Life-time inequality: The impact of taxation, transfers and pensions

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

July 8, 2015

Abstract

In this paper, we analyze income inequality over the life-cycle and study how taxation, transfer programmes and the pension system affect inequality. Importantly, in our life-cycle analysis we include both the working life and the retirement period of individuals and thereby we extend the previous literature on life-cycle inequality which has focused only on the working period. For our analysis we use a dynamic structural life-cycle model to derive earnings and net-income processes over the whole life-cycle. In the first part of the paper we use the model to understand the drivers of inequality in life-time gross earnings and life-time net income and study the impact of the different tax and transfer programs and the pension system on inequality between and within different educational groups. In the second part we use the structural model to analyze the distributional implications, both between and within educational groups, of several changes to the public 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

A growing literature studies cross-country and cross-time dimensions of inequality in life-time income, e.g, Kopczuk, Saez, and Song (2010), Corneo et al. (2015), Bowlus and Robin (2004), Bowlus and Robin (2012) or Flinn (2002). These studies are largely motivated by the observation that many individuals save and borrow to moderate the impact of transitory fluctuations in annual earnings, meaning that life-time earnings, rather than the underlying annual earnings, are the primary determinant of inequality in living standards.[1] Inequality in life-time earning has been be found to be markedly lower than inequality in annual earnings.[2] Results from this literature also point to different wedges between annual and life-time inequality measures depending on the countries and time periods under study.

The literature on life-time income inequality has so far only focused on inequality in gross earnings, and thus has not considered how income taxation, social transfer programs and the public pension system impact on life-time inequality.[3] This is perhaps surprising given the large literature documenting the powerful distributional effects of income taxation and transfer programs, see e.g. Piketty and Saez (2007)[4]. In addition, by focussing on life-time gross earnings, the previous literature has only considered the working-age period of the life cycle and has not included the retirement period .[5] However, the retirement period must also be considered for a fully understanding the nature and drivers of life-time inequality. Most important, as documented e.g. by Gustman and Steinmeier (2001) or Feldstein and Liebman (2002) public pension systems may lead to sizable distributional effects depending of the design of the system. In particular a progressive pension schedule or pension system with a minimum pension level induces large redistribution from individuals with high life-time earnings and this large pension contributions to those with low life-time earnings.

Therefore, in this paper we take the life cycle to include both the working age period and the retirement period and we make a distinction between gross life-time earnings and net life-time income, that is earnings adjusted for income taxation, welfare bene-

[1]With the same motivation several papers analyze inequality in annual consumption rather than annual earnings, see e.g. Blundell et. al 2008 or Krueger and Perri (2005).

[2]There exists an extensive literature on inequality in annual earnings, see e.g., the Handbook chapter by Katz and Autor (1999).

[3]One exception is Brewer et al. (2012) who study the effect of taxation and family related transfers on inequality during the working life.

[4]Blundell et al. (2015) compare dynamics in gross-earnings and net-income and document that income taxation and transfers significantly reduce the risk of permanent shocks

[5]E.g., Kopczuk, Saez, and Song (2010) (in their main specification) look at individuals aged 25-60 and Bowlus and Robin (2012) restrict their sample depending on the country to individuals between 16 to 60/65.

fits, Unemployment Insurance and the public pension system. In doing so, we extend on existing studies of life-time income inequality in three respects. First, we analyze inequality over the full life cycle, including both the working-age portion of the life-cycle and the retirement period. Second, we shed light on how the income taxation, welfare benefits, Unemployment Insurance and the public pension system impact on inequality in life-time net income. Third, we explore how counterfactual reforms to the fiscal system impact inequality in life-time net income. In particular, we compare reforms affecting the transfer system before retirement, i.e., reforms to Unemployment Insurance and welfare benefits, and reforms that impact the incomes of the retired population, i.e, reforms to the organization of pension benefits.

For our analysis we propose a dynamic structural life-cycle model and we use the model to derive earnings and net-income processes over the whole life-cycle including the working and retirement period.[6] The central motivation for using a dynamic life-cycle model for our analysis rather than exploiting administrative data with life-time information, such as Kopczuk, Saez, and Song (2010) or Corneo et al. (2015), is to account for behavioral responses when analyzing the effect of counterfactual reforms to the tax and transfer system and the public pension system on life-time inequality. In more detail, we develop a life-cycle model of labor supply, retirement and savings 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.[7]

In the first part of the paper we use the model to understand the drivers of inequality in life-time gross earnings and life-time net income. To this end we simulate life-time incomes and decompose the distributional effects of income taxation, social security contributions, welfare benefits and Unemployment Insurance during the working life and the public pension system. In more detail, we decompose the overall inequality in average life-time income by educational groups and analyze how the different aspects of the fiscal system affect the between- and within-group inequality. Further, we study how the different factors of within-group inequality, namely innate unobservable effects, employment risk, health shocks and longevity risk are insured by the different components of the fiscal system.

The second part of this paper focuses on the impact of hypothetical reforms to the

[6]Bowlus and Robin (2004), Bowlus and Robin (2012) or Flinn (2002) apply a similar simulation strategy, however their models differ in several dimensions from our life-cycle model. Most important in Bowlus and Robin (2004), Bowlus and Robin (2012) the authors develop a statistical model without structural foundation and Flinn (2002) uses a infinite horizon search model.

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

pension system. We use the structural model to analyze the distributional implications, both between and within educational groups, of several changes to the public pension system. In particular we investigate how reductions in the generosity of the pension system affect inequality related to innate unobservable effects, employment risk, health shocks and longevity risk. To this end we compare different reforms which either affect individuals with high or low life-time earnings, or which affect early retirement rules including disability pensions. For this analysis it is crucial to account for the behavioral responses of individuals to changes in the institutional environment; several previous studies have documented sizable employment and savings responses over the life-cycle to changes in the design of out of work programmes or the pension design and these responses affect the life-time income.

In the following we discuss the ..

2 Structural life-cycle model

We propose to analyze life-cycle inequality and the role of taxation, transfers and the pension system on inequality within a dynamic structural life-cycle model which accounts for labor supply, retirement and savings behavior of individuals. The structural life-cycle shares many similarities with French (2005), French and Jones (2011), or Haan and Prowse (2014). In particular in the model individuals make labor supply, retirement and consumption choices to maximize their utility over the life-cycle. Moreover the model accounts for endogenous accumulation of experience which impacts on wages, captures potential frictions on the labor market, i.e. stochastic job offers and involuntary separations, and includes different sources of risk related to unemployment, health shocks, longevity. Finally, we include a detailed specification of taxation, transfers and the pension system which amongst others depends on the accumulated experience over the working life. 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. In addition we allow for different unobservable types of individual types which affect the different processes.

2.1 Overview

We develop a discrete-time life-cycle model with finite horizon for male individuals. Individuals enter the labor force from education[8] and each period they make an employment and consumption decision. We distinguish between three different employment states:

[8]The timing of the transition from education into the labor force is assumed to be exogenous.

non-employment, full-time employment[9] and retirement, where retirement is an absorbing state. An individual's employment choice is restricted by job offer and separation rates that are estimated within the model. The offer rates depend on the employment history and capture persistence of the unemployment status. In more detail, 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. In general, retirement is only an option for individuals older than 60 year and the compulsory retirement age is 65; however individuals with a severe health condition can enter retirement earlier and they receive a disability pension. Below we discuss in more detail the pension rules and the related penalties individuals face when entering retirement before the compulsory retirement age.

As standard in structural life-cycle models, we assume that individuals are forward looking and conditional on the labor market frictions they make their period specific employment and consumption decisions subject to an intertemporal budget constraint to maximize the expected discounted life-cycle utility. In the following we discuss the functional form of the utility function, the intertemporal budget constraint which depends on the the structure of the tax and transfer and the pension system, the wage process, the dynamic health process. Finally we show how heterogeneity in life-expectancy enters the model.

2.2 Objective function

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

[9] In our sample, fewer than 5% of the male population works less than 30 hours per week. It is therefore reasonable to treat all employment as full-time work.

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.

2.3 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.

2.4 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, $q(\mathbf{s}_{nt+1}|s_{nt}, d_{nt})$ 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

[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.

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

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.

2.5 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):

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

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

2.6 Health transitions

Age-specific transition probabilities are 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.

2.7 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.

2.8 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 $1 - \sum_{m=1}^{M-1} \gamma_m$.

2.9 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.

2.10 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:

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

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.

3 Data and descriptive statistics

3.1 Estimation sample

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),[11] 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.[12] 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.[13]

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

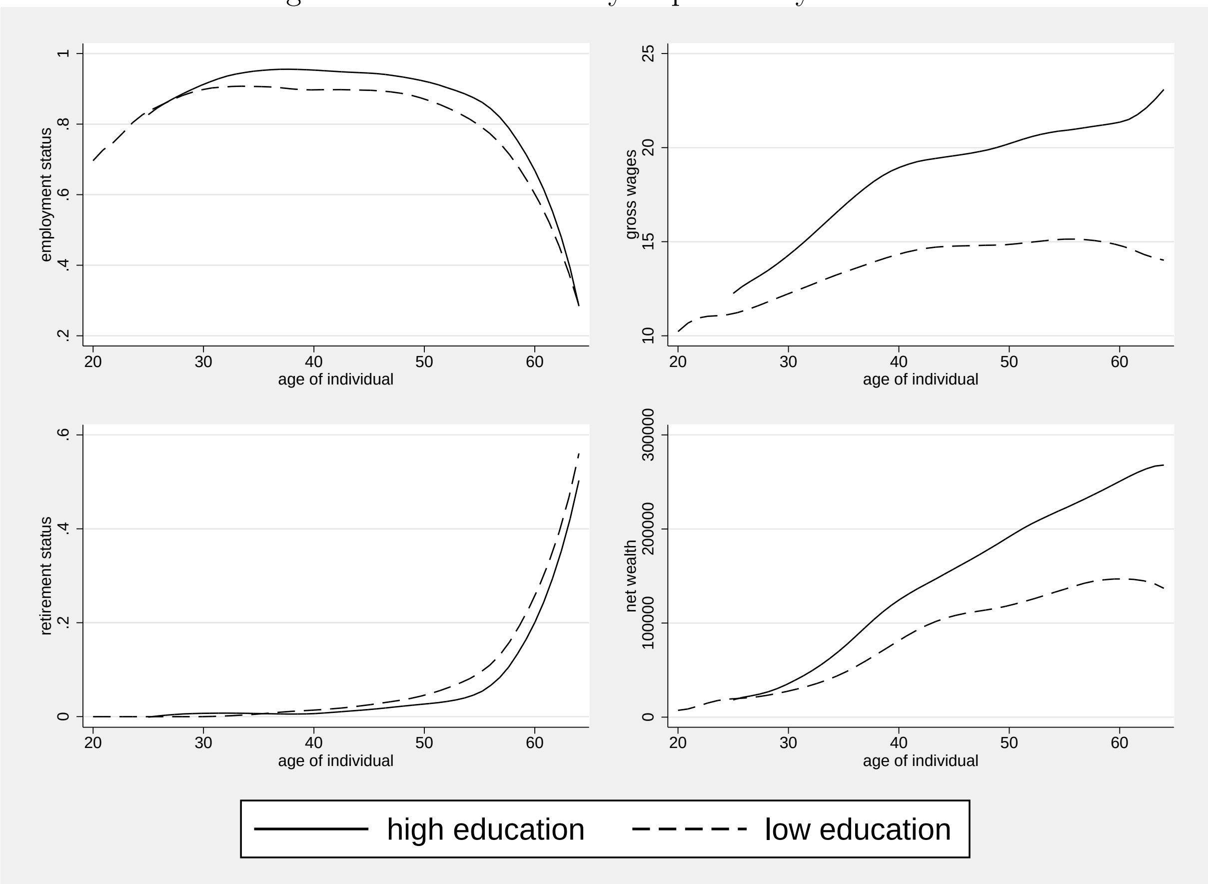
[11]Employment is defined as working at least 20 hours per week and median hours of work for the employees is 40. Note, only very few men work part-time and they are considered to be non-employed.

[12]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.

[13]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.[14] Real savings are defined as annual amortization payments.[15] 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.[16] 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 1: Estimated life cycle profiles by education



[14]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?"

[15]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.

[16]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.

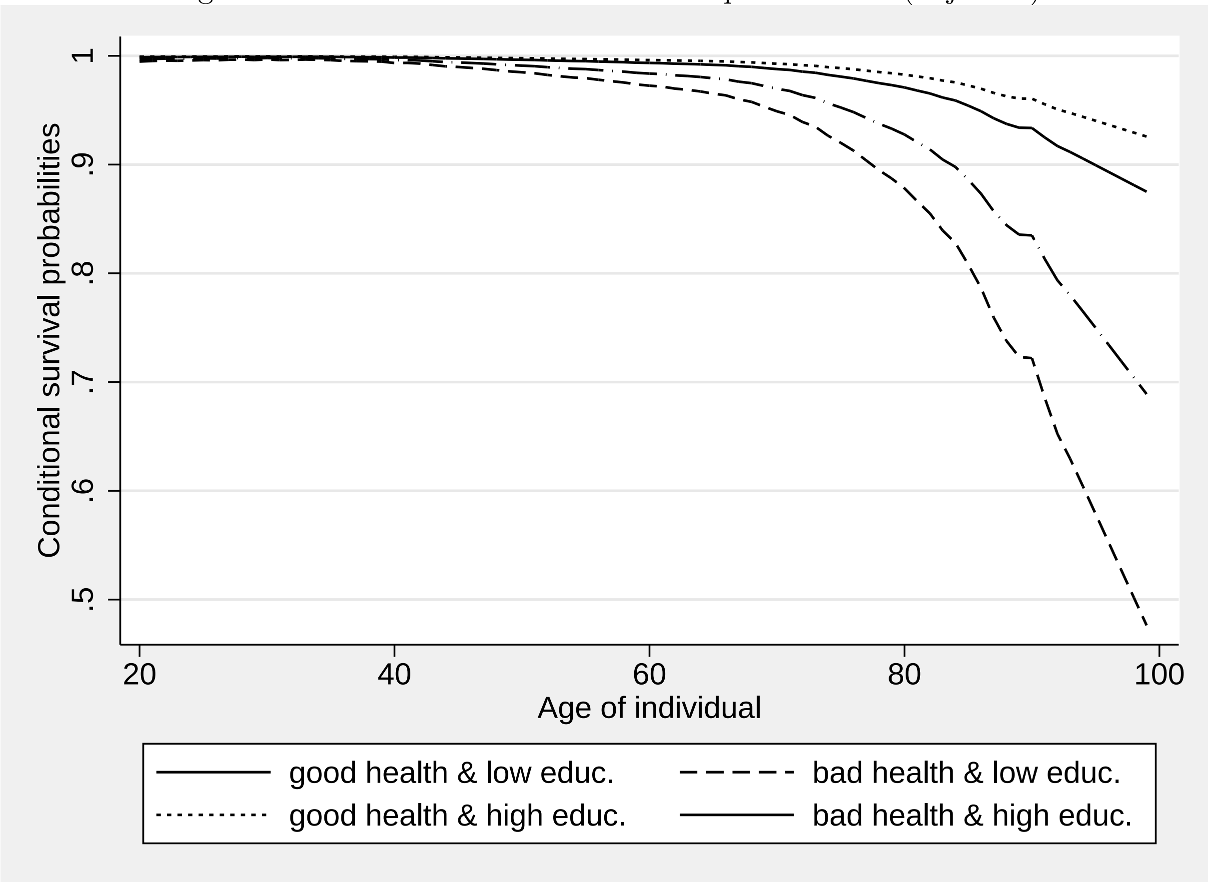
3.2 Conditional survival probabilities

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. As a consequence, we need estimates of the partial effects of both education and health on conditional survival probabilities.

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. This finding is insensitive to changes in the distributional assumptions of the proportional hazard model. The difference might be related to the fact that the follow-up survey on death cases may not have captured all the cases. 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 2 shows predicted conditional survival probabilities.

Figure 2: Predicted conditional survival probabilities (adjusted)



4 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.

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.

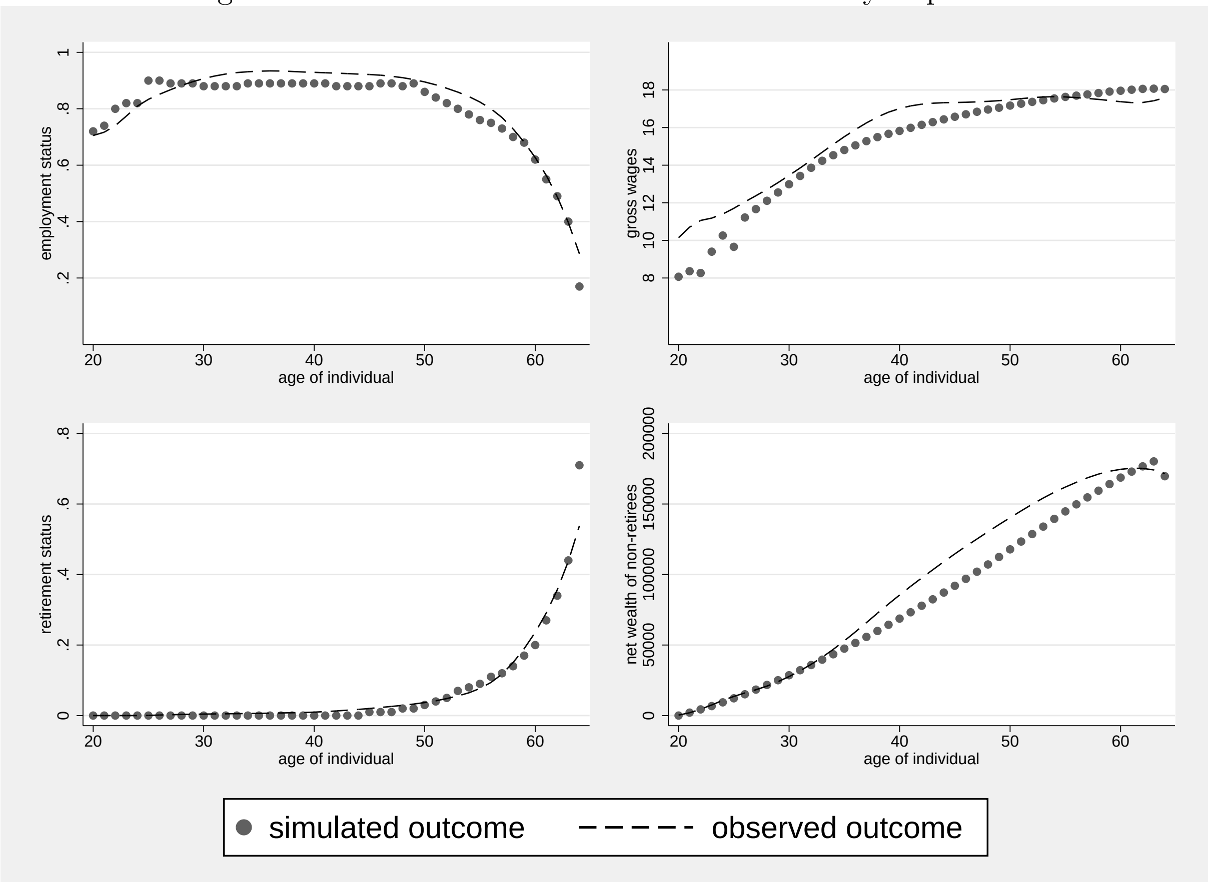
Table 2: Parameter estimates of FIML estimation

	Estimates	St.e.
α_1 (scaling factor)	1.673	(0.0432)
α_{21} (work, type 1)	-1.124	(0.0565)
α_{22} (work, type 2)	-0.678	(0.055)
ρ_1 (crra, type 1)	0.48	(0.0503)
ρ_2 (crra, type 2)	1.422	(0.0241)
Wage equation:		
κ_1 (constant, type 1)	1.699	(0.0129)
κ_2 (constant, type 2)	1.243	(0.0129)
δ_1 (years of education / 10)	0.562	(0.0071)
δ_2 (log(experience)*(educ$<$12))	0.169	(0.0026)
δ_3 (log(experience)*(educ$\geq$12))	0.203	(0.0026)
σ_μ (standard deviation)	0.23	(0.0009)
Job offers and separations:		
ϕ_1 (separation, constant)	-2.681	(0.1183)
ϕ_2 (separation, high educ)	-0.12	(0.1052)
ϕ_3 (separation, good health)	-1.233	(0.1212)
ϕ_4 (separation, 50$\leq$age$<$60)	0.409	(0.1399)
ϕ_5 (separation, age$\geq$60)	1.43	(0.1603)
ϕ_6 (offer, constant)	-1.815	(0.1405)
ϕ_7 (offer, high educ)	-1.115	(0.0858)
ϕ_8 (offer, good health)	2.153	(0.1463)
ϕ_9 (offer, 50$\leq$age$<$60)	-1.601	(0.1226)
ϕ_{10} (offer, age$\geq$60)	-2.497	(0.2412)
Type probabilities:		
γ_1 (prob. of type 1)	0.511	(0.0106)

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.

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.

Figure 3: Simulated outcomes and observed life cycle profiles



5 Consistency with previous findings

5.1 Behavioral responses

Similar to Low and Pistaferri (2010) we show that the implications of the model are consistent with previous literature exploiting pension reforms directly 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 are most relevant for our analysis. We consider three behavioral outcomes: the change in (a) retirement age and (b) accumulated net wealth at age 60. 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: +439€ p.a.). The income effects on individual behavior appear to be fairly small in comparison to the substitution effect. 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.

Table 3: Simulated treatment effects

Outcome	**Abolishment of early retirement disincentives**	
	Income & substitution effect	Pure income effect (€ 439 p.a.)
$\Delta\mathbb{E}$(retirement age)	-0.46	-0.06
$\Delta\mathbb{E}$(wealth at age 60)	€ +1,331	€ -209
	Pension age +1 year	**Life expectancy +5 years**
$\Delta\mathbb{E}$(retirement age)	0.63	0.17
$\Delta\mathbb{E}$(wealth at age 60)	€ -1,180	€ +1,346

In the third scenario, the statutory pension age is raised by one year from age 65 to 66 while other threshold values are kept constant. This induces an average postponement of retirement by 0.63 years. The 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 individual life cycle behavior. The simulations suggest that individuals postpone retirement by about 2 months and build up a larger wealth stock. 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.

5.2 Distributional outcomes

In this subsection, we compare distributional outcomes for gross labor earnings in our simulated data with the respective outcomes that are (1) directly estimated from the estimation sample (Ginis of pooled cross section outcomes) and (2) with recent findings of a study by Bönke, Corneo, and Lüthen (2015) that is based on high quality administrative data (Ginis of pooled cross sections and of life-cycle outcomes). These comparisons are highly relevant for the credibility of our analysis. We show that we have both a good internal fit in terms of distributional outcomes and that these outcomes are consistent with high quality estimates based on an administrative data set. Of course, we cannot check the internal validity with respect to the life-cycle outcomes because the SOEP data only provides a limited number of waves per individual.

Table 4: Ginis of gross earnings

	Own calculations			**Bönke et al. (2015)**
Sample	Simulated data	Estimation sample (SOEP)		Admin. data (VSKT)
	5000 life-cycles	Survey years 2005 to 2011		Cohorts 1935 to 1949
Cross-section	0.273	0.266 (def. 1)	0.282 (def. 2)	0.262–0.336
Lifetime	0.209	–	–	0.156–0.212

When checking the internal validity, we compute two Ginis on the pooled cross sections of the sample data. The first measure of gross labor earnings is fully consistent with the model assumptions (def. 1)[17] and the second measure (def. 2) simply uses the variable as it is provided by the SOEP. Generally, the model fit appears to be very good for the distributional outcomes. When comparing these Ginis to the estimates of Bönke, Corneo, and Lüthen (2015), our estimates are at the lower end of the range (estimates for the oldest cohorts). As good as our model may may be we probably miss some part of the cross-sectional variation because of the assumptions that we make e.g. regarding unsystematic annual variation in the wage outcomes. However, most of such inaccuracies presumably even out over life-cycles. Hence, the distributional outcomes for life-cycle earnings corresponds to the estimates of Bönke, Corneo, and Lüthen (2015) for the youngest cohorts that are also more representative for our estimation sample. These findings give credibility to our further analysis based on simulated distributions.

[17]By this definition, gross earnings are coded as zeros, if individuals are coded as non-working.

6 Inequality in life-time income

In the following we use the model to understand the drivers of inequality in life-time gross earnings and life-time net income. To this end we simulate life-time incomes and decompose the distributional effects of income taxation, social security contributions, welfare benefits and unemployment insurance during the working life and the public pension system. In more detail, we decompose the overall inequality in average life-time income by educational groups and analyze how the different aspects of the fiscal system affect the between- and within-group inequality. Further, we study how the different factors of within-group inequality, namely innate unobservable effects, employment risk, health shocks and longevity risk are insured by the different components of the fiscal system. Table 5 shows means and standard deviations of the simulated life-cycle outcomes. The first column starts with gross earnings, the second column adds gross capital income, and then we introduce (1) taxes, (2) unemployment insurance contributions, (3) unemployment insurance benefits, (4) social assistance benefits, (5) pension insurance contributions, and (6) pension insurance benefits. We decompose distributional effects by computing distributional outcomes for the respective income measures.

Table 5: Means and standard deviations of income measures (€ 1000)

	Gross E.	Gross I.	+Taxes	+UIC	+UIB	+SAB	+PIC	+PIB
Life-cycle:								
r=0	1,283.5	1,400.5	991.2	921.5	936.8	956.7	747.4	933.2
	(473.4)	(516)	(343.3)	(318.1)	(315.7)	(299.7)	(225)	(267.5)
r=0.02	856.2	911.4	648.9	602.9	612.9	625.9	487.9	554
	(290.9)	(311.4)	(207.6)	(192.3)	(190.1)	(179.7)	(134.2)	(144)
Annualized:								
r=0	23.8	25.8	18.2	16.9	17.2	17.5	13.7	16.6
	(10.8)	(11.3)	(7.6)	(7)	(7)	(6.7)	(5)	(4.6)
r=0.02	25.8	27.4	19.5	18.1	18.4	18.8	14.6	16.5
	(9.7)	(10.2)	(6.8)	(6.3)	(6.2)	(6)	(4.4)	(4.3)

The income measures from table 5 are computed both for life-cycle outcomes and for annualized outcomes and both without discounting and with an assumed real interest rate of 2%. The annualization accounts for the realized life-times of the individuals, where the annualization with discounting is done by computing an annuity based on the net present value of realized outcomes (discounted to age 20).

6.1 Decomposition based on Gini coefficients

Table 6: Ginis coefficients by income measures

	Gross E.	Gross I.	+Taxes	+UIC	+UIB	+SAB	+PIC	+PIB
Life-cycle:								
r=0	0.209	0.209	0.195	0.194	0.189	0.175	0.167	0.158
r=0.02	0.193	0.194	0.18	0.18	0.174	0.16	0.153	0.143
Annualized:								
r=0	0.256	0.25	0.236	0.234	0.23	0.217	0.208	0.158
r=0.02	0.215	0.213	0.199	0.198	0.193	0.18	0.171	0.147

Table 6 shows Gini coefficients for life-cycle and for annualized outcomes. The annualized outcomes account for the fact that heterogeneity in life-cycle outcomes that is only due to heterogeneity in realized lifetimes may not be very relevant from a welfare perspective. The progressive tax system reduces inequality. Unemployment insurance and social assistance benefits insure labor market risks and, thus, also reduce inequality. Interestingly, the pension exerts substantial distributional effects. This can be explained through the fact that it provides both longevity and disability insurance. As a consequence, policy makers should keep in mind the distributional implications of pension reforms. In the policy section, we consider different reforms of the pension system and investigate the effects on distributional outcomes.

Table 7: Changes in Ginis when life expectancies are homogeneous

	Gross E.	Gross I.	+Taxes	+UIC	+UIB	+SAB	+PIC	+PIB
Life-cycle:								
r=0	-0.002	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.016
r=0.02	-0.002	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.009
Annualized:								
r=0	+0.011	+0.009	+0.009	+0.009	+0.009	+0.01	+0.01	+0.002
r=0.02	+0.006	+0.005	+0.005	+0.005	+0.005	+0.006	+0.006	+0.001

Life expectancies vary in the population and may affect the distributional implications of the different institutions. In particular the pension system may exert a regressive effect on the life-cycle outcomes that is due to the higher life expectancies of the high educated. For this reason, it is interesting to consider a scenario where life expectancies are set to be homogeneous over educational groups while keeping constant the average life expectancy. Table 7 shows changes in Ginis under homogeneous life expectancies. While the pension system - unsurprisingly - only has small effect on annualized outcomes, it turns out to have a considerable effect on inequality in life-cycle outcomes. The simulations suggest that overall the introduction of the pension system reduces the Gini additionally by 0.02 without discounting and by 0.012 with discounting. This reflects the fact that the redistribution of the short-lived to the long-lived comes at the expense of the low educated when contributions and benefits are proportional. It follows that the pension system may reduce within-group inequality while rising between-group inequality.

6.2 Between- and within-group inequality

[VICTORIA]

7 Policy simulations

In the final part of this paper we focus on the impact of hypothetical reforms to the pension system. We use the structural model to analyze the distributional implications, both between and within educational groups, of several changes to the public pension system. In particular we investigate how reductions in the generosity of the pension system affect inequality related to innate unobservable effects, employment risk, health shocks and longevity risk. To this end we compare different reforms which either affect individuals with high or low life-time earnings, or which affect early retirement rules including disability pensions. For this analysis it is crucial to account for the behavioral responses of individuals to changes in the institutional environment; several previous studies have documented sizable employment and savings responses over the life-cycle to changes in the design of out of work programmes or the pension design and these responses affect the life-time income. We consider two kinds of simulations. The first type of simulations assumes that 10% of pension expenditures must be either cut through lower benefits or financed through higher contributions. Both is done first for a proportional change and, then, for a change that introduces progressivity. The second type of simulations raises the early retirement age for all pensioners (good or bad health) to 60, 63, and at last 65 (no early retirement at all), where the first simulation - of course - only affects the disability pension.

7.1 Expenditure cut versus rises in contributions

In these simulations, we compare a cut in pension benefits that reduces expenditures by 10% with a rise in contributions that generates a corresponding budgetary effect.

We consider both proportional and progressive changes in the contribution and benefit formulas. Furthermore, we consider simulations where behavior is kept constant and simulations where individuals adjust their behavior in response to the associated changes in incentives. When cutting benefits, this is not affecting the minimum pension level that corresponds to the social assistance benefits. When introducing progressivity in either benefits or contributions, this is done by allowing the respective rate to decrease / increase linearly from a certain threshold value onwards.

Table 8: Change in Ginis of net income after expenditure cut or rise in contributions

	Cut in expenditures				**Rise in contributions**			
	no behavior		with behavior		no behavior		with behavior	
	Prop.	Progr.	Prop.	Progr.	Prop.	Progr.	Prop.	Progr.
Life-cycle:								
r=0	-0.002	-0.006	-0.002	-0.008	0	-0.003	+0.001	-0.003
r=0.02	-0.001	-0.004	0	-0.005	0	-0.006	+0.001	-0.006
Annualized:								
r=0	+0.001	-0.005	+0.002	-0.006	-0.001	-0.008	0	-0.008
r=0.02	+0.001	-0.003	+0.001	-0.005	-0.001	-0.009	0	-0.009

The results in table 8 show that the overall distributional effects of proportional changes of both the benefits and contributions are small. This, however, may conceal relevant changes in within- and between-group inequality. For example, when proportionally decreasing pension benefits this reduces longevity insurance and, therefore, presumably raises within-group inequality. At the same time, it is expected to reduce between-group inequality because the low-income individuals are hardly effected because of the minimum pension level. When introducing progressivity, substantial changes in inequality can be realized, where individuals' behavioral responses tend to amplify the effects. If a policy maker is concerned about reducing inequality in annualized net income (which is most relevant from a welfare perspective), progressive increases in the contributions are predicted to have the largest effects.

[TO DO: EFFECTS ON BETWEEN- AND WITHIN-GROUP INEQUALITY]

7.2 Raising the early retirement age

The following simulations raise the early retirement age for all individuals (good or bad health) to 60, 63, and at last 65. The results in table 9 show that this is an effective means to raise the average retirement age and improve the financial situation of the pension system while the overall effects on inequality are comparatively small. The achieved per capita savings of the pension system are a bit larger than the net public returns that capture all revenues and expenditures of the government. This is due to the fact that many individuals will receive more unemployment insurance and social assistance benefits when being forced to retire later. But, there are also relevant effects on employment (rise in average years of work experience at age 65). Thus, individuals who would have retired otherwise on average work more and pay more taxes and social security contributions. The achieved savings of the pension system correspond to 7.39%, 9.21%, and 11.76% of expenditures, respectively.

Table 9: Changes in early retirement age for all individuals

	Age 60	**Age 63**	**Age 65**
$\Delta\mathbb{E}$(retirement age)	+1.32	+2.01	+2.69
$\Delta\mathbb{E}$(experience at age 65)	+0.29	+0.35	+0.43
$\Delta\mathbb{E}$(Gini of net life-cycle income)	+0.0015	+0.0017	+0.0019
$\Delta\mathbb{E}$(Gini of average net income)	+0.0004	-0.0001	-0.0007
$\Delta\mathbb{E}$(net pension returns)	€+13,778	€+17,187	€+21,925
	(7.39%)	(9.21%)	(11.76%)
$\Delta\mathbb{E}$(net public returns)	€+12,923	€+14,499	€+18,454

Even though overall effects on inequality are small, there may be substantial and countervailing effects on between- and within-group inequality. Since such reforms reduce the disability insurance provided by the pension system, within-group inequality presumably rises. Between-group inequality is expected to decrease because individuals with higher incomes benefit more from the disability insurance. Low income individuals often receive a disability pension at the minimum pension level anyway that corresponds to the social assistance benefits.

[TO DO: EFFECTS ON BETWEEN- AND WITHIN-GROUP INEQUALITY]

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.

Bönke, T., G. Corneo, and H. Lüthen (2015): "Lifetime Earnings Inequality in Germany," *Journal of Labor Economics*, 33(1), 171 – 208.

Bowlus, A. J., and J.-M. Robin (2004): "Twenty Years of Rising Inequality in U.S. Lifetime Labor Income Values," *Review of Economic Studies*, 71, 709743.

——— (2012): "An International Comparison of Lifetime Labor Income Values and Inequality," *Journal of the European Economic Association*, 10(6), 1236–1262.

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.

Flinn, C. (2002): "Labour Market Structure and Inequality: A Comparison of Italy and the U.S," *Review of Economic Studies*, 69, 611645.

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 securitybenefit 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.

Kopczuk, W., E. Saez, and J. Song (2010): "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937," *Quarterly Journal of Economics*, 125(1), 91–128.

Krueger, D., and F. Perri (2005): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 73(1), 163–193.

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.

Piketty, T., and E. Saez (2007): "How Progressive is the U.S. Federal Tax System: A Historical and International Perspective.," *Journal of Economic Perspectives*, 21(1), 3–24.

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