The simulation properties of microsimulation models
with static and dynamic ageing – a brief guide
into choosing one type of model over the other

Gijs Dekkers
Federal Planning Bureau, Belgium
Kunstlaan 47-49, 1000 Brussel
CEPS/INSTEAD, Luxembourg and
CESO, KU Leuven, Belgium
gd@plan.be

ABSTRACT: To assess possible distribution effects of alternative scenarios, including hypothetical future states, one can use either static ageing techniques, which age the population by reweighing and uprating, or dynamic ageing, which alter the relevant population by applying deterministic probabilities that a certain event may or may not occur. This paper makes the argument that, even though the two methods are technically completely different, they are not unlike in terms of their simulation properties. Starting from the thesis that under theoretical circumstances, both approaches are equivalent in terms of their simulation properties, the choice between the two archetypes of models comes down to assessing how far the actual and theoretical circumstances differ from each other. By highlighting the differences and resemblances between static and dynamic microsimulations in terms of their simulation properties, this short note will contribute to the debate in choosing between these two types of models, and can thus serve as an advice piece for someone contemplating the development of a microsimulation model.

KEYWORDS: dynamic ageing, static ageing, microsimulation

JEL classification: C63
1. INTRODUCTION

To make a projection of distribution effects of alternative realities, for example representing a future state of the starting dataset, or alternative fiscal or social measures, one needs models that can capture the impact of changes in policy parameters and/or changes in the behaviour of individual economic agents (individuals, households). This capturing can be done by either changing the characteristics of individual agents, or reweigh the contribution of unchanged agents to the aggregate result.

Thus, one can distinguish ageing techniques within the set of microsimulation models. To decide what kind of technique is best brought in action in a model, one should understand their simulation properties, properties that might be more alike than the technical properties of the two approaches suggest.

This note compares the two archetypes of discrete-step ageing techniques and investigates some of their simulation properties. It argues that, even though the two archetypes are technically very different, under theoretical circumstances, both approaches are equivalent in terms of their simulation properties. This implies that anyone having to decide whether to develop a model with static or dynamic ageing should take the first option, as it is the cheaper alternative. However, real and theoretical circumstances may differ. This paper argues that the choice between developing a model with dynamic or static ageing requires the assessment of how far the actual and theoretical circumstances differ from each other.

This conclusion, though stated differently, is by itself not new (e.g., O’Donoghue 2001, p. 9). This paper adds to the literature by raising a number of questions that should be addressed before taking the investment decision on the type of microsimulation model to be constructed. It may thus serve as an advice piece for someone contemplating the development of a dynamic microsimulation model.

In what follows, the discussion will be on using the two ageing techniques to mimic time, i.e. to have the microsimulation model arrive at a ‘future’ state. This, however, is merely for purposes of discussion because all the issues to be discussed could also be presented in relation to the difference between an ‘actual’ state (reflected by the starting dataset) and a ‘desired’ state.

2. SIMULATION PROPERTIES

Whether a microsimulation model applies dynamic or static ageing is a technical characteristic.
These technical characteristics obviously are very different. However, the question is to what extent they affect the 'simulation properties' of a model: the actual or potential research problems that a model can cover, as well as the implicit or explicit assumptions that a model makes when handling a specific research problem (Dekkers and Legros, 2006).

The question therefore is under what circumstances the different technical characteristics of the two archetypes may result in different simulation properties. This paper considers a number of possible or perceived problems which might cause one to choose a model with dynamic ageing over static ageing.

3. A BIRD’S EYE VIEW ON AGEING TECHNIQUES

The conceptually simplest ageing technique is “static-ageing”. A model that uses static ageing simulates time indirectly through “uprating” and “reweighting”. Uprating is the process through which monetary values are inflated or deflated to meet exogenous projected developments.

Suppose that individuals are described by \( n \) variables. At any point in time \( t \), an individual is a combination of characteristics of the \( n \) variables. By changing the weights of the individuals in the dataset, models with static ageing change the combination of characteristics, not the characteristics themselves. So the individuals in the dataset ideally are weighted by a joint \( n \times n \) reweighting matrix to transform the base dataset to the future year \( y \). However, one often does not have a \( n \times n \) reweighting matrix, especially as \( n \) gets large. In most cases, one only has a set of marginal distributions. The problem therefore is to empirically derive a new vector of sample weights that are as close as possible to the design weights while bringing the dataset as close as possible to the marginal target distributions (for a formal definition of the problem, and a solution for various distance measures, see Cai et al. (2005; section 2) and Creedy & Kalb (2006)).

A number of packages are available to solve this problem in practice. The German model MIKMOD-Est (Flory and Stöwhase, 2012) uses a package developed by Quinke, which the authors report to be quite similar to Merz’s “Adjust” reweighting algorithm (Merz, 1994). The GREGWT algorithm, developed by the Australian Bureau of Statistics (ABS) for spatial microsimulation, is used to reweight surveys to Australia wide and capital city benchmarks. NATSEM uses it for its SpatialMSM model (See Tanton and Vidyattama, 2011). Alternatives are CALMAR, developed by INSEE in France (INSEE, 2011), and Clan97 (Anderson and Nordberg, 1998, in Immervoll et al., 2005, 6). All three algorithms are developed in SAS. Algorithms developed in Stata include CALIBRATE (D’Souza, 2011), Reweight (Pacifico, 2011, in: De Blander et al., 2013, 8; revised as
SReweight, Pacifico, 2014) and Gomulka (1992, in: Brewer et al., 2011). Buddelmeyer et al. (2012, 43) use the approach by Cai (op. cit.) in a two-step reweighting process to adapt to demographic totals and employment and unemployment levels as provided by a CGE model. Finally, combinatorial optimisation (CO; Williamson et al., 1998) is an alternative approach that to our knowledge has not yet been used to ‘dynamize’ a static model or dataset. It involves selecting an appropriate combination of households from survey data to attain the known benchmark constraints. See Tanton and Vidyattama (2011) for a discussion and comparison with GREGWT.

In contrast to models with static ageing, models with dynamic ageing alter the contents of the dataset itself. So, “Individuals are aged and stochastically undergo transitions, as well as being subject to modified policy regimes” (Brown and Harding (2002) section 2.2). The data load of these models is considerably heavier than the one for models with static ageing. Since the states over time are interconnected in models with dynamic ageing, the data generating process is also more complicated than in the case of static ageing. In the latter case one lays down a predefined final distribution where the data are aligned with2. In dynamic ageing one also models some processes that generate the final distribution, like who marries whom or who gives birth to how many children. At the same time, the complication of modelling these intermediate processes implies the richness of models with dynamic ageing over those with static ageing.

Furthermore, since discrete-time3 models with dynamic ageing require the generation of intermediate results for each time state, the modeller can choose to sum past and current values of monetary variables into “life-time variables”. This way, it is possible to simulate the impact of fiscal or social policy on lifetime income of subsequent generations (see Creedy, 1999, and Nelissen, 1994, for applications). Though it is technically possible, this simulation of lifetime income is problematic in models with static ageing. First of all, and contrary to dynamic ageing models, it is conceptually meaningless to add up the incomes of the same individual in two years. This is because the observed individual represents two different individuals in two subsequent years. An individual observed with age 50 in the starting dataset of the year 2000 can be brought to, say, 2020 using static ageing techniques. But then the same individual represents an individual who is 50 in 2020, and who therefore was only 30 years of age in 2000. Hence the individual with the same identification number in two years conceptually is a completely different individual.

An obvious solution by which one could construct lifetime income in models with static ageing is by not following individuals but cohorts over time. Thus, we would add up the income of individuals aged 30 in 2000 with those aged 40 in 2010, with those aged 50 in 2020, and so on. This
would conceptually be possible but technically meaningless as one wants to simulate lifetime incomes to get rid of cross-sectional variation in earnings. However, by this approach one in fact introduces cross-sectional variation in the construction of lifetime income, since the incomes of subsequent observed cohorts—albeit reweighed and uprated—are added up. So the resulting lifetime incomes will not be better than the cross-sectional incomes. They both will suffer from cross-sectional variation so that their added value is limited. As static ageing models cannot produce credible estimates of lifetime income, the comparison will thus necessarily be limited to cross-sectional results for single years.

So far we have limited ourselves to identify some technical characteristics and basic simulation properties of microsimulation models with static and dynamic ageing. Which of these models is best brought in action to compute distributional indicators, depends on the question one wants to address.

4. SIMULATION PROPERTIES OF MODELS WITH STATIC AND DYNAMIC AGEING

In making abstraction from macroeconomic developments, models with static ageing mimic exogenous future demographic and labour market circumstances by reweighting the dataset. Models with dynamic ageing change the characteristics of the individuals in the dataset and do not consider the weights. In ideal circumstances, both models may be equivalent in the output they produce and in the range of policy assessments they can handle. This potential equivalence can be shown by a simple experiment of thought. Suppose a model with dynamic ageing that simulates all \( n \) individuals from a base-dataset in the year \( x \) to the future year \( y \), and suppose that individuals are described by a limited number of \( v \) variables. Then one can always find a joint \( n \times v \) reweighting matrix by confronting the simulated dataset at \( y \) with the base-data at \( x \). So, under these ideal circumstances, a model with static ageing can replicate the dataset \( y \) created by the model with dynamic ageing, by reweighting \( x \). Restated in the terminology of this paper: under ideal circumstances, the simulation properties of both models are the same, even though their technical characteristics are very different. Then the obvious choice would be to develop a model with static ageing and not a model with dynamic ageing, since the former is considerably cheaper in terms of data requirements, development and maintenance time, and effort.

However, circumstances often are not ideal. In this case, results from models with static and dynamic ageing may differ in such a way that one might want to prefer a model with dynamic ageing over static ageing. In the next paragraphs, possible problems with models with static ageing
will be discussed. The conclusion will be that these problems are limited or absent—so that one
may indeed choose to develop a model with static ageing—as long as the world at the future moment
in time \( t \) resembles starting dataset by close enough. Or, the problems with models with static
ageing remain limited as long as one does not simulate too far into the future.

This conclusion is by itself not new (see e.g. O’Donoghue, 2001, p. 9). This paper adds to the
literature by conceptually comparing the simulation properties of both archetypes of models. There
are various reasons why one might be tempted to choose to develop a model with dynamic ageing
over one with static ageing, notwithstanding the fact that the former is considerably more complex
and expensive in terms of development and updating. We deal with these arguments in more detail
in the remainder of this section. Each argument starts by describing a possible or perceived
problem that models with static ageing may have; next, we check whether this is actually a problem,
and—if so—whether it is possible to circumvent it.

4.1. New individuals

An obvious problem with models with static ageing is that they require that individuals with all
possible combinations of variables that will be present in the future, should already be present in
the base-dataset. In other words, if a model is to reweigh an individual to represent future
individual(s), the individual should exist in the starting dataset in the first place. If this is not the
case, then the reweighting procedure will fail. A first and most obvious reason for specific
dividuals not being present in the starting dataset is that they should be there, but, due to sampling
error, are not. Another reason might be that they should not even be present in the base dataset.
Indeed, it seems imaginable that behaviour has changed in such a way that those combinations of
variables that may arise in the future do not exist today. This becomes clear with an obvious
retrospective sample. Today, plenty of people of all age-categories cohabit without being married.
This is something which has arisen during the last decades. If one would have used a dataset of—
say- 1950, one might not have found cohabiting individuals. Immigration may also bring individuals
with unprecedented characteristics into a country. The problem of not having relevant individuals
in the base dataset increases as simulation is done further in the future.

Fortunately, this problem could be overcome by using Monte-Carlo techniques to create fictitious
individuals who have the yet unobserved combination of variables, or else to select actual
individuals from another data source, and include them in the starting dataset. The weight of these
individuals in the starting dataset obviously would be zero, and would only become positive when
exogenous constraints or hypothesis make these individuals emerge. I did not find applications of
this in the literature, so the practical consequences and problems with this approach are yet unknown.

4.2. The simulation of ‘future retrospective information’

Static ageing models ignore the processes that generated individual observations in the prospective sample (Zaidi and Rake, 2001, 2). This section and the next discuss the consequences of this technical property. Many countries have a Bismarckian social-security system, with benefits being a direct or indirect function of the previous career. Past information may affect future benefits in two ways:

- The relation between career characteristics (earnings and employment ‘track record’) may change. We call this “Bismarckian changes”. For example, the career length required for a full pension may increase over time.
- One shot events (cohort-specific effects such as political or demographical events) may affect career characteristics.

Discrete-time models with dynamic ageing simulate all intermediate steps between the starting dataset and the target future year $y$. Thus, current earnings in any year $t < y$ become retrospective earnings in $t + 1 \leq y$ and pension benefits are calculated for those entering retirement. Assuming that the relation between the fundamental variables (such as age and gender) and career variables (earnings, employment) remains unchanged, this makes models with dynamic ageing less vulnerable for one-shot events. Moreover, these models by definition allow for the inclusion of Bismarckian changes.

In its simplest form, static-ageing models like the Belgian model STATION (Dekkers, 2000) do not include retrospective information. These models therefore do not allow for the simulation of Bismarckian changes, and ignore the impact of past one shot events. But there is a straightforward solution for the first source of bias: if the starting dataset does include retrospective information tied to these observed benefits, one can use a static microsimulation model to change the way in which past earnings affect current pensions, and use static ageing techniques to bring the current simulation results into the future. A combination of a static model and static ageing techniques therefore does the trick. First, one uses uprating to bring individual earnings as well as the parameters of the pension system to the future year. Next, one uses the static model (including uprated parameters and the necessary technical changes) to replace the currently observed pension benefits by uprated-simulated pension benefits. Finally, the sample is reweighted to take into
account demographic and/or labour market trends between today and the future year. Note that the uprating-reweighting order is important here, because if uprating is not applied before the static model, then the combination cannot take into account the impact of differences between the development of earnings and the parameters of the pension system; a key element in the development of pension inequality (Dekkers, 2014)\(^7\).

Finally, none of this however solves for the vulnerability of models with static ageing to the possible impact of past one-shot events.

### 4.3. The simulation of a time-dependent policy measure

How do both archetypes of models handle a situation where a policy measure in itself is dynamic? The example discussed here is the gradual adjustment of the age of mandatory pension, a policy measure taken in various European countries and that may affect whether employees older than the mandatory retirement age can be observed to be working in a future year.

Formally, suppose the mandatory retirement age at time \(x\), \(a_x\), and suppose a dummy \(\text{ret} = [0,1]\) denoting whether the individual is mandatory retired or not. The condition then is \(p(\text{ret}|\text{age} \geq a_x) = 1\).

Now make this mandatory retirement age dynamic in that it can either increase or decrease by \(z\) years between \(x\) and \(x+1\). In other words, \(a_x \leq a_{x+1}\) or \(a_{x+1} \leq a_x\). In the case of models with dynamic ageing, the implementation is straightforward. Indeed, \(p(\text{ret}|\text{age} \geq a_{x+1}) = 0\) and \(p(\text{ret}|\text{age} + z \geq a_{x+1}) = 1\), with \(z\) any positive or negative discrete number. In the case of an increasing mandatory retirement age, this can be interpreted as that somebody of age \(a\) was eligible for retirement at \(x\), but no longer at \(x+1\).

In the case of a model with static ageing, the situation is analogous in that the above \(p(.)\) must be multiplied by a (current or future) weight \(w_x\) of the individual in the dataset. As a result, the weight of individuals of age \(a_x\) who meet the requirement in the base year \(x\) but no longer at \(x+1\) becomes zero. So, in the case of an increasing mandatory retirement age, retired individuals of age \(a_x\) with weight \(w_x(p(\text{ret}) = 1) = w\) in the base year \(x\), in the year \(x+1\) receive weight \(w_x(p(\text{ret}) = 0) = 0\). The weight of those working at age \(a_x\) is obviously adjusted upwards accordingly. In the opposite case of a decreasing mandatory retirement age, the weight of those working at \(a\) in the base year is multiplied by \([1 - (p(\text{ret} | \text{age} + z \geq a_{x+1}) = 1)] = [1 - 1] = 0\) in the year \(x+1\). The weight of those already retired before the mandatory retirement age is adjusted upwards accordingly. This immediately reveals when the model with static ageing will fail. If there are no observed individuals retired before the mandatory retirement age in the base dataset, the only way to compensate for the zero weights of the working
individuals is by creating them beforehand (see section 4.1) in the ‘observed’ dataset. Furthermore, this is an often trivial condition, certainly in the European context.

The conclusion therefore is that –save for one trivial condition for which an easy solution exists- the two archetypes of models seem equivalent in their simulation properties. There seems to be no a priori reason why one should choose one over the other.

5. CONCLUSION

The central thesis of this paper is that models with dynamic and static ageing might be more equivalent in terms of simulation properties than their technical differences suggest. The simulation properties of the two “archetypes of models” may under ideal circumstances be the same, but this may not be the case in practice. Next, we present and discuss several possible instances in which differences occur, and argue for some that a solution is available and may even be easy to implement.

Overall, we conclude that if the number of dimensions that have to be modified to capture ‘the future’ is limited, if future types of individuals and/or households are present in today’s dataset or can be simulated –in short, if the current population is an approximation of the future population- then the simulation properties of a model with static ageing are close to those of a model with dynamic ageing. Developing a model with static ageing then seems the better alternative.

If this assumption cannot be made, due to ‘one-shot events’ in the past or important societal or economic changes, then models with dynamic ageing might be preferable. This is certainly the case if one wishes to assess the impact of social policy on lifetime income of subsequent generations.

Another more practical argument that is relevant in choosing between one of the two archetypes is whether or not one has a arithmetic microsimulation model from the outset. If so, then this may be an argument in favour of choosing a model with static ageing, since the former may be used to overcome a possible handicap that the latter have in the simulation of time-dependent policy measures.

The ultimate choice rests with the developer, who should take into account actual and potential research problems and/or questions by policy makers. But the fundamental choice described in this paper deserves attention, not in the least because developing a dynamic microsimulation model is a very expensive project that will take several years. The cost of taking a poor decision is therefore important. This paper hopes to contribute to a deliberate choice, preferably before one starts

Dekkers, The simulation properties of microsimulation models with static and dynamic ageing – a guide into choosing one type of model over the other
developing.

ACKNOWLEDGEMENT

The author wish to thank Guy van Camp and Boris Majcen and two anonymous reviewers for their comments on a previous version of this paper. I also thank Cathal O'Donoghue for taking up the role of journal editor for an earlier and extended version, as well as for this note version of the paper.

REFERENCES


De Blander, Rembert, Ingrid Schockaert, André Decoster and Patrick Deboosere (2013) The impact of de la croissance...
of demographic change on policy indicators and reforms. FLEMOSI discussion paper 25, September 2013


Immervoll, Herwig, Klas Lindström, Esko Mustonen, Marja Riihelä and Heikki Viitamäki (2005), Static data “ageing” techniques. Accounting for population changes in tax-benefit microsimulation models. EUROMOD working paper, No EM7/05.

Liégeois, Philippe and Gijs Dekkers (2014), ‘Combining EUROMOD and LIAM tools for the development of dynamic cross-sectional microsimulation models: a sneak preview’. In: Gijs


Note that this discussion is about the ageing techniques. The notions of ‘static’ and ‘dynamic’ models are carefully avoided here, because of the inconsistency of this classification with ageing techniques (De Blander et al., 2013, 6).

This of course requires that the prospective auxiliary data is available in the first place (Zaidi and Rake, 2001, 2).

As one of the reviewers points out, a model that uses survival analysis, often in continuous time, built around a competing risk approach do not need to produce intermediate results either. See, for example, applications of the Lifepaths model (Statistics Canada, 2013) for the construction of lifetime incomes.

Note that the assumption that individuals are described by $v$ variables prevents bias in characteristics that are not controlled for (Immervoll et al., 2005, p. 9).

Of course, a vulnerability remains in that past events affect the data underlying the behavioural equations and transition tables in the dataset. But this problem is fundamental to all models and is therefore less relevant in making a choice between ageing techniques for microsimulation purposes.

Given that static ageing does not change the dataset but only the weights, this implies that these unknown one-shot events become perpetual in their impact on the simulation results for future periods.

Immervoll et al. (2005, footnote 6, p 5) argue that this same order is also relevant if the monetary variable is used to define subgroups that are relevant to the weighting process.

Some recent research (Tedeschi, 2011; Liégeois and Dekkers, 2014) however shows that combining a arithmetic model with a dynamic-aging model is also possible. Furthermore, development software such as LIAM2 now allow to easily move arithmetic modules back and forth between the ‘init’ and prospective phase of a model (Bryon et al., 2014).