WHAT ARE THE CONSEQUENCES OF THE AWG-PROJECTIONS FOR THE ADEQUACY OF SOCIAL SECURITY PENSIONS?

GIJS DEKKERS, HERMANN BUSLEI, MARIA COZZOLINO, RAFAEL DESMET, JOHANNES GEYER, DIRK HOFMANN, MICHELE RAITANO, VIKTOR STEINER, PAOLA TANDA, SIMONE TESCHI, FRÉDÉRIC VERSCHUEREN

ENEPRI RESEARCH REPORT NO. 65
AIM WP4

JANUARY 2009

ENEPRI Research Reports publish the original research results of projects undertaken in the context of an ENEPRI project. This paper was prepared as part of the Adequacy of Old-Age Income Maintenance in the EU (AIM) project – which has received financing from the European Commission under the 6th Research Framework Programme (contract no. SP21-CT-2005-513748). The views expressed are attributable only to the authors and not to any institution with which they are associated.

Available for free downloading from the ENEPRI website (http://www.cnepri.org) or the CEPS website (www.ceps.eu)
© Copyright 2009, Gijs Dekkers, Hermann Buslei, Maria Cozzolino, Raphael Desmet, Johannes Geyer, Dirk Hofmann, Michele Raitano, Viktor Steiner, Paola Tanda, Simone Tedeschi, Frédéric Verschueren
What are the consequences of the AWG-projections for the adequacy of social security pensions?

An application of the dynamic micro simulation model MIDAS for Belgium, Italy and Germany

Report of the Work Package 4 of The AIM project

September 2008

Gijs Dekkers\textsuperscript{1}, Hermann Buslei\textsuperscript{2}, Maria Cozzolino\textsuperscript{3}, Raphael Desmet\textsuperscript{1}, Johannes Geyer\textsuperscript{1}, Dirk Hofmann\textsuperscript{2}, Michele Raitano\textsuperscript{1}, Viktor Steiner\textsuperscript{2}, Paola Tanda\textsuperscript{1}, Simone Tedeschi\textsuperscript{2}, Frédéric Verschueren\textsuperscript{1}.

\textsuperscript{1} Federaal Planbureau (Federal Planning Bureau), FPB, Belgium. Contact: Gijs Dekkers, Federal Planning Bureau, Kunstlaan 47-49, 1000 Brussels, Belgium. Email: gd@plan.be. Tel. (02) 5077413.
\textsuperscript{2} Deutsches Institut für Wirtschaftsforschung (German Institute for Economic Research), DIW, Germany.
\textsuperscript{3} Istituto di Studi e Analisi Economica (Institute for Studies and Economic Analysis), ISAE, Italy.

Jel Classification - J14, D31, I32

Keywords - adequacy, pensions, microsimulation.
Executive Summary

Europe faces important demographic changes in the coming decades, changes that will have economic and budgetary consequences. The Economic Policy Committee (EPC) established the Ageing Working Group (AWG), one of whose tasks it is to assess the long-term sustainability of public finances in the long term. It does so by presenting a set of public expenditure projections for all member states, including on pensions. These projections are based on demographic forecasts provided by Eurostat and agreed assumptions on key economic variables.

To date, the projections that member states produce for the AWG include only a limited notion of pension adequacy, being the replacement rate. Other aspects, including the poverty risk among the elderly, and income inequality, are not considered. The assessment of adequacy of pensions is the work of the Indicator Subgroup (ISG) of the Social Policy Committee (SPC).

Even though the sustainability and adequacy of pensions are two sides of the same coin, the work of both committees is separated. This project aims to set a first step into integration by assessing the consequences of the AWG-projections and assumptions on the adequacy of pensions.

In the context of a European-funded sixth framework project called AIM, a dynamic microsimulation model MIDAS is being developed for Belgium, Germany and Italy. This is a joint effort by three institutions, the German DIW, the Italian ISAE and the Belgian FPB. This model simulates future developments of the adequacy of pensions, following wherever possible the projections and assumptions of the Ageing Working Group.

This paper starts by describing the model MIDAS in detail. It next presents and discusses some simulation results for Belgium, Germany and Italy. Finally, the simulation results of two alternative policy scenarios are presented and discussed.
Contents

1. Introduction and overview of this report ............................................................................................................. 1
   1.1. Ageing and the sustainability and adequacy of pensions ................................................................. 1
   1.2. Why microsimulation, and why MIDAS? .................................................................................. 2
   1.3. LIAM and its alignment properties ............................................................................................. 2
   1.4. An overview of MIDAS ................................................................................................................... 3
   1.5. Results for Belgium, Germany and Italy ....................................................................................... 3
   1.6. Acknowledgements ......................................................................................................................... 4
   1.7. References ......................................................................................................................................... 4

2. A classification and overview of micro simulation models, and the choices made in MIDAS .... 5
   2.1. Introduction ....................................................................................................................................... 5
   2.2. What are the ‘simulation characteristics’ of a model? .................................................................... 6
   2.3. Micro Simulation Models versus other simulation models .......................................................... 8
   2.4. A classification of microsimulation models .................................................................................... 11
         2.4.1. Static versus dynamic microsimulation models ................................................................. 11
         2.4.2. A classification of dynamic models .................................................................................... 13
   2.5. A focus on some relevant models ...................................................................................................... 15
   2.6. Micro simulation in the AIM-project: what kind of model is MIDAS, and why? ....................... 18
         2.6.1. The first objective: the prevention of old-age poverty ......................................................... 19
         2.6.2. The second objective: the preservation, at retirement, of a standard of living comparable to 
               that of the final part of the active life ......................................................................................... 20
         2.6.3. Finally: the simulation of demographic ageing and pensions ......................................... 20
   2.7. Conclusions ....................................................................................................................................... 22
   2.8. References ......................................................................................................................................... 23

3. A short overview of the MIDAS-version of The Life-Cycle Income Analysis Model (LIAM) .... 25
   3.1. Introduction ....................................................................................................................................... 25
   3.2. Objectives ......................................................................................................................................... 25
   3.3. Framework Features ....................................................................................................................... 26
         3.3.1. Structure of Framework ......................................................................................................... 26
         3.3.2. Data and Framework Data Structure .................................................................................... 27
         3.3.3. Linkages .................................................................................................................................... 28
         3.3.4. Parameterisation .................................................................................................................... 29
   3.4. Process Modules ............................................................................................................................... 31
         3.4.1. Transition Matrices (trap) ....................................................................................................... 31
         3.4.2. Regressions (regr) .................................................................................................................. 31
         3.4.3. Transformations (tran) .......................................................................................................... 32
         3.4.4. Marriage Market .................................................................................................................... 33
   3.5. Alignment .......................................................................................................................................... 34
   3.6. Conclusions ....................................................................................................................................... 36
   3.7. References ......................................................................................................................................... 37
   3.8. Appendix 1 ....................................................................................................................................... 38
4. The MIDAS model for Belgium, Germany, and Italy .................................................. 43
  4.1. Demographic module 43
    4.1.1. Overview 43
    4.1.2. Behavioural equations for Belgium 50
    4.1.3. Behavioural equations for Germany 53
    4.1.4. Behavioural equations for Italy 56
  4.2. Labour market module 58
    4.2.1. Overview 58
    4.2.2. Behavioural equations for Belgium 59
    4.2.3. Behavioural equations for Germany 73
    4.2.4. Behavioural equations for Italy 87
  4.3. The pension module 97
    4.3.1. The Belgian pension module 98
    4.3.2. The German pension module 101
    4.3.3. The Italian pension module 103
  4.4. References 107

5. The Simulation Results .......................................................................................... 108
  5.1. Belgium 108
    5.1.1. Demographics 108
    5.1.2. Labour market states and earnings 117
    5.1.3. Retirement 127
    5.1.4. Inequality, poverty and the adequacy of pensions 141
    5.1.5. Alternative scenarios 164
    5.1.6. Conclusions 182
  5.2. Germany 183
    5.2.1. Demographics 183
    5.2.2. Labour market states and earnings 192
    5.2.3. Retirement 208
    5.2.4. Inequality, poverty and the adequacy of pensions 218
    5.2.5. An alternative scenario 231
    5.2.6. Conclusions 234
  5.3. Italy 235
    5.3.1. Demographics 235
    5.3.2. Labour market states and earnings 243
    5.3.3. Retirement 251
    5.3.4. Inequality, poverty and the adequacy of pensions 260
    5.3.5. Alternative scenarios 268
    5.3.6. Conclusions 271
  5.4. References 272

6. Conclusions ........................................................................................................... 274
  6.1. References 277

7. Appendix 1: A short description of the AWG-projections ...................................... 279
  7.1. The AWG population scenario 279
  7.2. Participation, employment and unemployment rate assumptions 281
  7.3. Assumptions on productivity 283
  7.4. Other assumptions 284
  7.5. Results: sustainability of pensions in Belgium, Germany and Italy. 284

8. Appendix 2: Full description of country-specific pension systems .......................... 286
  8.1. Full Description of the Belgian pension system 286
    8.1.1. The wage-earners retirement schemes 286
    8.1.2. The civil servants retirement schemes 292
    8.1.3. The self-employed retirement scheme 297
2. A classification and overview of micro simulation models, and the choices made in MIDAS

Gijs Dekkers, Michele Belloni

2.1. Introduction

As politicians have become more aware of possible consequences of demographic ageing, there has been a growing need for policy support and the evaluation of (potential) measures in pension policy. As a consequence, many types of models have been developed or have had a new lease of life in the research on pensions and pension systems. These types of models include overlapping generations models, so-called ‘standard simulation models’, option value models, and various other types of models.

As a result of this growing analytical tool-box, models with quite different simulation characteristics are often used to address a common set of research problems. Even for the specialist reader, it is often difficult to see what the consequences are of choosing one type of model over another, or to foresee how a model, once developed, can be expanded to cover new research problems in the future. This is all the more relevant, because developing a new model in this field typically involves several years’ investment. As a result, model developers cannot remain idle until a politician comes by with a question for which a model is needed. Instead, public research agencies typically try to anticipate politician’s future questions, and invest in the development of such a model, designed to cover the largest range of potential questions and problems. To make this choice, it is imperative that one has an understanding of the fundamental characteristics of various models available, including an appreciation of their respective pros and cons, and what kinds of questions and extensions they are suitable for.

In order to evaluate a government program one may look at its effects between countries, industries or groups of individuals without analyzing its effects within these entities. However, one may also look at its effects at a more disaggregated level, such as the individual, the fiscal-unit or the household. In order to perform this second type of analysis, in fact, one needs micro-data, in the form of repeated cross-sections or panels. Existing datasets often do not suffice, either because the period they cover is too short, or because simulated ‘future’ micro-information is required for a priori evaluation. One therefore might need a model that can generate this data, and these models are called Micro Simulation Models (MSM).

The first aim of this second chapter is to discuss advantages and disadvantages of micro simulation models over other models, particularly macro models and – within models focusing at micro level -

---

2 Center for Research on Pensions and Welfare Policies (CeRP), University of Torino, Italy.
standard simulation models (section 2). Secondly, in section 3, we discuss some of the discerning characteristics of various types of micro simulation models. This will result in a classification of micro simulation models that is based on the fundamental simulation properties that they share. Contrary to earlier papers, this paper will extend the classification to include standard simulation models. Furthermore, it will discuss at length the order in which individuals and time periods are simulated in dynamic micro simulation models. This will further clarify the pros and cons of various types of models. Next, some important dynamic microsimulation models currently in use in Europe, will be discussed. Finally, the dynamic micro simulation model MIDAS that is simultaneously being developed for Belgium, Germany and Italy within the project AIM, will be presented and discussed. It will be linked to the above classification, and its fundamental simulation properties will be confronted with its ‘raison d’être’, namely the simulation of the adequacy of pensions.

2.2. What are the ‘simulation characteristics’ of a model?  

Define ‘simulation characteristics’ as those characteristics of a model that have consequences for the actual or potential research problems that can be covered by a model, as well as the implicit or explicit assumptions that a model makes when handling a specific research problem.

The first part of this definition limits characteristics to those that are relevant in the light of the (potential) research applications. Whether or not the model developer is married with two children and loves cats, or that the computer used for development and maintenance is a laptop, are not simulation characteristics because they do not say anything on the range of (potential) research problems which the model can be used for. For example, suppose that we want to assess inference aspects of a certain potential policy measure. Then the choice of what model to use should among other things be based on whether or not a model can simulate distributional effects. This feature then is a simulation characteristic.

But the definition of ‘simulation characteristics’ is more subtle than a mere description of what research problems a model can handle. For this would imply that two models that are used in the same research problem, by definition have the same simulation properties. This of course is not true. In fact, there are no a priori reasons why two models that are inherently different could not be used in tackling the same research problem. So, the definition of ‘simulation characteristics’ needs to be expanded beyond the range of potential research problems, towards the implicit and explicit assumptions a model makes in handling these research problems. This is the second part of the above definition. For example, both microsimulation models (henceforth called MSMs) and some Computational General Equilibrium models (henceforth called CGE-models), such as the Adelman and Robinson (1978) model, can be used to simulate inference-aspects of policy, but the underlying assumptions are very different. Indeed, where the MSMs base the distribution of income on a sample of individuals, the CGE in this case assumes a constant distribution of income within household types, meaning that changes in overall inequality are the results of redistribution between household types.

---

3 This section is based on Dekkers and Legros (2006).
Also, it is relevant to discern technical characteristics from simulation characteristics. For example, the programming language in which the model is written is a technical characteristic, and not a simulation characteristic, for one may envisage two models that are otherwise the same, to be developed in two different programming languages. However, it is also possible that technical characteristics determine the simulation characteristics of the model. For example, suppose that the choice of the language has consequences on what the model can be applied to, for instance due to the way the language handles data. Then the language is not a simulation characteristic itself, but the cause of a simulation characteristic. But the opposite, where a simulation characteristic causes a technical characteristic, is also possible. For example, the more complex a model is, i.e. the more simulation characteristics it has, the longer it takes for the model to run or to converge towards a steady state, and the more efficient the programming language therefore needs to be.

Simulation properties can either be the result of the fundamental characteristics of a model, or of deliberate extensions added to a model. For example, a microsimulation model by definition can simulate inferences, whereas a standard simulation model cannot (van Mechelen and Verbist, 2005). When choosing what type of model to develop, the modeller starts from the anticipated future research questions, and matches them with the simulation properties of model types available. Once a choice is made and a model of a certain type is developed, simulation characteristics can be changed by making extensions to the model, but these extensions remain within the ‘boundaries’ of the model type. These boundaries represent the fundamental simulation properties of a model, whereas the extensions are non-fundamental simulation properties.

As an example of such an extension, one can imagine that a microsimulation model as well as a standard model can be extended by a gross-net trajectory, allowing the simulation of fiscal effects. This will however not change the fundamental differences between the two types of models, and the effects of fiscal policy will be simulated for a sample of individuals (in the case of the MSM) and for typical fictitious individuals (standard models).

Finally, note that the above definition also covers the difference between ‘simulations’ and ‘projections’ for the latter is simply a special case of the former, namely a simulation under the assumption of an unchanged policy environment. Hence, simulations can refer to the measure of a proposed reform on a given population whereas projections are simulations of no reform. As a consequence, for a model to be able to generate simulations, it must explicitly include simulation properties. By contrast, a simple trend, a regression equation or a vector autoregression (VAR) model can perform equally well in making projections than many complex models do. They however lack any simulation possibilities. So, if a model has the properties that allow it to simulate an exogenous effect on an endogenous variable, then it can by definition make a projection (simulate the ceteris paribus clause) of this variable, but the opposite is not necessarily the case.

Now that the notion of simulation properties has been defined, the fundamental simulation properties of micro-level models can be used to make a classification. Before doing so, however, the advantages and disadvantages of micro level models must be put in comparative perspective. This will be done in the next section.
2.3. Micro Simulation Models versus other simulation models

When comparing micro and the macro simulation approach, the former has several advantages, but also some problems (Emmerson et al., 2004).

An important advantage of micro simulation models is that the level of modelling is in line with the level at which policy takes effect, especially in terrains such as public pensions, health care and other aspects of public finance. So, where macro simulation considers averages, a micro simulation model can simulate at the individual level, and therefore report the effects of policy on the income distribution, as well as poverty (often a function of the location of specific groups within this distribution). So, where macro economic models are specifically designed to consider financial consequences of a certain measure or development for the population as a whole, or for some subgroups, micro simulation models focus on redistributive impacts, and the adequacy of a social security scheme (in terms of preventing poverty and loss of welfare).

Furthermore, macro economic models do not consider the dynamics below the averages. Therefore, questions like “which types of individuals or households move up or down the income distribution over time?” are not considered by macro models but are key element in micro-models.

Caldwell and Morrison (2000, 201) describe several examples of research issues for which microsimulation models are particularly suited. These include analyses of projected winners and losers, exploration at the micro-level of the operation of social security programmes, quantification of incentives to work, to save or to retire, and longer-term consequences of societal trends in marriage, divorce and fertility.

Problems associated with the micro simulation approach are, first of all, that the theoretical underpinning of many micro simulation models is often scarce at best, though improvements are being made. It will be discussed later that longitudinal micro simulation models often have an underlying structural model of at least one key process such as the retirement decision or saving, cross-sectional models often have empirical ad-hoc solutions to many processes. One might therefore argue that structural models are a better alternative.

Structural models however often simulate one or two key processes for a couple of representative agents (and are therefore to be classified as standard simulation models, cf. infra). In contrast, most dynamic cross-sectional micro simulation models have a complex framework of a large sample of individuals of different characteristics, such as age gender and labour market status, where all these characteristics are simulated, taking into account parallel (if possible) or serial interaction, both between characteristics and between individuals.

Furthermore, many underlying assumptions of structural models, such as the functional form of the utility function and what exactly adds to utility, or the development of the endogenous macro economic environment, often have important impacts and might not improve the fit of the model.

---

4 Emmerson et al. (2004) discuss this in more detail, and make an explicit comparison between the UK-model PENSIM II and a structural model. This section of the text draws upon their work.
tural models also are rather complex, especially when the number of processes increases. This has so far discouraged the development of a structural cross-sectional micro simulation model, at least to our knowledge\(^5\). This might however change, as we learn from the development of longitudinal micro simulation models, which traditionally have a larger ‘structural component’, and as new and more powerful computers become available. But even then, the challenges in terms of finding the balance between simulation stability and empirical fit of the model on the one hand, and theoretical soundness on the other, are enormous and one might even wonder whether the development costs will ever be worth the benefits.

\(^5\) The exception is the CBOLT model, where a structural life-cycle model was developed to model saving and labour supply (Harris et. al., 2005). It should however be noted that there is a promising development in linking CGE models with static microsimulation models (see Peichl, 2008), which may in the future be extended to dynamic models.
Figure 1: A classification of empirical microeconomic models

start

Standard Simulation Models

Micro Simulation Models

Static

Dynamic

Static Ageing

Dynamic Ageing

Cross-sectional ageing

Longitudinal Ageing

Base-data: population

Base-data: cohort

Open

Closed
The following discussion and classification will follow the outline presented in Figure 1. Within the group of microeconomic models, some authors (see Van Mechelen and Verbiest, 2005 for a discussion) discern two broad categories. These are standard simulation models and micro simulation models. The standard simulation models simulate one or several – often synthetic - micro-unit, representing a category of micro-units. For instance, Dekkers (2006) simulates the effect of taxes and transfers on the expected future pension wealth of a single male or female blue or white collar worker. Micro simulation models by contrast simulate all micro-units in a representative cross-sectional sample of the population at a certain moment in time. So, they may simulate the effect of taxes and transfers on the level and distribution of income within the available dataset.

Standard simulation models are conceptually fairly simple – even though they can be technically complex- and there is no need for a large representative dataset of individual units. They are designed for the analysis of the ‘mechanics’ of a system of taxes, contributions and transfers. Given the characteristics of the synthetic individual, one may simulate income before and after implementation of a certain measure in the field of taxes, contributions and transfers, and simply compare the results. The drawback is that questions on the representativeness of the simulation results remain, for it is questionable to what extent simulations on one category of individuals can be used to draw general conclusions. Furthermore, given that one or several individuals are simulated, it is not possible to express the effects of policy measures in terms of changes of the sample moments of income. Put differently, it is not possible to simulate relative income poverty or inequality.

Micro simulation models start from a representative cross-sectional sample of the population, and then change this dataset to reflect an assumed future development, or the implementation of a certain policy measure, and so forth. Compared to standard simulation models, they usually are more complex, and therefore more expensive in terms of development and maintenance. They however are more representative for the population as a whole, and they are able to simulate poverty, income inequality and so forth. In the paragraphs to follow, the category of micro simulation models will be classified further.

2.4. A classification of microsimulation models

2.4.1. Static versus dynamic microsimulation models

The primary classification of MSM is based on whether and how they model time. Is the cross-sectional dataset simulated to reflect an assumed future development? If so, how? Static models do not include time, and therefore only simulate overnight effects of a change in policy. If time is not modelled, then the model is valid only for the cross-section data period (Merz, 1993, 4, 1994, 6). In several static models, however, the dataset that is the point of departure of the model is
‘aged’ to bring it up to date (Sutherland, 1995, 3). This will be discussed at length in the next section. The most well known European static model nowadays is EUROMOD, developed by an international group of researchers in the context of the fifth European framework. This model covers 15 pre-enlargement member states of the European Union, and its latest version has been extended to 4 new member states: Estonia, Hungary, Poland and Slovenia. EUROMOD is freely accessible\textsuperscript{a} and "has been used for a number of policy-related exercises, ranging from studies of the relationship of public spending on social benefits to poverty and the implications of a common European minimum pension, to the impact of welfare benefits on work incentives and the consequences of non-indexation of taxes and contributions" (Sutherland, 2001, 1). Sutherland (1995) discusses static models in Europe; Merz (1994) also discusses models developed in the US, Canada and Australia.

Dynamic models do include time, and the simplest ones are those where time is simulated indirectly, via the reweighing of the units dataset to mimic a process of demographic ageing. These models are referred to as dynamic MSM models with 'static ageing'. Basically, the technique used to update static models now becomes the way to mimic time. Instead of changing individual characteristics over time, dynamic models with static ageing use exogenous future aggregate data to adjust the sample (Merz, 1993, 4, 1994, 6). This process is described in Harding (1996, page 3 and further) and consists of two basic steps. The first step is reweighing. This involves changing the weight attached to each individual record in the micro data, usually to reflect demographic ageing i.e. the change of the relative size of the cohorts in the sample. The structure of the sample itself is therefore not modified. The second step is updating, where monetary values within the dataset are adjusted to meet exogenous future projected developments. An example of a simple model where both techniques are applied is STATION (Dekkers, 2000, idem, 2003). This model was designed to simulate the effect of full or partial linkage of pension benefits to the development of wages on poverty and inequality among pensioners in Belgium.

We now turn to dynamic micro simulation models with ‘dynamic ageing’ (henceforth referred to as dynamic MSM). In opposition to models categorized so far, dynamic MSM do not reweigh, but alter the contents of the dataset itself. It involves "updating each attribute for each micro-unit for each time-interval" (Caldwell, 1990, in Harding, 1996, 4). Taking a certain dataset, individuals face certain probabilities of a change in each of their attributes. In the modelling process, this is simulated by chance. The number of variables that can be modelled this way depends entirely on how much information on transition probabilities or risks are available (Dekkers, 2003, 183). A dynamic model builds up complete synthetic life histories for each individual in the dataset, including data on mortality, labour market status, retirement age, savings and so on (Emmerson, et al., 2004, 3).

\textsuperscript{a} A downloadable version can be found at http://www.iser.essex.ac.uk/msu/ emod/
Before discussing a classification of dynamic micro simulation models, let us take a quick look at the question what type of model to choose for what reason. If one is interested in the overnight effects of policy changes, or simulations that pertain only to the dataset used as a point of departure of the model, then one might choose a static model i.e. one without ageing. If one is interested in simulation results that can be analysed using cross-sectional analysis of current and ‘future’ simulated data, one might choose a dynamic model with static ageing. Finally, if policy analysis involves a panel data analysis, i.e. if it requires that the simulated units evolve over time, then one might opt for a model with dynamic ageing. On the more practical level, one needs to weigh complexity against applicability. Static models are less complex than dynamic models, which means that they take less time to develop and require less maintenance effort. On the other hand, the scope of dynamic models is much wider than static models, which means that the potential applicability of dynamic models exceeds that of static models.

2.4.2. A classification of dynamic models

Several characteristics can be used to classify dynamic MSMF. A first fundamental difference pertains to the dataset that is taken as the point of departure of the model. The dynamic population models involve the ageing and adjustment of a cross-sectional sample of an entire population. So, the point of departure is a dataset consisting of individuals of many age groups or cohorts. Dynamic cohort models, by contrast, age only one cohort and this from birth to death (Harding, 1996). A consequence of this difference is that the population model “will produce also many micro units with an incomplete life-cycle; some micro units are still living or have died in an earlier simulation” period (Merz, 1994, 9). Of course, since cohort models simulate just one cohort of individuals, cohort models cannot directly simulate demographic ageing, which after all is a change of the relative size of cohorts vis-à-vis each other.

A second characteristic has to do with the order in which individuals are simulated over time. This is the difference between cross-sectional models and longitudinal models. Suppose a model that is to simulate $N$ individuals from periods 1 to $T$. In cross-sectional models, all individuals are simulated for one year. In the first period, all $N$ individuals are simulated from period 1 to 2. Next, all individuals are simulated from period 2 to 3, and so forth. By contrast, longitudinal simulation models simulate one individual for all years. So, individual 1 is simulated from birth to death. Then, the same is done for individuals 2, 3, up to $N$.

The difference between the two types of models may seem trivial, as the result of both models is the same: a simulated data set of $N$ individuals for all $T$ years. However it has some important

---

7 A comprehensive and exhaustive overview of the characteristics of micro simulation models can be found in O’Donoghue (2001).

8 This is why they are sometimes referred to as “dynamic life-cycle models”, whereas the population models are called “cross-sectional models” (Merz, 1994, 9). This latter appellation may cause confusion with the typology based on the simulation order which is to be discussed next, so it is not used in this paper.
empirical consequences. A first consequence is that cross-sectional models allow for micro-interactions, i.e. interactions between individuals. If, for example, an individual experiences a certain change (he or she dies, to name one quite important change), this in real life affects the situation of other individuals (the partner becomes a widow/widower). In a cross-sectional model, this is easy to do. In a longitudinal model, this is more difficult since all individuals are simulated independently from each other: when the simulation of individual $y$ starts, the simulation of individual $x$ is ended, so that his or her whole future is already set. This however has as an advantage that the (future) life span of an individual is affected by a limited number of potential events, and this makes it possible to introduce forward-looking elements in the behaviour of the individual (see Sefton and van de Ven (2004) for an application). This makes the model theoretically appealing. A drawback of models with longitudinal ageing relative to those that use cross-sectional ageing, is that the former do not allow for the simulation of household income, whereas the latter do. As most measures of poverty risk are based on (equivalent) household income, models with cross-sectional ageing are the more useful when it comes to simulating poverty, (re)distribution and inequality. They however are less developed in terms of theoretical underpinning.

A second difference relates to the ability of both types of models to include 'life time decisions' such as savings. Both types of micro simulation models make it possible to model these decisions. However, longitudinal models simulate lifetimes separately, and this future is then 'frozen'. Hence, they are more suited for these kinds of decisions than cross-sectional models, which do not necessary simulate an entire life span of an individual. As a consequence, the theoretical foundation of longitudinal models usually is better developed than that of cross-sectional models, which concentrate on applicability and strengthening their policy-supporting role. In practice, therefore, one often sees that longitudinal models are developed for academic purposes, where cross-sectional models often have a policy-supporting role to play.

So far, a difference has been made between population models and cohort models, referring to the dataset they use as a point of departure, and between cross-sectional and longitudinal models, referring to the simulation order of the individuals in the dataset. This two-dimensional classification has been rarely considered in the literature so far. A possible explanation is that, to our knowledge, the combinations population-cross-sectional, and cohort-longitudinal are by far the most common, while the others are a minority. This may be why Harding (1993), O’Donoghue (2001) and others use a one-dimensional classification, The difference between the two types of models then becomes, in the words of O’Donoghue (2001, 17) "A cohort model is simply a model that ages a sample of unrelated individuals aged zero, while a population model ages a sample of individuals of different ages, some of whom are related". Space for our finer classification can be however found in the words of Harding (1996, 5), according to whom dynamic cohort models use exactly the same type of ageing procedures. We are aware that the usefulness of the two-dimensional classification is primarily on the theoretical ground. However, models that do not fit the one-
dimensional classification (population-longitudinal and/or cohort-cross-sectional models) exist already today, as will be shown below. Our classification might gain empirical relevance in the future, as more of such models are developed. Furthermore, separating these characteristics improves comprehension of the characteristics of a specific micro simulation model.

Another classification can be based upon whether the models are open or closed. This has to do with how marriage of individuals is modelled. A closed model generates new individuals in the case of birth or immigration only. So, when somebody in the model 'becomes eligible for marriage', his or her spouse is selected from the other living individuals in the dataset. In an open model, a 'synthetic individual' is created and linked to our marriage candidate. This of course is necessary when individuals are simulated independently from other individuals in the dataset, which is the case in cohort/longitudinal models. In population/cross-sectional models, however, pulling additional synthetic individuals out of a hat is unnecessary, for the simulation method allows for relations between existing individuals. Existing individuals therefore are often matched via a 'marriage market module' of some sort.

Another discerning factor between models has to do with how time is modelled. Here, discrete models stand in opposition to continuous models, a difference pertaining to the difference between discrete and continuous time hazards used in these models. Continuous models include continuous time hazards, defined with reference to a period of time (and not a probability) after which a certain event will occur. Discrete models by contrast include discrete time hazards: a probability that an event will occur in an interval of time. Emmerson et al. (2004, 10) explain the difference by saying that dynamic micro simulation models can simulate relevant life events either "year-on-year for the (starting) year $t$, $t+1$, $t+2$, in discrete time, or by starting at $t$, and predicting a life-event at $(t+n)$, where $n$ is positive and possibly non-integer (continuous time)". For a more elaborate discussion, the reader is referred to O'Donoghue (2001); we limit the discussion to that most models to date are of the discrete type, and one of the reasons for this is the lack of continuous data for the processes to be simulated.

2.5. A focus on some relevant models

In this section we provide an introductory description of some dynamic MSM's. We concentrate on the countries with more experience in microsimulation modelling and more specifically on the models which are more innovative or/and more relevant for pension issues. We choose MINT for the US, Pensim2 for the UK, and DYNAMITE for Italy.

The MINT (Modeling retirement Income in the Near Term) model (Panis and Lillard 1999, Butrica et. al. 2001, Toder et. al. 1999, O'Donoghue 2001) has been developed in the US by the Social Security Administration (Office of Research, Evaluation, and Statistics), with substantial assistance from the Brookings Institution, the RAND Corporation, and the Urban Institute. The
model projects the retirement income (social security and pension income, but also asset income and earnings of working beneficiaries) of individuals at their retirement age. The simulation period ranges from 1997 to 2031. The model has been applied to simulate several policy scenarios. Between them there are the analysis of the effects of social security benefits reforms on the level of benefits, retirement income and poverty, and the analysis of cohort differences in the sources of retirement income. Its detailed demographic component allows also simulating economic well-being in retirement.

MINT is a population model. Its base population – 113,000 individuals born between 1926 and 1965 – is obtained merging files coming from multiple sources. In particular, demographic information and marital histories come from the Census Bureau’s Survey of Income and Program Participation (SIIP, years 1990 to 1993), while earnings come from the SSA Summary Earnings Record dataset (SER, years 1951 to 1996).

Transitions into marriage and divorce (and to death) are modelled using a continuous time hazard specification. Using the estimated coefficients, the expected dates of these events are determined. Therefore the model is, at least with respect to these occurrences, continuous. Marriages are simulated in two steps. For each marrying candidate, the characteristics of the ‘ideal’ partner are first defined. Then, spouses in the sample are matched by means of a statistical matching algorithm which minimizes a ‘distance function’ defined in terms of individual characteristics. When the ‘ideal’ spouse is not found in the sample, because there is nobody with the desired characteristics, a ‘synthetic’ one is created. MINT is thus a mixed open and closed model.

The age of retirement, for those eligible to OASDI benefits, is determined by estimating a logit model. The probability of receiving social security benefits is explained by a set of individual characteristics like age, education, gender, race, marital status, earnings and non-housing wealth. Earning profiles are estimated using a fixed-effects specification (run separately for each gender and education level). The individual effect is also estimated, in order to predict earnings for out-of-sample years. MINT thus fully exploits the advantages of its panel data, when estimating transitional probabilities in the demographic module as well as in its economic modules. It fully takes into account that social security benefits depend not only on the pensioner’s earning history, but also on his marital histories and on the earnings history of his spouses(s). Earnings of the spouses are obtained in a different way depending on whether the spouse is in the sample. If the spouse is in the sample, they are predicted from the estimated earnings lifetime profile. If the spouse is ‘synthetic’, they are instead determined by first imputing a spouse from a pool of eligible donors, and then assigning her earnings to the main pension beneficiary.

Pensim2 (Emmerson et. al. 2004, Zaidi and Rake 2001) is the second version of the microsimulation model built by the British Government’s Department for Work and Pensions to analyze the distributional impact of pension policy reforms in the UK. The level of accuracy of many of its
modules makes it also suitable to quantify the effects of tax and benefits interventions, as well as tuition fees for higher education.

Pensim2 is a population model. It exploits information of several administrative and survey databases. Between them, the Lifetime Labour Market Database (LLMDB), the British Household Panel Survey (BHPS) and the Family Resource Survey (FRS) are the most broadly exploited. The model is discrete (many life events are modelled with a probit or a logit equation; some of them with nested logits) and runs up to the year 2050. The model is furthermore closed. The partnership module governs several events: new marriages, changes from cohabiting to married, new separations and custody of dependent children after it, divorces from separations. The matching process between individuals is determined by the 'order of decreasing differences' algorithm.

Pensim2 simulates occupation by first ranking the individuals according to the probability to be at work (using the BHPS), and then calibrates the transitional probabilities in BHPS using LLMDB. Labor supply is modelled in the household framework, although it is not jointly determined between the spouses. Decisions to work are in fact assumed to be taken in a sequential order: first the man decides and then the woman decides taking into account the choice of the man. Man's work status is therefore estimated without considering the work status of the woman, while woman's work status is estimated including as explanatory variable the man's one. Earnings are estimated using a random-effects specification and exploiting the BHPS panel. This specification has been preferred to a fixed-effects one, because it allows both to impute an individual effect for new entrants into the labour market and to quantify the effect of education on earnings.

A relevant modelling effort, within the pension module, has been dedicated to the attribution of public and private pension rights, both for current and future workers. Modelled events are whether an employee is offered the opportunity to join an occupational scheme, whether he chooses to join it and whether he joins a private pension scheme. To this aim, probit equations using data from the FRS are estimated.

DYNAMITE (Ando et al. 2000, Ando and Nicoletti Altimari, 2004, O'Donoghue 2001) has been built at the Bank of Italy, in collaboration with Albert Ando at the University of Pennsylvania. Its principal aim is to study the effects of demographic ageing (and of the evolution of the family structures) on aggregate saving in Italy, under the assumption of lifecycle behaviour. Further applications concern the effects of the 1995 reform of the Italian pension system - looking its steady state but especially at its transitional phase toward the new regime – and the distributional impact of tax reforms.

DYNAMITE is a population model. Its base population is generated starting from the 1993 wave of the Bank of Italy's Survey on Household Income and Wealth (SHIW). By means of a weighting
procedure based on non-response weights, the original sample of roughly 8,000 households was enlarged to 200,000 and made proportional to the Italian population. The unit of analysis is the household, and the model is closed. The demographic module is extremely fine and accurate. The simulation period is very long, more than 100 years.

Almost all the probabilities in the labour market module, such as participation, unemployment and transitions between sectors and occupational status, are calibrated to match aggregate characteristics. A grossing-up procedure is adopted in order to transform declared income, which are net of income taxes, into the taxable basis for social security contributions and pension benefits computation. A correction factor is applied for misreporting purposes, particularly relevant for self-employed. Lifetime income profiles are constructed by estimating a two-steps Heckman’s model on the pooled 1987-1995 SHIW cross-sections. In the estimation, cohort-specific effects on wages, which it was not possible to disentangle from age effects, are attributed exogenously assuming they are equal to the productivity growth.

The decision to retire is endogenous and depends on the financial incentives provided to the worker by the pension system. Incentives are measured by the ratio of the expected SSW of working up to age 60 to normal earnings. Two reduced-form behavioural equations are estimated, exploiting the information on the expected age of retirement included in SHIW. The first equation includes as explanatory variables, other than the incentive measure, a set of ‘time-invariant’ characteristics (e.g. year of birth and education). The second, instead, includes in addition some ‘time-variant’ characteristics (e.g. the family structure). The first equation is then used to predict the age of retirement of new occupied, while the second to predict it when the worker reaches age 50 and for each subsequent age. The model, in fact, takes into account that older workers may want to revise their expectations on the retirement age in light of updated information on the working career. The computation of social security wealth (SSW) is complex because of the evolving normative framework throughout the simulation period. Different formulae are applied depending on the type of workers (employee in the private sector, employee in the public sector, self-employed), cohort and simulation year.

2.6. Micro simulation in the AIM-project: what kind of model is MIDAS, and why?

In this section, the choices which have been made in the development of MIDAS are outlined. The model MIDAS (an acronym for ‘Microsimulation for the Development of Adequacy and Sustainability’) is developed within the AIM-project in order to simulate the adequacy of pensions in Italy, Germany and Belgium. The concept of adequacy was defined in the first chapter. Furthermore, it was translated into three objectives, two of which are the main objectives of a public pension system. These are the reduction of the risk of poverty in old age and the preservation, at retirement, of a standard of living comparable to that of the final part of the active life. In what fol-
lows, these objectives will be linked to the classification made in the previous chapter. Why should micro simulation models be used in the simulation of these objectives? Which class of micro simulation models seems the most appropriate in this context?

2.6.1. The first objective: the prevention of old-age poverty

Poverty is about a lack of welfare. As welfare is not directly measurable, social scientists often opt for an indirect measure, where poverty is based on the confrontation of the income of a household or family with a poverty line $z$. If this household income $x$, corrected for differences in size and composition of the household is below this line, then the individuals in this household or family are considered to be poor. Suppose $F(x)$ the distribution of household incomes $x$, and $f(x)$ the density function. The headcount ratio $HC$ can then be written as $F(z) = \int_{0}^{z} f(x) dx$. In its discrete form, this becomes $HC = p/n$, with $n$ being the population, and $p$ the number of individuals whose equivalent income is below the poverty line, or $(x-z)<0; p\leq n$. This is also denoted as reflecting the 'risk of poverty' in the population. Other, more sophisticated measures of poverty are based on the individual poverty gap, $(x-z)$ or $(x-z)/z$, reflecting the intensity of poverty among the poor. These will be discussed in more detail in chapter 5, the simulation results. How poverty is exactly measured is not relevant from the point of view of the modeller. In fact, any model that can simulate the equivalent household incomes $x$ in principle allows for the simulation of all kinds of income-based measures of poverty. It has been clear from the previous chapter that micro simulation models are designed to meet this demand. However, the question then is which category of micro simulation models is best suited for the task and hand?

Dynamic models with static ageing can simulate the future pension income of the elderly and they have been used to simulate poverty and income inequality among the elderly before (see, for instance, Dekkers, 2000). Dynamic cohort models with longitudinal dynamic ageing could in principle simulate income distribution among pensioners in a certain future year, but only within the same cohort. This makes them less attractive in the context of AIM, for a comparison of the relative income position between cohorts is not possible. To give an example, the degree to which pension benefits are indexed to the development of wages or prices has a strong effect on the income position of older retired cohorts relative to younger cohorts. This cannot be simulated by a cohort model, just because that it has one cohort as the point of departure. Furthermore, the longitudinal ageing process of these models does not allow for interactions between members of a household. This means that simulating household income is not possible, because the lifetime and earnings history of any individual in a typical household is given for any other individual.

Finally, do dynamic population models with cross-sectional ageing meet the requirements to simulate poverty among pensioners? Yes, for they simulate various cohorts and can therefore simulate different poverty risks of elderly versus the young in any future year. Furthermore, the cross-sectional ageing process implies that interactions between simulated individuals within a
household (and potentially within every group) are possible, and can be modelled. This implies that a meaningful simulation of household income is possible. In conclusion, both models with static ageing and models with dynamic ageing of the cross-sectional type meet the requirements of the first objective.

2.6.2. The second objective: the preservation, at retirement, of a standard of living comparable to that of the final part of the active life

The most straightforward measure of the preservation of income at retirement is of course the replacement rate. In the microeconomic sense of the word, the replacement rate is the ratio of one’s first pension income over one’s last salary. In order to simulate future replacement rate’s, one needs the pension benefit to be simulated within a model. And as the replacement rate is based on the information gathered in the year that an individual retires, the model must be capable of simulating this transition between work and retirement. It is clear that models with static ageing do not meet this requirement. Furthermore, the ‘Bismarckian character’ of the pension systems in most European member states is a relevant characteristic. For, the pension benefit one is entitled to when reaching the retirement age, may in some way depend on one’s earnings history. So, any model should incorporate and simulate this earnings history. This again implies that models with static ageing do not meet the demands, for this class of models do not have a ‘simulation memory’ in the sense that simulation results at the future time \( t+n \) are independent on the results in \( t+n-1 \). So there are two reasons why one may conclude that static models or models with static ageing are unsuitable for the simulation of individual future replacement rates. However, static models can be used to reproduce individual replacement rates, and hence to simulate the overnight effect of policy measures on this current replacement rate. Furthermore, if one adopts a more macroeconomic definition of replacement rate –as the average pension of young retired cohorts as a fraction of the average earning of the older working cohorts- then dynamic models with static ageing can simulate this replacement rate, even though they do not simulate the work-retirement transition itself. Using these kinds of models, the effect of partially linking pensions to the development of wages on poverty or inequality can be simulated.

Finally, both classes of models with dynamic ageing (cohort models with longitudinal ageing and population models with cross-sectional ageing) do simulate individuals over time. Microsimulation models of these categories can therefore simulate both macroeconomic and microeconomic replacement rates. They therefore meet the requirements of this second objective.

2.6.3. Finally: the simulation of demographic ageing and pensions

The first objective can be met by models with static ageing and population models with cross-sectional ageing. When we limit the second objective to simulating microeconomic replacement rates, then this second objective can only be met by the latter class of models. Only population
models with cross-sectional ageing therefore meet the requirements of both objectives. But there are two other important reasons why this class of models is the most relevant in the context of this project. The key stone of this project is that it concerns the consequences of ageing on the adequacy of pensions, and one therefore needs a model that allows for the simulation of demographic ageing. Let us consider whether the three classes of dynamic MSMs meet this demand.

A MSM with static ageing can in effect simulate demographic ageing by varying the weights associated with individuals from different cohorts. A cohort model can in principle not simulate ageing, for it has only one cohort as the point of departure. Finally, a cross-sectional model is best suited for the simulation of ageing, for it applies exogenous fertility and mortality rates that may change over time. This way, various subsequent cohorts are simulated at once, and ageing – essentially a change of the relative size of cohorts – is therefore inherent in the system. The very fact that the project assesses the consequences of ageing is a strong argument in favour of cross-sectional micro simulation models. Moreover, some important socio-demographic developments appear within households (such as the emergence of two-earner households), and this also requires cross-sectional micro simulation, for only these models allow for interaction between individuals. So, the conclusion is again that population models with cross-sectional ageing are the most suited given the research goals in the project AIM. It is for this reason that the model MIDAS fits into this class of MSMs.
2.7. Conclusions

In this paper, micro simulation models have been put in opposition to various other types of models, be it micro economic models such as standard simulation models, macro simulation models, option value models and generational accounting models. Furthermore, many characteristics that classify within the group of micro simulation models have been presented and discussed. This classification has been taken further to discern various characteristics of ‘dynamic micro simulation models with dynamic ageing’, often simply referred to as dynamic micro simulation models, where a difference has been made between cross-sectional and longitudinal ageing, and population models and cohort models. Finally, an overview of existing dynamic micro simulation models, for the largest part constructed from previous work of other authors, is presented and some general conclusions are drawn. First of all, most of the models to date are of the dynamic cross-sectional type, and this might be due to their larger scope. Furthermore, the number of models developed is increasing exponentially. This may for a part be caused by the rapid evolution of computers, allowing for more complex and more CPU-demanding models to be developed. But more importantly, it might reflect the growing interest of policy makers into not only the macroeconomic financial effects of social and fiscal policy, but to distributional effects as well.

Next, the choices made in the development of MIDAS are outlined and described. MIDAS aims at simulating the adequacy of pensions in Belgium, Germany and Italy. This requires that the model to be constructed allows for the simulation of old-age poverty, the measurement of a standard of living at retirement and the simulation of demographic ageing and pension system regulations. It is argued that population models with cross-sectional ageing are the best suited for this project.
2.8. References


REPORT


