n° 2019-06 April 2019

WORKING PAPERS

DyMH_LU: a simple tool for modelling and simulating the health status of the Luxembourgish elderly in the longer run

Anne-Sophie **GENEVOIS**[1]
Philippe **LIEGEOIS**[1]
María Noel **PI ALPERIN**[1]

[1] Luxembourg Institute of Socio-Economic Research (LISER), Luxembourg

DyMH_LU: a simple tool for modelling and simulating the health status of the Luxembourgish elderly in the longer run

Anne-Sophie Genevois
Luxembourg Institute of Socio-Economic Research

Philippe Liégeois
Luxembourg Institute of Socio-Economic Research

María Noel Pi Alperin[1]
Luxembourg Institute of Socio-Economic Research

ABSTRACT: We are facing one of the most important demographic events of the last decades in Europe: the population ageing process. This process will have significant economic effects particularly on health. As most diseases are age-related, this process might imply a proportionally higher share of individuals with declining health. Being able to forecast the health status of the population can help to deal with concerns about the financial and social sustainability of several public policies including health. In this paper, we present the *DyMH_LU* model, a dynamic microsimulation model focused exclusively on the health status of the Luxembourgish population. One of its major characteristics is that it simulates more than sixty different diseases and limitations in the activities of daily living. All this simulated information can be aggregated in order to compute, for each period, the overall health status of each individual, the marginal distribution of each disease among the total population and the global health status of the entire population. The starting point of the *DyMH_LU* model is the information collected in 2015 in the Wave 6 of the SHARE database that targets individuals aged 51 or older. The simulation period covers 2017 until 2045.

KEYWORDS: DYNAMIC MICROSIMULATION, HEALTH, SHARE, LUXEMBOURG

JEL classification: C1, C2, I1

[1] Corresponding author, Luxembourg Institute of Socio-economic Research (LISER), 11 Porte des Sciences, L-4366 Esch-sur-Alzette, G.D. Luxembourg, tel.: +352 58 58 55 416, e-mail: MariaNoel.PiAlperin@liser.lu

1. INTRODUCTION

The last century witnessed the beginning of a population ageing process that represents one of the most important demographic events of the last decades in Europe. The ageing of the population is expected to accelerate, not only because there is a decline in the mortality and fertility rates, but also because people live longer. Luxembourg also follows this trend and might have one of the highest life expectancies in Europe by 2040-2045, at around 84 years (Dlugosz, 2011).

This ongoing demographic process will have important economic effects. In particular, given the key role that the public sector currently plays in the within-period transfer of resources through the social security system from workers to retirees and from a healthier population to a less vigorous one, the population ageing process is inducing concerns about the financial and social sustainability of several public policies including pensions and health.

There is a growing interest to analyze the consequences of this demographic process on the economy in general and on public finances in particular. To do so, it is important to know more about the evolution of the health status of the population. Indeed, most diseases are age-related. In particular, the number of chronic illnesses and the degree of dependency are expected to grow fast over the coming years. In consequence, the ageing process might not only imply a higher share of old, non-active individuals to the total population, but also a proportionally higher share of individuals with a declining health.

During the last four decades, several simulation techniques, and in particular microsimulation models, have been developed for addressing health-related issues, such as the *CORSIM, DYNASIM, DYNAMOD, DYNOPTASIM, POHEM, LifePath, FEM and COMPAS* models.[2] However, most of these models are not focused on health status as such. In particular, they do not simulate the overall health status of individuals but only separate health outcomes. Lastly, up to now little has been done in this respect for Luxembourg – a largely open country with specific public finance challenges down the line.

---

[2] For more information about microsimulation models related to health issues, see Li and O'Donoghue (2013), Rutter, Zalasvsky and Feuer (2011), and Zucchelli, Jones and Rice (2010). A brief description of these models is presented in *Section 2*.

The aim of the present paper is to describe a modelling framework for the forecast of individuals' health status in Luxembourg. The *DyMH_LU* model is a dynamic simulation model grounded in the structure of *MiDAS_LU*[3], a tool developed for the analysis of the ageing of the population in Luxembourg at large.

The *DyMH_LU* model allows to simulate over time several aspects of health (reflecting different facets of the mental and the physical dimensions of health). Given a specific simulated period of time, these single items can be aggregated in order to describe the overall health status for each individual in the population. Additionally, this individual information can be aggregated in order to derive the global health status for the entire population. Therefore, these results can provide a more precise picture on how global health will evolve in the next future. This will later open an avenue for estimations of changes in the demand for healthcare, medicines and home care services.

The ageing process of the residents in Luxembourg, and its consequences on the individuals' and the population's health, are simulated starting from an instantaneous picture of the Luxembourgish population derived from the SHARE[4] database. One of the main characteristics of this survey is the richness of information that it contains in the domain of health. Thus, a large number of details about diseases and limitations in activities of daily living can be included in the model and analyzed further. Starting from the information in the 6th wave of SHARE collected in 2015, the *DyMH_LU* model simulates individual and aggregated health outcomes from 2017 until 2045.

This paper is organized as follows. *Section 2* gives an insight into previous attempts made to simulate health and health-related issues by means of microsimulation modelling. *Section 3* briefly describes the background literature on health and its determinants in Luxembourg that is used to construct the *DyMH_LU* model. *Section 4* presents the main characteristics of our model, including a brief description of the database used in the model. *Section 5* explains how the evolution of temporal variables is controlled for more stability and consistency with external macro-based projections or experts' previsions. *Section 6* gives a preview of simulation outcomes regarding the evolution of the

---

[3] The "MiDLAS" Project ("*Towards a dynamic Microsimulation model, administrative Data for microsimulation in Lxg, the comparative Analysis of tools pertaining to the economics of ageing and a better understanding of Stakeholders' perception*"). See Genevois and Liégeois (2015).

[4] *Survey of Health, Ageing and Retirement in Europe* (http://www.share-project.org).

health status of residents in Luxembourg who are aged 51 years or over. Concluding comments are presented in *Section 7*.

2. SIMULATING HEALTH AND HEALTH-RELATED ISSUES IN THE LONG RUN

During the last four decades, several simulation models have been developed to address different research issues. In particular, the microsimulation technique has been used for studying the consequences of an ageing population on: social security pensions (Dekkers, Conti, Desmet *et al.*, 2018, among others); the need for formal and informal age care (Nepal, Brown, Kelly *et al.*, 2011); government health expenditures (Bryant, Teasdale, Tobias *et al.*, 2004). Other studies use microsimulation to develop different policies for ageing and health and to evaluate different policy options to minimize the negative fiscal impact of an ageing population (ACIL Tasman, 2003).

As highlighted previously, health status is not always the central focus of most of these analyses and models (Spielauer, 2007). Some examples of models including health-related issues are: the *CORSIM* model (Strategic Forecasting, 2002) which makes use of four health risk factors: smoking, alcohol consumption, sugar consumption and diabetes. The *DYNASIM* "*Dynamic Simulation of Income Model*" (Orcutt, 1957) measures the health status by the number of limitations on activities of daily living. *DYNAMOD*, a "*Dynamic microsimulation model of the Australian population*" (King, Baekgaard and Robinson, 1999), introduces both disability and a mortality function. Four disability states are introduced ranging from no disability to severe disability in the *LifePaths* model, which is a dynamic microsimulation model developed at the Canadian Statistical Office (Statistics Canada, 2002) while the model of Bryant, Teasdale, Tobias *et al.* (2004) uses the age variable as a proxy for the distance to death.[5]

Nevertheless, a limited but growing number of dynamic microsimulation models mainly focus on health and health care expenditure. For example, Brown, Nepal, Booth *et al.* (2011) present the *DYNOPTASIM* model which is designed "…to simulate changing population characteristics and identify how broader socio-demographic, behavioral and biomedical factors impact on the progression

[5] The reader can refer to Spielauer, 2007; Gupta and Harding, 2007; Légaré and Décaire, 2011; Thurecht, Brown and Yap, 2011, for more examples.

of four major age-related disabilities: Cognitive impairment and dementia; mental health (depression), sensory impairment and mobility impairment. …" (pp. 7). However, this microsimulation model only simulates separated health outcomes rather than the overall health status of individuals.

A more recent example is the *Population Health Model* (*POHEM*) which is a dynamic microsimulation model initially developed by Statistics Canada as a sub-model of the *LifePath* model. *POHEM* allows to represent the lifecycle of the Canadian population. In particular, it simulates the evolution of a set of diseases (heart disease, diabetes, osteoarthritis and three different types of cancer), co-morbidities and the influence of risk factors (smoking, body mass index, cholesterol, blood pressure and mortality).

The *Future Elderly Model* (*FEM*) is a demographic and economic microsimulation model which covers the entire population aged 50 and over in the United States and ten European countries (Goldman, Shang, Bhattacharya *et al*., 2005). It simulates the health status, the use of medical resources as well as the work income, the labor supply and the retirement decisions of individuals. This model has been developed as an ex-ante policy evaluation tool used to produce realistic future projections of "status quo" trends and to test "what if" scenarios related to potential policy and program interventions.

As for the *COMPAS* health microsimulation model, it covers the population aged 30 and over in Quebec (Boisclair, Côté-Sergent, Duclos *et al*., 2015). This model simulates different dimensions of the health status of individuals, mortality and the use of medical resources related to the state of health of the simulated individuals.

The *DyMH_LU* model proposed in this paper is an exclusive health model aimed at simulating the global health status of the population living in Luxembourg aged 51 or over. It simulates more than 60 different diseases and limitation in the activities of daily living, which makes this model unique compared to other similar models such as *FEM* or *COMPAS*. The second innovation of this model is a double one. First, the way we compute both the overall health at individual level and the global health of the whole population in every simulated period is innovative in this kind of model. Second, the interaction between the impact that the present health status of an individual can have on their health status in the following period. This individual overall status may, in turn, play a role in several dimensions of health, still at the individual level, during the subsequent period, which is indeed another specificity of the present model. This is made more explicit in the next *Section*.

3. THE HEALTH STATUS OF THE LUXEMBOURGISH POPULATION AND ITS DETERMINANTS

One of the main characteristics of the *DyMH_LU* miscrosimulation model is that it simulates 61 different items of health (diseases and limitations on activities of daily living). Thus, computing the health status of individuals is one of the essential points in this model. All the health items simulated allow for the computing, for each period of time, of the overall health status of each individual, the global health status of the entire population, as well as unidimensional indicators for each dimension of health, or in other words, the marginal distribution of each health outcome among the total population. In order to do this, we follow the methodology proposed by Pi Alperin (2016) and construct, for each period, synthetic indicators of health that aggregate several dimensions reflecting different aspects of the mental and physical health status of individuals (see *Table 1*).

Table 1: Health dimensions considered in *DyMH_LU*

Global health	Dimensions of health	What is covered by the dimension of health?
Mental health	Depression	Twelve different aspects of symptoms of depression
	Memory	The ability of individuals to remember things (5 tests)
Physical health	Long term illness	Having one or more chronic illness (17 illnesses considered)
	Limitation on activities 1	Concerns mobility problems such as walking 100 meters (10 limitations)
	Limitation on activities 2	Difficulties with one or more basic daily activity (13 limitations)
	Eyesight	Having vision problems (distance and reading)
	Hearing	Quality of hearing (with or without hearing aid)

Source : Pi Alperin (2016)[6]

These indicators, based on the fuzzy set theory, allow to take into consideration different intensities of illness. In other words, for each dimension of health, there are healthy individuals, completely non-

[6] See *Annex 1* for a complete description of the health items included in the *DyMH_LU* model.

healthy individuals and individuals characterized by different degrees of health failure. In practice, the synthetic scores are calculated as a weighted mean of all the dimensions of health. The weight takes into consideration the existing correlation between the different dimensions.[7]

Another important issue when considering simulating health over time is to analyze what the determinants of health are. We based the selection of variables to be included as determinants on the work of Deutsch, Pi Alperin and Silber (2017) and Mussard and Pi Alperin (2016) on health inequality in Luxembourg. More specifically, the authors calculate the contribution of *circumstances* (which are beyond an individual's control), and the *effort* and *lifestyle* of the individual (which is assumed to be at least partly the consequence of personal choice) to inequality in health using Luxembourgish data.

Following their ideas, we have selected a number of variables to be included in our model that could explain the health status of the population and which reflect both the efforts and lifestyles (smoking, drinking alcohol, etc.) and the individual's circumstances (the individual's nationality, the nationality of parents, the longevity of parents, etc.). The list of the determinants simulated in this model will be presented in *Section 4.2*.

4. THE MODEL

As explained in *Section* 1, the analytical framework chosen in the present paper for the analysis of the evolution of the health status in Luxembourg is derived from basic items of health (reflecting different aspects of the mental and the physical dimensions of health), as well as, through aggregation of these items, an overall health status for each individual and a global one for the entire population for different periods of time.

Moreover, not only do the averages matter, but the distribution of such dimensions, as well as the identification of winners and losers if a new policy is implemented or the environment is modified (e.g. the causal chain between personal characteristics and risk of disease), are also taken into account.

[7] Within the *DyMH_LU* model, all the health synthetic indicators are computed using the *MDEPRIV* program (see Pi Alperin and Van Kerm, 2009).

This complexity of interactions and the need for an overview of distributive outcomes makes the microsimulation approach of relevance for such an analysis. Furthermore, the link between the underlying forces and parameters, and the individual and overall outcomes, is made explicit, which helps in validating the general procedure and analyzing the main drivers at stake. Additionally, by working with a representative sample of the population under study, the microsimulation approach allows to deliver results that can be interpreted as representative for the entire population.

The following s*ub-sections* will present the reasons why a dynamic framework was retained for the *DyMH_LU* model, the database used as starting point of the simulations and its internal structure.

4.1 A dynamic modelling framework

We opt for a dynamic microsimulation modelling framework for three main reasons. First, we consider mid- or long-term analyses. Second, the hypothesis of an invariant population, or general environment, over such a temporal horizon is weak (e.g. in case of general ageing, or with respect to several risks). Finally, both the history of events, or status, and the parental "heritage" are important for the analysis to be undertaken.

Two different options were evaluated within the dynamic framework: *static ageing* and (pure) *dynamic ageing*. Static ageing is a method that attempts to match the relative weights of several sub-groups in the initial population with external aggregates (age distribution, smoking behavior, etc.), when the latter are "known" for future times. In contrast, the dynamic ageing framework changes the characteristics of the individuals directly period after period, rather than altering their weights.[8]

Even if static ageing allows to limit the problems of drift in the projected population, and would have been simpler to undertake for this model and given our objectives and conceptual constraints, the dynamic ageing option was chosen for three main reasons. First, there is some missing information about essential marginal distributions for many key-variables in the longer run, especially in relation to several health items. The dynamic, stepwise building-up of individual histories may be preferable,

[8] The methods are evoked and compared in Dekkers (2015), Li and O'Donoghue (2013) and Immervoll, Lindström, Mustonen *et al.* (2005).

benefiting from "behavior answers" and "risk evaluation" that can be econometrically estimated based on present data. The progressive building-up of individual items may also help in dealing with the chronic nature of several kinds of illness in an easier way (this will be made more explicit in *Section 4.4*). The second motivation – which is indeed closely related to the previous one - for a dynamic implementation is the fact of having an outcome which makes explicit all individual histories, giving the possibly to have a richer source for validation and understanding of the results, in the subsequent stage of analysis. The third reason is purely path-dependent. More precisely, we could build this health model from an existing model written earlier for pensions. In other words, the analysis of health here implemented is largely based, in technical and conceptual terms for non-health parts, on the dynamic microsimulation model *MiDAS_LU* that embeds a derivation of basic individual socio-economic characteristics over time. As in *MiDAS_LU* , *DyMH_LU* is developed to benefit from LIAM2[9], an efficient and open-source modelling and simulation platform which makes the code of the model particularly transparent and easy to read.

4.2 The starting point in time: a picture to be derived from existing individual data

The analysis of health outcomes requires detailed information covering individual characteristics related to age, socio-economic status, mental and physical health and limitation of daily activities. More specifically, our starting point is a picture of the Luxemburgish population which should be as close as possible to the "real" one and developed at the individual level.

In order to be able to have a comprehensive picture of the population, such detailed information should be available for all individuals. This is not always possible and especially when the concern is about health and all its relevant related items. In practice, we have two possibilities: we can build a model based on administrative data, in other words, fully representative data but with relatively few relevant detailed data embedded; or, we can use some databases that are more exclusively dedicated to concerns about health, even if they target only part of the population.

[9] The reader will find on http://liam2.plan.be a comprehensive documentation, the (open-source) package and several useful pedagogical examples.

In *DyMH_LU* we mainly use information from the SHARE "*Survey of Health, Ageing and Retirement in Europe*" database. SHARE is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of more than 120,000 individuals aged 50 or older (more than 297,000 interviews). SHARE covers 27 European countries and Israel. The main objective of the survey is to better understand the ageing process and to examine the different ways in which people aged 50 and over live in Europe. In particular, SHARE is interested in physical health (objective and subjective health status, health behaviors), mental health (well-being and satisfaction), cognitive functions (memory, calculation), health care consumption (hospitalizations, visit to a specialist, visit to a general practitioner, home care services request, retirement home), the economic situation (occupation, occupational characteristics, job opportunities, retirement, education, sources and composition of current incomes, consumption and wealth, housing), mutual social assistance (family assistance, income or property transfers, social networks, activities), and the family situation (origin, children, parents).

We focus our analysis on people living in Luxembourg. The Luxembourgish part of SHARE was conducted for the first time in 2013. Two waves are currently available for this country: Wave 5 (collected in 2013) and Wave 6 (collected in 2015). To implement our model, it is necessary to have data at a time *t* but also at a previous moment (here *t- 2*). The input basis of the model therefore concerns the panel part of SHARE-Luxembourg database that contains all the individuals who have participated in both waves. Given that the waves are separated by two calendar years, our simulation proceeds the same way and produces an outcome every second year (which is a single so-called "period" in the model).

For logical reasons, we build our input dataset from the information in the SHARE survey. Therefore, we firstly expand the survey data[10], based on weights attached to each individual in the survey. In total, and taking into account those individuals' weights, our input database is composed of 179 879 individuals aged 51 and over.[11] More specifically, 52% of the population is female and 48% is male. Secondly, variables are renamed and their composition re-shaped for an easier and more efficient simulation procedure.

---

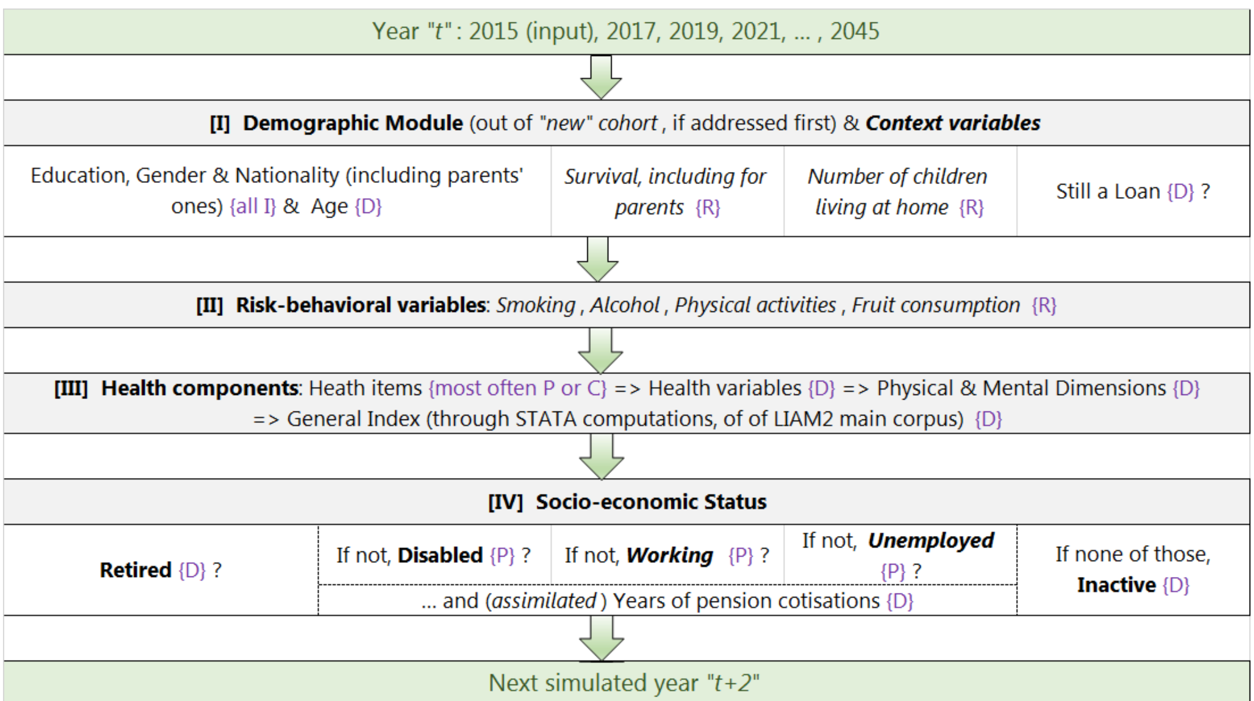
[10] "EXPAND" command, in STATA.

[11] The mention of a 51-year old cohort (rather than 50) will be clarified in Footnote *12*.

4.3 Structuring the simulation

The dynamic microsimulation model *DyMH_LU* is a first operational tool. As explained previously, its main objective is to address the question of the evolution of the health status of the Luxembourgish population which is steadily growing old. To do so, we compute over time different sets of variables (on health, socio-economic status, demography) that could have a direct or indirect impact on individual health status. In practice, *DyMH_LU* works in a similar way to the other socio-economic dynamic models. In particular, it starts in period *t* from a detailed and representative picture of the population at the individual level. Then, several additional so-called "simulated" characteristics are derived. These in turn lead us to a new picture of the population for the next period (year *t+2*). Starting from this again, the population in *t+4* is generated, and so on.

Figure 1: The Main envelope of *DyMH_LU* model



Legend: {I} Invariant / {D} Deterministic / {R} Purely random / {P} Probabilistic (logit or ordered logit) / {C} Continuous variable - Alignment
Source: *DyMH_LU*, version 226 (LISER, 2018)

Figure 1 lists all the actions to be undertaken, year after year, on individuals when carrying out the simulation with *DyMH_LU*. Specifically, starting from a known population (cf. input data for the year 2015), from the first period of simulation (2017 on), and for each person, we raise a series of questions relating to demography (row [I]: age, gender, nationality of the individual and their parents, educational level of individual and their parents, number of children in charge, financial situation at the age of 12, having a loan); the attitude towards a few specific health-related determinants (row [II]: smoking, drinking, physical activity, consumption of fruit and vegetables; body mass index); the health components (row [III][12]); and the socio-economic status (row [IV]: unemployed, pensioner, employment in private and public sector, invalid). *Figure 2* shows how the conceptual representation in *Figure 1* is translated into LIAM2 processes in a way that can be easily interpreted, an important characteristic of the model.

Of course, the order of the processes is relevant and its choice was part of the modelling process. In particular, the socio-economic module is set at the end of the periodic analysis, given that such a status was considered as strongly dependent on health items (and less vice versa). These latter are derived from risk-behavioral variables (e.g. attitude to smoking) and, still more basically, from demographic factors (e.g. education level).

Two different frameworks for *DyMH_LU* are proposed. First, the closed-population case, which implies that no new person is added to the initial sample as the ageing process is taking place, implying a progressive shrinking of the population over time. The second framework is the open-population one. In this case, new individuals are generated for each period of simulation, replacing the generation which is "disappearing" due to ageing. Specifically, the youngest one involves the 51- and 52-year-old cohorts.[13] Moreover, in order to take into account the announced change in size for this cohort, period after period, the number of individuals newly added is derived from external demographic sources (AWG projections, European Commission, 2012). The new members are cloned and randomly selected from existing ones observed during the period before and the simulation goes one step further, as in the "closed-population" case. This methodology makes the "new young" (51- and 52-year-olds) evolve over time (for example in terms of behaviors and health), compared to the corresponding cohorts as observed in the raw data (i.e. in 2015). However, it is expected that some relevant dimensions, for example nationality, will not properly evolve over time.

12 See Annex 1 for a full description of the health items included in *DyMH_LU*.

13 The individuals of 50 years old were not considered in the model as only a few individuals were observed in our database.

Figure 2: The "DyMH_LU" main Simulation Block in LIAM2 – An extract (*)

```
simulation:
    init:
    processes:

        - person: [

            ####
            # DEMOGRAPHY
            #############

            new_population,

            year,

            D_Age,
            …
            D_Father_Age,
            D_Mother_Age,
            …
            class_age_a,
            …

            alive_female,,
            alive_male,
            death_procedure,
            alive_parents,
            age_dead_parents,
            kill_dead_parents,

            nb_children_at_home,
            Loan,

            ####
            # HEALTH
            #########

            Physical_Activity,
            Alcool,
            Fruit,
            Smoking,

            Health_Components_Regressions,    # e.g. H_BMI
            Health_Components_Logits,         # e.g. H_Mind_Irratib / H_OthIll_Diabet
            Health_Components_OLogits,        # e.g. H_VisioFar
            Health_Composed_Variables,        # e.g. C_Depression

            ####
            # SOCIO-ECONOMIC STATUS & AFTER
            ################################

            Status_Pensioner,
            Status_Invalid,
            Status_Working,
            Status_Unemployed,
            Status_Miscellenaous,

            Pension_Cotisations,

            ]
```

(*) The processes involving an alignment (see *Section 5.1*) are outlined.
Source: *DyMH_LU*, version 226 (LISER, 2018).

4.3 Addressing transitions at the individual level

Putting aside invariant characteristics (e.g. gender) and a few continuous random ones (body mass index) that individuals face during their life path "transitions" from one status to another one (e.g. parenthood, illness, unemployment), the present modelling framework takes transitions as deterministic (e.g. age, retirement), purely random (death, smoking) or probabilistic (at work, diabetes). Probabilistic events may also be complementarily chronic, i.e. maintained at the individual level as soon as experienced earlier (health aspects such as Parkinson's disease). *Figure 1* gives an insight into the type of variable encountered: invariant, deterministic, purely random, probabilistic (logit or ordered logit) and continuous (regression).

The classification of an event in a specific type of transition (e.g. purely random rather than probabilistic), and additional constraints implemented, are obviously a matter of choices made by the modeler. The limitation in available information from input data or external sources can explain a specific classification. The technical difficulty to operationalize richer interactions may be another reason.

It is also worth mentioning that the necessary hierarchy of the different modules of the process modelling (emphasized in *Section 4.3*) implies that, at each level, the value of a variable can only be fixed based on: its value during the period(s) before; the contemporaneous value of other determinants if set earlier in the simulation for the same period; alternatively, the past values of such determinants; and/or other external parameters. This requirement must be kept in mind while proceeding with the econometric analysis needed for feeding the behavioral equations.

In the next section, we present two important strategies implemented in *DyMH_LU* and, in particular, the different ways of controlling the temporal evolution of the variables included and simulated in this model over time.

5. CONTROLLING FOR THE TEMPORAL EVOLUTION OF VARIABLES

Building-up a dynamic microsimulation model is a demanding task whose added value can be to provide the analyst both with more accurate information about possible distributional outcomes and more room for a better understanding of the underlying processes, taking into account the complexity of processes and a large number of persons over time.

Therefore, one of the main areas of added-value of dynamic microsimulation concerns individual outcomes and the evolution of their marginal distributions over time. However, microsimulation as such cannot tell the whole story. Alternative prospective approaches can be of interest, based on meso- and macro-modelling or external sources of information (including from experts). Moreover, the econometric risk - or behavioral - analyses feeding the microsimulation model (provided that it is even available) can be of poor explanatory value, e.g. when the quality of raw data made is limited given the objective, or when the time made available for development is counted. Therefore, we must keep in mind the necessity of a trade-off between resources mobilized and objectives in terms of analysis of distributions, on the one hand, and between outcomes, as freely resulting from the complex simulation undertaken and aggregate indications derived from external analyses and experts, on the other.

Such a trade-off may imply some additional control with respect to aggregate outcomes or the way individual histories and transitions are governed. The former types of controls are called *alignments*. The latter are more specific to the type of analysis undertaken, which is health-related. In what follows, both types of controls will be explained in more details, and the strategies retained will be made explicit.

5.1 Controlling the temporal evolution of variables through alignment

The objective, and nature, of the alignments is to ensure that output aggregates (e.g. share of population at work during the year 2031) can match with external macro-aggregates.[14] The alignments implemented in *DyMH_LU* with regard to a binary variable[15], the "*Sort By predicted Probability*", is carried out in two main steps. First, for a particular person and variable, the risk of facing a specific transition is determined based on behavioral equations previously estimated. Of course, this risk can include a

---

14 For a discussion about the methodological choices and different techniques see Dekkers and van Leeuwen (2010) and Li and O'Donoghue (2013).

15 Other alternatives are available, the general topic of alignment to external sources being a matter of debate among the developers.

random component. Second, the individuals are ranked according to their level of risk and only part of them are selected for the transition in question, the proportion of individuals being determined by external sources.

Figure 3: An extract of the *DyMH_LU* "Status_Working" process (*)

```
Status_Working:

            # LOGIT-TYPE TRANSITION, ALIGNED


                # INITIALIZE ...
                #################


                - score_value:  "-1000.0"
                - temp_1_float:  "0.0"
                - to_work: "False"

                # ... THEN PROCEEDING



                # SCORE FOR "FEMALES"

        …

                # SCORE FOR "MALES"

                - temp_1_float: 0.0939835 * D_NbYearsCot -1.993645 * C_LimAct12 + 0.3640511
                - score_value: if( MALE  and  not(PENSIONER)  and  not(INVALID) ,
                                   logit_score( temp_1_float ) ,
                                   score_value)

                # ALIGNEMENT FOR ALL

                - to_work:  align( score_value,
                                   (AL_ACTIVE_POP - AL_UNEMPLOYED) / AL_POPULATION,
                                   leave = ( not(AGE_ACTIVE) or
                                             PENSIONER or
                                             INVALID
                                           )
                                 )

                - D_Working: if( AGE_ACTIVE  and  not(PENSIONER)  and  not(INVALID) ,
                                 to_work,
                                 D_Working
                               )
```

Source: LISER, *DyMH_LU*, version 226 (LISER, 2018)

(*) the example shown here is indeed simplified, compared to the original code which takes into account, on top, the individual past working status

Such an alignment procedure is here illustrated with the example of the so-called *Working Status* process. The objective of the process is to determine whether an individual "selected" as neither retired nor invalid in *t-2*, can be considered as working or not in period *t*. *Figure 3* shows part of this alignment procedure, as implemented in LIAM2. In particular, the probability for a man to work during period *t* is given by an econometric analysis. Specifically, this probability is here positively correlated to the number of worked years up to period *t* (D_NbYearsCot) and negatively correlated with some limitations in daily living activities (C_LimAct12). Then, the alignment itself is based on the following proportions:

(AL_ACTIVE_POP - AL_UNEMPLOYED)/AL_POPULATION (1)

where all "AL_" expressions refer to exogenous parameters in the form of matrices/multidimensional tables mostly referring to numbers of persons; "AL_ACTIVE_POP - AL_UNEMPLOYED" being thus the matrix indicating the number of working individuals ("active" but not "unemployed") to be targeted, divided by the total population.

In *DyMH_LU*, the "macro" bases used for alignments are mainly (but not only) related to the European *Working Group on Ageing Populations and Sustainability* (AWG) economic and budgetary (including age-related) projections, hereafter referenced as "AWG Projections". In the present implementation of the model, the AWG projections are those of 2012, which cover the period 2010-2060 (European Commission, 2012).

All the tables used in the dynamic models cross (giving specific outcomes depending on the combination) the gender, the class of age and the period under consideration. Based on those factors, a proportion is derived which will govern the number of persons who will have to face the transition under consideration for the given period of simulation. As has been shown in the previous example of the *Working Status* process, the proportion to be extracted from external sources can itself result from some matrix-like computation, e.g. equation (1).

5.2 Controlling the temporal evolution of variables through thresholds, stepwise transitions and chronicity

In addition to the alignment procedure, other methods have been implemented in our model to control the evolution of certain variables over time. The first one is establishing *individuals' thresholds*. More precisely, the probability of a person having a particular health status is predicted using logit models. Then, a change in the status, e.g. contracting a specific disease, occurs when the predicted probability falls above a specific threshold. The threshold is selected in order to minimize a loss function[16] that is an equally weighted linear combination of classification errors (i.e. false positives and false negatives[17]). Therefore, compared to arbitrary selected thresholds, such an "optimal" selection procedure allows to minimize classification errors that would otherwise have been accumulated with each round of the simulation.

The second method implemented in the *DyMH_LU* model is to limit the jump between two different values for a given (random) variable. In other words, this means that for one variable with three or more potential modalities, the individuals can only change from one modality to the next one or the previous one. It is not possible to jump two modalities at once between two periods. For example, the variable "*number of children still at home*" for those families that still have charge of children, not considered as stable through time, is constrained to change by steps of maximum one unit for each period and is additionally enforced to decrease over time. Another example is the variable "*consumption of fruit*". This variable has three modalities: low, medium and high. Then, the variable is allowed to jump progressively from one modality to the closest ones (from "*low*" to "*medium*", or "*medium*" to "*high*" but not "*low*" to "*high*" at once).

The third and last method of control is to establish some chronic variables. We have to pay special attention to diseases that "intrinsically" persist for a long time: these are chronic. A chronic disease is one that lasts three months or more, cannot be generally prevented by vaccines or cured by medication, nor just disappears. This means in our model that as soon as one individual becomes ill in this way in year *t*, he/she will be ill for all subsequent periods of simulation. The list of chronic, and non-chronic, diseases included in the present model are exposed in *Annex 2*. In contrast, for non-chronic diseases, the risk of change in status with respect to that specific item (becoming sick, recovering from illness, or change in level) was computed using logits or ologits methods. Those risks can also depend on past values of health items, including the overall health status of the person as determined one period before, which a specificity of *DyMH_LU*.

Yet, the variable taking renal problems into consideration has been treated differently to the other

---

16 This is a commonly used technique for the development of early warning systems. See for example, Detken, Weeken, Alessi *et al.* (2014), Candelon, Dumitrescu and Hurlin (2012) and Kaminsky and Reinhart (1999).

17 The false positives are the number of persons wrongly classified as having the disease and false negatives account for those wrongly considered as healthy.

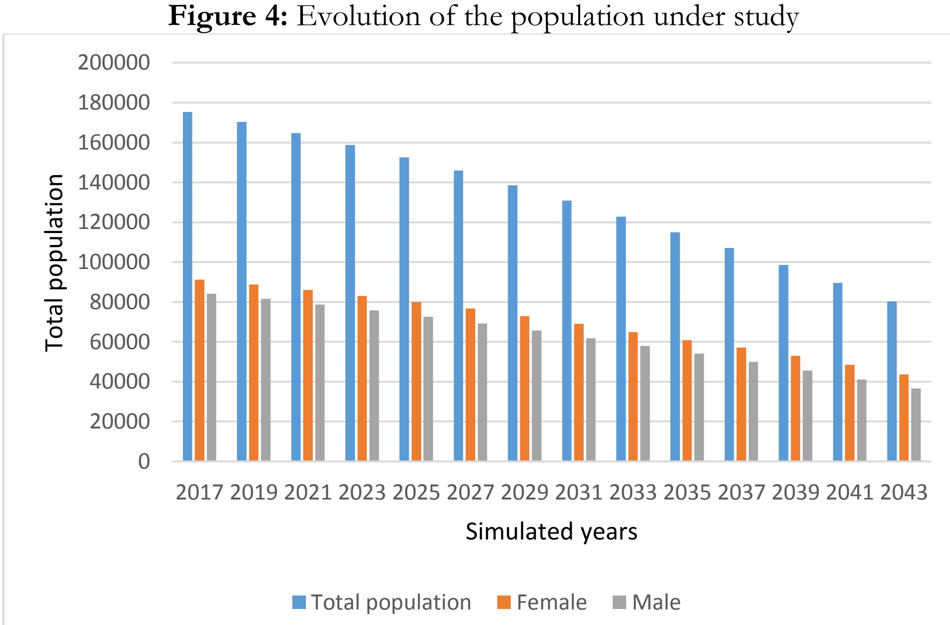
chronic variables. Indeed, this variable was introduced in the SHARE questionnaire only in Wave 6. It was not therefore possible for it to be considered and treated as chronic in the first period of simulation ($t = 2017$). After simulating the first period, we had enough information to treat this disease as a chronic one.

6. RESULTS

This section presents some results from our dynamic microsimulation model, *DyMH_LU*. In particular, we show the evolution of the global health status, the mental dimension of health and some specific health items from the model in its "closed population" version. In other words, no new individuals are introduced during the periods of simulation to compensate for the loss of individuals over time. We will analyze the health evolution only for those individuals that are aged 51 and over in 2015, 53 and older in 2017, and so on. Specifically, we only consider in our simulation those persons included in Wave 6 of SHARE.

Figure 4 shows the evolution of the population between 2017 and 2043. From a cohort of 175 380 individuals aged 51 years and over in 2017, the population size decreased to 80 306 individuals aged 77 and over in 2043.

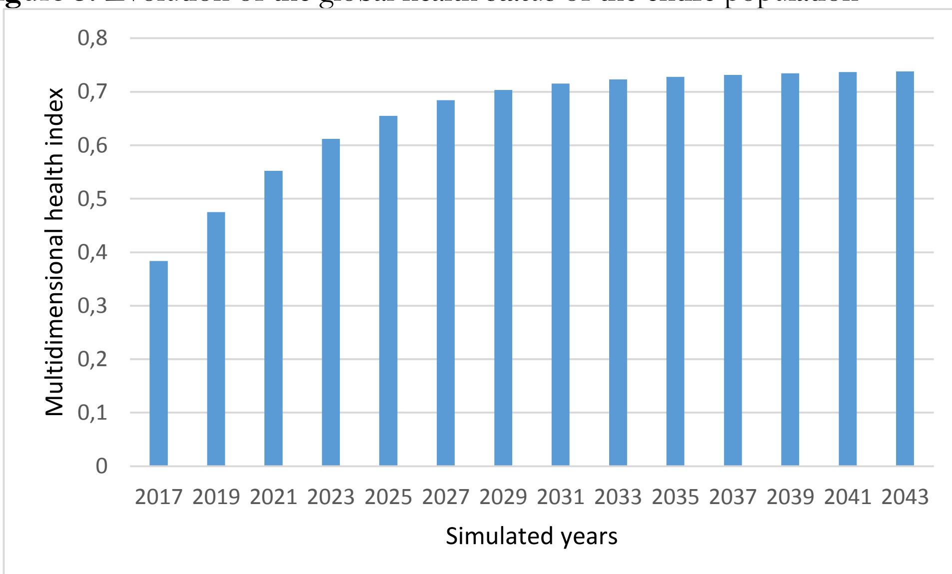
Figure 4: Evolution of the population under study



Source: *DyMH_LU*, version 226 (LISER, 2018)

As explained previously, for each period of time, we have constructed synthetic indicators of health that aggregate seven dimensions. These scores can have a value from 0 (a completely healthy population) to 1 (a completely unhealthy population). Between both values we have different intensities of global health. As we can see in *Figure 5*, the score value goes from 0, 38 to 0, 74 between 2017 and 2043, respectively. This result was expected as the individuals are younger and healthier in 2017 compared to 2043.

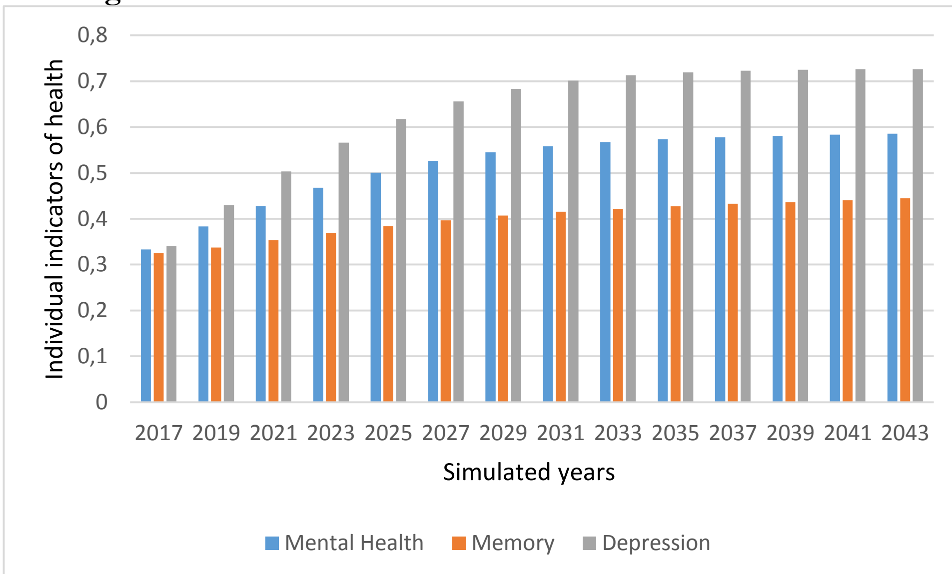
Figure 5: Evolution of the global health status of the entire population



Source: *DyMH_LU*, version 226 (LISER, 2018)
Note: "0" = completely *healthy* population / "1" = completely *unhealthy* population

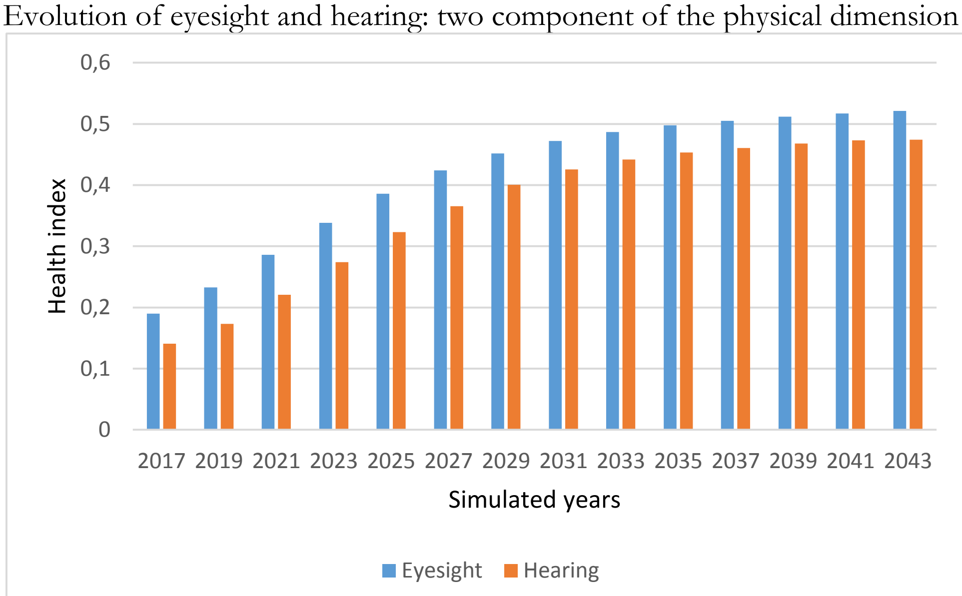
Figure 6 shows the evolution of the mental dimension of health and its components. This dimension is composed of two health *items*: memory and depression. If we consider the mental health score, it goes from 0, 32 to 0, 59. In 26 years, the score increases by 84% which is qualitatively consistent with the fact that the individuals have higher probabilities of memory loss and of becoming more depressed with age. Yet looking individually at these two components, it is also remarkable that while the memory index is growing by 38% during the total simulated period of time, the depression indicator more than doubles (115%). This implies that depression in ageing people is a very important issue to take into consideration when discussing the health problems associated with the third age, more so than memory problems.

Figure 6: Evolution of the mental dimension of health



Source: *DyMH_LU*, version 226 (LISER, 2018)
Note: "0" = completely *healthy* population / "1" = completely *unhealthy* population

Figure 7: Evolution of eyesight and hearing: two component of the physical dimension of health



Source: *DyMH_LU*, version 226 (LISER, 2018)
Note: "0" = completely *healthy* population / "1" = completely *unhealthy* population

Finally, concerning the physical dimension of health, *Figure 7* presents the evolution of two of its components: eyesight and hearing. These two problems are very present among the elderly. Their evolution is shown through the simulation to be rather similar across the whole period (even if the eyesight problem is more important than the hearing one). It is important to note that simulation outcomes can only suggest a possible path, given hypotheses and constraints embedded in the exercise.

7. CONCLUSION

The *DyMH_LU* model is the first operational tool to simulate the health status of the Luxembourgish population over time. This dynamic microsimulation model allows to forecast more than sixty different diseases and limitations in the daily activities. Thus, we can address the question of the evolution of the global health status of the population which is steadily growing old. One of the major characteristics of this model is that all these diseases cans be analyzed separately – e.g. the marginal distribution of Alzheimer's among the total population, – or aggregated in order to examine the mental and the physical dimensions of health separately, or the global health of the entire population, in every simulated period.

Future research - at the "extensive margin"- could be done by adapting this model to other countries that collect the SHARE database or other similar and harmonized databases. Another track for research – at the "intensive margin"- would be to deepen the analysis of global health and the contribution of the underlying components (including behaviors) throughout the timeline. Moreover, the composition of the new cohorts, all dimensions considered (demographic, behavioral, impact on basic aspects of health) would deserve some additional effort that is not specific to the present model. Finally, controlling the temporal evolution of key-variables through thresholds, stepwise transitions and chronicity is a matter of debate – not limited to the present model either - that might benefit from additional theoretical and modelling improvement. This is including a particularly demanding sensitivity analysis that will be a challenge given the number of items to be considered, separately or in combination.

ACKNOWLEDGEMENTS

This research is part of the HEADYNAP project supported by the National Research Fund, Luxembourg (contract FNR C12/SC/3977324/HEADYNAP/Pi Alperin) and by core funding from LISER from the Ministry of Higher Education and Research of Luxembourg.

We would like to show our gratitude to Gaston Andrés Giordana, for his helpful methodological support with respect to some concerns about threshold identification embedded in the present paper. Of course, any remaining errors, results produced, interpretations or views presented in this paper are exclusively the authors' own.

This paper uses data from SHARE Waves 5 and 6 (Börsch-Supan, 2017a and Börsch-Supan, 2017b), see Börsch-Supan, Brandt, Hunkler et al. (2013) for methodological details. The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N°211909, SHARE-LEAP: N°227822, SHARE M4: N°261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

REFERENCES

ACIL Tasman (2003). Using Microsimulation Techniques to Develop Policies for Ageing and Health. International Conference on Population Ageing and Health: Modelling Our Future Canberra, Australia, 2003, 8-12 December.

Boisclair, D., Côté-Sergent A., Duclos J-Y., Lekina A., Marchand S. & P-C Michaud (2015), COMPAS: Un modèle de microsimulation santé pour le Québec, Cirano, Série Scientifique 2015s-03.

Börsch-Supan, A. (2017a). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5*. Release version: 6.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w5.600.

Börsch-Supan, A. (2017b). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6*. Release version: 6.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.600.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S., & Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*. DOI: 10.1093/ije/dyt088.

Brown, L., Nepal, B., Booth, H., Pennec, S., Anstey, K., & Harding, A. (2011). Dynamic Modelling of Ageing and Health: The Dynopta Microsimulation Model, *NATSEM Working Paper*, No. 11/14, University of Canberra, Australia.

Bryant, J., Teasdale, A., Tobias, M., Cheung, J., & McHugh, M. (2004). Population Ageing and Government Health Expenditures in New Zealand, 1951-2051. *New Zealand Treasury Working Paper*, No. 04/14.

Candelon, B., Dumitrescu, E.-I., & Hurlin, C. (2012). How to Evaluate an Early-Warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods. *IMF Economic Review*, 60(1), 75-113.

Dekkers, G. (2015). The simulation properties of microsimulation models with static and dynamic ageing – a brief guide into choosing one type of model over the other. *International Journal of Microsimulation*, 8(1), 97-109.

Dekkers, G., Conti, R., Desmet, R., Sundberg, O., & Van den Bosch, K., (2018), What are the consequences of the AWG 2018 projections and hypotheses on pension adequacy? Simulations for three EU member states. *Federal Planning Bureau*, Brussels, REP_11732.

Dekkers, G., & Van den Bosch, K., (2016), Prospective microsimulation of pensions in European Member States. In Dekkers, G., & Mészáros, J., (eds), Applications of microsimulation modelling: a selection of papers presented during the 2016 European meeting of the

International Microsimulation Association in Budapest. Budapest: Társadalombiztosítási könyvtár.

Dekkers, G., & van Leeuwen, E., (2010), Editorial special issue on 'methodological issues in microsimulation'. *International Journal of Microsimulation*, (2010) 3(2) 1-2.

Detken, C., Weeken, O., Alessi, L., Bonfim, D., Boucinha, M., Castro, C., Frontczak, S., Giordana, G., Giese, J., Nadya, J., Kakes, J., Klaus, B., Lang, J.H., Puzanova, N., & Welz, P. (2014). Operationalising the countercyclical capital buffer: indicator selection, threshold identification and calibration options. *ESRB Occasional Paper Series, European Systemic Risk Board*, No 05.

Deutsch, J., Pi Alperin, M.N., & Silber, J. (2017). Using the Shapley Decomposition to Disentangle the Impact of Circumstances and Efforts on Health Inequality. *Social Indicators Research*. DOI: 10.1007/s11205-017-1690-5.

Dlugosz Z. (2011), "Population Ageing in Europe", Procedia Social and Behavioral Sciences, Vol. 19, pp. 47-55.

European Commission (2012). *The 2012 Ageing Report - Economic and budgetary projections for the 27 EU Member States (2010-2060)*, Brussels.

Genevois, A.S., & Liégeois, P. (2015). LuDMi - Dynamic Microsimulation Model for Luxembourg, Version 1. *Technical Report*, MiDLAS Project (EU-Grant VS/2013/0138), LISER, Luxembourg.

Goldman, D.P., Shang, B., Bhattacharya, J., Garber, A.M., et al. (2005). Consequences of health trends and medical innovation for the future elderly. *Health Affairs,* 24, W5R5.

Gupta, A., & Harding, A. (2007). *Introduction and Overview*. In Gupta and Harding (eds.), Modelling Our Future: Population Ageing, Health and Aged Care. Amsterdam: Elsevier.

Immervoll, H., Lindström, K., Mustonen, E., Riihelä, M., & Viitamäki, H. (2005). Static Data Ageing Techniques: Accounting For Population Changes In Tax Benefit Microsimulation Models. *EUROMOD Working Papers*, EM7/05.

Kaminsky, G.L., & Reinhart, C.M. (1999). The Twin Crises: The Causes of Banking and Balance-of-Payments Problems. *American Economic Review*, 89(3), 473-500.

King, A., Baekgaard, H., & Robinson, M. (1999). DYNAMOD-2: An Overview. *NATSEM Technical Paper*, No. 19, National Centre for Social and Economic Modelling, University of Canberra.

Légaré, J., & Décaire, Y. (2011). Using Statistics Canada LifePath Microsimulation Model to Project the Disability Status of Canadian Elderly. *International Journal of Microsimulation*, 4(3), 48-56.

Li, J., & O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure and methodology. *International Journal of Microsimulation*, 6(2), 3-55.

Mussard, S., & Pi Alperin, M.N. (2016). A Two-Parameter Family of Socio-Economic Health Inequality Indices: Accounting for Risk and Inequality Aversions. *Luxembourg Institute of Socio-Economic Research (LISER), Working Paper Series*, 2016-15.

Nepal, B., Brown, L., Kelly, S., Percival, R., Anderson, P., Hancock, R., & Ranmuthugala, G. (2011). Projecting The Need for Formal and Informal Aged Care in Australia: A Dynamic Microsimulation Approach. *NATSEM Working paper series*, No. 7, University of Canberra.

Orcutt, G. (1957). A New Type of Socio-economic System. *Review of Economics and Statistics,* 58, 773-797.

Rutter, C., Zalasvky, A., & Feuer, E. (2011), Dynamic Microsimulation Models for Health Outcomes: A Review. *Med Decis Making*, 31(1): 10-18.

Pi Alperin, M.N. (2016). A multidimensional approach to measure health. *Economics Bulletin*, 36(3), 1553-1568.

Pi Alperin, M.N., & Van Kerm, P. (2009). mdepriv – Synthetic indicators of multiple deprivation, v2.0 (revised March 2014). CEPS/INSTEAD, Esch/Alzette, Luxembourg.

Spielauer, M. (2007). Dynamic Microsimulation of Health Care Demand, Health Care Finance and the Economic Impact of Health Behaviors: Survey and Review. *International Journal of Microsimulation*, 1(1), 35-53.

Statistics Canada (2002). The LifePaths Microsimulation Model - An Overview. Statistics Canada, Ottawa.

Strategic Forecasting (2002). CORSIM: Analyst Documentation. Strategic Forecasting, Cornell.

Thurecht, L., Brown, L., & Yap, M. (2011). Economic Modelling of the Prevention of Type 2 Diabetes in Australia – The Diabetes Model. *International Journal of Microsimulation,* 4(3), 71-80.

Zucchelli, E., Jones, A.M., & Rice, N. (2010). The evaluation of health policies through microsimulation methods. *Health, Econometrics and Data Group Working Paper*, 10/03, University of York.

Annex 1: The list of all health items included in *DyMH_LU*

Global Health	Dimensions of health	Health items
Mental health	Depression	Depression; Concentration; Guilt; Loss of interest; Sleep; Irritability; Appetite; Stress; Pessimism; Suicide; Enjoyment; Tearfulness
	Memory	Orientation regarding: Date; Day of the week; Month; Year; and also their capacity to memorize a certain number of words
Physical health	Long term illness	Heart attack; Stroke; Cancer; Ulcer; Cataract; Fracture of the femur; Other fractures; Rheumatism; Hyper-tension; High-cholesterol; Diabetes; Pneumonia; Parkinson's; Alzheimer's; Anxiety; Arthrosis; Renal problems
	Limitation activities 1	Sitting for two hours; Climbing stairs; Reaching; Picking up a coin; Walking 100 meters; Getting up; Climbing several stairs; Kneeling; Pushing objects; Carrying weight
	Limitation activities 2	Getting out of or into bed; Using the toilet; Making phone calls; Dressing; Walking; Bathing; Eating; Using a map; Preparing a meal; Shopping; Taking medication; Cleaning; Managing money
	Eyesight	Farsighted; nearsighted
	Hearing	Hearing difficulties

Annex 2: The list of chronic and non-chronic diseases included in *DyMH_LU*

Non-chronic diseases		Chronic diseases	
Long term illnesses	Heart attack	**Long term illnesses**	Hyper-tension
	Stroke		High-cholesterol
	Cancer		Diabetes
	Ulcer		Pneumonia
	Cataract		Parkinson’s
	Fracture of the femur		Alzheimer’s
	Other fractures		Anxiety
	Rheumatism		Arthrosis
Limitation activities 1	Sitting two hours		Renal
	Climbing stairs	**Limitation activities 1**	Walking 100 mts
	Stretching arms		Up from chair
	Picking up a coin		Climbing several stairs
Limitation activities 2	Getting out of/into bed		Kneeling
	Using the toilet		Pushing objects
	Making phone calls		Carrying weight
Weight problems	BMI	**Limitation activities 2**	Dressing
Eyesight	Farsighted vision		Walking
	Nearsighted vision		Bathing
Depression			Eating
	Depression		Using a map
	Pessimism		Preparing a meal
	Suicide		Shopping
	Guilt		Taking medication
	Sleeping		Cleaning
	Interest		Managing money
	Irritability	**Memory**	Date
	Loss of Appetite		Month
	Stress		Year
	Concentration		Day of the week
	Enjoyment		
	Tearfulness		
	Words		
Hearing	Hearing		

LUXEMBOURG INSTITUTE OF
SOCIO-ECONOMIC RESEARCH

11, Porte des Sciences
Campus Belval
L-4366 Esch-sur-Alzette

T +352 58 58 55-1
F +352 58 58 55-700

www.liser.lu