Inequality and Defined Benefit Pensions when Life Expectancy is Heterogeneous

Peter Haan*, Daniel Kemptner†, and Victoria Prowse‡

February 12, 2015

Abstract

In this paper, we analyze how life expectancy-driven redistribution of income through a defined pension benefit system impacts on inequality in annual consumption. Our analysis combines a methodology that quantifies life expectancy-driven redistribution through the pension system with a structural life-cycle model in which labor supply, retirement and consumption decisions respond to changes in the pension system. Based on the estimated model, we show that the German pension system induces a large regressive redistribution of life-time income, and this redistribution increases inequality in average annual consumption. Behavioral responses to the pension system matter for the results. Increasing progressivity in pension contributions or pension benefits only partially offsets the life expectancy-driven redistribution via the pension system.

Key words: Defined benefit pension systems; Redistribution; Inequality; Life expectancy; Pension reform; Structural life-cycle models; Retirement; Labor supply.

JEL Classification: C61;D31;H31;H55;J26

*FU Berlin and DIW Berlin, phaan@diw.de

†DIW Berlin, dkemptner@diw.de

‡Department of Economics, Cornell University, prowse@cornell.edu

1 Introduction

Public defined benefit pension systems typically provide retired individuals with an annual pension that is linked to life-time earnings. While there exists substantial variety in the design of public defined benefit pension systems across countries, large redistributions between individuals are near universal. Reflecting this, there have been extensive discussions in academic, political and policy circles about the distributional effects of defined benefit pension systems. These discussions often focus on the distributional effects that arise directly from the design of the pension system. For example, Gustman and Steinmeier (2001) show that US Social Security has substantial progressive distributional effects that are due to the benefit formula, which deliberately favors low income groups. Contribution-independent minimum pension entitlements and child-related pension bonuses also induce important redistributions that favor, respectively, low income individuals and individuals with children (see, e.g., Feldstein and Liebman (2002)).

Defined benefit pension systems have additional redistributive effects that are driven by heterogeneity in life expectancy. Focusing on the redistribution of life-time income, i.e., the total sum of an individual's income over the entire life-cycle, Coronado and Glass (2000) and Gustman and Steinmeier (2001), among others, show that the positive relationship between life-time earnings and life expectancy reduces the progressivity of the US Social Security system.[1]

In this paper, we also examine potentially regressive, redistributive effects that are driven by heterogeneity in life expectancy. However, building on the previous literature, we shift the focus from life-time income to annual outcomes, and examine the link between the pension system and inequality in annual consumption. Our extension is motivated by the possibility that the previously-documented regressive redistributive effects may extend beyond providing longevity insurance for extra years of life, and actually work to increase differences in annual living standards. Any regressive effects on average annual consumption are likely to be an unintended feature of the pension system. Moreover, abstracting from interactions with other aspects of the fiscal system, regressive effects on average annual consumption are generally difficult to justify from a normative viewpoint.

The contribution of the paper is three fold. First, we propose a methodology that provides a general quantification of the life expectancy-driven distributive effects of defined benefit pension systems. The methodology relies on a life expectancy adjusted benchmark pension system, which eliminates life expectancy-driven redistribution but maintains all other features of the pension system, such as the benefits schedule and child-related bonuses. A comparison of the distribution of average annual consumption

[1] Numerous studies document a positive relationship between life-time earnings and life expectancy, e.g., Cutler, Deaton, and Lleras-Muney (2006). See Feldstein and Liebman (2002) for further discussion of the interplay between earnings, life expectancy of the redistribution of life-time income.

under a particular pension system and under the life expectancy adjusted pension system reveals how inequality in annual consumption is impacted by life expectancy-driven redistribution through the pension system.

Second, we present empirical analysis that quantifies the life expectancy-driven redistribution that is inherent in the current German public pension system. We show that the German pension system induces a large regressive redistribution of life-time income, due to differences in life expectancy. Moreover, the life expectancy-driven regressive redistribution of life-time income translates into an increase in the inequality of average annual consumption. Our results on life expectancy driven redistribution through the pension system are obtained by combining the life expectancy adjusted benchmark pension system with a life-cycle model of labor supply, retirement and consumption decisions. The life-cycle model includes labor market frictions, health shocks, a realistic specification of the tax, transfer and pension system, and heterogeneity in life expectancy arising from individual-level differences in health status and educational attainment.[2] Individuals are forward looking and therefore make labor supply, retirement and savings decision taking into account future pension benefits and life expectancy. The life-cycle model thus allows labor supply, retirement and consumption decisions to respond to pensions reforms, and provides a grounded prediction of individual behavior under the life expectancy adjusted pensions system. Our results reveal that behavioral responses are an important component of the distributional effects of pension reforms, thus justifying the operationalization of the benchmark scenario via a behavioral model. By using the life-cycle model to study distributive concerns related to differences in life expectancy, we extend the application of structural retirement models, which previously have been used to study the impact of Social Security on the timing of retirement (French, 2005), interactions between liquidity constraints and retirement (Rust and Phelan, 1997), fiscal sustainability in the face of increasing longevity (Haan and Prowse, 2014), household retirement decisions (van der Klaauw and Wolpin, 2008).

Third, our analysis becomes more policy orientated and we explore two reforms designed to reduce regressive redistribution present in the current German public pension system. One reform increases the progressivity of the contribution schedule and the other reform increases the progressivity of the benefit schedule. The behavioral effects of the two reforms differ, with the increase in the progressivity of contributions having more favorable effects on the labor supply and consumption outcomes of low-income individuals. Both reforms are only partially able to offset the life expectancy-driven redistribution via the pension system – a more effective reform would tie pension contributions or benefits

[2]This aspect of the specification is in line with regression-based evidence suggesting that the partial effect of income is fairly small after conditioning on socio-economic characteristics (see, e.g., Lleras-Muney (2005)).

directly to the socio-economic determinants of life expectancy (as occurs in the hypothetical benchmark pension system).

The paper is organized as follows. In section 2 we derive a general a methodology that quantifies the life expectancy-driven regressive redistribution in a given pension system. Then, using data from the federal statistical office in combination with the SOEP we provide empirical evidence about the heterogeneity in life-expectancy with a particular focus on education and health. In Sections 4 and 5, we provide information about the data, derive the structural life-cycle model and discuss estimation. Section 6 presents the estimation results of the model including a detailed discussion about the in-sample fit of the model and similar to Low and Pistaferri (2010) we show that the implications of the model are consistent with previous literature exploiting policy reforms for identification. In section 8 we present the simulation results and discuss the distributional effects of the current German pension system and we evaluate the distributional, behavioral and welfare effects of potential pension reforms. Finally, Section 9 summarizes the central results and concludes.

2 Life expectancy-adjusted scheme

In the following we derive a general a methodology that quantifies life expectancy-based redistribution in a given pension system. In particular, we discuss the life expectancy-related distributional effects using a life expectancy-adjusted benchmark scenario. In this benchmark we shut down life expectancy-based redistribution of life-time income while retaining other sources of redistribution (e.g. minimum pension provisions, a progressive schedule or family-related bonuses). In summary, we propose a distributional mechanism that redistributes expected extra years of pension income enjoyed by high life expectancy individuals because of their higher life expectancies proportional to annual pension benefits under the baseline rules. This mechanism guarantees that the relative annual claims of the individuals are maintained. A comparison of individual outcomes, such as life-time income, net pension benefits, or average annual consumption under the current pension formula with the respective outcomes under the adjusted formula provides estimates for the distributional analysis with respect to heterogeneity .

As a starting point we consider two types of individuals in the economy, individuals with high life expectancy (H) or with low life expectancy (L). This abstraction is sufficient to explain the distributional mechanism and alines with the assumptions of the structural model in which we approximate heterogeneity in life-expectancy by two educational types. We describe the methodology in several steps. First, we compute the sum of expected pension income that all the high life expectancy individuals receive extra thanks to their higher life expectancies (pension pot):

$$pot = \sum_{n=1}^{N_H} \sum_{t=1}^{T} [S(t|1, type_n = H) - S(t|1, type_n = L)]\, P_{n,t}^{old}(Hist_n, RAge_n)$$

where N_H is the number of high life expectancy individuals, T is the last period until which individuals may possibly live up, $S(t|1, type_n)$ is the survival probability until period t in period 1 for a given type, being either H or L. $P_{n,t}^{old}(Hist_n, RAge_n)$ is the pension benefits that individual n receives in a certain period t under the old pension formula. The benefits are a function of an individual's wage and employment histories (summarized by $Hist_n$) and retirement age $RAge_n$. Before retirement, $P_{n,t}^{old}(Hist_n, RAge_n) = 0$ because no pension benefits are received. The dynamic incentives of the pension system affect pension benefits through employment and retirement choices.

In a second step we redistribute the pension pot among all individuals that are in retirement according to their relative annual pension claims under the old formula. Hence, the individuals' shares of the pension pot are given by

$$share_n = \frac{P_{n,R}^{old}(Hist_n, RAge_n)}{\sum_{n=1}^{N_R} P_{n,R}^{old}(Hist_n, RAge_n)}$$

where N_R is the number of individuals that survive until retirement and $t = R$ indicates a period where the individual is retired.

Finally, an individual's life expectancy-adjusted annual pension benefits are given by the following formula:

$$P_{n,R}^{new}(Exp_n, RAge_n) = \frac{\left[LE(RAge_n, type_n = L) \times P_{n,R}^{old}(Hist_n, RAge_n)\right] + [Share_n \times pot]}{LE(RAge_n, type_n)}$$

where $LE(RAge_n, type_n = L)$ indicates the life expectancy of individual n at retirement assuming that the individual is low educated. On top of the pension income that the individual would have received in expectation under a low life expectancy, the individual receives its share of the pension pot. The total expected pension income, then, is annuitized by dividing through the actual life expectancy of individual n at retirement.

Since a change in the pension formula affects the dynamic incentives with respect to employment and retirement behavior, public revenues with respect to pension contributions and taxes also change. In order to ensure that the new pension formula does not affect the government budget, a further constraint is introduced that completes the adjustment formula:

$$BD^{old} = FSurplus^{new} + [Contr^{new} + \tau] - \sum_{n=1}^{N} \sum_{t=1}^{T} P_{n,t}^{new}(Hist_n, RAge_n) S(t|1, type_n)$$

where BD^{old} is the budget deficit under the old pension formula, $FSurplus^{new}$ is the new fiscal surplus, and $Contr^{new}$ are the new pension contributions. The last term indicates the new sum of expected pension benefits. τ captures changes in the pension contributions that are due to changes in the respective formula. The contribution formula is adjusted in order to make the constraint binding and to ensure budget neutrality of the life-expectancy adjusted pension formula.

3 Heterogeneous life expectancies

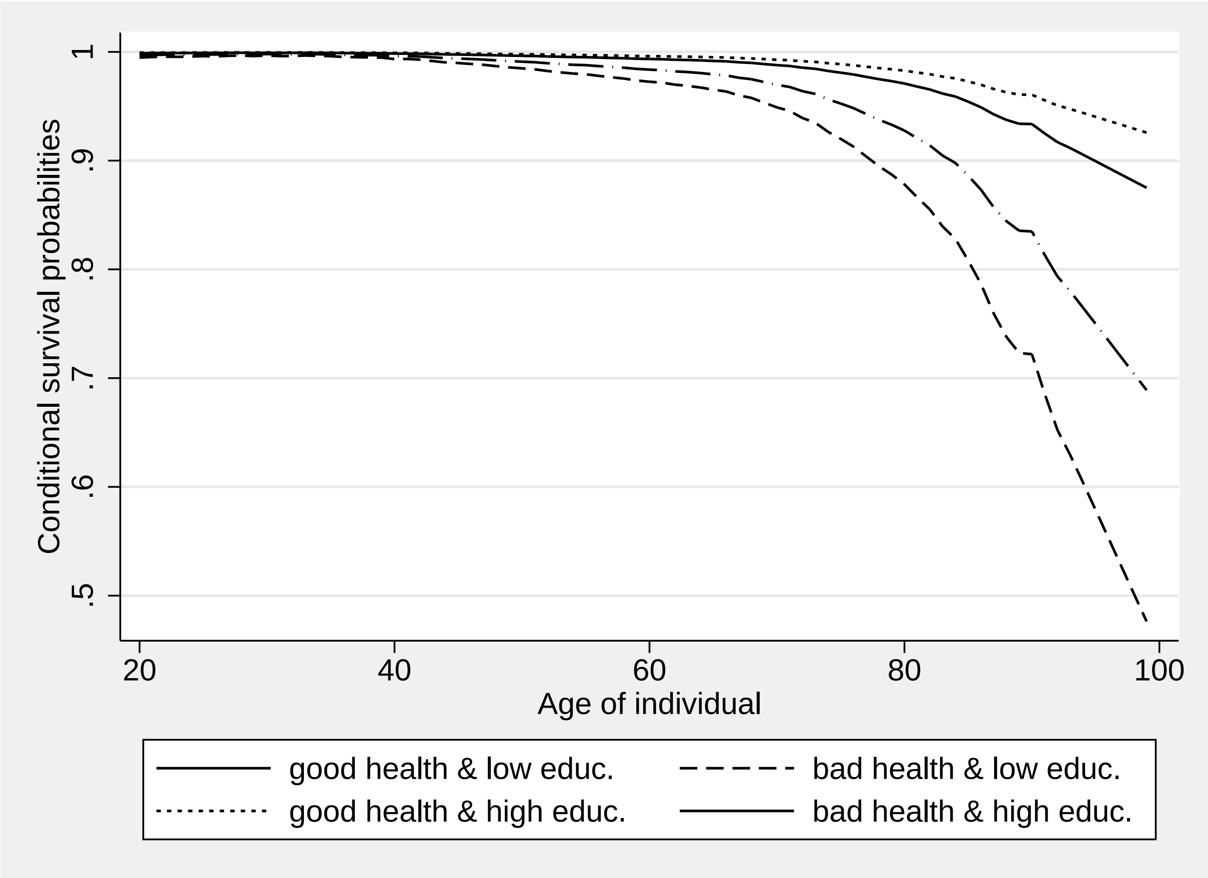
Previous literature has documented heterogeneity in life-expectancy and numerous studies have estimated the effects of socio-economic determinants on life-expectancy, see e.g. Cutler, Deaton, and Lleras-Muney (2006). In particular Lleras-Muney (2005) provides evidence that education, controlling for income, has a sizable effect on life expectancy. Based on this evidence we introduced heterogeneity by education in the life-cycle model. Education is assumed to affect life expectancy both directly and indirectly through the health risks also account for heterogeneity by education. As a consequence, we need estimates of the partial effects of both education and health on conditional survival probabilities.

In the following we combine data from the SOEP with information from the life tables of the federal statistical office to document the heterogeneity in life-expectancy by education. Moreover, the empirical estimates serve as an input for the structural life-cycle model to construct education- and health-specific survival probabilities. We estimate conditional survival probabilities using a proportional hazard model with a Weibull distribution based on an unbalanced panel of male individuals taken from the SOEP the survey years 2004-2008 (48,921 observations). Furthermore, we exploit information from a follow-up survey on death cases of individuals who have left the SOEP study. This survey has been collected for the last time in 2008 and allows reducing attrition bias in the estimates for the conditional survival probabilities (because individuals are likely to leave the study before they die). This leaves us with 467 death cases in our sample.

Hazard rates are estimated by health status (good/bad) and level of education (years of education ≥ 12 or < 12). Good health is defined as neither having legal disability status nor assessing own health as bad or very bad (see next section for more details). Education affects predicted life expectancies both directly and indirectly. The indirect effect results from education also affecting the probability of good health status. We find substantial heterogeneity both by health status and level of education. The conditional effect of health status seems to be more important. When computing average life expectancy at age 20 based on these estimates, we under predict the average life expectancy that is calculated by the federal statistical office by about 3 years.[3] This difference might be related to the fact that the follow-up survey on death cases may not have captured all the cases.

[3]This finding is insensitive to changes in the distributional assumptions of the proportional hazard model.

Figure 1: Predicted conditional survival probabilities (adjusted)



Therefore we combine the information from the statistical office with the evidence from the SOEP. In more detail, we use the estimates from the SOEP with respect to the relative changes that are induced by health and education and adjust the predicted conditional survival probabilities such that the average probabilities equal the respective average probabilities that are presented in the life tables. Figure 1 shows predicted conditional survival probabilities.

4 Data and descriptive statistics

For the estimation of our structural life cycle model, we use longitudinal data from the SOEP. We construct an unbalanced panel covering the years 2004 to 2012. The sample is restricted to males aged 20-64 years in West Germany and excludes self-employed and civil servants. We consider the age cohorts 20 to 64 because the individuals' behavior is only modeled until the statutory retirement age of 65 years. The final sample consists of 14,552 observations on 3,128 individuals.

Table 1: Descriptive statistics

Variable	Mean			Median	Std.
	Full sample	Low education	High education	Full sample	
Age	44.8	46.3	43.4	44	9.6
Employed	0.87	0.83	0.91	1	0.34
Retired	0.05	0.07	0.03	0	0.22
Hourly wage (€)	16.7	14.4	18.7	15.5	6.3
Years of education	12.3	10.5	14	12	2.47
Good health	0.84	0.8	0.88	1	0.37
Work experience	21.7	24.6	19	22	10.3
Savings before retirement (€)	4,334	3,303	5,289	3,101	6,263
Net wealth before retirement (€)	105,639	92,283	118,000	52,102	123,540

For our analysis, we use information on employment (full-time or non-employment),[4] retirement status, gross wages, work experience, years of education, binary health status, net wealth, and savings. Education is measured as years of education and we use this variable to define two groups: years ≥ 12 and years < 12.[5] Work experience is defined as years of full-time experience, where one year of pre-sample part-time experience is counted as half a year of full-time experience. Wealth information is contained in the SOEP only every 5 years. In 2007, the information comprises market values of real estates, financial assets, building loan contracts, private insurances, business assets, tangible assets, consumer debts, and overall debts. We compute net wealth by combining the information on gross wealth and debts. The variable is imputed for the other survey years when it is unobserved.[6]

We follow the approach of Schündeln (2008) defining total savings are defined as the sum of financial and real savings. The SOEP participants indicate their financial savings

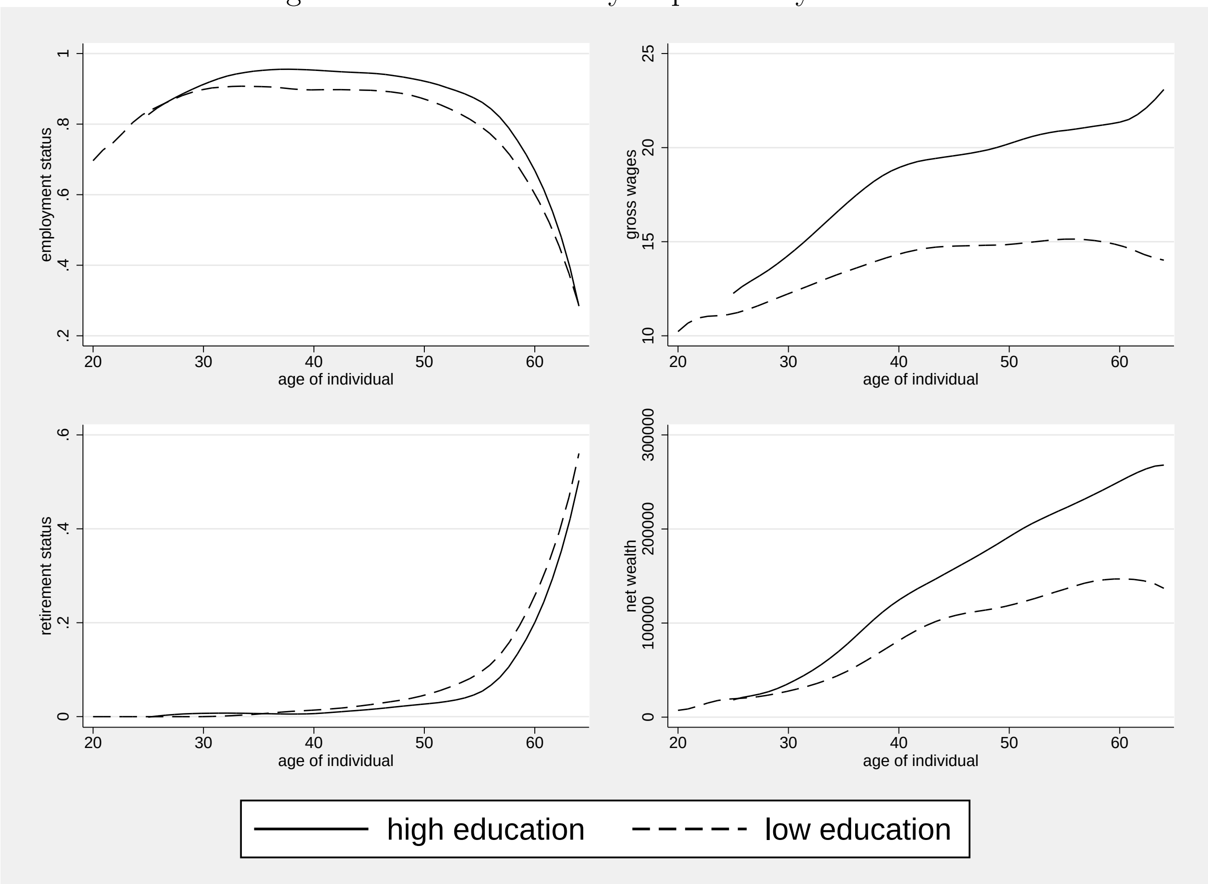
[4]Employment is defined as working at least 20 hours per week and median hours of work for the employees is 40. Note, only very few men work part-time and they are considered to be non-employed.

[5]The SOEP constructs the years of education variable from respondents' information on the obtained level of education and adds some time for additional occupational training.

[6]This is done by using information on saving behavior and carrying forward net wealth under some assumptions from the year 2007 to the other survey years. We assume that individuals borrow at a real interest rate of 6% and receive a real interest rate of 2% on both their financial and real savings. Moreover, we take into account observed capital losses. In order to make the wealth measure consistent with our model assumptions, we introduce a censoring such that individuals always have non-negative wealth and can have at most as much wealth as they could possibly have accumulated within our life cycle model until their respective ages.

annually by answering a question about the “usual” amount of monthly savings.[7] Real savings are defined as annual amortization payments.[8] Since saving information in the SOEP is left-censored (dissavings are unobserved), we assume that working individuals aged 20 to 64 have non-negative net savings over the period of a whole year and make assumptions on the dissavings of the unemployed and retirees. Unemployed individuals are assumed to dissave in the case that they are not eligible for unemployment insurance benefits and fail the means test required for social assistance benefits.[9] Retirees are assumed to dissave according to the value of an actuarially fair life annuity that could be bought with their accumulated wealth at retirement.

Figure 2: Estimated life cycle profiles by education



[7]Question: “Do you usually have an amount of money left over at the end of the month that you can save for larger purchases, emergency expenses or to acquire wealth? If yes, how much?”

[8]Since the SOEP question asks for the sum of amortization and interest payments, the share of interest payments must be derived from information on the amount of debts.

[9]These individuals receive an income at the minimum income level (social assistance benefits) that is deducted from their net wealth, where €10,000 are exempted from the means test that is required for social assistance benefits. The exemption level of €10,000 is assumed because the actual rules are very complicated and enforcement of these rules is unobserved.

Our analysis focuses on redistribution that arises between the high and the low educated through the heterogeneity in their life expectancies. Therefore, our life cycle model must be set up such that it allows accounting for relevant differences in the life cycle profiles between these two groups. In particular, the model should capture differences with respect to employment, retirement, wages, and wealth accumulation. The respective life cycle profile are presented in figure 2. These profiles are computed by separately estimating local polynomial regressions of the outcome variables on age for the low and the high educated. It turns out that the low educated are less employed at all ages, retire earlier, and accumulated substantially less net wealth until the age of 65. Wage profiles are diverging between the two groups over the life cycle. This suggests that the low educated receive lower returns to their work experience.

5 Structural life-cycle model

We propose to analyze the distributional effects related to the heterogeneity in life-expectancy within a dynamic structural life-cycle model which accounts for labor supply, retirement and consumption responses to changes in the pension system. Therefore in our analysis we can incorporate sizable behavioral responses which are related to the distributional effects. Moreover the structural model allows the evaluation of counterfactual pension reforms which we consider in the last section of the paper.

The structural life-cycle shares many similarities with French (2005), or French and Jones (2011). In particular in the model individuals make labor supply, retirement and consumption choices to maximize their utility over the life-cycle; crucially we allow for heterogeneity in life-expectancy. Moreover the model accounts for potential frictions and several sources of risk related to stochastic job offers and involuntary separations, and a stochastic wage and health process. Finally, we include a detailed specification of taxation, transfers and the pension system. A key variable in the model is education which affects directly the wage and health process, frictions on the labor market and most important life-expectancy.

5.1 Objective function

We propose a discrete-time life-cycle model with finite horizon. Discrete time is indexed by t (individual's age), and the final period of an individual is denoted with T. Each individual n receives a utility flow $\mathbb{U}(\mathbf{s}_{nt}, d_{nt})$ in period t where $\mathbf{s}_{nt}$ is a vector of state variables, and d_{nt} indicates the individual's choice. Every period t, an individual n observes the state variables $\mathbf{s}_{nt}$ and selects the choice d_{nt} from the choice set $D(\mathbf{s}_{nt})$ that maximizes expected lifetime utility:

$$\mathbb{E}\left[\sum_{j=0}^{T-t} p(t+j, \mathbf{s}_{nt})\beta^j \mathbb{U}(\mathbf{s}_{nt+j}, d_{nt+j})\right]$$

where β is the time discount factor (set to be 0.97) and $p(t+j, \mathbf{s}_{nt})$ is the conditional survival probability for period t+j given survival until period t. $p(t+j, \mathbf{s}_{nt})$ varies between individuals by health status and level of education (high/low) which introduces heterogeneity in life-expectancy. The choice set is restricted by the eligibility requirements for early retirement which are related to age and health and by job offer and separation rates that are estimated within the model. Before retirement, individuals choose between working, not working and early retirement. Furthermore, they make a saving decision.[10] Individuals retire no later than the statutory pension age which is 65 in our sample period. After retirement, individuals consume according to the value of an actuarially fair life annuity to decumulate their stock of wealth.

5.2 Utility function

Individuals have preferences over consumption and leisure time that are represented by the following time separable random utility function:

$$\mathbb{U}(\mathbf{s}_{nt}, d_{nt}) = \alpha_n \frac{c(\mathbf{s}_{nt}, d_{nt})^{(1-\rho_m)} - 1}{(1-\rho_m)} + \epsilon_{nt}(d_{nt})$$

$$\alpha_n = \alpha_1 + \alpha_{2m} work(d_{nt})$$

where $\epsilon_{nt}(d_{nt})$ is assumed to be type 1 extreme value distributed. $c(\mathbf{s}_{nt}, d_{nt})$ is the level of consumption associated with state $\mathbf{s}_{nt}$ and choice d_{nt}. $work(d_{nt})$ indicates employment. α_n is a consumption weight that depends on a constant (scaling factor) and the parameter α_{2n} reflects unobserved heterogeneity in the disutility for work. ρ is the coefficient of relative risk aversion. In line with e.g. Attanasio and Weber (1995) we account for non-separability between consumption and leisure time. The vector $\boldsymbol{\theta}_U = (\alpha_1, \rho, \alpha_{2n})$ contains the parameters of the utility function.

5.3 Value function

Individuals' beliefs about future states are captured by a Markov transition function $q(\mathbf{s}_{nt+1}|\mathbf{s}_{nt}, d_{nt})$ that indicates the transition probabilities. In particular, $q(\mathbf{s}_{nt+1}|s_{nt}, d_{nt})$

[10] As standard in the literature we approximate the continues savings decision. In line with the empirical distribution, the approximation differs by employment status and captures the fact that some unemployed individuals are dissaving. In more detail we define 5 grid points for the employed and 3 grid points for the unemployed. The results are insensitive to an increase in the number of grid points and to the location of these points.

captures expectations about the transitions of the health status and the expectations of unemployed individuals to receive a job offer and of employed individuals to face a job separation in the following period (see below). For state variables that evolve deterministically for given choices, the probability of the determined state is one while it is zero for all other possible states of the variable (e.g. net wealth, work experience). The value function $\mathbb{V}_t(\mathbf{s}_{nt})$ can be represented recursively as

$$\begin{aligned}\mathbb{V}_t(\mathbf{s}_{nt}) = \max_{d_{nt}\in\mathbb{D}(\mathbf{s}_{nt})} & \mathbb{U}(\mathbf{s}_{nt}, d_{nt}) + \\ & p(t+1, \mathbf{s}_{nt})\beta \int_{\epsilon} \left[\sum_{\mathbf{s}_{nt+1}} \mathbb{V}_{t+1}(\mathbf{s}_{nt+1}) q(\mathbf{s}_{nt+1}|\mathbf{s}_{nt}, d_{nt}) \right] g(\epsilon_{nt+1})\end{aligned}$$

where $g(\cdot)$ is the probability density function of the unobserved random components of the utility function. $\mathbb{D}(\mathbf{s}_{nt})$ is the choice set available to individual n in period t. The choice set is restricted by eligibility requirements for early retirement and by job offer and separation rates.

5.4 Job offer and separation rates

An individual's choice of employment is restricted by job offer and separation rates that are estimated within the model. The offer rates capture persistence of the unemployment status. Individuals who have been unemployed in the previous period may only choose employment if they receive a job offer in the current period. Analogously, individuals who have been employed in the previous period may only choose employment if they do not face a job separation in the current period. The rates are estimated differentially by level of education (high/low), health status, and age ($50 \geq age < 60$, and age≥ 60):

$$\begin{aligned}\Pr(offer_{nt} = 1) &= \Lambda(\phi_1 + \phi_2 educ_n^{high} + \phi_3 health_{nt} + \phi_4 age_{nt}^{50-59} + \phi_5 age_{nt}^{60+}) \\ \Pr(separation_{nt} = 1) &= \Lambda(\phi_6 + \phi_7 educ_n^{high} + \phi_8 health_{nt} + \phi_9 age_{nt}^{50-59} + \phi_{10} age_{nt}^{60+})\end{aligned}$$

where $\Lambda(\cdot)$ is the logistic distribution function. The parameters for the job offer and separation rates are contained by the vector $\boldsymbol{\phi} = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, \phi_8, \phi_9, \phi_{10},)$.

5.5 Health transitions

Age-specific transition probabilities estimated in a first stage. The transition probabilities are computed by estimating local polynomial regressions of health status on age first using the sample of individuals in good health status and, then, using the sample of individuals in bad health status. We do this separately for the high and the low educated in order to account for differential health risks by education.

5.6 Wage equation

The logarithm of gross wages is modeled as

$$\log(wage_{nt}) = \delta_1 educ_n + \delta_2 \log(exper_{nt}) \times (educ_n < 12) + \\ \delta_3 \log(exper_{nt}) \times (educ_n \geq 12) + \kappa_m + \mu_{nt}$$

where $educ_n$ is years of education, $exper_{nt}$ is years of work experience, κ_n is time-constant unobserved heterogeneity, and μ_{nt} is i.i.d. $N(0, \sigma_\mu)$. It is due to the DPDC framework that individuals take into account the human capital accumulation process when making their employment choice. Hence, work experience is an endogenous variable in the model. In the interaction terms between work experience and education account for heterogeneous returns to work experience for the high and the low educated (as reflected by the diverging wage profiles). The correlation between individual-specific leisure preferences and the unobserved component, κ_n, in the wage equation accounts for selection into the labor market. When computing gross labor earnings, I assume that individuals work the median number of hours, which is 40 in the sample. The vector $\boldsymbol{\theta}_w = (\delta_1, \delta_2, \delta_3, \kappa_n, \sigma_\mu)$ contains the parameters of the wage equation.

5.7 Unobserved heterogeneity

Unobserved heterogeneity is modeled semi-parametrically by allowing for a finite number of unobserved types $m \in 1, ..., M$ in the population. The probability that individual n is of type m is given by γ_m, where γ_M is normalized to $1 - \sum_{m=1}^{M-1} \gamma_m$.

5.8 Budget constraint

Individuals face a budget constraint when making their saving/consumption choice. The constraint comprises three equations:

$$\begin{aligned} c(\mathbf{s}_{nt}, d_{nt}) &= \mathbb{G}\left(\mathbf{s}_{nt}, d_{nt}\right) - savings(d_{nt}) \\ wealth_{nt+1} &= (1 + r_t)\left(wealth_{nt} + savings(d_{nt})\right) \\ wealth_{nt} &> 0 \end{aligned}$$

where $c(\mathbf{s}_{nt}, d_{nt})$ is the level of consumption associated with state $\mathbf{s}_{nt}$ and choice d_{nt}, and $G(\cdot)$ indicates net income by applying the rules and regulations of the German tax and transfer system and of the statutory pension insurance. The budget constraint's first equation defines the possible levels of consumption in period t, the second equation describes the wealth accumulation process, and the third equation is a non-negativity constraint. We assume that the forward looking individuals do not expect future changes in the institutional framework. $wealth_{nt}$ is period t's net wealth, r_t is the real interest rate that is set to be 0.02, and $savings(d_{nt})$ is the amount of savings associated with state $\mathbf{s}_{nt}$ and choice d_{nt}. Pension claims are a deterministic function of retirement age, work experience, and past wages that are reconstructed using the wage equation.

5.9 Solving the model

Given the finite horizon of the individual's optimization problem, it can be solved recursively. The expected value function, $v_t(\mathbf{s}_{nt}, d_{nt})$, for period T is simply given by this period's expected utility flow:

$$
\begin{aligned}
&v_T(\mathbf{s}_{nT}, d_{nT}) = u(\mathbf{s}_{nT}, d_{nT}) \\
&v_t(\mathbf{s}_{nt}, d_{nt}) = u(\mathbf{s}_{nt}, d_{nt}) + p(t+1, \mathbf{s}_{nt+1})\beta\times \\
&\quad \sum_{\mathbf{s}_{nt+1}} \log\left[\sum_{d_{nt+1}\in\mathbb{D}(\mathbf{s}_{nt+1})} \exp(v_{t+1}(\mathbf{s}_{nt+1}, d_{nt+1}))\right] q(\mathbf{s}_{nt+1}|\mathbf{s}_{nt}, d_{nt}) \\
&\quad \forall t = 1, \ldots, T-1
\end{aligned}
$$

where $v_t(\mathbf{s}_{nt}, d_{nt})$ is the expected value function (Rust, 1987). The computation of the expected value functions for periods t=65,...,T is comparatively simple because individual choices are only modeled for t=40,...,64. Rust (1987) shows that under the assumptions of additive separability and conditional independence, the conditional choice probabilities have a closed form solution (mixed logit probabilities):

$$
\Pr(d_{nt}|\mathbf{s}_{nt}) = \frac{\exp(v_t(\mathbf{s}_{nt}, d_{nt}))}{\sum_{j\in\mathbb{D}(\mathbf{s}_{nt})} \exp(v_t(\mathbf{s}_{nt}, j))}
$$

When computing choice probabilities, we take into account that the choice of employment is restricted by the job offer and separation probabilities. The expected value functions are computed for a discretized state space in order to save computational time (Keane and Wolpin, 1994). As a consequence, interpolation methods must be used to approximate the functions at the observed values of the state variables. For each of these variables, we define five grid points. The results are insensitive to an increase in the number of these grid points or the choice of interpolation function.

6 Estimation results and model fit

We estimate our model by the method of maximum likelihood. In a first step, we implement a sequential and inefficient Expectation-Maximization algorithm in order to obtain good starting values for a subsequent full information maximum likelihood (FIML) procedure (as proposed by Arcidiacono and Jones (2003)). Using good starting values, the maximum of the log-likelihood function can be found easily by conventional optimization routines supplying a numerical gradient and a BHHH Hessian (see Appendix for details). Table 2 shows the estimates of the efficient FIML estimation.

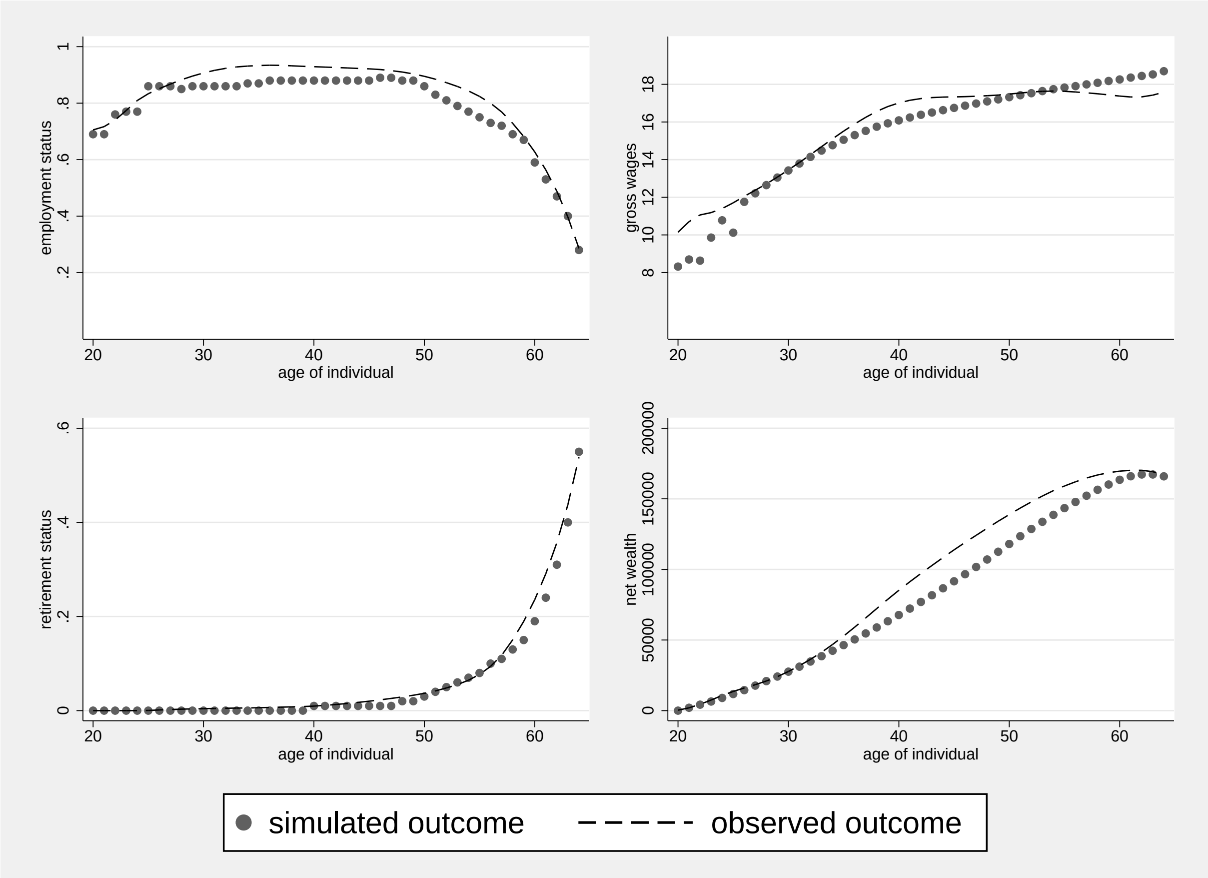
Table 2: Parameter estimates of FIML estimation

	Estimates	St.e.
α_1 (scaling factor)	1.153	(0.0259)
α_{21} (work, type 1)	-0.696	(0.0229)
α_{22} (work, type 2)	-0.108	(0.0298)
ρ_1 (crra, type 1)	0.279	(0.0293)
ρ_2 (crra, type 2)	0.882	(0.0134)
Wage equation:		
κ_1 (constant, type 1)	1.727	(0.013)
κ_2 (constant, type 2)	1.262	(0.0128)
δ_1 (years of education / 10)	0.571	(0.0071)
δ_2 (log(experience)*(educ$<$12))	0.158	(0.0026)
δ_3 (log(experience)*(educ$\geq$12))	0.19	(0.0026)
σ_μ (standard deviation)	0.228	(0.0009)
Job offers and separations:		
ϕ_1 (separation, constant)	-2.753	(0.1226)
ϕ_2 (separation, high educ)	0.085	(0.1017)
ϕ_3 (separation, good health)	-1.347	(0.121)
ϕ_4 (separation, 50$\leq$age$<$60)	0.497	(0.1387)
ϕ_5 (separation, age$\geq$60)	1.598	(0.1535)
ϕ_6 (offer, constant)	-2.052	(0.1412)
ϕ_7 (offer, high educ)	-1.09	(0.0846)
ϕ_8 (offer, good health)	2.264	(0.1477)
ϕ_9 (offer, 50$\leq$age$<$60)	-1.442	(0.1193)
ϕ_{10} (offer, age$\geq$60)	-2.293	(0.2444)
Type probabilities:		
γ_1 (prob. of type 1)	0.511	(0.0106)

The estimates suggest significant unobserved heterogeneity in both the coefficient of relative risk aversion and the disutility of work. The more productive individuals of type 1 (larger constant in the wage equation) are estimated to be less risk averse. The returns to one additional year of education are 5.7% and the high educated are estimated to have larger returns to work experience than the low educated. Health status strongly affects the job offer and separation rates while education only exerts a significant effect on the probability of receiving a job offer. For individuals with age ≥ 50 and even more for age

≥ 60, the probability of job separations rises and the probability of job offers decreases. The estimated probability of being of type 1 suggests that about half of the individuals in the population are of type 1, while the other half is of type 2.

Figure 3: Simulated outcomes and observed life cycle profiles



Using the point estimates of the parameters, we simulate a sample of 5,000 synthetic individuals. The simulations start between age 20 and 26 (depending on education). Initial conditions are drawn from the empirical distribution of education and the estimated distribution of unobserved types. Individuals are assumed to be in good health status and employed when entering the labor force. Choices and random transitions of state variables are based on the respective probabilities and pseudo-random draws from the uniform distribution. Life cycle paths of state variables, social security contributions, tax payments, and received benefits are saved. Figure 3 displays a comparison of simulated outcomes and observed life cycle profiles. The good model fit suggests that our simulations are representative for our sample population.

7 Consistency checks

Similar to Low and Pistaferri (2010) we show that the implications of the model are consistent with previous literature exploiting policy reforms for identification. Table 3 shows the simulated effects of five counterfactual scenarios. These simulation outcomes allow checking the model's consistency with respect to the behavioral margins that matter most for our analysis. We consider three behavioral outcomes: the change in (a) retirement age, (b) work experience at age 65, and (c) accumulated net wealth at age 65. Unlike the policy simulations that we present in the following section, these simulations are not made budget neutral by adjusting e.g. the social security contribution rate because they merely aim at checking the consistency of the behavioral responses.

First, we simulate the abolishment of early retirement disincentives that penalize individuals who opt for early retirement by up to 18 % of their annual pension benefits (0.3% reduction per month of early retirement). This induces substantial behavioral responses (composed of a substitution and income effect). Second, we simulate the pure income effect of the abolishment of the early retirement disincentives by giving individuals a lump sum increase on their pension benefits that equals the average rise in pension benefits that individuals would have enjoyed without behavioral responses to the abolishment of the disincentives (average income effect: +407€). We simulate the lump sum transfer (a) before the means-test and (b) after the means-test. This differentiation matters in particular for low income individuals. The behavioral effects are smaller when the transfer is made before the means-testing. As expected, this difference is largest for the low educated. The income effects on individual behavior appear to be fairly small in comparison to the substitution effect (compare scenarios (I) and (II)). This finding is consistent with a recent study by Manoli et al. (2011). Relying on policy changes for identification, they estimated social security wealth and accrual elasticities in individuals' retirement decisions in Austria.

In the fourth scenario, the statutory pension age is raised by one year from age 65 to 66. This induces a rise in overall lifetime employment by 0.46 years and results in an average postponement of retirement by 0.72 years. These predictions are in line with a study by Mastrobuoni (2009). Mastrobuoni exploits a policy change in the U.S. that increased the national retirement age (NRA) from 65 to 67 and raised the penalty for claiming retirement benefits before the NRA. He concludes that an increase in the NRA by 2 months delays effective retirement by around 1 month. Of course, the German institutional framework differs in some regards from the situation in the US. The general setup where individuals incur penalties for early retirement and receive a respective bonus when retiring after age 65 is, however, similar in the two countries.

At last, we simulate an increase in individuals' life expectancies by five years. This scenario is meant to give an idea with respect to the link between life expectancy and

Table 3: Simulated treatment effects for consistency check

Outcome	**(I)** **No retirement disincentives**	**(II)** **Lump sum +407€ benefits p.a.** before means-test	after means-test
$\Delta\mathbb{E}$(retirement age)	-0.84	-0.1	-0.11
$\Delta\mathbb{E}$(experience at age 65)	-0.41	-0.04	-0.06
$\Delta\mathbb{E}$(wealth at age 65)	€-6,105	€-978	€-1,182
	(III) **Pension age +1 year**	**(IV)** **Life expectancy +5 years**	
$\Delta\mathbb{E}$(retirement age)	+0.72	+0.22	
$\Delta\mathbb{E}$(experience at age 65)	+0.46	+0.15	
$\Delta\mathbb{E}$(wealth at age 65)	€+430	€+4,342	

individual life cycle behavior. The simulations suggest that individuals postpone retirement by about 2 months and increase labor market participation over their life-cycles. Furthermore, they build up a larger wealth stock. The net returns to the pension contributions increase substantially. Of course, this is a very artificial scenario because such an increase in life expectancy in the population would have to go along with either a rise in pension contributions or a decrease in pension benefits in order to maintain the financial sustainability of the pension system.

8 Policy simulations

In our policy simulations, we focus on average outcomes by unobserved type and level of education (low/high). In particular, we consider (1) average retirement age and work experience at age 65, (2) average accumulated wealth at age 65, (3) average sum of net pension benefits minus the sum of contributions to the pension system, (4) average of average annual lifetime consumption, and (5) the Gini coefficient of average annual lifetime consumption. All these outcomes can be interpreted as expected outcomes at age 20. When comparing simulations, changes in retirement age, work experience, and accumulated wealth summarize the individuals' behavioral responses. A change in the difference between the sum of net pension benefits and the sum of contributions to the pension system captures changes in the net returns of individuals' pension contributions. The average of average annual lifetime consumption indicates changes in per-period consumption and relates to common welfare measures. A change in the Gini coefficient of average annual lifetime consumption sheds light on distributional effects.

8.1 Life expectancy-based redistribution

Now, we use our life expectancy-adjusted benchmark scenario to address the question who is benefiting in the current German pension system. We simulate the outcomes under the life expectancy-adjusted scheme and compare them to the respective outcomes under the current scheme. First, we simulate the changes in the outcomes (3)-(5) without behavioral responses to the life expectancy-adjusted scheme. Then, we allow behavioral responses and consider changes in the outcomes (1)-(5). In both cases, we assure budget neutrality of the simulations by adjusting the pension contribution rate, where neutrality is targeted with respect to the overall public budget.[11] When implementing the simulations with behavioral responses, we need two loops. The inner loop converges if behavioral responses to the redistribution rule are consistent with respect to the pension pot that is available for redistribution. The outer loop takes the fix point of the inner loop as given and adjusts the pension contribution rate until budget neutrality is achieved.

Considering the behavioral implications of the life expectancy-based redistribution, we see that there are relevant effects on the lifetime employment for both the low and the high educated individuals. Under the life expectancy-adjusted scheme, the low educated individuals of type 1 have 0.14 more years of work experience at age 65 and the less productive individuals of type 2 even have 0.32 more years. This is driven by both an income and substitution effect. The larger effect on the type 2 individuals relates to the means-test. Under the life expectancy-adjusted scheme these individuals are simply less likely to end up with pension claims below the social security minimum at retirement such that employment and wealth accumulation pays off more for them. The high educated individuals are induced to postpone retirement by 0.38 and 0.14 years respectively. However, this only translates for the type 1 individuals into a higher lifetime employment. These effects follow from the substantial reduction in pension claims for the high educated and the change in the pension contribution rate. There are non-negligible effects on the level of accumulated net wealth at retirement. Overall, the simulations suggest that the behavioral responses allow for more consumption in the sample population.

The simulations show that the low educated gain substantially under the life expectancy-adjusted scheme in terms of their net pension returns. Or put differently, the low educated loose on average under the current scheme €17,305 for type 1 individuals and €6,079

[11] Even without behavioral responses budget neutrality is not ensured mechanically. For low income individuals with low education the transfer may be eaten up by the means-test if the resulting pension level is still below the social security minimum.

Table 4: Effects of regressive redistribution

Outcome	**Low education**		**High education**	
	Type I	Type II	Type I	Type II
without behavioral adjustment:				
$\Delta\mathbb{E}$(NPBs-contributions)	€+10,061	€+3,140	€-13,668	€-5,644
$\Delta\mathbb{E}$(average annual cons.)	€+198	€+75	€-212	€-86
Δ Gini (average annual cons.)	-1.00% [baseline coefficient: 0.17]			
Δ 90-to-10 ratio (average annual cons.)	-0.02 [baseline ratio: 2.26]			
Adjustment τ to contribution rate	-0.0033 [new contribution rate: $(0.195+\tau)/2$]			
with behavioral adjustment:				
$\Delta\mathbb{E}$(retirement age)	-0.03	+0.05	+0.38	+0.14
$\Delta\mathbb{E}$(experience at age 65)	+0.14	+0.32	+0.18	+0.02
$\Delta\mathbb{E}$(wealth at age 65)	€+367	€+2,545	€+2,663	€-940
$\Delta\mathbb{E}$(NPBs-contributions)	€+17,305	€+6,079	€-10,330	€-1,166
$\Delta\mathbb{E}$(average annual cons.)	€+311	€+163	€-161	€-71
Δ Gini (average annual cons.)	-1.36% [baseline coefficient: 0.17]			
Δ 90-to-10 ratio (average annual cons.)	-0.03 [baseline ratio: 2.26]			
Adjustment τ to contribution rate	-0.0074 [new contribution rate: $(0.195+\tau)/2$]			

for type 2 individuals in net pension returns. Comparing these numbers to the scenario with no behavioral responses (€10,061 and €6,079 respectively) shows that individual behavior matters substantially when investigating the distributional implications of the pension scheme. This finding suggests that there is substantial life expectancy-based redistribution of lifetime income. For the low educated, the losses in net pension returns translate into reductions in average annual consumption of €311 for individuals of type 1 and of €163 for individuals of type 2 in average annual consumption. On the other hand, the high educated gain on average €161 and €71 of average annual consumption, respectively. The change in the Gini coefficient by -1.36% under the life expectancy-adjusted scheme suggests that the overall distributional effect of the heterogeneity in the life expectancies is regressive. Hence, the poor individuals in terms of average annual consumption pay for the higher life expectancy of the richer individuals.

Table 5: Introduction of progressivity (Gini: -1.21%)

Outcome	**Low education**		**High education**	
	Type I	Type II	Type I	Type II
progressive contributions:				
$\Delta\mathbb{E}$(retirement age)	-0.00	-0.01	-0.06	-0.00
$\Delta\mathbb{E}$(experience at age 65)	+0.01	+0.12	-0.05	+0.02
$\Delta\mathbb{E}$(wealth at age 65)	€+100	€+1,256	€-500	€+718
$\Delta\mathbb{E}$(NPBs-contributions)	€+464	€+4,878	€-8,694	€+5,462
$\Delta\mathbb{E}$(average annual cons.)	€+9	€+74	€-101	€+55
Adjustment τ to contribution rate	-0.0079 [new contribution rate: $(0.195+\tau)/2$]			
progressive pension benefits:				
$\Delta\mathbb{E}$(retirement age)	+0.06	-0.00	+0.20	-0.00
$\Delta\mathbb{E}$(experience at age 65)	-0.00	0.06	+0.12	+0.00
$\Delta\mathbb{E}$(wealth at age 65)	€+253	€+821	€+1631	€+640
$\Delta\mathbb{E}$(NPBs-contributions)	€+2,585	€+3,440	€-7,182	€+4,616
$\Delta\mathbb{E}$(average annual cons.)	€-6	€+48	€-107	€+38
Adjustment τ to contribution rate	-0.0056 [new contribution rate: $(0.195+\tau)/2$]			

8.2 Pension reforms

We evaluate two pension reforms that reduce the redistribution from lifetime poor to lifetime rich by targeting Gini coefficient of the life expectancy-adjusted scheme. The first reform introduces progressivity to pension contributions and the second reform introduces progressivity to pension benefits. This analysis addresses two questions: (1) Are conventional redistribution mechanisms like progressivity in either contributions or pension benefits suitable means to offset life expectancy-based redistribution? (2) Are there potential advantages or drawbacks to either of the two mechanisms? Based on the structural life-cycle model we evaluate the reforms by targeting a the Gini coefficient of the life expectancy-adjusted pension scheme in the simulations. As before, we ensuring budget neutrality by adjusting the pension contribution rate. Table 5 shows the behavioral effects and the implications for net pension returns and average annual consumption.

Both reforms induce moderate employment effects. While progressivity in the pension benefits rather stimulates employment of the high productivity individuals, the results suggest that progressive pension contributions may have favorable effects on employment of the low-income individuals (low education and type 2). Under both schemes the high educated individuals of type 1 loose substantially. It turns out that mainly the type 2 individuals with both low and high education benefit from the reforms. It is due to the positive employment effects on the low-income individuals that this group enjoys a higher level of average annual consumption when progressivity is introduced to the pension contributions. But, the rise in average annual consumption even for this group is still substantially lower than it is under the life expectancy-adjusted scheme (€ +74 with progressive pension contributions versus € +163). While both schemes are set up in order to match the Gini coefficient of the life expectancy-adjusted scheme, the transfers to the low educated are not as effective with respect to both net pension returns and average annual consumption. In particular, the low educated of the more productive type 1 benefit very little. In general, the low educated only benefit if their low level of education translates into a comparatively low income. The high educated of type 2 benefit on average not only from life expectancy-based redistribution, but also from progressivity in either pension contributions or pension benefits. We conclude from these findings that a reform that aims at targeting more effectively the issue of life expectancy-based redistribution via the pension system needs to tie pension contributions or benefits directly to the socio-economic determinants of life expectancy.

9 Conclusion

In this paper, we also examine potentially regressive, redistributive effects that are driven by heterogeneity in life expectancy. However, building on the previous literature, we shift the focus from life-time income to annual outcomes, and examine the link between the pension system and inequality in annual consumption. Our extension is motivated by the possibility that the previously-documented regressive redistributive effects may extend beyond providing longevity insurance for extra years of life, and actually work to increase differences in annual living standards.

The contribution of the paper is three fold. First, we propose a methodology that provides a general quantification of the life expectancy-driven distributive effects of defined benefit pension systems. The methodology relies on a life expectancy adjusted benchmark pension system, which eliminates life expectancy-driven redistribution but maintains all other features of the pension system, such as the benefits schedule and child-related bonuses. A comparison of the distribution of average annual consumption under a particular pension system and under the life expectancy adjusted pension system reveals how inequality in annual consumption is impacted by life expectancy-driven redistribution through the pension system.

Second, we present empirical analysis that quantifies the life expectancy-driven redistribution that is inherent in the current German public pension system. We show that the German pension system induces a large regressive redistribution of life-time income, due to differences in life expectancy. Moreover, the life expectancy-driven regressive redistribution of life-time income translates into an increase in the inequality of average annual consumption. Our results on life expectancy driven redistribution through the pension system are obtained by combining the life expectancy adjusted benchmark pension system with a life-cycle model of labor supply, retirement and consumption decisions. Our results reveal that behavioral responses are an important component of the distributional effects of pension reforms, thus justifying the operationalization of the benchmark scenario via a behavioral model.

Third, our analysis becomes more policy orientated and we explore two reforms designed to reduce regressive redistribution present in the current German public pension system. One reform increases the progressivity of the contribution schedule and the other reform increases the progressivity of the benefit schedule. The behavioral effects of the two reforms differ, with the increase in the progressivity of contributions having more favorable effects on the labor supply and consumption outcomes of low-income individuals. Both reforms are only partially able to offset the life expectancy-driven redistribution via the pension system – a more effective reform would tie pension contributions or benefits directly to the socio-economic determinants of life expectancy (as occurs in the hypothetical benchmark pension system).

References

Arcidiacono, P., and J. B. Jones (2003): "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71(3), 933–946.

Attanasio, O. P., and G. Weber (1995): "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, 103(6), 1121–57.

Coronado, J.L., D. F., and T. Glass (2000): "The progressivity of social security," *NBER Workingpaper 7520.*

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The Determinants of Mortality," *Journal of Economic Perspectives*, 20(3), 97–120.

Feldstein, M., and J. Liebman (2002): *Social Security*pp. 2246 – 2324. Handbook of Public Economics, Volume 4.

French, E. (2005): "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour," *Review of Economic Studies*, 72(2), 395–427.

French, E., and J. Jones (2011): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, Vol. 79, No. 3, 693–732.

Gustman, A. L., and T. L. Steinmeier (2001): "How effective is redistribution under the social security benefit formula?," *Joural of Public Economics*, 82, 1–28.

Haan, P., and V. Prowse (2014): "Longevity, life-cycle behavior and pension reform," *Journal of Econometrics*, 178, 3, 582–601.

Keane, M. P., and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76(4), 648–72.

Lleras-Muney, A. (2005): "The Relationship Between Education and Adult Mortality in the United States," *Review of Economic Studies*, 72, 189–221.

Low, H., and L. Pistaferri (2010): "Disability risk, disability insurance and life cycle behavior," Discussion paper, National Bureau of Economic Research.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65(4), 781–832.

van der Klaauw, W., and K. I. Wolpin (2008): "Social Security and the Retirement and Savings Behavior of Low-Income Households," *Journal of Econometrics*, 145(1-2), 21–42.