

# The ESRC-Sage Model

A dynamic microsimulation model.

Base data derived from 1991 SARs.

Individuals are aged each year.

Transitions:

Demographic - mortality, fertility,  
partnership changes, health;

Labour Market - employment &  
earnings;

Pensions and benefits in  
retirement.

The ESRC-Sage team is studying the lifecourse and developing a dynamic microsimulation model to project future experiences in later life, with particular emphasis on income in retirement and long-term care.

Our model is a full population model. We start with a representative sample of the population of Great Britain based on the 1991 Census. Each year, the model ages each individual, changing characteristics such as partnership status and employment status in accordance with transitions observed in panel studies, and accumulating histories such as pension contributions.

The model will be an important tool for policy makers who wish to examine the likely effects of population changes, economic scenarios, and various policy regimes on income and care in later life. It could also be adapted to study other aspects of the lifecourse, such as the impact of different forms of student financing on family finances during the child-rearing years.

# Data Requirements

Representative and comprehensive base data, including histories.

Transition probabilities - often history/duration dependent.

Rules for pension schemes and benefits.

Problems:

- No single source of all data;

- Biases in estimating transitions from different data sets;

- Inadequate sample sizes for full models;

- Inadequate panel lengths for history dependence;

- Unpredictable period and cohort effects.

Dynamic microsimulation has heavy data requirements. We need representative and comprehensive base data, which include the historical attributes used in the model. We also need to be able to estimate transition probabilities that accurately represent the variety of trajectories followed by individuals.

In order to analyse the predictors of, for example, future incomes, we need rich longitudinal data on labour market participation and earnings covering the whole of working life. Ideally, all data would come from one vast source, and we could do an unbiased joint estimate of all transition probabilities.

However, this is not possible and we need to combine data sources. This can lead to biases in estimating interdependent transitions from different data sets. We also find that sample sizes may be too small to estimate detailed models that incorporate sufficient population heterogeneity. Often panel lengths are inadequate to estimate the full influence of history on transitions.

Even when adequate data are available, they relate to particular cohorts and periods. It is helpful to be able to compare different cohorts and periods, but future cohort and period effects are essentially unpredictable.

# Sources Used

1991 SARs for base data, with imputations from BHPS, QLFS, FWLS.

BHPS for demographics & earnings.

ONS LS & GAD for mortality.

QLFS for employment & health in working life.

Pensions and benefits based on typical schemes and scenarios.

Other studies important for theory and background.

Our base data are taken from residents of households and institutions in the 1991 SARs, in order to get a fully representative sample of the GB population. This was felt to be important because biases in representativeness could tend to be magnified over time in the simulation. Variables not available in the SARs were imputed from other sources.

The BHPS proved to be the best source for most of the transition probabilities required. It has rich data, including some retrospective histories, covers all ages, and allows annual transitions to be directly observed. However, the sample sizes for important subgroups are small, the panel is subject to attrition and as yet covers only a fraction of each lifecourse, and retrospectively collected data are less reliable.

We used registration-based sources for mortality, and preferred the Quarterly Labour Force Survey 5-quarter longitudinal datasets for labour market transitions, since these have a very large sample size and could better capture frequent status changes.

I should also mention that numerous other studies have been important to the project in informing our decisions on model processes through empirical and theoretical work.

# Key Issues

Maintaining panel and cohort studies.

Comparability of questions and responses between and within studies.

Sample size.

Time series are also useful.

Accessibility of micro-data.

Smaller, qualitative studies are important to inform statistical analysis.

The most important issue for us is the maintenance of existing panel and cohort studies. The longer a panel, the more useful it is for lifecourse studies.

Changes in question frequency and wording have limited the number of years of surveys that can be used. However, the changes have usually been improvements.

Sample size is often a restriction on the detail of analysis that can be performed. Any augmentation of the BHPS would be very welcome.

Time series obtained from repeated cross-sectional surveys and administrative data can be valuable for evaluating and aligning the results of longitudinal analysis.

For micro-simulation modelling, access to micro data is essential. For example, access to the DWP's LLMDB would have been very useful to us. Access only in a secure environment is not adequate for data to be used as base data, and arrangements would have to be made to be able to export derived datasets that did not pose a disclosure risk.

Finally, although major funding should be directed to maintaining the main large studies, the role of smaller, qualitative studies in informing statistical analysis should not be neglected.