990s. See above for a web site listing recent books. Rivlin and Wiener (1988) has a straightforward description of the Brookings-ICF Long-Term Care Financing Model, which uses data from ICF's Pension and Retirement Income Simulation Model (PRISM) and several other data sets.

- The World Wide Web sites maintained by almost every organization that maintains a microsimulation model. These are listed above. Of particular note is the Mathematica Policy Research site, which is

particularly useful in introducing microsimulation and explaining exactly what the MATH model offers, and both Statistics Canada and MPR, which offer copies of their models for interested users.

In addition, a series of papers presented at the Dagstuhl Seminar on Social Science Microsimulation: Tools for Modeling, Parameter Optimization, and Sensitivity Analysis" are available at http://www.uni-koblenz.de/~kgt/Dagstuhl9719.html.

- Journals and conferences. The *Journal of Artificial Societies and Social Simulation*, whose web site is indicated in the list of references, has articles on all aspects of simulations. The annual meeting of the Association for Public Policy Analysis and Management usually has several papers using microsimulation data; these are sometimes available online (http://qsilver.queensu.ca/appam). Every few years one of the major centers for microsimulation has a conference, with the results published; the latest are Harding (1996), a volume of selected papers from the International Association for Research into Income and Wealth's conference on microsimulation in Canberra in 1993, and Troitzsch, Mueller, Gilbert and Doran (1996), which includes papers from a conference of both social and computer scientists on a variety of approaches to simulation.

**V. Steps Involved in Transforming A Survey into a Simulation**

Surveys are not simulations. Surveys can be transformed into simulations. Usually this process involves one or more of the following adjustments of the "host" database (the microsimuation) using data from one or more "donor" databases:

*Routine recoding and transformations.* This editing is necessary so that the codes on the survey variables will match what the simulation has been using in the past. Income, for example, must be in months in order to simulate the operation of welfare programs.

*Exact matches with other data bases.* In the U.S., these are most often thought of with reference to either Social Security or Internal Revenue Service files, but privacy concerns and other priorities for the outside agencies have meant that such matches have occurred only rarely. In this circumstance an identification variable would be used to match the record in the survey with the record in the outside database, and the variables in the outside database could be added as needed to the survey. Some exact matching is possible because of sample design, e.g., with the U.S. Current Population Survey, for which any given respondent is interviewed once a month for four months, and then again for the same time period one year later.

*Statistical matching.* Statistical matches link records from two databases where the databases contain different sets of respondents and different variables. In a statistical match, certain variables are common across the two surveys or databases. These become, in one method of doing statistical matching, the basis for a distance or difference function that is minimized across the two data bases. The observations that match the closest can then be merged, with the necessary variables from the donor database being added to the host database. This technique has been considered relatively expensive, especially with the large databases that are often used in microsimulations (the U.S. Current Population Survey includes over 57,000 households and 120,000 individuals monthly).

*Statistical imputation.* Statistical imputation is more commonly used to add data to the host database. A variety of techniques exist, ranging from the simple to the elaborate. In the former category are bivariate crosstabulations from the donor database producing a mean value for a third variable. The crosstabulations are then used to assign the mean value for the third variable to the host database. More elaborate methods use econometric techniques to estimate one or more variables that are then added to the host database.

Some imputations are done as a matter of course in processing survey data. The U.S. Census Bureau documentation for the Current Population Survey, for example, makes reference to "hot deck" imputation, used for income questions to "assign missing responses to sample persons with similar informaiton from matched sample persons with similar demographic and economic information who answered these questions" (U.S. Bureau of the Census, 1995, 2-3). And other recoding does not fit neatly into the above categories. The U.S.

Census Bureau, for example, assumes for the Current Population Survey that all children in a household with a householder or spouse covered by Medicaid were also covered by Medicaid, since the questionnaire does not ask specifically about Medicaid coverage for those under the age of 15. "All adult AFDC recipients and their children, and SSI recipients living in States which legally require Medicaid coverage of all SSI recipients, were also assigned [Medicaid] coverage" (U.S. Bureau of the Census, 1995, 10-8).

Thus, the key differences between a large-scale social survey and a microsimulation are two-fold: first, the microsimulation takes the survey data and adjusts those data so that the simulation can produce financial and participant estimates that are valid by other, external criteria. Second, the survey is useful for testing hypotheses and models from theoretical social science; the microsimulation has as its chief goal to produce financial and participant estimates for the policy process, for real-world decision makers who are going to make decisions tomorrow.

**VI. Issues in Microsimulations of Public Policy**

Several characteristics about microsimulations of public policy make them significantly different from the average social science enterprise.

First, they require a substantial institutional commitment to establish and keep up-to-date. This commitment can be considerably more than the average social science survey. The Office of the Assistant Secretary for Planning and Evaluation in the U.S. Department of Health and Human Services, for example, has supported the development and updating of the Transfer Income Model (TRIM2) at The Urban Institute for approximately 30 years, and the current level of support is in the range of a million dollars (US) per year. Other simulations have required similar levels of support, and many simulations over the years have had either certain modules or the entire model atrophy from lack of use or lack of an institutional sponsor.

Second, microsimulations of public policy have provided decision makers, chiefly at the national level, in many countries, with information that, for over 20 years, they have considered a "given" in the consideration of policy changes to tax and transfer income programs. Citro and Hanushek (1991, 24, 41-42) note that consideration of a substantial broadening of the Aid to Families With Dependent Children program in the US in 1987 ended altogether as a result of several factors, one of them being a microsimulation-based estimate of the extra cost that showed the extra cost would have been substantial. Over the years, decision makers have become used to obtaining--relatively quickly--the sort of detailed policy analyses that microsimulations can produce.

Third, understanding and using microsimulations of public policy involves a substantial amount of staff commitment and energy, so much so that many of the simulations are associated with the research groups or individuals who have developed and maintained them for long periods (more than a decade in most cases) of time. For individual researchers not associated with consulting organizations or these research groups, the opportunity to use and understand these microsimulations is just beginning to become available, and the change is a healthy one. Several microsimulations are now or will shortly be available for downloading or operation through the World Wide Web. I suspect the microsimulation community, which until now has been a fairly small circle of very applied researchers, is about to receive a substantial infusion of new blood. In the past, validation of microsimulations and non-policy, non-real world research on microsimulations has been limited by the priorities of sponsors. Now, more research that is *not* for tomorrow's policy appetite can be done and published.

Fourth, not only do researchers have to spend a substantial amount of time learning how to use microsimulations, agency decision-makers and staff have had the same problem. This situation will not change in the near future.

Fifth, the microsimulation community needs more outlets to publish its research, and the development of a World Wide Web based-journal just for research on these kinds of microsimulations would be a major step in the right direction. In researching this paper, I was struck by the lack of links from the web sites of each microsimulation or research group to the web sites of others. The social science simulation community has expanded immensely with the revival of interest in social science simulations in the 1990s; more interest is clearly being expressed in microsimulations of public policy as well.

## VII. Conclusions

For public managers, the lessons are significant:

1. Managers may be called upon to support research at a substantial level without understanding that research, without an important and signiicant policy payoff during the current fiscal year. Supporting the models over time--maintaining them when there are no immediate uses during the current fiscal year--is important. It is virtually impossible to take a neglected model after some years and estimate some parameter quickly, if only because the people who developed the model and knew it thoroughly will now have moved on to different endeavors.
2. Some microsimulation models can do state level estimates, something which will be even more importrant as states have more control over programs like TANF with the devolution revolution.
3. Public managers should be familiar with these kinds of models as there are other areas where they could be developed and have not been because of lack of knowledge as to their potential.
4. Managers should make a concerted effort to make these models, their data, their fundamental assumptions, and results available for outside users to use either directly, over the World Wide Web, or by downloading data and software. The more users, the stronger the user community, the stronger the disciplinary ties and links and the better quality of the research produced.

## References

Bailey D, Cohen L, Iams H. 1998. "Comparison of MIMEs Between CORSIM and SIPP Matched to Administrative Data." Presented at the Association for Public Policy Analysis and Management annual research conference, New York, October. http://qsilver.queensu.ca/appam/NY98/abs/bailey.htm

Conte R., Hegselmann R, Terna P, Eds. 1997. Simulating Social Phenomena. Berlin: Springer Verlag.

Gaylord RJ, D'Andria L. 1998. Simulating Society - A Mathematica Toolkit for Modeling Socioeconomic Behavior. New York: TELOS/Springer Verlag.

Giannarelli L, Moore M. 1997. Using TRIM3: An Introduction for Program and Policy Analysts. Urban Institute Report.

Giannarelli L. 1992. An Analyst's Guide to TRIM2, The Transfer Income Model, Version 2. Washington, D.C.: Urban Institute Press.

Gilbert N. 1995. "Simulation: An Emergent Perspective." Lecture given at the conference on New Technologies in the Social Sciences, October. Available at http://www.soc.surrey.ac.uk/research/simsoc/tutorial.html.

Gilbert N., Troitzsch K. 1999. Simulation for the Social Scientist. London: Open University Press. Available (partially) online at http://www.uni-koblenz.de/~kgt/Learn/Textbook/Book.html

Harding A. 1996. Microsimulation and Public Policy. Amsterdam: North Holland, Contributions to Economic Analysis 232.

Hegselmann R. Review of Gaylord and D'Andria (1998). Journal of Artificial Societies and Social Simulation, June, 1998. http://www.soc.surrey.ac.uk/JASSS/1/3/review1.html

Journal of Artificial Societies and Social Simulation. http://www.soc.surrey.ac.uk/JASSS/

Lewis GH, Michel RC, Eds. 1990. Microsimulation Techniques for Tax and Transfer Analysis. Washington, D.C.: Urban Institute Press.

National Research Council, Panel to Evaluate Microsimulation Models for Social Welfare Programs. 1991. Improving Information for Social Policy Decisions, The Uses of Microsimulation Modeling, Vol. I, Review and Recommendations. C. F. Citro and E. A. Hanushek, Eds. Washington, D.C.: National Academy Press. See also: http://www.nap.edu/bookstore/isbn/030904541X.html

Rivlin A, Wiener J. 1988. Caring for the Disabled Elderly, Who Will Pay? Washington, D.C.: Brookings.

Terna P. 1998. Simulation Tools for Social Scientists: Building Agent-Based Models with SWARM. Journal of Artificial Societies and Social Simulation. Vol 1, no. 2. March. http://www.soc.surrey.ac.uk/JASSS/1/2/4.html

Trippe C, Stavrianos M. 1998. "Microsimulation Modeling: A Powerful Tool for Estimating the Impact of Welfare Reform on the Food Stamp Program." Presented at the Association for Public Policy Analysis and Management annual research conference, New York, October. http://qsilver.queensu.ca/appam/NY98/abs/trippe.htm

Troitzsch KG, Mueller U, Gilbert GN, Doran JE. 1996. Social Science Simulation. Heidelberg: Springer Verlag.

U.S. Bureau of the Census. 1995. "Current Population Survey, March 1995. Technical Documentation, CPS 95. Washington, D.C.: Bureau of the Census.