Microsimulation General Introduction

Martin Spielauer
martin.spielauer@dms-c.com

Organization

- What is Simulation?
- What is microsimulation?
- Why and when does it make sense?
- Modern research perspective: life course paradigm
- Policy maker’s perspective: virtual world
- Data context
- Components
- Typical application fields
- Development context
- Limitations
- Technology
- Prospects

What is simulation?

Simulation = creation and use of models pursuing a purpose

- Exploration
- Prediction
- Problem solving
- Training
- Raising consciousness
- Theory building

- Many policy questions can only be answered using simulation

What is micro-simulation?

Computer-simulation of a society or economy in which the population is represented by a large sample of its individual members and their behaviors.

- Static: tax-benefit accounting
- Dynamic: behaviors over time

- Agent based simulation: theory driven
- Data driven microsimulation: based on statistical models

When does it make sense? 1/3

Population heterogeneity:
If individuals are different and differences are given importance.

- No single representative agent (or too many homogeneous groups)
- Research focus on distributions: e.g. Winners and losers of reform

When does it make sense? 2/3

If behaviors are more stable or better understood on micro level.

- Composition effects
- Non-linear tax and benefit rules formulated on micro level

When does it make sense? 3/3

If individual histories matter

- Modeling of Non-Markov processes: agents possess a memory of their past which shapes the future
- Social insurance: Individual contribution histories matter

Origin

First introduced in Social Sciences 1957 by Guy Orcutt: “A new type of socio-economic system”:

- Improves “limited predictive usefulness” of economic models concerning policy effects;
- Predicts not only aggregates but distributions;
- Use of knowledge about decision-making units;
- No restrictive assumptions of “absurdly simple relationships about elemental decision-making units” in order to be able to aggregate.
- “.. is intelligible to people of only modest mathematical sophistication”

A 1956 Computer

- 1956 IBM 305 RAMAC: first computer with a disk drive, the IBM 350 Disk File.
- The 350 Disk File consisted of a stack of fifty 24" discs. Total capacity: 4.4 MB
- Annual lease: $35,000

Advances

- Impressive improvement of computer power
- Availability of data: Longitudinal surveys, administrative data, etc.
- New research paradigms:
  - From macro to micro …to multilevel integration
  - Life course paradigm
- Wealth of analysis of individual behavior
  … still need of synthesis: Microsimulation
- New demands
  - Detailed projections
  - Longitudinal policy analysis

Life course perspective

- Agency: decisions and events modeled at level where they take place and in individual context
- Life-course interactions between life domains: education, work, family, health
- Linked lives: Interactions between individuals
  - Family and social networks
  - Transfers and transmission processes

Synthesis of life-course analysis bridging micro-macro gap

Policy makers' perspective

- Virtual world:
  - Test and fine-tune new policies in a virtual world
  - Identify winners and losers of reform
- Detail:
  - Detailed projections for planning purposes
  - Policy simulations at any level of detail
- Longitudinal perspective:
  - Sustainability issues in the context of demographic change
  - Study of distributional aspects both in cross-sectional way and over time: re-distribution over life-course, between cohorts and generations

Data perspective

Microsimulation creates data to answer policy questions which cannot be answered by information available in a single existing dataset

- Synthesis: Combination of information from various sources
- Projections into the future

Use of synthetic data

Many micro-simulations start from a cross-sectional starting population which is built by combining information from various sources

- Quality: combination of data sources can increase consistency
  - Taxes and benefits add up to administrative totals
  - Tails of distributions better covered as by surveys
- Accessibility: synthetic data have no confidentiality problems
- Relevance: more information in a single consistent dataset

Generation of synthetic data

Dynamic micro-simulation models create synthetic data of the future but can also model the past

- Projection:
  - Simulated future cross-sectional data
  - Simulated longitudinal data
- Imputation of past information
  - Microsimulation can be used to restore past data
  - Synthetic population models model whole life-courses from birth

Components

- A database of a population represented by individuals
  - Individuals have a list of characteristics
  - Individuals can be linked to families or households
- Parameter tables
- Models for the dynamic update of characteristics
  - Discrete time models: Updates at fixed steps of time
  - Continuous time models: driven by events which can happen at any moment in time: e.g. finding/losing a job, pregnancy, school enrolment, death.
- Policy and accounting rules
  - Taxes, contributions, benefits
- Table and micro-data output

Typical application fields

- Static tax-benefit models: standard tool available in many countries to identify winners and losers of tax reform and to calculate tax revenues
- Population projection models: detailed projections with fine-grained geography and socio-demographic characteristics
- Dynamic policy models: study of policies with a longitudinal component
  - Education: return to investments, student loans
  - Pension systems: sustainability and adequacy
  - Health: public health, treatment, finance
  - Care: health care, elderly care, care networks

Development context

- Population heterogeneity
  - Developing countries often display high levels of inequality
  - No single `representative agent`
  - Many policies target specific vulnerable populations
- Highly dynamic populations
  - Fast and complex socio-demographic changes
  - Composition effects overlay behavioral changes
  - Potentially strong demographic policy effects
- Path-dependencies and high policy impact
  - Decisions on policy designs have long-term consequences
  - Policy decisions have potentially important consequences on individual lives and life-course decisions

Development context: Applications

- Demographic projections
  - More detailed population projections for planning
  - Modeling of the potentially strong demographic effects of policies, e.g. educational investments
- Poverty and vulnerability profiling: who is mostly affected by crises, shocks or policies (e.g. trade liberalization)
- Decomposition analysis: better understanding which factors lead to differences between population groups and which factors drive change and development.
- Ex-ante evaluation of policies and cost-benefit analysis in more detailed context including down-stream effects
- Exploration of social insurance options (new systems)

Development context: Benefits

- Supports planning
  - More detailed projections
  - Early Identification of problems
- Supports policy and program design
  - Distributional analysis in various dimensions
  - Flexibility in measures: data can be aggregated in many ways allowing output of a variety of indicators
  - Flexibility in accounting: private versus social return
  - Ex-ante evaluation of policies and cost-benefit analysis in more detailed context including down-stream effects
- Supports policy decisions by better understanding of options
- Supports validation and monitoring of programs

Limitations

- More detailed projections come at the price of higher randomness lowering the prediction power of results.
- Limited theoretical foundation and “mechanical” statistical modeling (without market response) limit economic projections

-> More complex applications frequently require combining strengths of various approaches by aligning some results to macro projections or by combining microsimulation with macro models.

- High data demands: fast progress incl. in synthetic data generation and availability of administrative data.
- Price tag of implementation: recent availability of efficient software and computational power make microsimulation more accessible

Technology

The recent availability of powerful programming technologies make microsimulation increasingly feasible

Examples:

- LIAM: (open source) language for discrete time models
- Modgen: (developed at Statistics Canada) fully generic and especially powerful for continuous time models
- openM++: open source implementation of Modgen

Prospects

Microsimulation is a logical next step in data analysis increasing the policy relevance of data and bridging the micro-macro gap in research.

Increasing demand for:

- detailed projections
- distributional analysis
- longitudinal analysis modeling processes over time
- What-if analysis for policy planning and evaluation