REPORT

Theoretical foundation of the dynamic general equilibrium model with overlapping generations, heterogeneous households and family networks

November 2023

Miguel Sanchez-Romero
David Zettler
Alexia Prskawetz

FutuRes

Towards a Resilient Future of Europe

Document History		
Version	Date	Comments
1.0	30.11.2023	
2.0	25.01.2023	Correction of the retirement decision, Inclusion of additional explanatory figures and diagrams as suggested by Jakub Bijak and Emily Barker

Funded by
the European Union
Grant Agreement n° 101094741

The FutuRes project has received funding from the European Union's Horizon 2022 Research and Innovation Programme under Grant Agreement No 101094741.

Contents

1 Introduction

The aim of the technical note is to explain the set up of the General Equilibrium model with Overlapping Generations (GE-OLG) that will later be calibrated to historical data for selected European countries and applied to quantify the resilience of various pension reforms.

Our model is comprised of three economic agents (cf. Figure 1): households (blue circle), firms (green circles), and a government (gray circle). To include heterogeneous households, we build on Sánchez-Romero et al. (2023), where agents initially differ by learning ability and effort of schooling (red cloud). These two initial characteristics produce heterogeneity in education choices, labor force participation, income, wealth and retirement age. In addition to the initial characteristics, we extend the model of Sánchez-Romero et al. (2023) by introducing stochastic shocks along the life cycle. In particular, individuals face unemployment, health and mortality risks that are recurring over the life cycle of individuals (as indicated by the purple circles contained in the yellow area on the left-hand side in Figure 1), and a one-time shock of fertility (purple circle) that determines the family structure. Education, labor force participation and retirement choice are endogenously determined by our heterogeneous households and will be affected by these life cycle shocks. Hence, our model allows for the fact that households are heterogeneous in terms of their life cycle decisions, which in turn implies that they will be differently affected by private and public policies and will also differently react to the various life cycle risks. In addition, our framework not only allows to better quantify distributional aspects at the macro economic scale (income distribution, etc.), but also allows to investigate how various welfare reforms will affect different groups in our society.

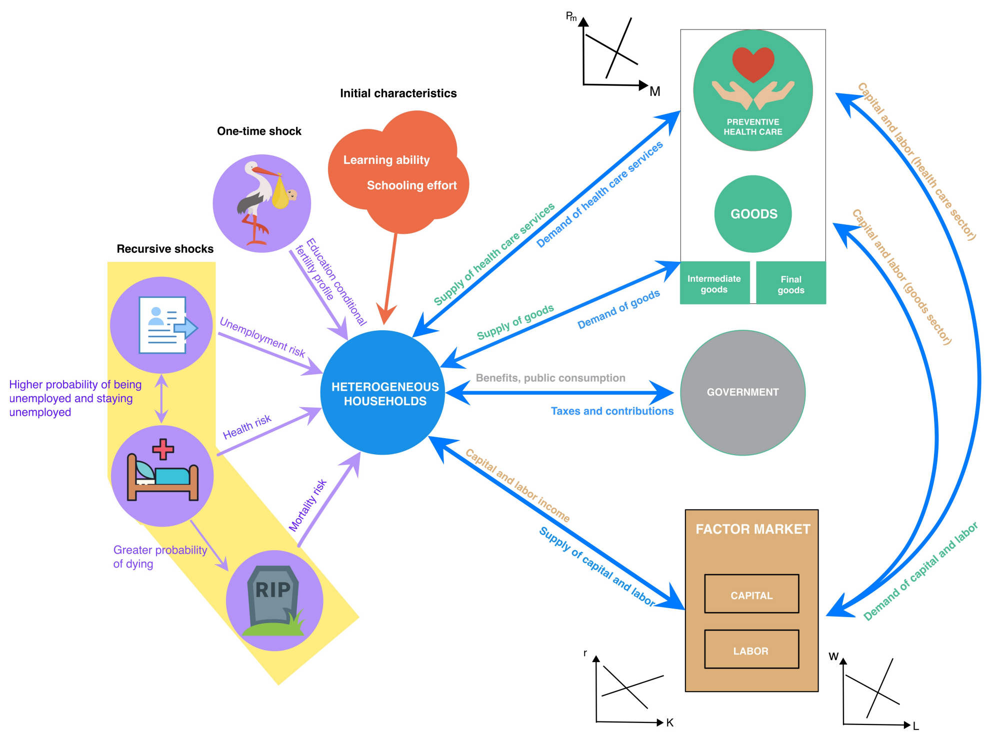

Figure 1: Flow diagram (macro level).

Note: Our economic model contains three types of agents: households (blue), firms (green), and government (gray). Households are heterogeneous based on initial differences in learning ability and schooling effort (red cloud). Moreover, households are affected by recurrent stochastic shocks, including unemployment, health, mortality risks (purple circles), and a one-time fertility shock (purple circle). Households will make endogenous decisions on education, labor force participation, capital accumulation, consumption of final and intermediate goods, health care, and retirement. These decisions will differ depending on the initial heterogeneities but also depend on the shocks experienced throughout the life cycle. Prices on wages, interest rates, and health care will be set in competitive markets.

Individuals are constrained by their income and time when they optimally choose their consumption, leisure, childcare, home production and medical care along their life cycle. The household framework is explained in detail in Section 3.[1] Households supply labor and capital in exchange for capital income and labor income (as indicated by the blue arrow pointing from the household to the factor market in Figure 1). Households have to pay taxes and social contributions to the government, out of which the government finances public benefits and public consumption (as highlighted by the blue arrow between the household and government sector in Figure 1). Households and firms are also connected through the supply and demand of goods (intermediate and final) and health care services (indicated by the two arrows connecting the household and the firms). We choose the price of the final good as the numeraire (i.e. its price is equal to one). The price of health care services P_M relative to the price of the final good sector will be determined in the market as indicated in the price-quantity diagram to the left of the health care sector with M denoting the amount of health care services.[2] The arrows on the right-hand side depict the link between the demand of capital and labor by the production sector of final goods and health care services and the supply of these factors in the factor market. In equilibrium supply and demand of capital (K) and labor (L) will be equal and determine the prices of labor (wages w) and capital (interest rate r) as indicated in the price-quantity diagrams to the left and right of the factor market box.

Based on the theoretical setting summarized in Figure 1, our model will allow us to investigate various research questions raised in the FutuRes project, for instance:

- Which groups in our society are more resilient to labor market and health shocks?
- How and by which support systems can we best target those groups that are most vulnerable to these shocks?

Moreover, the model can be used for studying:

- Ageism. In particular, we can assume that firms discriminate towards younger and older workers by lowering their chances of being hired.
- How resilience is related to various family types.
- How different transition probabilities of employment status and health status may impact on the economy (sensitivty analyses).

We have organized this report into six additional sections. Section 2 introduces the notation for demographic variables and demonstrates how to account for unobserved heterogeneity when it influences the survival of individuals. It continues by presenting the population balancing equation for each gender and explaining how the kinship structure for each individual is constructed. Section 3 provides the micro-foundation of the economic model. It details the idiosyncratic risks faced by individuals, their inter-temporal budget constraints, and their optimal decision-making process. Section 4 specifies the budget constraints faced by the public sector. Section 5 introduces the economic problem for the representative firms in the final goods sector and in the health are sector. Section 6 closes the model by specifying its equilibrium conditions. Finally, Section 7 concludes the technical report. An extended appendix with additional mathematical details is included.

2 Population

2.1 Population characteristics

We set up a discrete time stochastic-dynamic general equilibrium-OLG model with heterogeneous households. The model is populated by overlapping generations or birth cohorts. Each birth cohort is comprised of a

[1] Our model follows the tradition of OLG models that use a heterogenous population and endogenize the retirement age (French (2005), Sánchez Martín (2010), French and Jones (2011), Fehr et al. (2012), Fehr et al. (2013), Laun et al. (2019), Sánchez-Romero et al. (2023), Börsch-Supan et al. (2023)). However, only a few papers combine stochastic shocks along the life cycle with a retirement choice as implemented in our set up. We thereby also account for the fact that the retirement age could potentially be readjusted by the individual in reaction to shocks experienced prior to the retirement (French and Jones (2011), Laun et al. (2019)).

[2] Note that we assume a real economy model, where inflation is set to zero.

number of heterogeneous individuals that initially differ according to their gender $g \in \{f, m\}$, where f denotes females and m males, and by a set of permanent unobservable characteristics $\theta \in \Theta$. The set Θ denotes all possible values of the unobservable characteristics. Similar to Sánchez-Romero et al. (2023), we assume θ is comprised of two variables (learning ability level and effort of schooling) and its distribution, $p(\theta)$, is the same across all birth cohorts. The learning ability level influences the labor income earned by individuals in the labor market, while the effort of schooling reflects the unobservable personal circumstances that prevents individuals to attain the educational level that maximizes their lifetime labor income (Sánchez-Romero et al., 2016). These two characteristics, together with the changing demographic and economic environment, will allow us to replicate the observed distributions on income and education over time and across cohorts.

In addition to the initial heterogeneity in terms of gender and unobservable characteristics, individuals will face four different risks: mortality risk, health risk, unemployment risk, and an education specific realization of the number of children. Regarding the latter assumption, we assume that individuals choose their educational level and to each educational level we assign a conditional distribution (based on empirical data) of the total number of children. By choosing their education, individuals take into account the expected number of children they will have. Once they have chosen their education, the actual number of children is revealed by randomly drawing a profile from the distribution of the total number of children conditional on their educational decision. See Sections 2.3 and 3 for a detailed explanation of the set of alternative risks. These risks will be taken into account by each representative individual of type $(g, \theta) \in (\{f, m\} \times \Theta)$ to make their decisions.

2.2 Population dynamics

Given that our population is comprised of heterogeneous individuals that differ according to unobservable characteristics, in this section we explain how we build up the demography of our heterogeneous individuals to be consistent with the real population data. Firstly, we define the total population and, secondly, we explain the dynamics of the population. In the following we slightly deviate from the standard demographic notation for the sake of applying a consistent notation for all stochastic variables.

Given the gender g and the set of characteristics $\theta \in \Theta$ for each birth cohort, we define the total population of gender g in period t, $\mathbf{N}_t^g$, as the sum of the population of gender g across all ages in period t[3]

$$\mathbf{N}_t^g = \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}_{jt}^g = \sum\nolimits_{j=0}^{J_\Omega} \int_\Theta N_{jt}^g(\theta) d\theta, \tag{1}$$

where $\mathbf{N}_{jt}^g$ is the total population size of gender g at age j in period t, $N_{jt}^g(\theta)$ is the total population size of gender g at age j in period t with initial characteristics θ, and J_Ω is the maximum age that an individual can live. To show that our population is comprised of heterogeneous agents with different survival probabilities, we decompose $N_{jt}^g(\theta)$ into the following three terms: (i) the total number of newborns in period $t-j$ of gender g (denoted by $\mathbf{B}_{t-j}^g$), (ii) the probability that the newborns of gender g in year $j-t$ with characteristics θ survive to age j (denoted by $S_{jt}^g(\theta)$), and (iii) the probability that the newborns have characteristics θ (denoted by $p(\theta)$). Thus, using the fact that $N_{jt}^g(\theta) = \mathbf{B}_{t-j}^g S_{jt}^g(\theta) p(\theta)$, we can rewrite Eq. (1) in terms of the total number of births and the survival probability as follows

$$\mathbf{N}_t^g = \sum\nolimits_{j=0}^{J_\Omega} \int_\Theta N_{jt}^g(\theta) d\theta = \sum\nolimits_{j=0}^{J_\Omega} \mathbf{B}_{t-j}^g \int_\Theta S_{jt}^g(\theta) p(\theta) d\theta. \tag{2}$$

This equation clearly shows that individuals with different initial characteristics may face different mortality schedules, which are represented by $S_{jt}^g(\theta)$. In the GE-OLG model, the survival probability will depend not only on the year of birth, but also on the educational attainment, and on the average history of health outcomes. However, we do not need to express the survival probability in terms of the educational attainment and the average history of health, because the model setting implies that there exists a correspondence between the unobservable set θ and the other observables. Notice that we can still compute the total population of gender g with standard aggregate demographic variables as follows

$$\mathbf{N}_t^g = \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}_{jt}^g = \sum\nolimits_{j=0}^{J_\Omega} \mathbf{B}_{t-j}^g \mathbf{S}_{jt}^g, \tag{3}$$

[3]We define aggregate variables (that do not depend on individual characteristics/heterogeneities) in bold letters.

where $\mathbf{S}^g_{jt} = \int_\Theta S^g_{jt}(\theta)p(\theta)d\theta$ is the (observed) probability of surviving from birth to age j in period t.

To obtain the dynamics of our population, we first define the conditional probability of the population of gender g to survive from birth to age j in period t with characteristics θ as $\pi^{Sg}_{jt}(\theta)$. Then, we use Eq. (1) in period $t+1$ and relate it to the total population of gender g in period t using the following steps:

$$\begin{aligned}
\mathbf{N}^g_{t+1} &= \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{j+1,t+1} + \mathbf{B}^g_t \\
&= \sum\nolimits_{j=0}^{J_\Omega} \int_\Theta N^g_{j+1,t+1}(\theta)d\theta + \mathbf{B}^g_t \\
&= \sum\nolimits_{j=0}^{J_\Omega} \int_\Theta N^g_{jt}(\theta)\pi^{Sg}_{jt}(\theta)d\theta + \mathbf{B}^g_t \\
&= \sum\nolimits_{j=0}^{J_\Omega} \mathbf{B}^g_{t-j} \int_\Theta S^g_{jt}(\theta)\pi^{Sg}_{jt}(\theta)p(\theta)d\theta + \mathbf{B}^g_t \\
&= \sum\nolimits_{j=0}^{J_\Omega} \mathbf{B}^g_{t-j}\mathbf{S}^g_{jt} \int_\Theta \frac{S^g_{jt}(\theta)}{\mathbf{S}^g_{jt}}\pi^{Sg}_{jt}(\theta)p(\theta)d\theta + \mathbf{B}^g_t \\
&= \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{jt} \int_\Theta \frac{S^g_{jt}(\theta)}{\mathbf{S}^g_{jt}}\pi^{Sg}_{jt}(\theta)p(\theta)d\theta + \mathbf{B}^g_t \\
&= \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{jt}\boldsymbol{\pi}^{Sg}_{jt} + \mathbf{B}^g_t
\end{aligned}$$

where $\boldsymbol{\pi}^{Sg}_{jt}$ is the conditional survival probability to age j in period t. By multiplying and dividing by $\mathbf{N}^g_t$ we get

$$\begin{aligned}
\mathbf{N}^g_{t+1} &= \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{jt}\boldsymbol{\pi}^{Sg}_{jt} + \mathbf{B}^g_t \\
&= \mathbf{N}^g_t \sum\nolimits_{j=0}^{J_\Omega} \frac{\mathbf{N}^g_{jt}}{\mathbf{N}^g_t}\boldsymbol{\pi}^{Sg}_{jt} + \mathbf{B}^g_t \\
&= \mathbf{N}^g_t\boldsymbol{\pi}^{Sg}_t + \mathbf{B}^g_t, \qquad (4)
\end{aligned}$$

where $\boldsymbol{\pi}^{Sg}_t$ is the average mortality rate for the population of gender g in period t.

Second we calculate the total number of birth in period t, $\mathbf{B}_t$, multiplying the total number of women at each age by the corresponding fertility rate:

$$\mathbf{B}_t = \sum\nolimits_{j=0}^{J_\Omega} \int_\Theta N^f_{jt}(\theta)f_{jt}(\theta)d\theta = \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^f_{jt} \int_\Theta \frac{S^f_{jt}(\theta)}{\mathbf{S}^f_{jt}} f_{jt}(\theta)p(\theta)d\theta = \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^f_{jt}\mathbf{f}_{jt} \qquad (5)$$

where $f_{jt}(\theta)$ is the fertility rate at age j in period t with characteristics θ and $\mathbf{f}_{jt}$ is the fertility rate at age j in period t.[4] For simplicity, we will assume that the fraction of females at birth (f_{fab}) is constant an equal to 0.4886 (Preston et al., 2000). Hence, the total number of females born in period t is $\mathbf{B}^f_t = \mathbf{B}_t f_{fab}$ and the total number of males born in period t is $\mathbf{B}^m_t = \mathbf{B}_t(1 - f_{fab})$.

2.3 Kinship structure

Family structure is among the most important factors which influences economic decisions of households. In the context of our model, the number of children living in a household influences the household heads labour supply, savings decisions and retirement plans. Therefore it is important to generate fertility profiles that resemble the household structure.

We model the age-profile of the number of children living in a household as a one-time stochastic endowment, which starts at the end of mandatory education level (currently age 14 in our model). Thereby the exact realisation of the profile is random, but its distribution depends on the parents education. Similar

[4] Note that the fertility rate at ages outside the reproductive ages are zero.

approaches in the literature are Fehr et al. (2017), where the total number of children is also a one time stochastic endowment, but without age dependent profiles, or in Sánchez-Romero et al. (2013), where the household size varies by age, but is deterministic. The novelty is a combination of these two approaches, i.e. allowing for uncertainty in the fertility outcome but at the same time also implementing an age specific schedule of fertility.

At age 14 (when individuals complete the mandatory lower secondary education), every individual is assigned the total number of children he or she may have during their lifetime. Thereby, individuals will be assigned an age-dependent profile for the household size. The distribution of profiles for the household size depends on education (e) and age (j) of the household head, as well as the period (t) and the total number of children (Z^c). For a given level of education (e), fertility data provides us with an age and period specific fertility transition matrix $\Pi_{jt}(e)$. For a given educational level e, using the transition matrix Π_{jt}, we generate N random histories (over the age of the household head) of the number of children living in a household by Monte Carlo simulation. We take into account the following factors:

- Children leave the household at age 18.
- Children may die according to child mortality rates.

In the following we will explain in detail how to calculate the age-dependent profiles for the household size by averaging Monte Carlo simulations. The random histories we receive from Monte Carlo sampling are not yet conditioned on the total number of children. But since this is the random variable we want to draw on the household level, we first have to group the simulated histories by total number of children. After we have collected them in separate bins, we average them.

To this end, let the N simulated histories be denoted as $(n_{jt;1}, n_{jt;2}, \ldots, n_{jt;i}, \ldots, n_{jt;N})_{j,t}$, where $n_{jt;i}$ represents the number of children living in household i with a household head of age j, at time t. To average the profiles, we condition on the total number of children $Z^c = k$. First we collect all simulations that belong to $Z^c = k$:

$$I_k^{e,t} = \left\{ i \ : \sum_{j=1}^{J_\Omega} (n_{j-1,t;i} - n_{jt;i})_{\geq 0} = k \right\}. \tag{6}$$

Note, that $I_k^{e,t}$ gives the total number of children leaving the household which of course is then equivalent to the total number of children having ever been born in a household.

We build the age and time specific number of children living in a household of k ever born children and educational level e by averaging across the simulations that belong to the set $Z^c = k$:

$$\overline{n}_{jt}(Z^c = k, e) = \frac{1}{|I_k^{e,t}|} \sum_{i \in I_k^{e,t}} n_{jt;i} \tag{7}$$

From the average profile $\overline{n}_{tj}$ we calculate the average household size as:

$$H_{jt}(Z^c = k, e) = \sqrt{1 + \overline{n}_{jt}(Z^c = k, e)} \tag{8}$$

and the distribution of Z^c conditioned on e:

$$\pi_t^c(Z^c = k|e) = \frac{|I_k^{e,t}|}{N}. \tag{9}$$

Advantages of our approach are a substantial decrease in computational time compared to models which resample fertility from period to period. Moreover our approach also offers a straightforward way to implement time transfer profiles (which we directly get from National Transfer Time Accounts (NTTA) data, see agenta-project.eu): The total available time $T(Z^c, e)$ will simply be a function of the total number of children and the education choice.

The main drawback is that we will not obtain an age dependent fertility path for each individual.

3 Households

For notational simplicity, we skip in this section the indexes for time and gender. Figure 2 illustrates the timeline of a representative agent with a given set of characteristics (or endowments). An agent is born at age 0 and spends the childhood period within the household of a representative parent until age $\underline{a}$. At age $\underline{a}$, the agent receives an initial set of characteristics θ, leaves the parent's home, and settles a new household. From age $\underline{a}$ on, the agent starts making decisions. The decision made by each agent is heterogeneous in multiple ways. This is so because agents differ in terms of their initial characteristics: learning ability and the effort of schooling. In other words, agents are equipped with different initial endowments that give them different preferences towards which education and employment path to choose. Secondly, even if the agents have similar initial endowments, our model creates dynamic heterogeneity. To achieve this, we add stochastic components (unemployment shocks, health shocks), which evolve over time. This implies that agents make decisions under uncertainty. The exact realisation can only be anticipated by the agent, modeled through discrete time Markov chains, but is not known in advance.

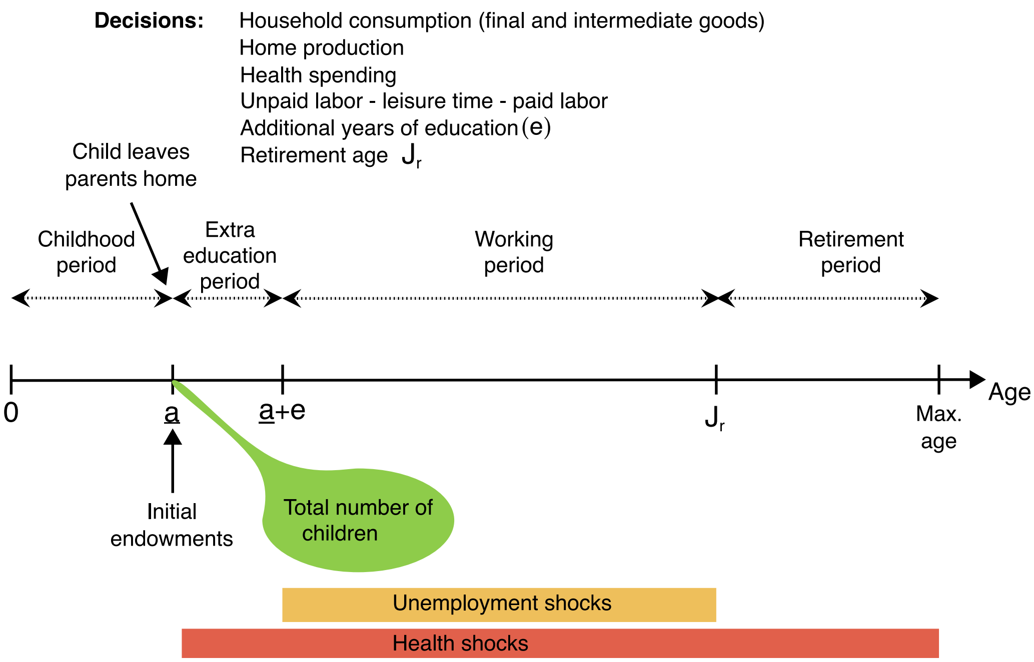

Figure 2: Agents' timeline

The first decision made by our agent is about the additional years of education, which will be chosen based on expectations about the labor employment history, the evolution of the health status, and the expected number of children. Once the education is chosen, the agent is randomly assigned, conditional on the education decision, an age-profile of the number of children. The total number of children will not only determine the evolution of the household size over the lifespan of the agent, but also the number of offspring that may assist the agent once that she will need familial support. Familial support will be considered in a second stage of the project, in which we plan to incorporate long term care. The long term care demand will be based on the health care characteristics of individuals and their remaining life expectancy. The long term care will be provided either within the family (offspring) or outside of the family, by hiring the services on the market. To determine the number of surviving offspring linked to each parent, we will employ the formulas developed in Sánchez-Romero et al. (2018). From period $\underline{a}$ on, our agent will decide over the consumption of the household, how much to spend on health care services, the retirement age, the time devoted to housework, and leisure. While the agent is not retired, the market labor supply in each period will be indirectly determined by the difference between the total available time (after childcare) and the sum of housework and leisure. All the decisions will be made facing the risk of experiencing bad health as well as being unemployed, while on the labor market. Next, we detail how the the stochastic variables are modeled.

3.1 Stochastic Variables

We model dynamic heterogeneity via Markov chains. Our individuals are subject to unemployment and health shocks. Figure 3 shows the transition graph for these two stochastic processes. The figure is divided in two panels. Panel A shows the states and the transition probabilities associated to employment, whereas Panel B shows the states and the transition probabilities associated to health.

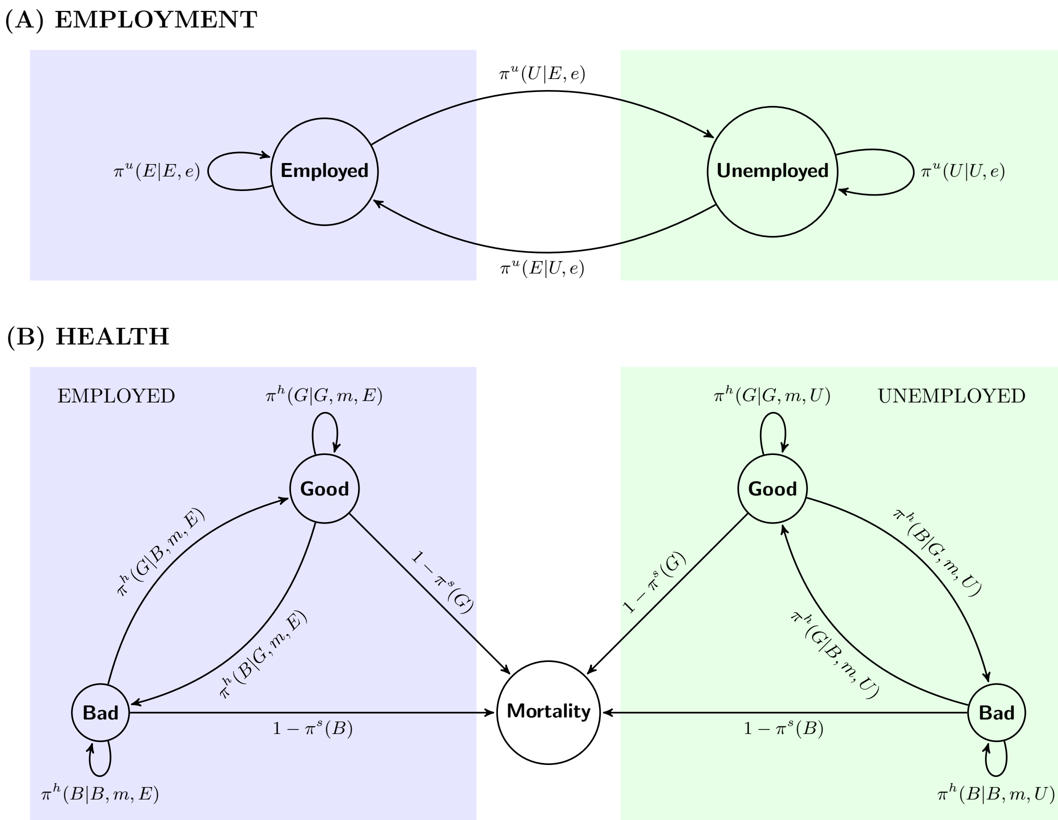

Figure 3: Transition graph.

Note: Figure 3.A shows the states and the transition probabilities associated to employment, whereas Figure 3.B shows the states and the transition probabilities associated to health. The purple area indicates that the individual is employed, whereas the green area indicates that the individual is unemployed.

Employment. There are two different employment states (Z^u): Employment (E) and Unemployment (U). This approach is a simplified version of the one used in Heijdra et al. (2017) or Kindermann and Krueger (2022), where the authors also use Markov chains to model labour states, including unemployment. We restrict ourselves to two states in order to keep the model computationally feasible. The unemployment transition probabilities π_j^u will depend on education, e, and age, j:

$$\begin{bmatrix} \pi_j^u(E|E,e) & \pi_j^u(E|U,e) \\ \pi_j^u(U|E,e) & \pi_j^u(U|U,e) \end{bmatrix} \tag{10}$$

where $\pi_j^u(z_{j+1}^u|z_j^u,e)$ denotes the probability of the transition from labour state z_j^u at age j, to z_{j+1}^u at age $j+1$ and given the educational level e.

Health. We follow Fonseca et al. (2021) (while excluding additional risk factors) to model the health of our agents. There are two different health states (Z^h): Good (G) and Bad (B) health. We assume the transition probabilities π_j^h to depend on healthcare expenditures m_j, the employment status Z_j^u and on age j:

$$\begin{bmatrix} \pi_j^h(G|G, m_j, Z_j^u) & \pi_j^h(G|B, m_j, Z_j^u) \\ \pi_j^h(B|G, m_j, Z_j^u) & \pi_j^h(B|B, m_j, Z_j^u) \end{bmatrix} \tag{11}$$

where $\pi_j^h(z_{j+1}^u|z_j^h, m_j, Z_j^u)$ is the probability to transit from health state z_j^h at age j to z_{j+1}^h at age $j+1$ for health expenditures m_j and employment status Z_j^u. We furthermore assume:

- Unemployed individuals have a higher probability to stay in and transit into the bad health state, i.e.

$$\pi_j^h(B|B, m_j, U) > \pi_j^h(B|B, m_j, E) \quad \text{and} \quad \pi_j^h(B|G, m_j, E) < \pi_j^h(B|G, m_j, U) \tag{12}$$

- Spending money on health care m_j has a preventive effect, so individuals can increase their probability to stay in good health:

$$m > \tilde{m} \implies \pi_j^h(G|G, m, Z^u) > \pi_j^h(G|G, \tilde{m}, Z^u) \tag{13}$$

  and it also increases the probability to recover from bad health:

$$m > \tilde{m} \implies \pi_j^h(G|B, m, Z^u) > \pi_j^h(G|B, \tilde{m}, Z^u) \tag{14}$$

We assume the conditional survival probability depends on the health status of individuals. In particular, household heads with bad health status face a lower conditional survival probability π_j^s than household heads with good health state; i.e. $\pi_j^s(G) > \pi_j^s(B)\, \forall t$. We may assume that those individuals who become unemployed face a negative health shock.

Considering the existing literature on modelling retirement decisions and pensions, several models also include health shocks (French (2005), French and Jones (2011), Fehr et al. (2013), Laun et al. (2019)). What sets our approach apart is that we allow individuals to influence their health through health spending. A different approach of modelling health in the context of pension reforms can be found in Börsch-Supan et al. (2023), where the authors assume the agents to be a priori of different health, and hence they exclude transitions between the health states over the life time. More details on how we model health transitions can be found in section D in the appendix.

Besides the age dependent shocks of employment status and health, individuals receive a draw of their total number of children Z^c dependent on their educational attainment, i.e. individuals choose their educational attainment $e^* \in E$ and afterwards they are randomly assigned a total number of children. Let the probability of having $Z^c = k$ children conditional on the educational attainment $e \in E$ be $\pi^c(Z^c = k|e)$ (details on kinship structure can be found in the previous section). Finally let us denote by $Z_j = (Z_j^u, Z_j^h, Z^c)$ the vector of the three random variables.

As it is common practice in probability theory we will use the notation $(\Omega, \mathcal{A}, P)$ to capture the entire randomness of our model: $\omega \in \Omega$ will denote specific realisation of the stochastic components, $P(\omega)$ its probability and $\mathcal{A}$ the set of measurable events. This notation will be useful as soon as we need to calculate averages in order to transfer from the microeconomic level to the entire economy (e.g. labour and capital supply).

3.2 Household Optimisation Problem

Individuals make their decisions in the following order: At the beginning of period j they are aware of their state variables (W_j, B_j), where W_j denotes wealth and B_j the amount of pension benefits at the beginning of period j. Then they face a health shock $Z_j^h \in \{G, B\}$. If individuals are employed, they may face an unemployment shock, i.e. $Z_j^u \in \{U, E\}$ is realized. Taking their health and employment state into account, individuals then maximize their expected remaining lifetime utility by choosing consumption (c_j), intermediate goods for household production (i_j), health-care services (m_j), unpaid labor supply (h_j), and leisure (ℓ_j). If they are above the minimum retirement age $\underline{J}$, they also decide on whether to retire or not.

The retirement choice is described by ζ_j, which takes the value 0 for being in the labour force and 1 for being retired. Once individuals are retired, i.e. $\zeta_j = 1$, they cannot return to the labour market, so automatically $\zeta_{j+1} = 1$ and there is no more retirement choice to be made by the individual. We will denote by m_j and $x_j = (c_j, i_j, h_j, \ell_j)$ respectively, individual health spending and the vector of control variables other than health at age j.

We recursively solve the household optimisation problem at each age $j \in \{\bar{a}, \cdots, J_\Omega\}$, given each initial state, using the Bellman equation[5]

$$V_j(W_j, B_j, Z^h_j, Z^u_j, \zeta_j | Z^c, e, \theta) = \max_{m_j, x_j, \zeta_{j+1}} \left\{ \begin{array}{l} U_j(x_j|Z_j, \zeta_j, e, \theta) + \\ \beta\, \mathbb{E}_{m_j}\left[\pi^s_{j+1}(Z^h_{j+1}) V_{j+1}(W_{j+1}, B_{j+1}, Z^h_{j+1}, Z^u_{j+1}, \zeta_{j+1}|Z^c, e, \theta) \big| Z_j\right] \end{array} \right\}$$

subject to

$$W_{j+1} = \tilde{R}_j W_j + (1-\zeta_t)[\tilde{w}_j y_{Lj}(Z_j, e, \theta) + \mathbf{1}_U(Z^u_j)\mathrm{tr}_j] + \zeta_t B_t - \tilde{p}^c_j(c_j + i_j) - \tilde{p}^m_j\, m_j$$

$$B_{j+1} = \begin{cases} \widehat{R}_j \phi^S_{j+1}(e) B_j + \varphi_{j+1}(e)\tau^S_j w_j y_{Lj}(Z_j, e, \theta) & \text{for } \zeta_j = 0, \\ B_j & \text{for } \zeta_j = 1, \end{cases} \tag{15}$$

and the boundary conditions $W_0 = W_{J_\Omega+1} = B_0 = 0$. The term $\pi^s_{j+1}(Z^h)$ denotes the conditional survival probability, which depends on the current state of health Z^h, and

$$\tilde{w}_j y_{Lj}(Z_j, e, \theta) = \tilde{w}_j \mathrm{gap}^g_j \epsilon_j(e, \theta)\mathbf{1}_E(Z^u_j)(1 - \delta^h \mathbf{1}_B(Z^h_j))(T(Z^c, e) - \ell_j - h_j) \tag{16}$$

is the labor income received at age j. Labor income at age j is compound by several factors. The term $\tilde{w}_j$ is the effective wage rate at age j net of contributions and taxes, gap^g_j denotes the gender wage gap at age j and $\epsilon_j(e, \theta)$ is the age– and education–specific productivity of an individual of type θ. Further we assume that individuals with bad health are less productive by a factor of δ^h and receive a transfer tr_j when unemployed. An important fact is that the expectation operator $\mathbb{E}_{m_j}$ depends on the amount of health spending m_j, since individuals can reduce their probability to be in bad health by investing m_j. B_j are the pension benefits claimed at age j. The terms $\tilde{p}^c_j$ and $\tilde{p}^m_j$ denote the prices for market good and purchased health care services, respectively.

While individuals are not retired (i.e. $\zeta_j = 0$), their pension benefits increase for three reasons. First, pension benefits increase because they are capitalized according to the rate of return of the pension system $\widehat{R}_j$. Second, pension benefits may increase because of changes in the pension replacement rate formula due to working one additional year, which is captured by the function $\phi^S_{j+1}(e)$. And third, pension benefits also increase at a rate $\varphi_{j+1}(e)$ because of additional contributions paid to the pension system. The term τ^S_j is the social contribution rate at age j, and w_j is the effective wage rate at age j. Once individuals are retired, their pension benefits remain constant.

3.3 Productivity

For each individual with education level e and initial characteristics $\theta \in \Theta$, their productivity at age j is denoted by $\epsilon_j(e, \theta)$, and it is defined as follows:

$$\epsilon_j(e, \theta) = \exp\left(\beta_0(e, \theta) + \beta_1(e)(j - J_e) + \beta_2(e)(j - J_e)^2\right) \tag{17}$$

Here, $J_e = \bar{a} + e$ represents the age at which an individual enters the labor market. The returns to education β_0 will depend on the educational choice e and the initial endowments θ, whereas the returns to experience parameters β_1 and β_2 only depend on the educational choice e.

3.4 Time Constraint

The total available time $T_j(Z^c, e)$ of an individual is the amount of time left after childcare. Therefore $T_j(Z^c, e)$ is a function of age, gender, the number of children and education.[6] The sum of home production $h_j \geq 0$ and leisure $\ell_j \geq 0$ must always be smaller or equal to $T_j(Z^c, e)$:

[5]The Bellman equation is used for solving optimization problems over time where decisions are made sequentially, and the optimal decision at each step depends on the state of the system Heer and Maußner (2009).

[6]Note that for notational convenience, in this section we skip the indexes time and gender.

$$\text{Time Constraint: } h_j + \ell_j \leq T_j(Z^c, e) \tag{18}$$

The time spent in the labor market l_j is indirectly calculated as the difference between the total time available and the sum of the time devoted to unpaid labor and leisure

$$l_j = T_j(Z^c, e) - \ell_j - h_j \tag{19}$$

Whenever $T_j(Z^c, e) = \ell_j + h_j$ holds, a corner solution, where the individual completely withdraws from the labour market, is realized.

3.5 Preferences

An agent with a given set of initial characteristics $\theta \in \Theta$ and education (e) is assumed to have preferences in state Z_j over the consumption of final goods (c), home-produced goods (ch), and leisure (ℓ). For simplicity, we assume preferences are separable and logarithmic in consumption. This assumption is imposed to guarantee the existence a steady state when individuals have preferences for consumption and leisure (p. 427, Barro and Sala-i Martin, 2004). The utility is assumed to decrease because agents incur a cost for attending schooling $v_j^E(e, \theta) = \mathbf{1}_{\{j<14+e\}}\eta$ (Oreopoulos, 2007; Restuccia and Vandenbroucke, 2013; Le Garrec, 2015; Sánchez-Romero et al., 2016), where $\eta \in \theta$ is an initial endowment. On the contrary, individuals are assumed to gain utility from being retired $v_j^R(e, \theta)$. We assume the utility from being retired to depend on the education and the characteristics of individuals. The period utility of an agent has the following functional form:

$$U_j(x_j|Z_j, \zeta_j, e, \theta) = H_j(Z^c, e) \log \left[\left(\frac{c_j}{H_j(Z^c, e)} \right) \left(\frac{\text{ch}_j}{H_j(Z^c, e)} \right)^{\alpha_c} \right] + \alpha_\ell \log \ell_j - v_j^E(e, \theta) + \zeta_j v_j^R(e, \theta), \tag{20}$$

where $H_j(Z^c, e)$ is the household size (see Eq. (8) in section 2.3).[7] We have chosen this particular way of representing the household size because it takes into consideration economies of scale that are decreasing in the household size. The parameters α_c and α_ℓ denote preferences for home-produced goods and leisure relative to final market produced-goods, respectively. To produce home-goods agents combine intermediate goods (i) and unpaid labor (h) using a Cobb-Douglas technology:

$$\text{ch}_j = i_j^\gamma (h_j)^{1-\gamma}, \tag{21}$$

where γ is the intermediate goods share in the total home-production. Similar to Sánchez-Romero et al. (2023) the parameter affecting the cost of attending schooling $v_j^E(e, \theta)$, together with other parameters, will be calibrated to replicate the educational distribution of the specific country that is studied. It represents the marginal cost of each additional year of schooling and can be considered as a proxy for the socioeconomic background of the agent.

3.6 Household Decisions

3.6.1 Consumption of Market Goods, Home–Produced Goods, Unpaid Labor and Leisure

Optimal values for the control variables are solved period-wise via first order conditions. This leads to the standard Euler condition under uncertainty

$$\beta \tilde{R}_{j+1} \frac{\tilde{p}_j^c}{\tilde{p}_{j+1}^c} \mathbb{E}_{m_j} \left[\pi_{j+1}^s (Z_{j+1}^h) \frac{\frac{\partial}{\partial c_{j+1}} U_{j+1}(x_{j+1}|Z_{j+1}, \zeta_{j+1}, e, \theta)}{\frac{\partial}{\partial c_j} U_j(x_j|Z_j, \zeta_j, e, \theta)} \middle| Z_j \right] = 1 \tag{22}$$

[7]In an economic setting in which agents can insure all their risks and the (net) market interest rate exceeds the subjective discount factor (β), the chosen functional form of the utility from consumption guarantees that the consumption of the household head increases monotonically with age.

the equations for optimal health investment and intermediate goods:

$$0 = \frac{d}{dm_j}\mathbb{E}_{m_j}\left[\pi^s_{j+1}(Z^h_{j+1})V_{j+1}(W_{j+1}, B_{j+1}, Z^h_{j+1}, Z^u_{j+1}, \zeta_{j+1}|Z^c, e, \theta)\big|Z_j\right] \tag{23}$$

$$\frac{\partial}{\partial i_j}U_j(x_j|Z_j, \zeta_j, e, \theta) = \frac{\partial}{\partial c_j}U_j(x_j|Z_j, \zeta_j, e, \theta) \tag{24}$$

and equations for paid labor, unpaid labor, and leisure:

$$\frac{\partial}{\partial h_j}U_j(x_j|Z_j, \zeta_j, e, \theta) = \frac{\partial}{\partial \ell_j}U_j(x_j|Z_j, \zeta_j, e, \theta) \tag{25}$$

$$\frac{\frac{\partial}{\partial h_j}U_j(x_j|Z_j, \zeta_j, e, \theta)}{\frac{\partial}{\partial c_j}U_j(x_j|Z_j, \zeta_j, e, \theta)} = \frac{\tilde{w}_j \mathrm{gap}^g_j \epsilon_j(e,\theta)(1-\delta^h \mathbf{1}_B(Z^h_j))}{\tilde{p}^c_j} \qquad \text{for } \zeta_j = 0 \text{ and } Z^u_j = E, \tag{26}$$

$$\ell_j = T_j(Z^c, e) - h_j \qquad \text{otherwise.} \tag{27}$$

The first order conditions indicate that health investments are proportional to the value of life and that higher wage rates reduce leisure and unpaid labor. In contrast, a higher wealth increases leisure and unpaid labor. In section E in the appendix the derivation of the policy functions of households is explained in more detail.

3.6.2 Retirement Decision

Following the existing literature we endogenize the age at retirement (Sánchez Martín, 2010; Fehr et al., 2012, 2013; Laun et al., 2019; Sánchez-Romero et al., 2023; Börsch-Supan et al., 2023). One period before reaching the minimum retirement age $j \geq \underline{J} - 1$ and while not yet being retired, individuals have the possibility to choose whether to retire in the next period or not. They compare their expected remaining lifetime utility when retired, i.e. $\zeta_j = 1$ for all $j \geq \underline{J}$, to the case when not retiring ($\zeta_j = 0$). Individuals retire as soon as there is a an age where the discounted expected lifetime utility when retiring exceeds the discounted expected lifetime utility when not retiring. When reaching the maximum retirement age $j \geq \overline{J}$, individuals have to retire mandatorily.

3.6.3 Education Decision

Individuals decide upon their education by taking into account that it influences their fertility. A longer education period may lead to later parenthood and a lower total number of children. Individuals are aware of these correlations and consider them when making their education decision. Therefore the following expected lifetime utility is maximized:

$$e^*(\theta) = \arg\max_{e \in E} \sum\nolimits_{z=0}^{N} V_0(W_0, B_0, Z^h_0, Z^u_0, \zeta_0|z, e, \theta)\pi^c(Z^c = z|e) \tag{28}$$

where the superscript * denotes that the decision is optimal and $\pi^c(Z^c = z|e)$ is the conditional probability to have z children during life when choosing education e. It is important to note here that the random variable Z^c (number of children) is realised after education $e^*(\theta)$ is completed. Put differently, individuals only have expectations about the role of education on the number of children, but the exact number of children (i.e. the realization of the stochastic fertility process) is unknown to them until they completed education. It could therefore potentially be the case that their education choice is not optimal with respect to the realisation of Z^c, although it was optimal with respect to their expectations.

3.7 Supply of Capital and Labour

Let N^g_{jt} be the population size of gender g at age j in year t. The total private wealth ($\mathbf{W}_t$) is then given by

$$\mathbf{W}_t = \sum\nolimits_{g \in \{f,m\}} \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{j,t} \int_\Theta \frac{S^g_{jt}(\theta)}{\mathbf{S}^g_{jt}} \overline{W^g_{jt}}(\theta) p(\theta) d\theta \tag{29}$$

where

$$\overline{W^g_{jt}}(\theta) = \int_\Omega W^g_{jt}(\theta,\omega)dP(\omega) \tag{30}$$

is the average wealth of an individual of gender g and characteristics θ at period t and age j. To calculate the average wealth we integrate $W^g_{jt}(\theta,\omega)$ over all possible realisations of the stochastic components ω with respect to the distribution after private policy functions are taken into account, denoted by $P(\cdot)$. This probability measure captures the entire observed randomness of our model.

In an open economy, we have $\mathbf{W}_t = K_t + D_t + F_t$, where K_t is the capital stock, D_t is the public debt, and F_t is the net foreign asset position. Analogously labor supply is given as

$$L_t = \sum\nolimits_{g\in\{f,m\}} \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{j+1,t+1} \int_\Theta \frac{S^g_{j+1,t+1}(\theta)}{\mathbf{S}^g_{j+1,t+1}} \overline{yL^g_{jt}}(e^*(\theta),\theta)p(\theta)d\theta \tag{31}$$

where

$$\overline{yL^g_{jt}}(e^*(\theta),\theta) = \int_\Omega (1-\zeta^g_{jt}(\theta,\omega))yL^g_{jt}(e^*(\theta),\theta,\omega)dP(\omega)$$

is the average labor income of individuals of gender g with initial characteristics θ at period t and age j.

4 Public Sector

The government will provide in each period t public goods and services (G_t), subsidies ($\mathcal{S}_t$), unemployment benefits (B^u_t), and pension benefits (B^p_t). The total amount of unemployment benefits and pension benefits claimed by individuals are

$$\mathcal{B}^u_t = \sum\nolimits_{g\in\{f,m\}} \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{j+1,t+1} \int_\Theta \frac{S^g_{j+1,t+1}(\theta)}{\mathbf{S}^g_{j+1,t+1}} \overline{\mathrm{tr}^g_{jt}}(\theta)p(\theta)d\theta. \tag{32}$$

$$\mathcal{B}^p_t = \sum\nolimits_{g\in\{f,m\}} \sum\nolimits_{j=0}^{J_\Omega} \mathbf{N}^g_{j+1,t+1} \int_\Theta \frac{S^g_{j+1,t+1}(\theta)}{\mathbf{S}^g_{j+1,t+1}} \overline{B^g_{jt}}(\theta)p(\theta)d\theta. \tag{33}$$

where $\overline{\mathrm{tr}^g_{jt}}(\theta) = \int_\Omega (1-\zeta^g_{jt}(\theta,\omega))\mathrm{tr}^g_{jt}(\theta,\omega)dP(\omega)$ is the average unemployment benefit received by a person of gender g with initial characteristics θ at age j in period t and $\overline{B^g_{jt}}(\theta) = \int_\Omega \zeta^g_{jt}(\theta,\omega)B^g_{jt}(\theta,\omega)dP(\omega)$ is the average pension benefit of an agent of gender g with initial characteristics θ at age j in period t. Notice in Eq. (33) that if individuals who receive higher benefits also have a greater survival, the pension system becomes more costly, since the ratio $S^g_{j+1,t+1}(\theta)/\mathbf{S}^g_{j+1,t+1}$ is greater (resp. lower) than one for $\overline{B^g_{jt}}(\theta)$ greater (resp. lower) than the average. To finance all the unemployment benefits, the government collects contributions to finance the total cost of the unemployment insurance

$$\tau^U_t w_t L_t = \mathcal{B}^u_t, \tag{34}$$

where τ^U_t is the unemployment contribution rate in period t, w_t is the wage rate per effective hour worked, and L_t is the total labor supply measured in effective units of labor. Since in some countries a fraction of the total pension benefits are financed by general taxes (Sánchez-Romero et al., 2023), we assume the government collects contributions to finance a fraction $\phi_B \in [0,1]$ of all the pension benefits claimed

$$\tau^S_t w_t L_t = \phi_B \mathcal{B}^p_t, \tag{35}$$

where τ^S_t is the social contribution rate in period t. To finance all the public goods and services, subsidies, and the fraction $1-\phi_B$ of the remaining pensions claimed, the government levies taxes and can also issue debt. The budget constraint of the government in period t is

$$D_{t+1} = R_t D_t + G_t + \mathcal{S}_t + (1-\phi_B)\mathcal{B}^p_t - T_t \tag{36}$$

where D_t is the debt in period t, R_t is the compound interest rate in period t paid on debt, and T_t is the total tax base. The total tax base is comprised of taxes levied on financial wealth, consumption, and labor income

$$T_t = \tau_t^K r_t \mathbf{W}_t + \tau_t^C C_t + \tau_t^L (w_t L_t - (1-\phi_B)\mathcal{B}_t^p) \tag{37}$$

where $(\tau_t^K,\ \tau_t^C,\ \tau_t^L)$ is the set of taxes on the interests from holding financial wealth, on consumption, and on labor income, respectively.

5 Firms

5.1 Production: Final Good

We assume one representative firm that produces a final good by combining capital (K^f) and effective labor (L^f). Final goods can either be saved or consumed. The production function, that exhibits constant returns to scale, takes the following form

$$Y_t^f = (K_t^f)^{\alpha_f}(A_t L_t^f)^{1-\alpha_f}, \tag{38}$$

where Y_t^f is output of the final good, α_f is the capital share of the final good, and A_t is labor-augmenting technology, whose law of motion is $A_{t+1} = (1+g_t^A)A_t$ with g_t^A being the productivity growth rate. Aggregate capital stock evolves according to the law of motion $K_{t+1}^f = K_t^f(1-\delta_K) + I_t^f$, where δ_K is the depreciation rate of capital and I_t^f is aggregate gross investment. The quantities L_t^f and K_t^f are calculated by aggregating the labour and capital supply at the individual level across all individuals currently living in our economy.

We assume our representative firm maximizes the net cash flow by renting capital and hiring labor from households in competitive markets at the rates r_t and w_t respectively. Capital and labor inputs are chosen by firms according to the first-order conditions:

$$\frac{\partial Y_t^f}{\partial K_t^f} = \alpha_f \frac{Y_t^f}{K_t^f} = r_t + \delta_K, \tag{39}$$

$$\frac{\partial Y_t^f}{\partial L_t^f} = (1-\alpha_f)\frac{Y_t^f}{L_t^f} = w_t. \tag{40}$$

5.2 Production: Health Care

Like in the final good sector, we assume one representative firm that provides health care services by combining capital (K^m) and effective labor (L^m). Unlike final goods, health care services can only be consumed. Let the price per unit of health care service be p_t^m. The production function, that exhibits constant returns to scale, takes the following form

$$Y_t^m = (K_t^m)^{\alpha_m}(A_t L_t^m)^{1-\alpha_m}, \tag{41}$$

where Y_t^m is the total health care services provided, α_m is the capital share in the health care sector. Similarly, aggregate capital stock evolves according to the law of motion $K_{t+1}^m = K_t^m(1-\delta_K) + I_t^m$, where δ_K is the depreciation rate of capital and I_t^m is aggregate gross investment.

We assume our representative firm maximizes the net cash flow by renting capital and hiring labor from households in competitive markets at the rates r_t and w_t respectively. Capital and labor inputs are chosen by firms according to the first-order conditions:

$$p_t^m \frac{\partial Y_t^m}{\partial K_t^m} = p_t^m \alpha_m \frac{Y_t^m}{K_t^m} = r_t + \delta_K, \tag{42}$$

$$p_t^m \frac{\partial Y_t^m}{\partial L_t^m} = p_t^m (1-\alpha_m)\frac{Y_t^m}{L_t^m} = w_t. \tag{43}$$

5.3 Labor and capital supply across sectors

Individuals can freely work in both sectors, which implies that the wage rate and the interest rate in both sectors coincide. Hence, the following relationship is satisfied:

$$\frac{r_t + \delta_K}{w_t} = \frac{\partial Y_t^f / \partial K_t^f}{\partial Y_t^f / \partial L_t^f} = \frac{\partial Y_t^m / \partial K_t^m}{\partial Y_t^m / \partial L_t^m} = \frac{\alpha_f}{1-\alpha_f}\frac{L_t^f}{K_t^f} = \frac{\alpha_m}{1-\alpha_m}\frac{L_t^m}{K_t^m}.$$

To calculate the capital shares (α_f, α_m), we collect information on the share of the final good sector in total production Y_t^f/Y_t and the share of the health care spending relative to the total production $p_t^m M_t/Y_t$. See the appendix C for the derivation of the input shares.

6 Recursive competitive equilibrium

Let $(\Omega, \mathcal{A}, \mathcal{P})$ be the probability space of all possible realizations of our random processes. Let $\mathcal{E}$, $\mathcal{G}$, $\mathcal{J}$, $\mathcal{T}$, Θ be the sets of educational groups, genders, ages, time periods, and of initial endowments, respectively.

Given an initial population distributions $\{N_{jt_0}^g(e)\}_{j\in\mathcal{J},g\in\mathcal{G},e\in\mathcal{E}}$ and time-series of mortality and fertility rates $\{\pi_{jt}^g(e), f_{jt}(e)\}_{e\in\mathcal{E},g\in\mathcal{G},j\in\mathcal{J},t\in\mathcal{T}}$, a recursive competitive equilibrium is defined as

- a set of government policy functions $\Psi_t = \{G_t, \tau_t^C, \tau_t^L, \tau_t^S, \tau_t^K\}$ that satisfies (32)–(37);
- sets of households decision rules

  $$\{W_{jt}^g(\theta,\omega), B_{jt}^g(\theta,\omega), c_{jt}^g(\theta,\omega), i_{jt}^g(\theta,\omega), m_{jt}^g(\theta,\omega), h_{jt}^g(\theta,\omega), z_{jt}^g(\theta,\omega), \zeta_{jt}^g(\theta,\omega), e_{t-j}^g(\theta)\}_{g\in\mathcal{G},j\in\mathcal{J},\theta\in\Theta,\omega\in\Omega},$$

  that satisfy (22)–(27);
- a set of prices $\{r_t, w_t, p_t^m\}$ that equal the marginal productivities (39)-(40) in the final good sector and the marginal productivities (42)-(43) in the health care sector;
- a stock of capital and a stock of effective labor

$$K_t = \sum_{g\in\{f,m\}} \sum_{j=0}^{J_\Omega} \mathbf{N}_{j,t}^g \int_\Theta \frac{S_{jt}^g(\theta)}{\mathbf{S}_{jt}^g} \left(\int_\Omega W_{jt}^g(\theta,\omega) dP(\omega) \right) p(\theta) d\theta - D_t - F_t \tag{44}$$

$$L_t = \sum_{g\in\{f,m\}} \sum_{j=0}^{J_\Omega} \mathbf{N}_{j+1,t+1}^g \int_\Theta \frac{S_{j+1,t+1}^g(\theta)}{\mathbf{S}_{j+1,t+1}^g} \left(\int_\Omega (1-\zeta_{jt}^g(\theta,\omega)) y_{L_{jt}^g}(e_{t-j}^*(\theta),\theta,\omega) dP(\omega) \right) p(\theta) d\theta, \tag{45}$$

that clear the market of final goods and services (health) at all $t \in \mathcal{T}$,

$$Y_t^f + p_t^m Y_t^m = C_t + G_t + I_t, \tag{46}$$

where C_t is aggregate consumption in period t

$$C_t = \sum_{g\in\{f,m\}} \sum_{j=0}^{J_\Omega} \mathbf{N}_{j+1,t+1}^g \int_\Theta \frac{S_{jt}^g(\theta)}{\mathbf{S}_{jt}^g} \overline{C_{jt}^g}(\theta) p(\theta) d\theta \tag{47}$$

with

$$\overline{C_{jt}^g}(\theta) = \int_\Omega \left(c_{jt}^g(\theta,\omega) + i_{jt}^g(\theta,\omega) + p_t^m m_{jt}^g(\theta,\omega) \right) dP(\omega),$$

G_t is the aggregate consumption of publicly financed goods in period t, and $I_t = I_t^f + I_t^m$ is the total investment in period t.

7 Conclusion

To study the resilience of pension systems we need to account for the increasing heterogeneity in aging populations together with the fact that individuals are exposed to several risks over their life cycle.

In this technical note we present a discrete time stochastic-dynamic general equilibrium model populated by overlapping generations that extends our previous work Sánchez-Romero et al. (2023) (where we have already accounted for heterogeneous agents) in several dimensions.

First, we assume that individuals face mortality risk and idiosyncratic risks with respect to their employment status and health status. Second, to prevent experiencing negative health care shocks, which reduce labor productivity and decrease survival, individuals can invest in health care. Third, individuals are randomly assigned an age-specific fertility rate profile that depends on their educational attainment. Fourth, given that individuals face idiosyncratic risks, the retirement age is contingent on the realized life cycle path. Fifth, at the macro level, we assume two production sectors: a final good sector and a health care sector. Sixth, the economy will be open to foreign capital and the government is allowed to issue debt.

Beyond economic considerations, our population includes two genders, and individuals exhibit diverse kinship structures. These two characteristics will enable us to better study the economic impact of alternative policy reforms on each gender.

Acknowledgement

Funded by the European Union's Horizon Europe Programme, Grant Agreement No 101094741. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

This publication reflects only the authors' views and the European Union is not liable for any use that may be made of the information contained therein.

We thank Jakub Bijak and Emily Barker for their helpful comments.

References

Barro, R. and Sala-i Martin, X. (2004). *Economic growth, second edition*. Cambridge MA.: The MIT Press.

Börsch-Supan, A., Härtl, K., Leite, D. N., and Ludwig, A. (2023). Preventing reforming unequally. *Journal of Population Economics*.

Fehr, H., Kallweit, M., and Kindermann, F. (2012). Pension reform with variable retirement age: a simulation analysis for Germany. *Journal of Pension Economics and Finance*, 11(3):389–417.

Fehr, H., Kallweit, M., and Kindermann, F. (2013). Should pensions be progressive? *European Economic Review*, 63:94–116.

Fehr, H., Kallweit, M., and Kindermann, F. (2017). Families and social security. *European Economic Review*, 91:30–56.

Fonseca, R., Michaud, P.-C., Galama, T., and Kapteyn, A. (2020). Accounting for the rise of health spending and longevity. *Journal of the European Economic Association*, 19(1):536–579.

Fonseca, R., Michaud, P.-C., Galama, T., and Kapteyn, A. (2021). Accounting for the rise of health spending and longevity. *Journal of the European Economic Association*, 19(1):536–579. Epub 2020 Mar 12.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2):395–427.

French, E. and Jones, J. B. (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica*, 79(3):693–732.

Heer, B. and Maußner, A. (2009). *Dynamic general equilibrium modeling: Computational methods and applications*. Springer-Verlag GmbH Germany, 2 edition.

Heijdra, B. J., Kindermann, F., and Reijnders, L. S. (2017). Life in shackles? The quantitative implications of reforming the educational financing system. *Review of Economic Dynamics*, 25:37–57. Special Issue on Human Capital and Inequality.

Kindermann, F. and Krueger, D. (2022). High marginal tax rates on the top 1 percent? Lessons from a life-cycle model with idiosyncratic income risk. *American Economic Journal: Macroeconomics*, 14(2):319–66.

Laun, T., Markussen, S., Vigtel, T. C., and Wallenius, J. (2019). Health, longevity and retirement reform. *Journal of Economic Dynamics and Control*, 103:123–157.

Le Garrec, G. (2015). Increased longevity and social security reform: Questioning the optimality of individual accounts when education matters. *Journal of Population Economics*, 28(2):329–352.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91(11-12):2213–2229.

Poole, D. and Raftery, A. E. (2000). Inference for deterministic simulation models: the bayesian melding approach. *Journal of the American Statistical Association*, 95(452):1244–1255.

Preston, S., Heuveline, P., and Guillot, M. (2000). *Demography: Measuring and modeling population processes*. Wiley-Blackwell.

Raftery, A. E. and Bao, L. (2010). Estimating and projecting trends in HIV/AIDS generalized epidemics using incremental mixture importance sampling. *Biometrics*, 66(4):1162–1173.

Restuccia, D. and Vandenbroucke, G. (2013). A century of human capital and hours. *Economic Inquiry*, 51(3):1849–1866.

Sánchez-Romero, M., d'Albis, H., and Prskawetz, A. (2016). Education, lifetime labor supply, and longevity improvements. *Journal of Economic Dynamics and Control*, 73:118–141.

Sánchez-Romero, M., Sambt, J., and Prskawetz, A. (2013). Quantifying the role of alternative pension reforms on the Austrian economy. *Labour Economics*, 22:94–114.

Sánchez-Romero, M., Schuster, P., and Prskawetz, A. (2023). Redistributive effects of pension reforms: Who are the winners and losers? *Journal of Pension Economics and Finance*, pages 1–27.

Sánchez-Romero, M., Wrzaczek, S., Prskawetz, A., and Feichtinger, G. (2018). *Does demography change wealth inequality?*, volume 687, pages 349–375. Springer, Cham.

Sánchez Martín, A. R. (2010). Endogenous retirement and public pension system reform in Spain. *Economic Modelling*, 27(1):336–349.

A Demographic Notation

$\mathbf{N}_{jt}$	Total population of age j in period t
$\mathbf{N}_{jt}^{g}$	*idem* of gender g
$N_{jt}^{g}(\theta)$	*idem* and characteristics θ
$\mathbf{B}_t$	Total number of births in period t
$\mathbf{B}_t^g$	*idem* of gender g
$\mathbf{S}_{jt}$	Probability of surviving from birth to age j in period t
$\mathbf{S}_{jt}^{g}$	*idem* of gender g
$S_{jt}^{g}(\theta)$	*idem* and characteristics θ
$\boldsymbol{\pi}_t^s$	Total mortality rate in period t
$\boldsymbol{\pi}_{jt}^s$	Conditional probability of surviving to age j in period t
$\boldsymbol{\pi}_{jt}^{Sg}$	*idem* of gender g
$\pi_{jt}^{sg}(\theta)$	*idem* and characteristics θ
$\mathbf{f}_{jt}$	Fertility rate at age j in period t
$f_{jt}(\theta)$	*idem* and characteristics θ

B Stochastic Variables Notation

Z_j^h	Health state at age j
Z_{jt}^h	Health state at age j in period t
Z_j^u	Employment state at age j
Z_{jt}^u	Employment state at age j in period t
Z^c	Total number of children
$Z_j = (Z_j^h, Z_j^u, Z^c)$	Vector of stochastic states at age j
Z_{jt}	Vector of stochastic states at age j in period t
$\pi^u(\ \cdot\ \mid\ \cdot\ , e)$	Transition probabilities of Z^u conditional on education e
$\pi^h(\ \cdot\ \mid\ \cdot\ , m_j, Z^u)$	Transition probabilities of Z^h conditioned on health spending m_j and employment Z^u
$\pi^c(\ \cdot\ \vert e)$	Distribution of Z^c
$P(\ \cdot\ )$	Common distribution of all stochastic variables (after private policy functions)
ω	realisation of entire randomness of the system
Ω	state space of ω
$\mathcal{A}$	σ-algebra on Ω

C Production sector: Capital shares

Combining the first-order conditions (FOCs) and dividing by the total production in the economy Y_t, the following relationships are satisfied:

$$\frac{w_t L_t^f}{Y_t} = \frac{1}{\gamma_f - \gamma_m}\left(\frac{(r_t + \delta_K)K_t}{Y_t} - \gamma_m \frac{w_t L_t}{Y_t}\right) = \frac{\alpha_Y - \gamma_m(1-\alpha_Y)}{\gamma_f - \gamma_m}$$
$$\frac{(r_t + \delta_K)K_t^f}{Y_t} = \gamma_f \frac{w_t L_t^f}{Y_t} = \frac{\gamma_f \alpha_Y - \gamma_f \gamma_m (1-\alpha_Y)}{\gamma_f - \gamma_m}$$

where $\gamma_f = \frac{\alpha_f}{1-\alpha_f}$, $\gamma_m = \frac{\alpha_m}{1-\alpha_m}$, and α_Y is the capital share in the total gross domestic product. Adding both shares in the final good sector gives

$$\frac{Y_t^f}{Y_t} = (1+\gamma_f)\frac{w_t L_t^f}{Y_t} = (1+\gamma_f)\frac{\alpha_Y - \gamma_m(1-\alpha_Y)}{\gamma_f - \gamma_m} \tag{48}$$

The share of health care spending on the total output is

$$\frac{q_t M_t}{Y_t} = \frac{1}{1-\alpha_m}\left(\frac{w_t L_t}{Y_t} - \frac{w_t L_t^f}{Y_t}\right) = \frac{1}{1-\alpha_m}\left(1-\alpha_Y - \frac{\alpha_Y - \gamma_m(1-\alpha_Y)}{\gamma_f - \gamma_m}\right). \tag{49}$$

Thus, using the last two equations we can determine the values of α_c and α_m that correspond to the studied economy.

D Health Transitions

To model the health status transitions we follow Fonseca et al. (2020). For simplicity let Z^c and Z_t^h behave independently. Therefore the transition probabilities are given by:

$$\pi_j^h(z_{j+1}^h = k | z_j^h = i, Z_j^e, m_j) = \frac{e^{\delta_{0ik} + \delta_{1k} j + \delta_{2k}\log(m_j) + \delta_{3k}\mathbf{1}_U(Z_j^e)}}{\sum_{k'} e^{\delta_{0ik'} + \delta_{1k'} j + \delta_{2k'}\log(m_j) + \delta_{3k'}\mathbf{1}_U(Z_j^e)}}, \tag{50}$$

where $i, k \in \{\text{Good}, \text{Bad}\}$ and j is the age of the individual. Note that only δ_0s account for the origin and the destination. Differentiating the probability of transiting from state i to state k at age t with respect to m_j gives

$$\frac{\partial \pi_j^h(z_{j+1}^h = k | z_j^h = i, Z_j^e, m_j)}{\partial m_j} = \frac{1}{m_j}\left(\delta_{2k} - \delta_{2-k}\right)\pi_{ik,j}^h(1-\pi_{ik,j}^h), \tag{51}$$

Note: To shorten the expression we abbreviated $\pi_j^h(z_{j+1}^h = k|z_j^h = i, Z_j^e, m_j) = \pi_{ik,j}^h$ and by $-k$ the opposite state of k (e.g. if $k = G$, then $-k = B$).

E Solving the Household Problem

Since we will solve the problem recursively, we will denote the total expected utility of the agent at t with capital W_j, pension benefits B_j, state of the Markov chain $Z_j = (Z_j^h, Z_j^e)$, conditional on being in state ζ_j in the current period, with the total number of children Z^c, and initial characteristics θ by $V_j(W_j, B_j, Z_j, \zeta_j | Z^c, e, \theta)$. The maximization problem is divided in three stages:

- $j \leq \underline{J} - 1$ (below minimum retirement age):

$$\begin{aligned} V_j(W_j, B_j, Z_j, \zeta_j = 0 | Z^c, e, \theta) = \max_{m_j, x_j} \big\{ & U_j(x_j, \zeta_j = 0 | Z^c, e, \theta) \\ & + \beta\, \mathbb{E}_{m_j}\left[\pi_{j+1}^s(Z_{j+1}^h) V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} = 0 | Z^c, e, \theta) \middle| Z_j\right] \\ & + \nu_j^h h_j + \nu_j^\ell \ell_j + \nu_j^l (T_j(Z^c, e) - \ell_j - h_j) \big\}. \end{aligned} \tag{52}$$

- $\underline{J} \leq j < \overline{J}$ and $\zeta_j = 0$ (not yet retired):

$$\begin{aligned} V_j(W_j, B_j, Z_j, \zeta_j | Z^c, e, \theta) = \max_{m_j, x_j, \zeta_{j+1}} \big\{ & U_j(x_j, \zeta_j | Z^c, e, \theta) \\ & + \beta\, \mathbb{E}_{m_j}\left[\pi_{j+1}^s(Z_{j+1}^h) V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} | Z^c, e, \theta) \middle| Z_j\right] \\ & + \nu_j^h h_j + \nu_j^\ell \ell_j + \nu_j^l (T_j(Z^c, e) - \ell_j - h_j) \big\} \end{aligned} \tag{53}$$

- $\underline{J} \leq j \leq \overline{J}$ and $\zeta_j = 1$ (individual is already retired) or $j \geq \overline{J}$:

$$\begin{aligned} V_j(W_j, B_j, Z_j, \zeta_j = 1 | Z^c, e, \theta) = \max_{m_j, x_j} \big\{ & U_j(x_j, \zeta_j = 1 | Z^c, e, \theta) \\ & + \beta\, \mathbb{E}_{m_j}\left[\pi_{j+1}^s(Z_{j+1}^h) V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} = 1 | Z^c, e, \theta) \middle| Z_j\right] \\ & + \nu_j^h h_j + \nu_j^\ell \ell_j + \nu_j^l (T_j(Z^c, e) - \ell_j - h_j) \big\}. \end{aligned} \tag{54}$$

The optimization problem is subject to the inter temporal budget constraint:

$$W_{j+1} = \tilde{R}_j W_j + (1-\zeta_j)[\tilde{w}_j y_{Lj}(Z_j, e, \theta) + \mathrm{tr}_j(Z_j^u)] + \zeta_j B_j - \tilde{p}_j^c(c_j + i_j) - \tilde{p}_j^m\, m_j \tag{55}$$

and the dynamics of pension benefits:

$$B_{j+1} = \begin{cases} \widehat{R}_j \phi_{j+1}^S(e) B_j + \varphi_{j+1}(e)\tau_j^S w_j y_{Lj}(Z_j, e, \theta) & \text{for } \zeta_j = 0, \\ B_j & \text{for } \zeta_j = 1. \end{cases} \tag{56}$$

The remaining constraints are given by

$$\begin{aligned} h_j &\geq 0 \\ \ell_j &\geq 0 \\ T_j(Z^c, e) - \ell_j - h_j &\geq 0 \\ \nu_j^h h_j &= 0 \\ \nu_j^\ell \ell_j &= 0 \\ \nu_j^l (T_j(Z^c, e) - \ell_j - h_j) &= 0 \\ \nu_j^h, \nu_j^\ell, \nu_j^l &\geq 0 \end{aligned}$$

which results from the limited total time available, and $\nu_j^h, \nu_j^\ell, \nu_j^l$ denote the Kuhn-Tucker multipliers necessary to solve the constrained optimization problem. Further, income is given by

$$y_{Lj}(Z_j, e, \theta) = \mathrm{gap}_j^g \epsilon_j(e, \theta) \mathbf{1}_E(Z_j^u)(1 - \delta^h \mathbf{1}_B(Z_j^h))(T(Z^c, e) - \ell_j - h_j) \tag{57}$$

We derive first-order conditions (FOCs) by differentiating (52) with respect to every control variable. Before the individual reaches the minimum retirement age, i.e. $j < \underline{J}$, the FOCs only have to be solved for $\zeta_j = 0$. As soon as $j \geq \underline{J}$, FOCs for both $\zeta_j = 0$ and $\zeta_j = 1$ have to be calculated and both sets of optimal control variables have to be compared. Then the set of controls with highest expected lifetime utility is chosen. If $\zeta_{\tilde{j}} = 1$ for some $\tilde{j}$, then $\zeta_j = 1$ for all $j \geq \tilde{j}$, so the FOCs have only to be calculated for $\zeta_j = 1$.

$$\begin{aligned}
m:\quad & \sum_{Z_{j+1}^u \in \mathbf{L}} \sum_{Z_{j+1}^h \in \{G,B\}} \frac{\partial \pi^h(Z_{j+1}^h, Z_j^h | Z_j^u, t, m_j)}{\partial m_j} \pi^u(Z_{j+1}^u | Z_j^u) \pi_{j+1}^s(Z_{j+1}^h) \\
& \quad \times V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}^u, Z_{j+1}^h, \zeta_{j+1} | Z^c, e, \theta) \\
& = \tilde{R}_j \tilde{p}_j^m \mathbb{E}_{m_j} \left[\pi_{j+1}^s(Z_{j+1}^h) \frac{\partial V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} | Z^c, e, \theta)}{\partial W_{j+1}} \middle| Z_j^h, Z_j^u \right] \\
c:\quad & U_c(x_j, \zeta_j | Z_j, e, \theta) = \tilde{R}_j \tilde{p}_j^c \beta\, \mathbb{E}_{m_j} \left[\pi_{j+1}^s(Z_{j+1}^h) \frac{\partial V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} | Z^c, e, \theta)}{\partial W_{j+1}} \middle| Z_j^h, Z_j^u \right] \\
i:\quad & U_i(x_j, \zeta_j | Z_j, e, \theta) = \tilde{R}_j \tilde{p}_j^c \beta\, \mathbb{E}_{m_j} \left[\pi_{j+1}^s(Z_{j+1}^h) \frac{\partial V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} | Z^c, e, \theta)}{\partial W_{j+1}} \middle| Z_j^h, Z_j^u \right] \\
& U_h(x_j, \zeta_j | Z_j, e, \theta) = (1-\zeta_j) \mathrm{gap}_j^g \epsilon_j(e, \theta) \mathbf{1}_E(Z_j^u)(1 - \delta^h \mathbf{1}_B(Z_j^h)) \\
h:\quad & \quad \times \beta\, \mathbb{E}_{m_j} \left[\begin{array}{c} \pi_{j+1}^s(Z_{j+1}^h) \frac{\partial V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} | Z^c, e, \theta)}{\partial W_{j+1}} \tilde{w}_j + \\ + \pi_{j+1}^s(Z_{j+1}^h) \frac{\partial V_{j+1}(W_{j+1}, B_{j+1}, Z_{j+1}, \zeta_{j+1} | Z^c, e, \theta)}{\partial B_{j+1}} \varphi_{j+1}(e) \tau_j^S w_j \end{array} \middle| Z_j^h, Z_j^u \right] \\
& \quad - \nu_j^h + \nu_j^l
\end{aligned} \tag{58}$$

$$U_\ell(x_j,\zeta_j|Z_j,e,\theta) = (1-\zeta_j)\text{gap}_j^g \epsilon_j(e,\theta)\mathbf{1}_E(Z_j^u)(1-\delta^h\mathbf{1}_B(Z_j^h))$$

$$\ell: \quad \times\, \beta\, \mathbb{E}_{m_j}\left[\begin{array}{c} \pi^s_{j+1}(Z^h_{j+1})\frac{\partial V_{j+1}(W_{j+1},B_{j+1},Z_{j+1},\zeta_{j+1}|Z^c,e,\theta)}{\partial W_{j+1}}\tilde{w}_j + \\ +\,\pi^s_{j+1}(Z^h_{j+1})\frac{\partial V_{j+1}(W_{j+1},B_{j+1},Z_{j+1},\zeta_{j+1}|Z^c,e,\theta)}{\partial B_{j+1}}\varphi_{j+1}(e)\tau_j^S w_j \end{array}\middle| Z_j^h, Z_j^u\right]$$

$$-\,\nu_j^\ell + \nu_j^l$$

$$\zeta: \quad \zeta^*_{j+1} = \arg\max_{\zeta\in\{0,1\}}\left(U_j(x_j,\zeta|Z^c,e,\theta) + \beta\,\mathbb{E}_{m_j}\left[\pi^s_{j+1}(Z^h_{j+1})V_{j+1}(W_{j+1},B_{j+1},Z_{j+1},\zeta|Z^c,e,\theta)\middle|Z_j\right]\right) \quad (59)$$

Further we obtain the envelope condition (EC) by differentiating (52) with respect to W_j:

$$W: \quad \frac{\partial V_j(W_j,B_j,Z_j,\zeta_j|Z^c,e,\theta)}{\partial W_j} = \tilde{R}_j\,\beta\,\mathbb{E}_{m_j}\left[\pi^s_{j+1}(Z^h_{j+1})\frac{\partial V_{j+1}(W_{j+1},B_{j+1},Z_{j+1},\zeta_{j+1}|Z^c,e,\theta)}{\partial W_{j+1}}\middle|Z_j\right] \quad (60)$$

$$B: \quad \frac{\partial V_j(W_j,B_j,Z_j,\zeta_j|Z^c,e,\theta)}{\partial B_j} = \begin{cases} \widehat{R}_j\,\phi^S_{j+1}(e)\,\beta\,\mathbb{E}_{m_j}\left[\pi^s_{j+1}(Z^h_{j+1})\frac{\partial V_{j+1}(W_{j+1},B_{j+1},Z_{j+1}|\zeta_{j+1},Z^c,e,\theta)}{\partial B_{j+1}}\middle|Z_j\right] & \text{for } \zeta_j = 0 \\ \beta\,\mathbb{E}_{m_j}\left[\pi^s_{j+1}(Z^h_{j+1})\frac{\partial V_{j+1}(W_{j+1},B_{j+1},Z_{j+1},,\zeta_{j+1}|Z^c,e,\theta)}{\partial B_{j+1}}\middle|Z_j\right] & \text{for } \zeta_j = 1 \end{cases} \quad (61)$$

F Parametrization and Calibration

Most parameters of the model will be estimated combining information from the existing literature, the optimal conditions, and data provided by Work Package 3. To estimate the parameters governing the unobservable variables, the model will be structurally calibrated. In particular, we will replicate the educational distribution and the income distribution of several European countries. The calibration will be done applying the Bayesian melding method (Poole and Raftery (2000)) with the incremental mixture importance sampling (IMIS) algorithm (Raftery and Bao, 2010). Within our model framework we will then study how different pension reforms may induce a redistribution across different SES groups.

The model will be implemented with economic and demographic data from various sources, including SHARE, EU-SILC, National (Time) Transfer Accounts, National accounts, WIC Human Capital Explorer, Eurostat, and others, for a specific group of European countries.