Progressive Pensions as an Incentive for Labor Force Participation

Fabian Kindermann
University of Regensburg,
CEPR, HCEO, and Netspar

Veronika Püschel
University of Regensburg

August 9, 2023

Abstract

In this paper, we challenge the conventional idea that an increase in the progressivity of old-age pensions unanimously distorts the labor supply decision of households. Building on the literature on the optimal design of income transfer programs for low income workers, we propose to introduce a pension system that subsidizes employment. In particular, we study an Earned Income Pension Credit that allows for a disproportionately high accumulation of pension claims for low earners. A quantitative evaluation shows that the employment gains of such a policy can be sizeable, especially for low productivity workers. Increased labor force participation mitigates the overall efficiency costs from intensive margin labor supply distortions. As the Earned Income Pension Credit effectively reduces income inequality at old age and insures labor productivity risk, it comes with aggregate efficiency gains.

JEL Classification: D15, H31, H55, J21, J22

Keywords: progressive pensions, labor supply, employment incentives

We thank participants at the University of Regensburg lunch seminar, the 2020 annual meeting of the Verein für Socialpolitik, the Public Sector Economics 2020 Conference, the Netspar International Pension Workshop 2021, the RGS Doctoral Conference 2021, the SED 2021 Meeting, the 2021 Public Economics Research Seminar at LMU Munich, the IIPF Congress 2021, the EEA Annual Congress 2021, the 12th CESifo Norwegian-German Seminar, the 25th T2M and the SMYE 2022 as well as Johannes Becker, Hans Fehr, and Jana Schuetz for helpful comments. We gratefully acknowledge financial support from the Fritz-Thyssen-Stiftung (Grant: 10.19.1.014WW).

Contact Information:
Kindermann: fabian.kindermann@ur.de, Püschel: veronika.pueschel@ur.de

1 Introduction

Pension systems in Western economies have undergone major reforms over the course of the last decades. To guarantee their sustainability, most countries have adopted a policy mix between increasing normal retirement ages and tying pension payments to the evolution of life expectancy. Some 20 years later, the reforms of the early 2000s start taking their first effects, and with them comes another political debate: the question of whether all individuals are adequately able to provide enough funds for their retirement, or the debate about old-age poverty. One policy measure to counteract income inequality of the elderly is to introduce a progressive component into the pension formula. This weakens the link between pension contributions and pension payments and narrows the distribution of retirement benefits across income groups.

In this paper, we quantify the incentive and welfare effects of progressive pension systems. Our main contribution is to show that a well-designed progressive pension formula can limit economic costs in the form of labor supply distortions on workers, while still providing adequate benefits to poor pensioners. We build on a literature starting with Saez (2002) that analyzes optimal income transfer programs for low income workers. When labor supply responses are concentrated along the extensive margin, as it is empirically plausible for low earners (see e.g. Meyer, 2002), an optimal labor tax policy explicitly subsidizes employment. In the context of pensions, a public pension system therefore ideally links pension payments to both individual earnings and an individual's employment status. While pension insurers typically have detailed records of an individual's employment history available, one might nevertheless fear that substantial employment subsidies may cause households to extensively engage in minimum hours contracts or even in fictitious contracts to just become eligible for social security.[1] A second best policy hence looks quite similar as the Earned Income Tax Credit (EITC) in the US. An EITC-style policy links transfer payments solely to individual earnings, which are observable by the government. Earnings below a threshold are partially matched by the government, which increases the return-to-work and pulls low income workers into employment.

We evaluate the effects of employment based progressive pension policies in a quantitative stochastic overlapping generations model, in which households face an explicit labor force participation decision and can on top choose their working hours. Households can partially self-insure through saving in a riskless asset. They make decisions under the presence of persistent shocks to labor productivity and longevity risk as well as shocks to individual life expectancy. A government collects progressive taxes on labor earnings and taxes on consumption to finance government expenditure, and operates a pay-as-you-go pension system that is financed by payroll taxes.

The starting point of our analysis is a situation with a *proportional pension system*,

[1] In the end, the government still is bound by not being able to observe individual productivity.

in which old-age pension benefits are directly proportional to lifetime earnings. In our reform scenarios, we increase the progressivity of the pension system by allowing for a disproportionately high accumulation of pension claims for earnings-poor working households, which comes at the expense of a cut in pension claims for high-earnings individuals. We do so in two steps. First, we introduce a progressive pension component that is directly linked to the individual employment decision. Through this component, households acquire pension claims for every year they were employed, irrespective of how much they earned. We use this *employment-linked progressive pension system* (ELS) as a benchmark case, with the implicit assumption being that the public pension insurer is perfectly informed about an individual's employment status. In a second step, we appreciate the fact that the government may have problems in observing employment and study an EITC-like progressive pension which we call the *Earned Income Pension Credit* (EIPC). This system solely relies on individual earnings as a measure to calculate pension benefits. The pension formula has a phase-in and a phase-out region. Households with earnings less than a certain threshold accumulate disproportionately high pension claims for each year they are in employment. Consequently, the system sets incentives for both labor force participation and higher labor hours at the lower end of the earnings distribution.

Our quantitative simulation model is calibrated to the German economy, which currently features a proportional pension system. To adequately describe individual labor market risk over the life cycle, we use administrative data from the German pension insurance system. We find that individuals are exposed to a significant amount of labor productivity risk, much richer than the standard AR(1) process for log-labor earnings would predict. Most importantly, individuals face a serious portion of *low earnings episodes*. A typical worker in such an episode only makes about ten percent of average labor earnings in a year. Low earnings episodes significantly impact on life-time earnings and make individuals marginally attached to the labor force. We estimate a first-order Markov process for labor productivity that captures the salient features of low earnings episodes over a household's working life.

Using the calibrated simulation model, we first quantify the effects of employment based progressive pension reforms on individual labor force participation and labor hours. The positive employment effects can be sizable and are concentrated among workers with adverse productivity shocks. In the long run, the overall employment rate increases by 1.3 percentage points. Most of the employment gains stem from high-school educated workers, but college-educated workers react positively, too. As an example, the introduction of an employment-linked progressive pension leads the least productive 35-year-old high-school and college-educated workers to increase their employment by 14 and 3 percentage points, respectively. An Earned Income Pension Credit is somewhat less efficient in stimulating employment, as it can not directly tackle the individual employment decision. Nevertheless, the projected employment gains are still substantial, in the order of 1.1 percentage points for the working population at large. Along the intensive margin, the labor supply decision of households is mostly distorted downwards, leading to an overall decline of about 0.9 hours per week. Intensive margin distortions also affect high productivity workers, but positive employment effect are concentrated among low productivity individuals. As a result, aggregate labor input declines by roughly 1 percent.

All progressive systems we consider substantially reduce old-age income inequality and provide insurance against labor productivity shocks. By stimulating employment at the lower end of the productivity distribution, they also alter the risk-properties of labor earnings during working life. The reduced need for self-insurance leads to a decline in aggregate savings along a transition path. The reforms also induce a drop in aggregate consumption. In the initial periods of the transition path, consumption falls by about 0.8 percent. As private assets shrink along the way to the new long-run equilibrium, the consumption decline becomes even more pronounced.

Finally, we evaluate the welfare and efficiency effects of progressive pension reforms. Our preferred measure of household welfare is ex-ante expected life-time utility. We calculate the consumption equivalent variation that each cohort affected by a pension reform[2] experiences. As the welfare effects of pension reforms can vary a lot across different generations, we also derive an aggregate measure of the economic efficiency effect that takes into account the welfare changes of all affected cohorts. The introduction of progressive pensions increase welfare of almost all cohorts, except for the already retired at the time of the reform. The latter experience a small welfare loss from a rise in the consumption tax rate. The aggregate efficiency effect of introducing an employment-linked progressive pension is positive. It amounts to a permanent rise in consumption of 0.73 percent. EIPC systems are less effective tools, as they can not directly condition on the individual employment decision. Yet, they can still recover around 90 percent of the efficiency gains of an employment-linked progressive pension. Positive welfare effects predominantly stem from high-school workers. College graduates on average experience welfare losses.

Related Literature Methodologically, our paper is related to a strand of literature that uses quantitative general equilibrium models with heterogeneous agents to analyze the incentive effects and welfare implications of redistributive fiscal policy. Popular themes of papers in this field include the optimal progressivity of the income tax code or the optimal taxation of capital income, see for example Domeij and Heathcote (2004), Conesa and Krueger (2006), Conesa et al. (2009), and Kindermann and Krueger (2022).

Huggett and Ventura (1999) were among the first to evaluate the welfare implications of redistributive social security in a quantitative macroeconomic model. Nishiyama and Smetters (2008), Fehr and Habermann (2008), Fehr et al. (2013),

[2]I.e., the initial cross-section of households at the time of reform as well as all new-born generations along the transition path

and Brendler (2021) study the optimal design of the social security benefit formula. They generally find progressive pensions to be desirable both from a long-run welfare and an aggregate efficiency point of view. Yet, their models only feature intensive margin labor supply decisions. O'Dea (2018) finds support for the introduction of means-tested old-age income programs that substantially reduce the variance of life-time consumption. Nam (2022) points to the fact that progressive pensions are an adequate measure for counteracting heterogeneity in job stability over the life cycle. In contrast to these papers, we look at both intensive and extensive labor supply choices and point to the fact that considering these margins together has significant implications for the optimal design of progressive pension formulas.

The literature presenting evidence on the labor supply incentives of social security design is much scarcer. Wallenius (2013) and Lalive et al. (2023) point to the fact that social security reforms can have an impact on the timing of retirement. French et al. (2021) exploit the switch from a defined benefit to a defined contribution system in Poland to estimate the labor supply reactions to pension reforms. Their study demonstrates that pension reforms can have an impact on individual labor supply decisions already many years prior to retirement. These results are reassuring in the sense that the observed labor supply reactions in our simulation analysis are also likely to occur in reality. Moreover, compared to the Polish reform, our proposed pension formula is much simpler and the implied incentives are easier to understand for households, which could result in even stronger labor supply reactions.

The design of optimal income support payments to the poor is at the core of Saez (2002) and Bierbrauer et al. (2023), who generally argue in favor of having an Earned Income Tax Credit. Hansen (2021) provides explicit conditions under which an optimal income tax codes contains an EITC component. The incentive and welfare effects of the EITC are also studied, for example, in Chan (2013), Athreya et al. (2010), and Ortigueira and Siassi (2022).

Finally, our paper is related to a literature that is concerned with other features of social security that might lead to a progressive or regressive distribution between households. Breyer and Hupfeld (2010), Goda et al. (2011), Coronado et al. (2011), Bagchi (2019), and Jones and Li (2022) point to the positive correlation between income and life expectancy and study its implications for the social security system. Nishiyama (2019) quantifies the impact on spousal and survivor benefits on labor supply and welfare.

The remainder of our paper is structured as follows: Section 2 investigates the structure of life-time labor earnings inequality using administrative data from the German pension system. Section 3 illustrates the basic economic mechanisms at work in a tractable, two-period analytical model. In Section 4, we present our full quantitative simulation model, and discuss its calibration in Section 5. In Section 6, we show simulation results for life-cycle choices, macroeconomic performance and welfare along a transition path. The last section concludes.

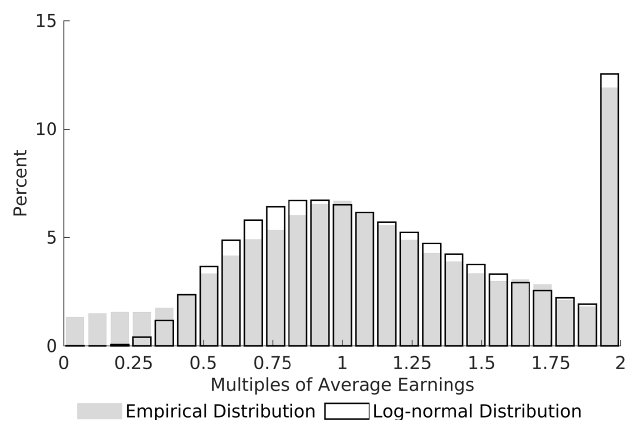
2 Facts about labor earnings inequality

Providing a proper model for the household's life-cycle labor earnings process is crucial if one wants to assess the benefits of fiscal redistribution and insurance. To this end, we document salient facts on labor earnings inequality and risk over the life cycle. Our discussion is based on data from Germany, as the German public pension insurance system (Deutsche Rentenversicherung) offers an adminstrative dataset with detailed information on the earnings histories of a subsample of all insured households. What we find in this data is consistent with recent research from other countries, especially the US. In particular, we will argue that a simple log-normal AR(1) process is not a good description of the dynamics of individual labor earnings, a fact also supported by the work of e.g. Guvenen et al. (2015), Busch and Ludwig (2020), de Nardi et al. (2020) or Halvorson et al. (2020).

Our administrative dataset, the scientific use file of the Versichertenkontenstichprobe 2017, contains information from the insurance accounts of 69,520 individuals actively insured under the public mandatory German pension scheme.[3] Next to information on age, gender and education, insurance accounts record a monthly history of accumulated pension claims together with an indicator of the source these claims were accumulated from (like labor earnings, unemployment, child care, etc.). Note that in the German pension system, pension claims that stem from regular employment are proportional to individual earnings.[4] Hence, they are a good indicator for estimating earnings processes.

Figure 1 shows a histogram of raw individual annual earnings (gray bars) of men aged between 25 and 60,[5] expressed as multiples of average labor earnings of the total population. The figure reveals two salient features of the data: First, the

Figure 1: Histogram of pension claim distribution



[3]The German pension scheme covered 38 million actively insured individuals in 2017.

[4]We adjust the data in the case of earnings from so-called mini- and midi-jobs, which are subject to reduced social security contributions.

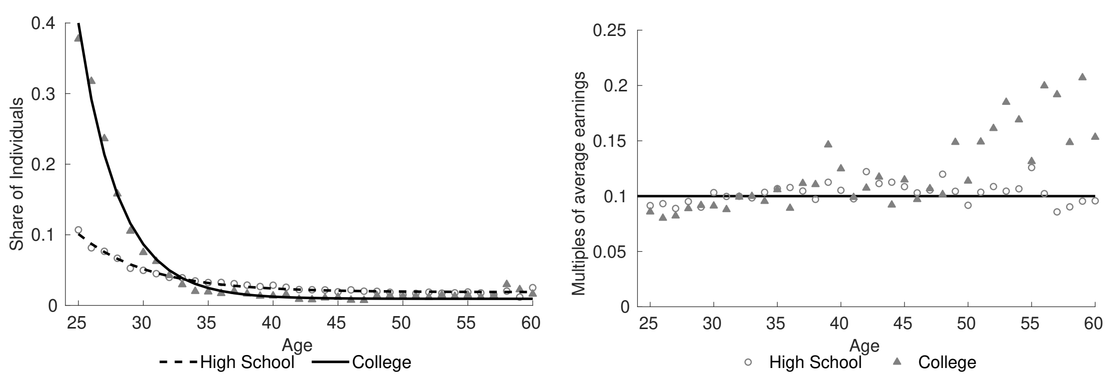
[5]In Appendix A.1, we provide a more detailed description of our sample and sample selection.

data are top-coded at about two times the average earnings. This is owing to the presence of a contribution ceiling in the German pension system. Second and more importantly, there is a substantial mass at values below 0.25, which is atypical under the usual assumption of log-normally distributed earnings. To strengthen this point, the framed bars in Figure 1 show the histogram of a log-normal distribution that provides the best fit to our data. Under log-normality, the share of households at the lower end of the earnings distribution is almost zero. Our sample hence looks stratified and using the assumption of a common log-normal distribution to describe individual earnings seems invalid.

To deal with this feature of the labor earnings data, we will split the dataset in two parts for our empirical analysis. We define the earnings threshold that separates the two groups as 6 months of full-time work at the German minimum wage, or 0.23 times average labor earnings. All individuals with labor earnings above the threshold are henceforth said to have *normal labor earnings*. The earnings process of such workers is well described by the standard assumptions about life-cycle labor earnings. In particular, the data shows the typical life-cycle labor earnings profiles, a significant college wage premium, and a high auto-correlation of earnings, which can be approximated by an AR(1) process in logs.

Individuals with earnings below the threshold are called *low earnings individuals*.[6] The left panel of Figure 2 shows the fraction of individuals by age that are members of the low earnings group (circles for high school and triangles for college educated workers), the right panel shows average labor earnings of low earning individuals. Low earnings episodes significantly impact on life-time earnings and

Figure 2: Life-cycle dynamics of low labor earnings



make individuals marginally attached to the labor force. A typical low-earnings worker only makes about ten percent of average labor earnings in a year. The dynamics of low earnings episodes over the individual life cycle are quite distinct across education groups. College educated workers predominantly experience low

[6]Low earnings can, for example, arise from having some months of temporary unemployment or non-employment throughout a year or being marginally employed (i.e. having a so-called mini-job).

labor earnings early in their career, for example when doing internships or while working in addition to studying in college. At later ages, the share of individuals in the low earnings region converges to almost zero. For high school workers, on the other hand, experiencing a low earnings episode is a phenomenon that is more equally distributed across ages. To deal with these particular features of the life-cycle labor earnings data, we estimate a first-order Markov process that captures the salient features of low income shocks over a household's working life and use the estimation results to inform our quantitative model. The calibration section as well as Appendix A provide the technical details and estimation results.

3 Building Intuition: A Two-Period Framework

Before setting out our large-scale simulation model, we want to build some intuition for the main mechanisms at work using a much simpler and stylized framework. Households in this framework live for two periods $j = 1, 2$. They can supply labor only in the first period of life, in the second period they are retired. The interest rate r as well as the wage rate w for effective labor are exogenous.

The labor supply decision consists of two stages: an extensive and an intensive one. Households first have to decide whether to work or not. We denote the choice to be non-employed or employed by $e \in \{0, 1\}$. Once they joined the labor force, agents choose their optimal number of labor hours ℓ. Individuals derive utility from consumption c_j in each period and suffer disutility from working. For analytical tractability, we assume that preferences are quasi-linear in consumption and that the time discount rate equals the interest rate r.[7] More specifically, we let preferences be represented by the utility function

$$U(c_1, c_2, \ell, e) = c_1 + \frac{c_2}{1+r} - \frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e. \tag{1}$$

Consistent with household choices, disutility from labor is due to an intensive and an extensive margin component. The former is primarily governed by the Frisch elasticity of labor supply χ. The latter kicks in through a utility costs of employment ξ. Note that households only have the capacity to earn income by providing hours ℓ if they formally joined the labor force ($e = 1$).

Households maximize utility in (1) subject to the present value budget constraint

$$c_1 + \frac{c_2}{1+r} = (1-\tau_p)y + \frac{p}{1+r}. \tag{2}$$

Households pay contributions to the pension system in the form of a payroll tax τ_p on their total labor earnings $y = we\ell$. As a reward for their contributions, they receive a pension payment p.

[7]We relax all of these assumptions later on in our quantitative model in Section 4.

Following our previous discussion about the design of employment based progressive pensions, we analyze two progressive pension systems. The first system creates redistribution based on an indidividual's employment decision e. We refer to this system as the Employment-Linked System (ELS). The second system is closely related to the Earned Income Tax Credit in that it wants to set employment incentives. In contrast to the ELS, however, it is entirely based on individual earnings y. We refer to this system as the Earned Income Pension Credit (EIPC).

3.1 The Employment-Linked System

In the ELS, we assume that pension payments p are calculated from two components: First, the household's employment status in period 1, and second, her individual labor earnings. Specifically, we let

$$p = \kappa \times \Big[\lambda \bar{y} e + (1-\lambda) y\Big], \tag{3}$$

where $\bar{y}$ denotes average labor earnings of the employed. Households receive a fixed pension reward whenever they are employed, which is indexed to average earnings and independent of the household's own income position. In addition, they get an earnings-tied pension. The left panel of Figure 3 depicts this system graphically. On the horizontal axis, the figures shows the earnings of an individual relative to the average earnings of the population $\frac{y}{\bar{y}}$. On the vertical axis, we indicate a worker's pension benefit normalized by the pension replacement rate $\frac{p}{\kappa}$. The dashed line indicates a proportional pension system, while the solid line illustrates the ELS with a value of $\lambda = 0.5$.

The factor λ indicates the strength of the employment component relative to the earnings-related component. Since the size of the employment component is independent of individual income, λ is also a measure for the progressivity of the pension system. If $\lambda = 0$, the ELS collapses to a purely proportional pension system. Note, that redistribution within the pension system is limited to the employed, since households do not acquire any pension claims when they are not in employment in the first period ($e = 0$).

In the following, we deliberately assume that $\tau_p = \frac{\kappa}{(1+r)}$.[8] Combining the household budget constraint (2) with the pension formula (3) as well as the return assumption on pension payments, we can write the budget constraint as

$$c_1 + \frac{c_2}{1+r} = \Big[1 - \underbrace{\lambda \tau_p}_{=:\tau_p^{\text{imp}}}\Big] y + \underbrace{\lambda \tau_p \bar{y}}_{=:\tau_p^{\text{sub}}} e. \tag{4}$$

The pension system influences the household budget constraint in two ways. On the one hand, it imposes an *implicit tax* on intensive labor supply $\tau_p^{\text{imp}} = \lambda \tau_p$,

[8] In a closed economy model, this assumption implies that the population growth rate of the economy, which defines the implicit return on pension contributions, has to be equal to the interest rate on financial investments, i.e. $r = n$. In Appendix B, we provide some general equilibrium foundations for this and also investigate the case of $r \neq n$.

which rises in the degree of pension progressivity λ. The implicit tax is equal to zero when the pension system is fully proportional ($\lambda = 0$). In this case, any additional Euro a household contributes to the system pays the same return as a financial investment. In the other extreme case where $\lambda = 1$, an increase in intensive margin labor supply has no effect on the size of the household's pension. Consequently, we have $\tau_p^{\text{imp}} = \tau_p$, meaning that all of the pension contribution is perceived as a tax. On the other hand, the pension system comes with a *subsidy to employment* $\tau_p^{\text{sub}} = \lambda \tau_p \bar{y}$. This subsidy emerges when the pension system pays benefits that are solely linked to the employment status of a household, and not to individual earnings. A larger λ implies a greater importance of the employment component, and therefore leads to a higher employment subsidy.

Summing up, a higher pension progressivity λ has two opposing effects: it distorts labor supply on the intensive margin by imposing a higher implicit tax rate on households, but it encourages employment by providing a greater participation subsidy.

3.2 The Incentive Effects of the ELS

The above conclusion becomes even clearer when we look at the households' labor supply choices. Maximizing utility in (1) subject to the household budget constraint (4) yields

$$\ell(z|e=1) = \left[(1-\tau_p^{\text{imp}})w\right]^{\chi}.$$

In the absence of income effects, the intensive margin labor supply choice is immediately determined by the implicit tax rate τ_p^{imp} of the pension system. A higher progressivity λ that leads to a rise in the implicit tax rate τ_p^{imp} unanimously distorts the intensive margin labor supply decision downwards.

As for the employment choice at the extensive margin, the household has to compare her utility from working to the utility from not working. She will choose to be employed whenever

$$U(e=1) - U(e=0) = \frac{\left[(1-\tau_p^{\text{imp}})w\right]^{1+\chi}}{1+\chi} + \tau_p^{\text{sub}} - \xi \geq 0.$$

Appendix B.5 shows that this utility difference increases with λ, whenever a worker's earnings are smaller than the average earnings of the workforce, i.e., $y < \bar{y}$. Hence, the ELS provides positive employment incentives for the earnings poor.

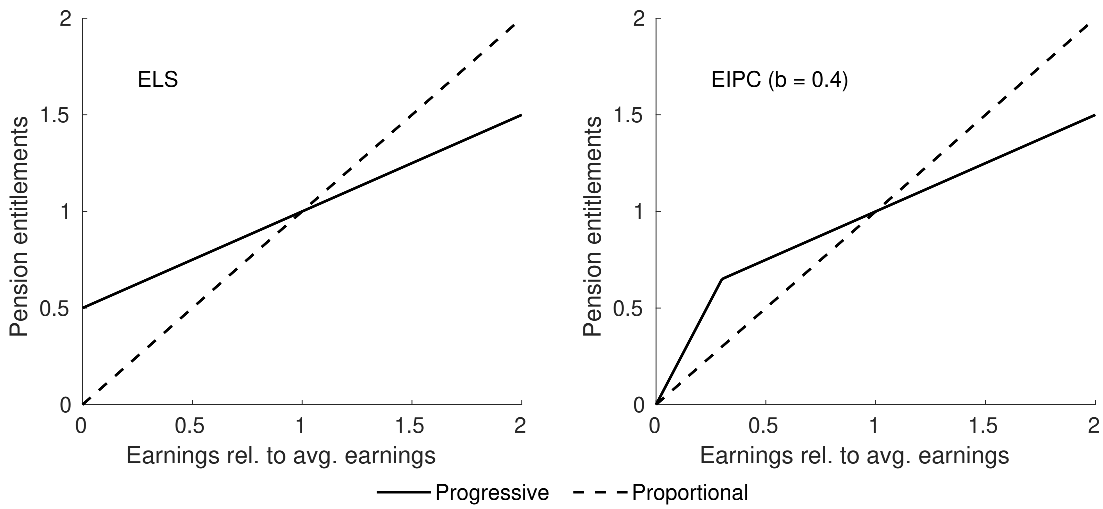
3.3 The Earned Income Pension Credit

The EIPC can only grant pension payments on the basis of individual earnings y. To make the system consistent with the ELS, we propose the functional form

$$p = \kappa \times \begin{cases} \lambda \frac{y}{b} + (1-\lambda)y & \text{if } y < b\bar{y} \text{ and} \\ \lambda \bar{y} + (1-\lambda)y & \text{otherwise,} \end{cases} \tag{5}$$

where $b \in (0, 1)$. The right panel of Figure 3 indicates the shape of this pension formula. The dashed line again indicates the proportional pension system, while the solid line illustrates the EIPC pension formula proposed in (5) with values $\lambda = 0.5$ and $b = 0.3$.

Figure 3: Progressive Pension Formulas



The phase-in and phase-out structure in the spirit of the Earned Income Tax Credit becomes immediately apparent from the figure. The proposed pension formula defines a threshold level b as a fraction of average earnings $\bar{y}$ at which a worker can enjoy the maximum employment subsidy $\tau_p^{\text{sub}} = \tau_p \lambda \bar{y}$. The incentive effects of this system are in fact identical to those of the ELS for all workers with earnings greater than the threshold level $b\bar{y}$. For workers with labor earnings less than the threshold, we can write the budget constraint as

$$c_1 + \frac{c_2}{1+r} = \Big[1 - \underbrace{\lambda \tau_p}_{=:\tau_p^{\text{imp}}}\Big] y + \underbrace{\lambda \tau_p \frac{y}{b}}_{=:\tau_p^{\text{sub}}}. \tag{6}$$

The size of the employment subsidy is now conditional on a worker's labor earnings y. The more earnings an individual can generate from working, the higher the implicit employment subsidy will be, and the more likely that individual is to participate in the labor market. Additionally, low-earning individuals have a strong incentive to increase their intensive margin labor supply up to the bend point b, as their pension benefits increase disproportionately with any additional Euro earned. It should be noted that in this case, we cannot strictly separate the employment subsidy from the implicit tax rate. Note further that as b approaches zero, the EIPC looks more and more like the ELS, with all the benefits and problems.

Summing up, the proposed progressive pension systems can be expected to affect aggregate labor supply in two ways. While they distorts labor supply along the intensive margin for the earnings rich, they provide incentives for taking up employment and (potentially) for expanding labor hours for the earnings poor. The effect on total labor earnings is therefore ambiguous and depends on the exact choices of the intensive margin labor supply elasticity, the shape and distribution of participation costs and the distribution of labor earnings in the population. What is more, a progressive pension system not only influences households' labor supply decisions. It also alters the distribution of household income at old-age by redistributing between households with different life-time incomes and by providing insurance against productivity fluctuations over the life cycle. We quantify the importance of labor supply distortions, redistribution and insurance for aggregate welfare and economic efficiency in a quantitative simulation model in the next section.

4 The Quantitative Simulation Model

Our full quantitative simulation model is based on the previous theoretical considerations and informed by the empirical facts regarding income risk. In particular, we employ a general equilibrium overlapping generations model with survival risk in the spirit of Auerbach and Kotlikoff (1987). Households draw persistent shocks to their labor productivity, like in Conesa et al. (2009), and have to decide about whether to be employed, how many hours to supply and about how much to consume and save. In addition, individuals face shocks to their life expectancy. The government operates a (potentially progressive) pay-as-you-go pension system financed by payroll taxes and collects resources through a consumption tax and a progressive tax on labor earnings in order to finance general government expenditure. We consider an open economy framework, so that the prices for capital and labor are fixed, but government parameters adjust in order to keep the fiscal tax and transfer systems balanced. Our simulations start from a long-run equilibrium calibrated to the German economy. Any reform to the pension system puts the economy on a transition path to a new steady state. We calculate this entire transition path and measure the welfare effects on different cohorts and households along the transition.

4.1 Demographics

The economy is populated by overlapping generations of heterogeneous individuals.[9] At each point in time t, a new generation of size N_t is born. We assume that the population grows at a constant rate n. Households start their economic life at age $j = 20$ and live up to a maximum of J years, after which they die with certainty. They can supply labor to the market until they reach the mandatory retirement age j_R. Throughout their entire life, individuals are subject to idiosyncratic survival risk. Specifically, we denote by $\psi_{j,h}$ the conditional probability of an agent to survive from period $j-1$ to period j, with $\psi_{20,h} = 1$ and $\psi_{J+1,h} = 0$. Survival probabilities, and hence life expectancy, depend on the individual health status h, discussed in more detail below.

As population grows with a constant rate n, a long-run equilibrium in this economy is characterized by all aggregate variables growing at this very same rate. To make aggregates stationary again, we express all variables in per capita terms of the youngest generation at a certain date t. We denote by m_j the time-invariant relative size of a cohort aged j at any point in time.

4.2 Technology

A continuum of identical firms produce a single good Y_t under perfect competition. They hire both capital K_t at price r_t and labor L_t at price w_t on competitive spot markets. Firms operate a constant returns to scale technology

$$Y_t = \Omega K_t^{\alpha} L_t^{1-\alpha}. \tag{7}$$

Ω denotes the aggregate level of productivity, whereas α is the elasticity of output with respect to capital. In the process of production, a fraction δ of the capital stock depreciates. Given the assumptions about competition and technology, we can safely assume the existence of a representative firm that takes prices as given and operates the aggregate technology in (7). In addition to employing factor inputs, the firm has to invest I_t into its capital stock. The law of motion for the capital stock reads

$$(1+n)K_{t+1} = (1-\delta)K_t + I_t.$$

4.3 Preferences and Endowments

Preferences Households have preferences over stochastic streams of consumption $c_{j,t} \geq 0$, labor supply $\ell_{j,t} \geq 0$ and employment $e_{j,t} \in \{0, 1\}$. They maximize a discounted, generalized recursive, expected utility function

$$U_{j,t} = u(c_{j,t}, \ell_{j,t}, e_{j,t}) - \beta\psi_{j+1,h} E_t \left[(-U_{j+1,t+1})^{1+\gamma}\right]^{\frac{1}{1+\gamma}}.$$

[9] We use the terms individual, household and agent synonymously.

Our preference formulation follows Swanson (2018) and is a generalization of Epstein and Zin (1989) that allows to separate intertemporal substitution from risk aversion. Individuals form expectations with respect to future labor productivity and health and incur a utility loss from being employed. They discount the future with the constant time discount factor β as well as their individual survival rate. For the sake of notational ease, we deliberately drop the time index t on all household level variables.

Labor productivity Households are ex ante homogeneous, but differ ex post in their labor productivity $z(j, s, \eta)$. At the beginning of life, they draw one of two education levels: high-school education ($s = 0$) or college education ($s = 1$); the probability to draw $s = 1$ is ϕ_s. All individuals of education s share a common deterministic age-specific labor productivity profile $\theta_{j,s}$.

Throughout their working life, households' labor productivity is due to idiosyncratic shocks η. For individuals with *normal labor earnings*, we assume that their productivity follows a standard, education-specific AR(1) process in logs

$$\eta^+ = \rho_s \eta + \varepsilon^+ \quad \text{with} \quad \varepsilon^+ \sim N(0, \sigma^2_{\varepsilon,s}), \tag{8}$$

where innovations ε^+ are iid across households.

The evidence provided in Section 2 has shown that a simple AR(1) process is not enough to describe the earnings distribution of households. To cope with the fact that a significant part of workers experiences *low earnings episodes* we proceed as follows: We assume that, knowing their education level, households divide into two groups $m \in \{0, 1\}$. m is a permanent state that indicates whether an individual faces a stable career path ($m = 0$) or an unstable career path ($m = 1$). The probability to draw the state $m = 1$ is denoted by ϕ_m. The labor productivity dynamics of workers with stable careers is described solely by the AR(1) process shown above. On top, agents with an unstable career can be hit by an additional persistent (but not permanent) *low productivity shock*, regardless of their current productivity. When exiting the low productivity state, agents revert to normal AR(1) productivity. We provide details on the exact parameterization of low productivity shocks in the calibration section.[10]

We denote by $\pi_\eta(\eta^+|\eta, j, s, m)$ the probability distribution of next-period's productivity η^+, conditional on current labor productivity η, age j, education s and career stability m. Finally, the wage an individual faces equals the product of the wage rate per efficiency unit of labor and her individual labor productivity $w_t \times z(j, s, \eta)$.

[10] This approach is consistent with empirical evidence from the labor literature that starts with Hall (1982). More recently, Kuhn and Ploj (2020) investigate the importance of career instability for heterogeneity in household wealth. Nam (2022) analyzes the consequences of career instability for the optimal progressivity of the pension system. The approach also follows Castaneda et al. (2003) or Kindermann and Krueger (2022), who augment standard AR(1) processes for labor productivity with additional shocks to paint a realistic picture or top 1% earnings and wealth heterogeneity in the US.

Budget constraint Markets are incomplete. Like in Bewley (1986), Imrohoruglu (1989), Huggett (1993), and Aiyagari (1994), households can only self-insure against fluctuations in individual labor productivity by saving in a risk-free asset a with return r_t. Savings are subject to a tight borrowing constraint, so that household wealth needs to satisfy $a \geq 0$. Households' resources are composed of their current wealth (including returns), their income from working $y = w_t z(j, s, \eta) e\ell$, intergenerational transfers b,[11] as well as pension payments p. They use these resources to finance consumption expenditure $(1 + \tau_{c,t})c$ (including consumption taxes) and savings into the next period a^+, contributions to social security $T_{p,t}(y)$ as well as progressive income taxes $T_t\Big(y - T_{p,t}(y) + p\Big)$. Households can deduct social security contributions from gross income for the purpose of taxation. In turn, all pension benefits are liable for taxation.

Individual life expectancy A household's savings behavior is shaped by the interest rate, the discount factor, productivity risk and individual life expectancy. As for the latter, we assume that individual survival probabilities are defined by some health state h. Each health level is associated with a set of age specific survival probabilities $\psi_{j,h}$ that lead to a certain life expectancy. An agent's health status can change over the life cycle according to the probability distribution $\pi_h(h^+|h, j, s, \eta)$. Future health h^+ hence is conditional on current health, age, education and individual labor productivity.

Dynamic optimization problem The current state of a household is described by a vector $\mathbf{x} = (j, s, m, \eta, h, a, ep)$ that summarizes the household's age j, education s, career stability m, her current labor productivity shock η, health h, her wealth position a as well as the amount of already accumulated pension claims ep. The dynamic optimization problem of an individual then reads

$$v(\mathbf{x}) = \max_{c,\ell,e,a^+,ep^+} \; u(c, \ell, e) - \beta\psi_{j+1,h} E\left[\Big[-v_{t+1}(\mathbf{x}^+)\Big]^{1+\gamma} \;\middle|\; \mathbf{x}\right]^{\frac{1}{1+\gamma}} \tag{9}$$

with $\mathbf{x}^+ = (j + 1, s, m, \eta^+, h^+, a^+, ep^+)$. Households maximize (9) subject to the borrowing constraint $a^+ \geq 0$, the budget constraint

$$(1 + \tau_{c,t})c + a^+ + T_{p,t}(y) + T_t\Big(y - T_{p,t}(y) + p\Big) = (1 + r_t)a + y + p + b$$
$$\text{with } \; y = w_t z(j, s, \eta)e\ell,$$

the accumulation equation for pension claims ep^+ discussed in Section 4.4 as well as the laws of motion for labor productivity π_η and health π_h. The result of this dynamic program are policy functions c, ℓ, e, a^+, and ep^+ that all depend on the household's current state $\mathbf{x}$. We derive the first-order conditions in Appendix C.1.

[11]Intergenerational transfers consist only of accidental bequests that households might leave if they die before the terminal age J. We assume that the total of those accidental bequests is distributed lump-sum to all working age households.

4.4 The Pension System

The pension system has a contribution ceiling equal to two times average labor earnings of the employed. We therefore define *pension relevant earnings* as

$$y_p = \min\Big(wz(j,s,\eta)e\ell,\ 2\bar{y}_t\Big).$$

Households pay payroll taxes at rate τ_p on relevant earnings. In reward for their contributions, they earn pension claims ep. We can write

$$T_{p,t}(y) = \tau_{p,t} \times y_p \quad \text{and} \quad ep^+ = ep + f_t(y_p), \tag{10}$$

where the function f_t determines the relationship between relevant labor earnings and pension claims. In the initial equilibrium denoted by $t = 0$, we assume that the pensions system is purely proportional (as it is in Germany) and therefore set $f_0(y) = y$.

Finally, individual pension benefits $p(ep)$ are calculated from the life-time average of earned pension claims as

$$p(ep) = \kappa_t \times \frac{ep}{j_R - 20},$$

where κ_t is the replacement rate.

The pension system operates on a pay-as-you-go basis. In the initial equilibrium, total pension contributions hence need to be equal to the total amount of pension payments. Letting Φ_t denote the cross-sectional measure of households over the state space,[12] we require

$$\tau_{p,0} \times \underbrace{\int y_p \, d\Phi_0}_{\text{contribution base}} = \underbrace{\int p(ep) \times \mathbb{1}_{j \geq j_R} \, d\Phi_0}_{\text{total pension claims}}. \tag{11}$$

We will depart from the notion of period-by-period budget balance along the transition path in order to smooth the costs and benefits of pension reforms over multiple generations. We provide more details in Section 6.2.

4.5 The Tax System and Government Expenditure

The government collects proportional taxes on consumption expenditure and progressive taxes on labor earnings net of social security contributions as well as pension payments. In addition, it can issue debt B_t. Fiscal revenue is used to

[12] Φ_t is a measure and indicates the mass of households on each subset of the state space. We require that for each age j, Φ_t sums up to the total mass of households in a cohort m_j. A detailed analytical description of Φ_t can be found in Appendix C.3.

finance (wasteful) government spending as well as debt services. The government budget constraint reads

$$\tau_{c,t} \times C_t + \int T_t\Big(y - T_p(y_p) + p\Big)\; d\Phi_t + (1+n)B_{t+1} = G_t + (1+r_t)B_t$$
$$\text{with } \; y = w_t z(j,s,\eta)e\ell. \quad (12)$$

C_t denotes aggregate consumption and T_t the progressive income tax schedule. We assume that government consumption is fixed per capita. Consequently, we adjust the tax system to keep the fiscal system in balance.

4.6 Capital Markets, Trade and Equilibrium

We model a small open economy that freely trades capital and goods on competitive international markets. All private savings that are not absorbed by the domestic production sector or the government are invested abroad at the international interest rate $\bar{r}$. The capital market equilibrium reads

$$K_t + B_t + Q_t = A_t,$$

where A_t are aggregate private savings and Q_t is the country's net foreign asset position. As the economy grows at rate n, the net foreign asset position increases over time such that the capital account is $Q_t - (1+n)Q_{t+1}$. Net income from abroad, on the other hand, amounts to $\bar{r}Q_t$. According to the balance of payments identity, we therefore have a trade balance of

$$TB_t = (1+n)Q_{t+1} - (1+\bar{r})Q_t. \quad (13)$$

The economy's interest rate is then equal to the world-wide interest rate $r_t = \bar{r}$.

We assume that the government collects all accidental bequests and redistributes them in a lump-sum way among the surviving working-age population. Consequently,

$$b_{j,t} = \frac{\int \frac{1-\psi_{j,h}}{\psi_{j,h}} \times (1+r_t)a\; d\Phi_t}{\int \mathbb{1}_{j<j_R}\; d\Phi_t} \quad \text{if } j < j_R. \quad (14)$$

Given an international interest rate and the fiscal policy parameters, a *recursive competitive equilibrium* of this model is a set of household policy functions, a measure of households, optimal production inputs, factor prices, accidental bequests, a net foreign asset position and a trade balance that are consistent with individual optimization and market clearance. A formal definition of the equilibrium is available in Appendix C.2.

5 Calibration

This section discusses our choices of functional forms and parameters. We pay particular attention to the labor supply decision of households along the extensive and the intensive margin. We calibrate our model to the German economy, which currently features a proportional pension system in line with the one described in the previous section. Germany therefore serves as a good benchmark for reforms that aim at introducing progressivity into the pension formula.

5.1 Demographics

We assume a population growth rate of $n = 0.0$, which is a compromise between the average growth rate of 0.4% reported in the period 2012 to 2017 for the German population at large, and the fact that most of German population growth came from refugee migration, see German Statistical Office (2020).[13] We let households start their economic life at the age of 20 and allow for a maximum life span of 99 years. Mandatory retirement is at the age of 64, which equals the current average retirement age of the German regular retirement population, see Deutsche Rentenversicherung Bund (2019).

With regards to life expectancy, we extract the 2017 annual life tables for men from the Human Mortality Database (2020) to calculate average survival probabilities $\bar{\psi}_j$ of the overall population. We assume that all households share these common survival probabilities throughout their working life. Upon entering retirement, each individual draws one out of eight permanent health shocks $h \in \{0, \ldots, 7\}$. A health shock is associated with a set of survival probabilities $\psi_{j,h}$ that we choose such that (i) life expectancy at the lowest health shock $h = 0$ is ten years below average, (ii) life expectancy at the highest health shock $h = 7$ is ten years above average and (iii) life expectancy evolves linearly with health shocks h. The left panel of Figure 15 in the Appendix shows the respective survival probability profiles.

The probabilities $P(h|s, \eta)$ to draw a certain health shock upon entering retirement depend on the individual's education s and on the last labor productivity shock η prior to retirement. They are chosen to match a 2.5 year life-expectancy gap between high-school and college workers (Luy et al., 2015) and a gap of around 7 years between individuals in the top and the bottom life-time labor earnings decile (Haan et al., 2020). The details of how we exactly calibrate these shocks can be found in Appendix D. Our model features one single health shock that individuals are exposed to right before entering retirement. After the individual health status is revealed, households retain their health level for the rest of their life. While agents share a common set of survival probabilities during their entire working life, they still form expectations with respect to their survival chances

[13] In fact, the growth rate of the native population was -0.2% in the same time period.

at retirement. Hence, the need for old-age savings differs across individuals of different education levels and labor productivities.

5.2 Technology

On the technology side we choose a depreciation rate of $\delta = 0.07$, leading us to a realistic investment to output ratio of 21 percent. We set the capital share in production at $\alpha = 0.3$ and normalize the technology level Ω such that the wage rate per efficiency unit of labor w_t is equal to 1. Finally, we assume an international interest rate of $\bar{r} = 0.03$, which constitutes as mix between the (in 2017) very low interest rates on deposits and long-run investment opportunities that offer higher returns.

5.3 Preferences and Endowments

5.3.1 Preferences

We let the period utility function be

$$u(c, \ell, e) = \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \nu_s \frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi_s e.$$

We choose an intertemporal elasticity of substitution σ of 0.8. The choice of σ has important implications for the size of the income effect of wage changes on labor supply and therefore for the life-cycle profiles of participation and labor hours, see Section 6.1. Our choice of σ ensures that our model is able to match empirical life-cycle profiles. Our preferred value for the Frisch elasticity is $\chi = 0.4$, which is a medium range value, see for example Keane (2011). We choose the education-specific level parameters of intensive labor supply $\nu_1 = 46.55$ and $\nu_2 = 32.8$ so as to target a 38.1 hour and a 40.0 hour work week for high-school and college-educated employed workers, respectively. According to Swanson (2018), the relative risk aversion with respect to fluctuations in individual consumption in our utility formulation is approximately equal to

$$R_c \approx \frac{1}{\sigma+\chi} + \frac{\gamma(1-\sigma)}{\sigma + \frac{1-\sigma}{1+\frac{1}{\chi}}}.$$

We set $\gamma = 9.286$ so that relative risk aversion is equal to 3.[14] Finally, we set the time discount factor to $\beta = 0.9835$ so that all capital and public debt is entirely absorbed by private savings in the initial equilibrium, and net foreign assets as well as the trade balance are zero.

[14]Note that in the absence of additional curvature, consumption risk aversion would only be 0.83.

The micro Frisch elasticity χ only is an intensive margin elasticity and does not incorporate extensive margin choices. The macro labor supply elasticity, which incorporates both intensive and extensive margin choices, is typically larger, see the discussion in Keane and Rogerson (2011) or Peterman (2016). The extensive margin labor supply reaction to a change in wages is to a large degree determined by the probability density of the utility costs of employment ξ. Our calibration strategy for the distribution of participation costs ξ is the following: We assume that ξ is iid across households and independent of the household's labor productivity $z(j, s, \eta)$. We let ξ follow a log-normal distribution with education-specific mean $\mu_{\xi,s}$ and a common variance σ_ξ^2. The means are set so as to target employment-to-population ratios for the 25 to 54 year old by education level. The variance is chosen to target evidence on participation elasticities in Bartels and Pestel (2016), see Appendix D.3 for further details.

5.3.2 Labor Productivity

In Section 2, we already sketched the dynamics of labor earnings using administrative data on the German working population. However, in our quantitative model we need to parameterize labor productivity, which differs from labor earnings when individual labor hours vary across ages and states.

We parameterize the age-productivity relationship using the functional form

$$\theta_{j,s} = b_{0,s} + b_{1,s}\frac{\min(j, j_{M,s})}{10} + b_{2,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^2 + b_{3,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^3. \quad (15)$$

This functional form is flexible enough to capture both a hump-shaped ($j_{M,s} = \infty$) and a stagnating ($j_{M,s} < j_R$) life-cycle labor productivity profile. Note that in the case of a stagnating profile, labor productivity is constant from age $j_{M,s}$ onward. We let labor productivity risk of workers be guided by a standard first-order autoregressive process with parameters ρ_s and $\sigma^2_{\varepsilon,s}$ as in (8). In addition, we assume that workers with an unstable career path ($m = 1$) are exposed to an additional first-order Markov process of the form

$$\Pi^s_{low} = \begin{bmatrix} 1 - \pi^s_{low,0} & \pi^s_{low,0} \\ 1 - \pi^s_{low,1} & \pi^s_{low,1} \end{bmatrix} \quad \text{with initial distribution} \quad \begin{bmatrix} \omega^s_{low} \\ 1 - \omega^s_{low} \end{bmatrix}. \quad (16)$$

This process governs the transition into and out of the low earnings state, in which individuals face a labor log-productivity of η_0. $\pi^s_{low,0}$ consequently denotes the probability to receive a low-earnings shock, while $\pi^s_{low,1}$ is an indicator of the persistence of the low earnings state.

To provide a suitable calibration for the labor productivity process, we first set the share of college educated workers to $\phi_s = 0.2373$ in accordance with the data and assume $\phi_m = 0.5$. We then estimate a subset of parameters directly from the earnings data, see Tables 9 and 10 in Appendix A for details. This includes the autocorrelation ρ_s of normal labor productivity risk, the initial distribution ω^s_{low}, and the probabilities $\pi^s_{low,0}$ and $\pi^s_{low,1}$ of the low labor productivity shock process.

This leaves us with a total of 13 parameters that need to be calibrated:

1. the 10 parameters $b_{i,s}$ and $j_{M,s}$ of the polynomials in (15) for high school and college educated workers;
2. the innovation variances $\sigma^2_{\varepsilon,s}$ of the normal labor productivity processes for each education level;
3. the labor productivity η_0 of low productivity workers.

We calibrate these parameters within our simulation model such that the model-implied statistics for labor earnings match their empirical counterparts. In particular, we target the following statistics:

1. the results of an age fixed-effects regression for labor earnings, see Figure 4 for a comparison between empirical and model implied life-cycle earnings;
2. the variance of normal labor earnings in Table 9 in Appendix A;
3. average labor earnings of low productivity individuals as shown in the right panel of Figure 2.

Figure 4: Empirical and model implied average life-cycle earnings profiles

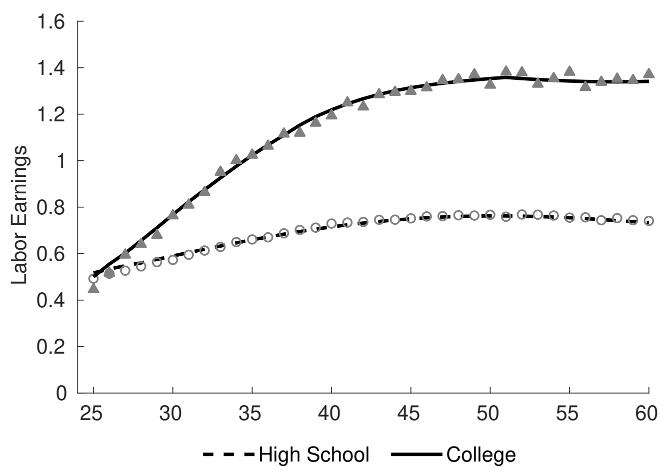


Table 1 summarizes the parameters of labor productivity profiles and risk. More details on the calibration process as well as the formulation of the productivity process in model terms can be found in Appendices A and D.

5.4 Government Policies

We set the pension contribution rate at $\tau_p = 0.187$, the current statutory rate of the German pension system in 2017. In equilibrium, our choice of τ_p results in a

Table 1: Parameter values of labor productivity profiles and risk

	High School $s = 0$	College $s = 1$
Normal labor productivity		
Intercept $b_{0,s}$	−2.0732	−6.4829
Linear age term $b_{1,s}$	0.6238	3.6932
Quadratic age term $b_{2,s}$	−0.0595	−0.7130
Cubic age term $b_{3,s}$	0.0000	0.0467
Stagnation threshold $j_{M,s}$	∞	51
Autocorrelation ρ_s	0.9881	0.9900
Innovation Variance $\sigma^2_{\varepsilon,s}$	0.0045	0.0042
Low labor productivity		
Productivity level $\exp(\eta_0)$	0.0675	0.0675
Initial share of low productivity earners ω^s_{low}	0.2022	0.8005
Probability to transition to low productivity $\pi^s_{low,0}$	0.0064	0.0052
Probability to stay low productivity earner $\pi^s_{low,1}$	0.8374	0.7282

value of $\kappa = 0.455$, the gross replacement rate of the system, which is close to the gross standard replacement rate of 48.3 percent in Germany in 2017, see Deutsche Rentenversicherung Bund (2020).

In our initial economy, we fix government consumption at 19 percent of GDP. We employ the statutory German progressive income tax code, see Appendix D for details. We thereby account for the fact that a substantial number of households consist of married couples, who enjoy a tax advantage (income splitting). Finally, we set the consumption tax rate at $\tau_c = 0.207$ to balance the fiscal budget. Table 2 summarizes the parameters of our model.

6 Simulation Results

In this section, we present simulation results from our quantitative model. We start by showing the central features of our initial equilibrium economy. We then turn to counterfactual policy simulations, in which we introduce progressive components into the pension formula.

6.1 The Initial Equilibrium

Table 3 summarizes central macroeconomic aggregates of our initial equilibrium economy with a proportional pension system as outlined in Section 4.4 and compares it to data from the German economy in 2017. We calibrated the discount

Table 2: Summary of model parameters

Exogenous parameters	Value	Endogenous Parameter	Value
Share college educated ϕ_{Col}	0.237	Depreciation rate δ	0.070
Share unstable careers ϕ_m	0.500	Technology level Ω	0.923
Population growth rate n	0.000	Disutility of labor hours ν_{HS}	46.55
Retirement age	64	Disutility of labor hours ν_{Col}	32.80
Pension contribution rate τ_p	0.187	Mean disutility empl. $\mu_{\xi,HS}$	1.013
International interest rate $\bar{r}$	0.030	Mean disutility empl. $\mu_{\xi,Col}$	0.590
Capital share in production α	0.300	Var. disutility empl. σ_ξ^2	0.138
Intert. elasticity of substitution σ	0.800	Discount factor β	0.984
Frisch elasticity of labor supply χ	0.400	Consumption tax rate τ_c	0.207
Expected utility curvature γ	9.286	Replacement rate κ	0.455

factor such that private savings cover total demand by firms and the government. In reality, private savings are somewhat higher than capital plus public debt. However, a substantial part of these assets come from the top 1 percent wealth holders, a particular group that we do not include in our model. As a result, the German economy holds net foreign assets worth about 45 percent of GDP.

Table 3: Macroeconomic aggregates

Variable	Value (HS/Col)	Data 2017
Private Assets	360.00	433.09
Capital Stock	300.00	305.24
Public Debt	60.00	64.60
Net Foreign Assets	0.00	44.25
Private Consumption	60.00	52.11
Government Consumption	19.00	19.84
Investment	21.00	20.96
Trade Balance	0.00	7.09
Labor Tax Revenue	8.38	8.35
Consumption Tax Revenue	12.42	8.74
Average Work Week of Employed 25-54 (in hrs)	38.2/40.1	38.1/40.0
Employment-to-Population Ratio 25-54 (in %)	84.4/95.1	84.4/95.1

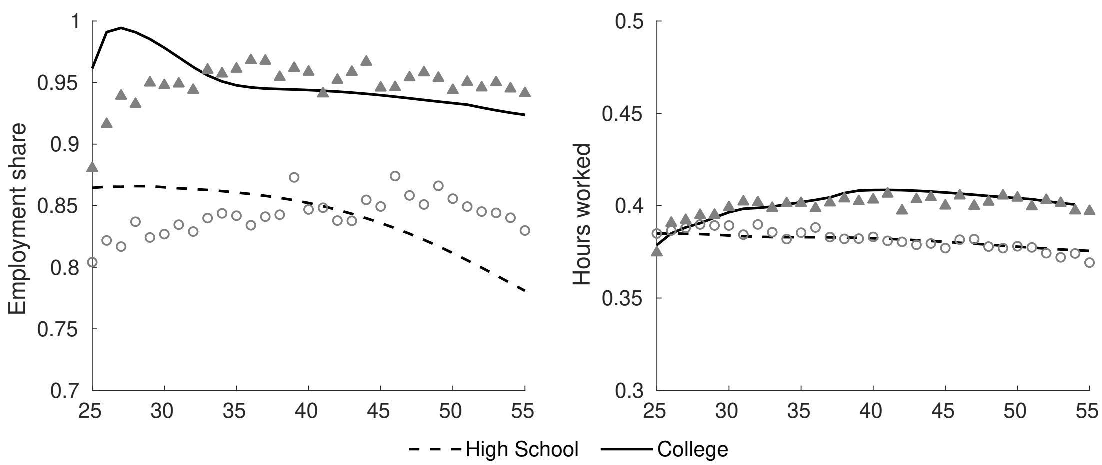
Variables in percent of GDP if not indicated otherwise.

Data sources: PA: Alvaredo et al. (2022), CS: German Statistical Office (2020), PD, NFA: Deutsche Bundesbank (2022), PC, GC, I, TB: German Statistical Office (2020), LTR, CTR, AWW, EtP: RDC (2017)

On the goods market, government consumption and investment almost perfectly match their empirical counterparts. The trade balance in our model is zero, like the net foreign asset position, which implies private consumption to be higher than in the data.[15] Labor tax revenue is close to that of the German economy. Consumption tax revenue is somewhat higher, as we ignore other taxes like capital income or corporate taxes. The average work week of prime aged workers is equal to 38.2 hours for high school and 40.1 hours for college educated workers, just like in data from the German Microcensus (RDC, 2017). The employment-to-population ratio is at 84.4 and 95.1 percent, respectively.

The left panel of Figure 5 compares the labor force participation profiles of high-school (dashed/circles) and college (solid/triangles) workers with their empirical counterparts derived from RDC (2017). The right panel shows life-cycle labor hours by education level. Overall, our model fits the data decently. Yet, as

Figure 5: Labor force participation and hours over the life-cycle



households start their life with zero assets, the employment share is somewhat too high early in life. As households become older and have accumulated some wealth, they successively withdraw from the labor force. Note that the life-cycle labor productivity profile of high school workers is much flatter than that of college graduates, see Figure 4. As a result, labor force participation of the former drops faster than that of the latter. The model-implied labor hours profiles, on the other hand, match the data almost perfectly.

[15]Note that Germany has both a positive trade balance and a positive net foreign asset position. In a long-run equilibrium, this is impossible to achieve without a permanently positive balance of payments. Hence, we decided to strike a balance by having both the net foreign asset position and the trade balance equal to zero.

6.2 The Thought Experiment

We present results from counterfactual policy analyses arising from the introduction of either an employment-linked progressive pension system (ELS) or an Earned Income Pension Credit (EIPC). For our analysis, we selected a medium-range progressivity parameter of $\lambda = 0.5$. This means that 50% of pension payments are proportional to earnings, while the other 50% are subject to redistribution. We conduct simulations for the EIPC with bend points $b \in \{0.2, 0.4\}$ to investigate how much of the efficiency gains inherent in a genuinely employment-linked pension system can be recovered.

To ensure comparability between simulations, we use the same set of structural parameters, but fix per-capita government consumption over time. We assume that the contribution rate of the pension system remains at the initial equilibrium level. In doing so, we ensure that the size of the pension system remains constant for all reforms relative to total labor hours. We use the replacement rate κ to balance the pension budget.[16]

We calculate full transition paths. Starting from an initial long-run equilibrium (indicated by $t = 0$), we assume that the economy is surprised by the reform of the pension formula and therefore enters a transition path at date $t = 1$. It then converges towards a new long-run equilibrium. We allow the government to smooth the benefits and costs of the pension reform over time. To this end, we let the consumption tax rate balance the *intertemporal budget* of the government. The balancing consumption tax rate τ_c can be calculated from

$$\tau_c \cdot \sum_{t=1}^{\infty} R_t C_t + \sum_{t=1}^{\infty} R_t \int T_t\Big(\cdot\Big)\, d\Phi_t = \sum_{t=1}^{\infty} R_t G_t \quad \text{with} \quad R_t = \left[\frac{1+n}{1+\bar{r}}\right]^t .$$

We choose the same approach to calculate a replacement rate κ that balances the intertemporal budget of the pension system. All instantaneous budget imbalances are financed by issuing or repaying public debt.

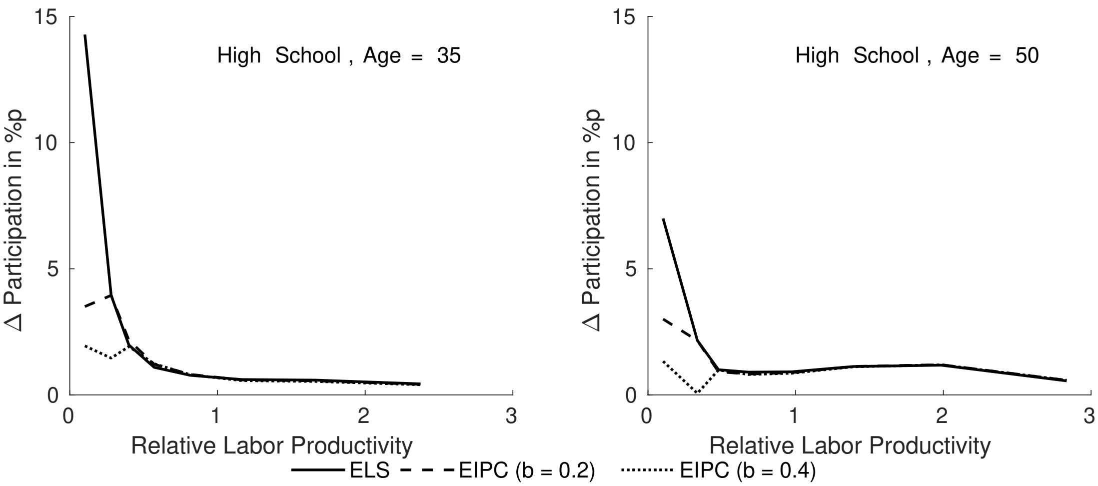
6.3 Labor Supply Effects of Pension Progressivity

Before studying full transitional dynamics, we illustrate some of the long-run effects of progressive pensions that are important for understanding welfare and efficiency effects. Figure 6 shows the long-run employment effects induced by the progressive pension reforms. The horizontal axis denotes an agent's labor productivity relative to the average labor productivity of the working-age population. On the vertical axis, we plot the change in employment between the initial proportional system and the new progressive pension system in percentage points. We

[16] Note that alternatively, we could fix total expenditure of the pension system at the initial equilibrium level. This is, however, counterfactual to the nature of a pay-as-you-go system. With fixed total expenditure, an increase in labor force participation or labor hours would lead to a decline in per capita pension payments and therefore lead to a cut in pension benefits which would counteract the positive effects of our pension reforms.

show the employment effects for 35- and 50-year old high-school workers. Results for the college-educated workforce are qualitatively identical and can be found in Figure 17 in Appendix E. Employment changes are evaluated at the average distribution of wealth and pension claims of an agent in a respective age and education bin.

Figure 6: Employment changes and labor productivity

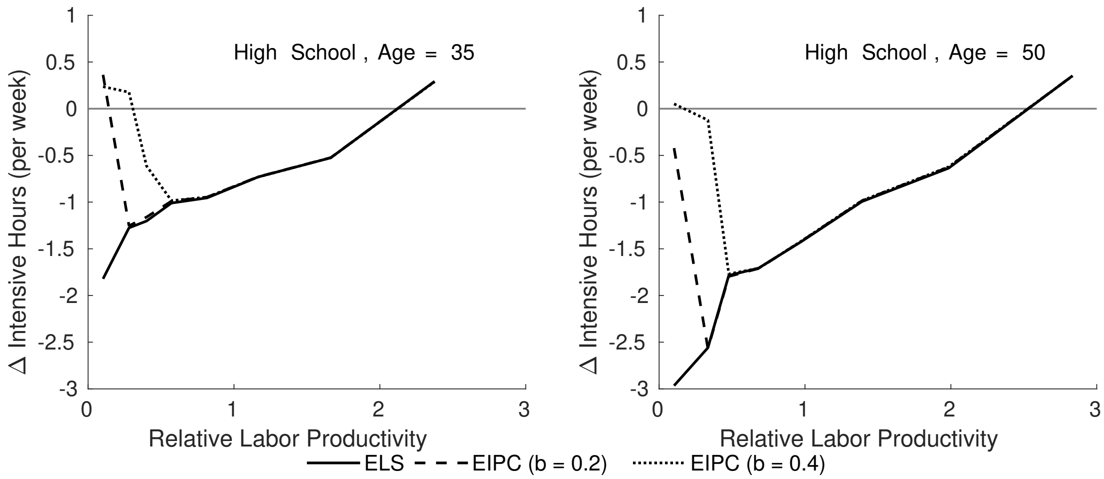


We first focus on the solid lines, which represent the employment effects that come with the introduction of an ELS. Regardless of age, all households experience an increase in labor force participation. The effects are most pronounced for the productivity poor, as they experience the highest implicit employment subsidy, see (4). A rising labor force participation of households with high productivity, on the other hand, is the result of a negative income effect stemming from increased pension progressivity. At young ages, where individuals do not have a lot of wealth, the employment effect is quite high for the productivity poor. It fades out somewhat for older workers, as individual wealth increases. Yet, employment of the productivity-poorest 50-year-olds still increases by a remarkable 7.0 percentage points.

The dashed and the dotted lines indicate the employment effects of an EIPC with bend points $b = 0.2$ and $b = 0.4$, respectively. As we discussed in Section 3, under such purely earnings-related pension systems, the full employment subsidy only unfolds for individuals with relative earnings equal to or greater than the threshold level b. Hence, it is not surprising that the employment effects are almost identical to those of the ELS for individuals with a higher labor productivity. For the productivity poorest individuals, the employment subsidy increases with earnings. Consequently, these systems are less effective in stimulating employment at the lowest end of the earnings distribution. Yet, we still see an employment effect equal to about one third of the ELS for $b = 0.2$.

Figure 7 shows the intensive margin labor supply response to increased pension progressivity.[17] The structure of this figure is the same as the previous one, though

Figure 7: Intensive margin labor supply changes and labor productivity



on the vertical axis we show the change in intensive margin labor hours of employed individuals. For the ELS, the picture is almost inverse to the previous one. Weakening the link between accumulated pension claims and individual earnings leads to a higher implicit tax rate of the pension system, see (4). Hence, increased pension progressivity comes with negative labor supply incentives at the intensive margin, and especially so for earnings poorer households. The negative incentive effect, however, only kicks in for individuals with labor earnings below the contribution ceiling of $2\bar{y}$. Once a household's income is greater than this ceiling – which happens if labor productivity is large – any additional Euro of income earned is not subject to the payroll tax anymore. Therefore, the negative intensive labor supply effect fades out with productivity. For the richest households, there is even a slight increase in hours, which stems from the negative income effect of higher pension progressivity.

The picture is again different for the EIPC. As we noted in Section 3, the EIPC is less effective at stimulating employment at the lower end of the productivity distribution, since the implicit employment subsidy increases with earnings up to the bend point b. However, this increasing subsidy does stimulate intensive margin labor supply for low-earning individuals. This can be observed directly in Figure 7.

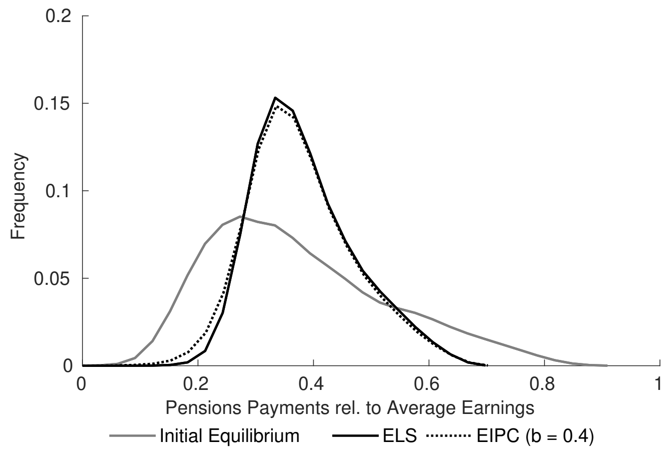
6.4 Progressivity and the Distribution of Pension Claims

Increased pension progressivity not only comes with labor supply effects, it also alters the distribution of pension claims a household accumulates over her working

[17]Results for the college educated can be found in Figure 18 in Appendix E.

life. Figure 8 shows the distribution of pension payments relative to average labor earnings at the retirement age j_R under different pension systems.

Figure 8: Distribution of pension claims



The gray line displays the distribution of pension payments in the initial equilibrium. As pension claims are perfectly earnings related, this distribution is closely linked to the lifetime earnings distribution of households. Recall that the replacement rate is $\kappa = 0.455$ in the initial equilibrium. However, the mode of the pension payment distribution is somewhat lower at around 0.25. This is owing to potential interruptions in the individual's employment history and the fact that the accumulation of pension claims is capped at twice the average earnings.

The distribution of pension claims is much more concentrated with an ELS, as shown by the black line in Figure 8. In fact, the mass of individuals with a pension of less than 20 percent of average earnings shrinks to almost zero. The dotted line finally indicates the distribution of pension payments under an EIPC with bend point $b = 0.4$.[18] The system is slightly less efficient in mitigating inequality in pension payments compared to the ELS, but the differences are small.

6.5 A Macroeconomic Evaluation

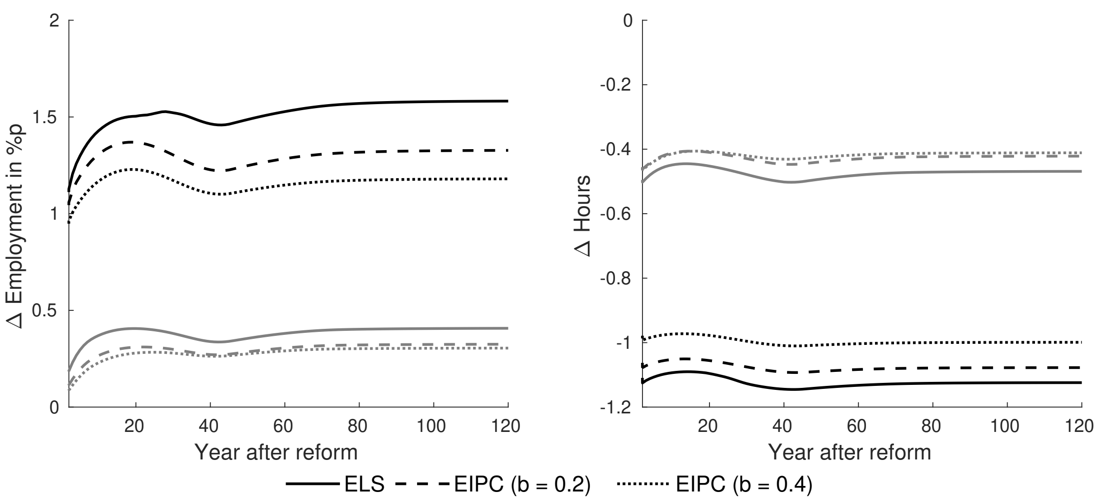
In this section, we evaluate the macroeconomic consequences of progressive pension reforms along a full transition path. Recall that we indicate the initial long-run equilibrium by $t = 0$. The pension reform comes at a surprise in period $t = 1$ and induces a transition path to a new long-run equilibrium. We assume that the reform leaves the existing pension claims of individuals that already lived in the initial equilibrium untouched. Only new pension claims that are accumulate after

[18] The distribution of a pension system with bend point $b = 0.2$ would obviously lie in the middle. But since the differences are small anyway, we don't show the results for such a system in this graph.

the reform are due to the new ELS or EIPC pension formula. Computational details are discussed in Appendix C.4.

Figure 9 shows the employment and intensive labor supply effects for high school (black) and college (gray) educated workers. Overall, the effects are quite evenly distributed over time. Employment of college educated workers rises by about 0.40 percentage points on average. The employment effect is much larger for high school workers, with a peak effect of 1.6 percentage points. The adverse impact on working hours can be seen in the right panel of Figure 9. The pattern is quite similar, with a smaller decline in hours for those with a college degree and larger effects for high school workers. Additionally, the impact on both extensive and intensive margin labor supply is most pronounced under the ELS. When looking at the EIPC, we find that a higher bend point reduces both the overall positive employment effects as well as the negative intensive margin distortions.

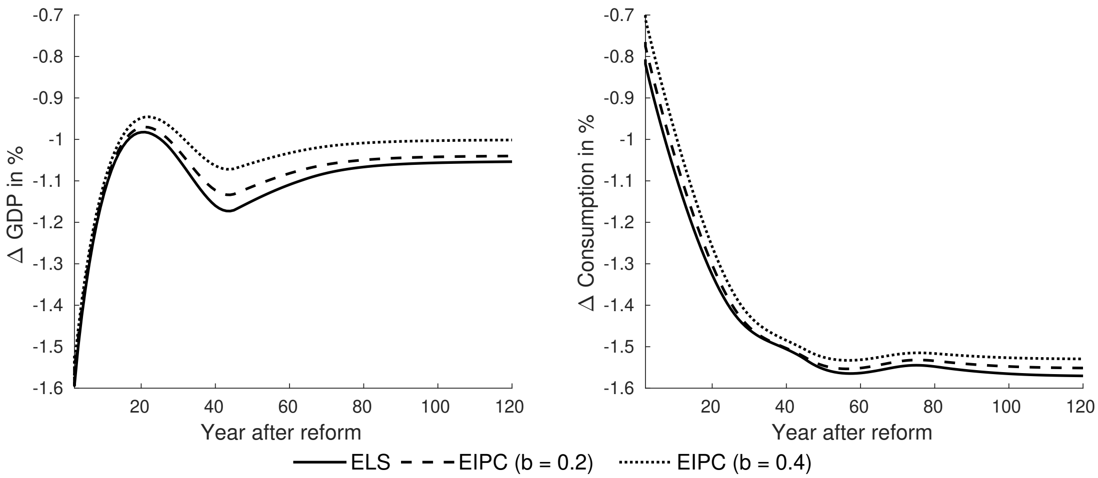
Figure 9: Aggregate Labor Supply Effects



Combining the extensive and intensive margin labor supply responses, we find that overall labor input declines. This can be seen from the left panel of Figure 10, which shows the evolution of GDP over time. Recall that we consider a small open economy setting. Hence, aggregate capital, labor input and GDP all move synchronously. The overall decline in labor input is not surprising in light of the fact that the employment effect is most pronounced for low-productivity workers, but the negative intensive margin distortions are distributed more evenly across productivity types. The drop in labor hours and therefore GDP is the strongest in the period directly after the reform ($t = 1$). It is mitigated somewhat by a decline in aggregate savings over time, see below.

The right panel of Figure 10 shows the consequences of our pension reforms for aggregate consumption. Not surprisingly, the decline in GDP causes aggregate consumption to drop immediately as we introduce progressive pensions into the

Figure 10: GDP and Aggregate Consumption



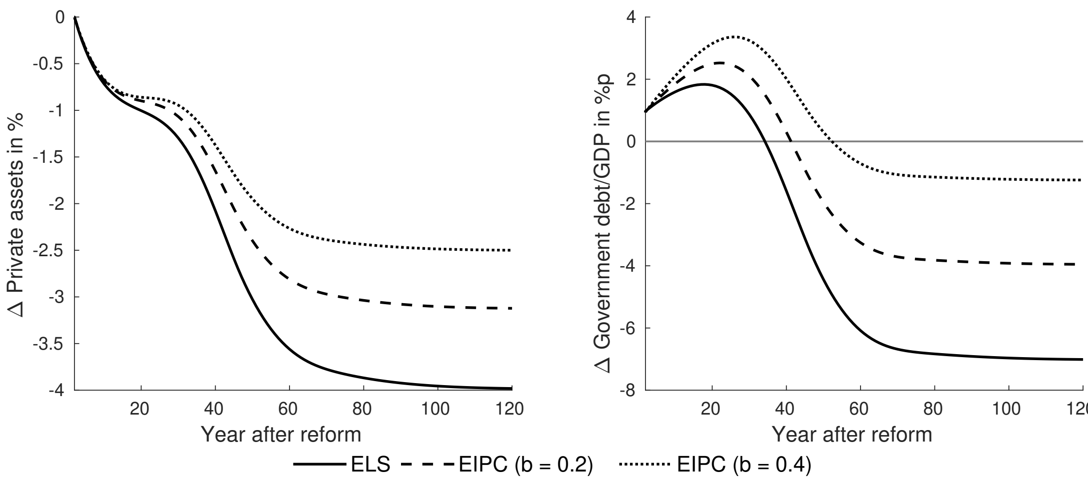
economy. In the short-run, individuals can still live on their private assets, which damps the immediate consumption response. As private assets melt down over time, however, the consumption response becomes more pronounced. In the long-run, aggregate consumption declines by approximately 1.6 percentage points.

The left panel of Figure 11 illustrates the gradual meltdown of private assets along the transition. As opposed to the previous graphs, we can see a substantial difference in the long-run effects of different pension system designs. The decline in private assets is most pronounced for the ELS, while it is much more moderate for an EIPC with bend point $b = 0.4$. The differences can be explained by the risk properties inherent in the different systems. The introduction of an ELS has the largest employment effect for productivity-poor individuals. As labor income in the poorest productivity states rises, the need for precautionary savings to insure a short-fall in labor earnings upon adverse productivity shocks declines. This impacts on private asset accumulation. The EIPC system with bend point $b = 0.4$ is much less successful in stimulating employment at the bottom end of the productivity distribution. Hence, households rely more on precautionary savings which damps the asset meltdown over time.

Finally, the right panel of Figure 11 shows the evolution of public debt over time. In the short-run, a strong decline in labor hours causes both a short-fall in tax revenue and pension contributions. To cope with the resulting budget imbalance, the government has to issue additional debt. As labor supply stabilizes in the medium- and long-run, however, the government is even able to reduce its debt level by about 7 percentage points. This comes with a relief for future generations.

Summing up, our simulation results indicate that the macroeconomic consequences of the proposed pension reforms are generally negative. The stimulation of labor force participation at the lower end of the productivity distribution somewhat mitigates the burden from larger labor supply distortions at the intensive mar-

Figure 11: Capital and Public Debt



gin. Yet, GDP and aggregate consumption still decline by 1 and 1.6 percent, respectively. However, we also see that the introduction of progressive pensions alters the risk properties of labor earnings risk, which affects aggregate savings. It also mitigates differences in pension income and therefore reduces consumption inequality especially at old age. To jointly evaluate the negative level and positive distributional consequences, we next take a look at aggregate welfare and economic efficiency.

6.6 Welfare Analysis

We now evaluate the welfare and efficiency effects of progressive pensions. Our preferred measure of household welfare is ex-ante expected life-time utility EV_t before any information about the household's education level or labor productivity has been revealed. We calculate ex-ante utility for any generation that is affected by a progressive pension reform, i.e., the initial cross-section of households at the reform date $t = 1$ as well as all new-born generations along the transition path. We distinguish affected generations by their birth date $t \in \{-J+1, \ldots, 0, 1, \ldots, \infty\}$. We compare the utility measures of these generations to the utility level $\overline{EV}$ of a generation that was born and has lived entirely through the initial equilibrium with a proportional pension system. To give the welfare numbers a meaningful interpretation, we calculate the corresponding consumption equivalent variation CEV_t. The consumption equivalent variation indicates by how many percent we would have to increase or decrease the consumption level of households at each age and each potential state in the initial equilibrium in order to make them as well off as in a reform scenario with progressive pensions. A positive value for CEV_t indicates that a progressive pension system increases welfare of a particular cohort t, and that households of this cohort would be willing to pay a positive amount of resources in order to live in a world with progressive pensions.

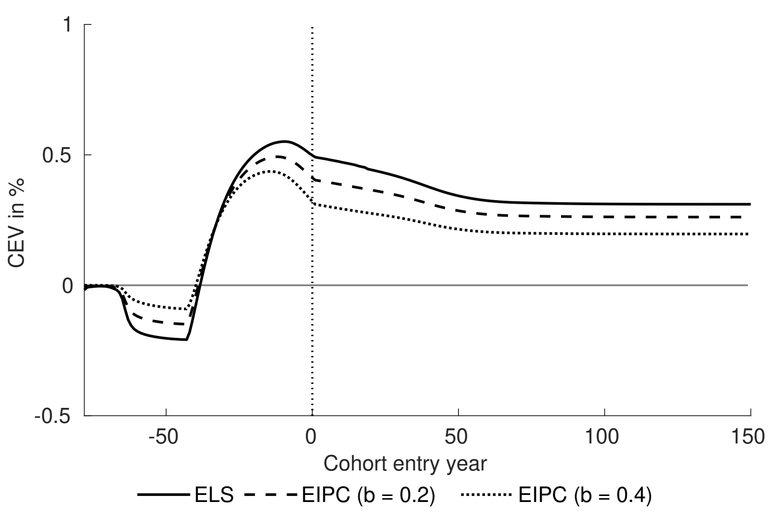
The welfare effects of pension reforms can vary a lot across different cohorts because of intergenerational redistribution. To derive a meaningful measure of the economic efficiency effect of pension reforms, we have to find a way to aggregate different welfare changes across cohorts to one aggregate efficiency measure. Our method follows Fehr and Kindermann (2015) and Kindermann and Krueger (2022). We calculate the monetary transfer Ψ_t that each affected generation would have to pay in order to be indifferent between living along the reform path and in the initial equilibrium. We then derive the present value of all of these transfers, which gives us a wealth-based measure W of economic efficiency. To turn this into a consumption based measure C, we convert the wealth-based measure into an annuity that pays out a constant stream along the transition path and in the new long-run equilibrium:

$$W = \sum_{t=-J+1}^{\infty} \left[\frac{1+n}{1+\bar{r}}\right]^t \Psi_t \quad \text{and} \quad C = W \times \left[\sum_{t=1}^{\infty} \left[\frac{1+n}{1+\bar{r}}\right]^t\right]^{-1}.$$

We express the resulting time-invariant welfare gain C in percent of the initial equilibrium consumption level C_0.

Figure 12 shows the welfare effects of our different reform scenarios. The vertical dotted line separates cohorts that were already alive in the initial equilibrium and that were surprised by the reform at some date in their life cycle from those cohorts born along the transition path $t > 0$. All three reforms generally come

Figure 12: Welfare effects along the transition path



with positive welfare effects for current working cohorts as well as all newborns. Only generations that were already retired at the time of the pension reform lose slightly from an increase in consumption taxes. Young workers as well as newborns at the reform date experience the highest welfare gains, in the order of 0.5 percent of life-time consumption. As the transition proceeds and private assets melt down, welfare again declines but remains positive even in the long-run.

Welfare effects differ across reform scenarios. The introduction of an ELS results in the highest welfare gains for current workers and future generations, but current retirees experience the highest welfare losses. The effects are more moderate for an EIPC system with bend points $b = 0.2$ and $b = 0.4$. There are two possible explanations for this. First, as we saw in Figure 11, the asset melt-down is least pronounced for $b = 0.4$, which suggests a lower degree of intergenerational redistribution under this scenario. Second, a higher value for the bend-point b reduces the effectiveness of the progressive pension system in stimulating employment and leads to fewer risk-sharing opportunities, resulting in lower economic efficiency.

To disentangle intergenerational redistribution from economic efficiency, the first row of Table 4 shows the aggregate efficiency effects of the different progressive pension systems. The ELS turns out to be the most efficient system to implement. It generates a permanent increase in welfare worth 0.73 percent of aggregate consumption. This is not surprising in light of the fact that this system operates under the assumption that the government can condition pension payments on the individual employment decision. If the government is bound by informational constraints and can only condition pension payments on income, aggregate efficiency has to deteriorate. However, even under the "second best" policies with bend points $b = 0.2$ and $b = 0.4$, the government can still recover 90% and 78% of the original efficiency effect, respectively.

Table 4: Welfare effects of increased pension progressivity

		EIPC	
Variable	ELS	b = 0.2	b = 0.4
Change in aggregate efficiency	0.73	0.66	0.57
Change in ex-ante long-run welfare	0.31	0.26	0.20
Long-run Welfare by Permanent Types			
– for high school with unstable career	0.31	0.22	0.14
– for high school with stable career	0.52	0.57	0.52
– for college with unstable career	−0.14	−0.22	−0.26
– for college with stable career	−0.19	−0.14	−0.18
Long-run Welfare Decomposition			
– average utility of consumption	−0.26	−0.26	−0.25
– average disutility of labor	−0.15	−0.09	−0.07
– risk sharing possibilities	0.71	0.61	0.51

Table reports CEV over initial equilibrium in percent.

The second panel of Table 4 shows the welfare consequences for the four different permanent types $(s, m) \in \{0, 1\} \times \{0, 1\}$. The major beneficiaries of progressive pensions are high school workers, as they tend to be the recipients of employment subsidies. The college educated lose because of higher labor supply distortions and a reduction in their pension benefits. Within the group of high-school workers, it is those with a stable career path that experience the highest welfare gains. The welfare gains of workers with an unstable career path are only half the size. The reason is that workers with a stable career tend to have a higher labor market attachment. As such, they can enjoy the full benefits of redistribution through the progressive pension without incurring any major extra cost. Those with an unstable career are motivated to stay attached to the labor force even when they experience a low labor productivity shock. They do so to enjoy the employment subsidy embedded in the progressive pension. However, labor market participation comes at a higher utility participation cost, which damps the welfare benefits of such households. This effect becomes even more pronounced under an EIPC system with bend point $b = 0.2$ or $b = 0.4$ that reduces the employment subsidy for households with very low labor earnings.

Household welfare gains can stem from (i) increases in average life-cycle consumption, (ii) a decline in the average disutility of labor or (iii) increased risk-sharing possibilities that lead to a decline in the variance of consumption and/or labor hours. In the last panel of Table 4, we decompose the long-run welfare gain into effects coming from exactly these three components. The welfare gains of progressive pensions are entirely due to improved risk-sharing possibilities. As we already discussed before, average life-cycle consumption declines and the utility costs of labor force participation increase. Both effects reduce long-run welfare. However, a declining variance of consumption and labor hours within age-groups overcompensates these negative effects and leads to an overall welfare increase.

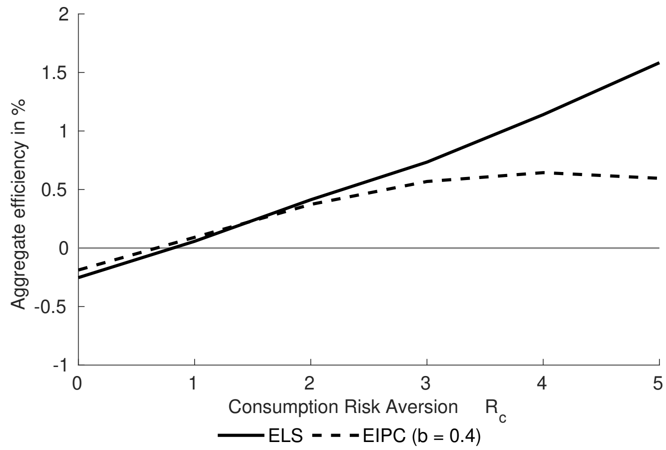
6.7 Sensitivity Analysis

This section provides sensitivity checks with respect to two central elements of our quantitative model: individual risk aversion and the structure of the labor market.

6.7.1 Risk Aversion

Figure 13 shows how the aggregate efficiency effects of progressive pensions depend on household risk aversion. In our preferred calibration, we use a consumption risk aversion of $R_c = 3$. A higher risk aversion leads to additional welfare gains from increased social insurance. The effect is quite strong for the ELS, as this system comes with the best risk-sharing possibilities for households. With a risk aversion of $R_c = 5$, aggregate efficiency increases to a remarkable 1.6 percent. An EIPC system with bend point $b = 0.4$ is much less successful in insuring labor productivity risk, especially for households with very low productivity shocks.

Figure 13: Sensitivity analysis



This is directly reflected in the aggregate efficiency numbers.

In addition to showing that a higher risk aversion raises the size of efficiency gains from redistribution and social insurance, there is another point of interest in Figure 13. When we choose a value of $\gamma = -3.57$, individual risk aversion drops to zero. In this case, the gains from redistribution are absent and the efficiency effects from progressive pensions emerge solely from labor supply distortions. Aggregate labor supply distortions are most pronounced for the ELS. While this system sets the highest employment incentives, it also comes with a positive implicit tax rate for all working individuals. An earnings based system with $b = 0.4$, on the other hand, sets additional positive labor supply incentives at the intensive margin for the productivity poor, which limits aggregate efficiency losses.

6.7.2 Structure of the labor market

Table 5 displays the aggregate efficiency consequences of introducing an earnings-based progressive pension with bend point $b = 0.4$ for different assumptions about the structure of the labor market. In our benchmark scenario, we assumed that 50 percent of the population is exposed to low productivity shocks, while the other half faces stable career paths. To check the importance of this assumption, we let the whole population be exposed to low productivity shocks $\phi_m = 1$ and recalculate the respective shock process to guarantee consistency with the data. As the results in Table 4 reveal, the consequences for both aggregate efficiency and long-run welfare are only minor.

In the last row of the table, we report the results from simulations in which we assume away participation costs. Without participation costs, households are always employed, regardless of their productivity shock. Hence, setting extensive margin employment effects can not improve economic efficiency by definition.

Table 5: Sensitivity Analysis

	EIPC (b = 0.4)	
	agg. efficiency	long-run welfare
Benchmark simulation	0.57	0.20
Career Stability: $\phi_m = 1$	0.49	0.18
No extensive margin costs ($\xi = 0$)	0.44	0.17

Table reports CEV over initial equilibrium in percent.

Consequently, the aggregate efficiency effect of introducing progressive pensions shrinks by about 25 percent as compared to our benchmark scenario.

7 Conclusion

This paper studies reforms of the pension system aimed at increasing progressivity. We quantify the effects of progressive pension systems in a stochastic overlapping generations model with labor supply responses at the intensive and at the extensive margin. Our focus is on the extensive margin labor supply reactions to progressive pension reforms. A pension system with an employment-linked component increases labor force participation and hence mitigates negative labor supply effects at the intensive margin. Aggregating the resulting welfare effects along the transitional path and in a new long-run equilibrium shows that such a reform is efficiency improving. We address potential feasibility concerns and propose a second reform, the Earned Income Pension Credit, which redistributes pension claims solely based on earnings. Our simulation results indicate that a substantial share of the efficiency gains from the employment-linked system can be restored with the EIPC.

Although our model covers some important dimensions of individual heterogeneity, it is silent about the effects for women. Women represent the largest risk group when it comes to old-age poverty, and especially so when they are single mothers. Moreover, they face considerable income risk over the life-cycle and exhibit higher elasticities at both margins of response. Hence, we expect that a large fraction of women would benefit from the proposed pension reforms. Our welfare results can thus be regarded as a conservative estimate.

8 Bibliography

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics 109*(3), 659–684.

Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez (2022). World inequality database.

Athreya, K. B., D. Reilly, and N. B. Simpson (2010). Earned income tax credit recipients: income, marginal tax rates, wealth, and credit constraints. *Economic Quarterly 96*(3Q), 229–258.

Auerbach, A. and L. Kotlikoff (1987). *Dynamic Fiscal Policy*. Boston: Cambridge University Press.

Bagchi, S. (2019). Differential mortality and the progressivity of social security. *Journal of Public Economics 177*, 104–044.

Bartels, C. and N. Pestel (2016). Short- and long-term participation tax rates and their impact on labor supply. *International Tax and Public Finance 23*(6), 1126–1159.

Bayer, C. and F. Juessen (2012). The life cycle and the business cycle of wage risk: A cross-country-comparison. *Economics Letters 117*(3), 831–833.

Bewley, T. F. (1986). Stationary monetary equilibrium with a continuum of independently fluctuating consumers. In W. Hildenbrand and A. Mas-Collel (Eds.), *Contributions to Mathematical Economics in Honor of Gerard Debreu*, pp. 79–102. North Holland, Amsterdam.

Bierbrauer, F., P. C. Boyer, and E. Hansen (2023). Pareto-improving tax reforms and the earned income tax credit. *forthcoming: Econometrica.*

Brendler, P. (2021). Rising earnings inequality and optimal social security. Working Paper, University of Bonn.

Breyer, F. and S. Hupfeld (2010). On the fairness of early-retirement provisions. *German Economic Review 11*(1), 60–77.

Busch, C. and A. Ludwig (2020). Higher-order income risk over the business cycle. Barcelona GSE Working Paper No. 1159.

Casanova, M. (2013). Revisiting the hump-shaped wage profile. Working Paper, University of California, Los Angeles.

Castaneda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003). Accounting for the u.s. earnings and wealth inequality. *Journal of Political Economy 111*(4), 818–857.

Chan, M. K. (2013). A dynamic model of welfare reform. *Econometrica 81*(3), 941–1001.

Conesa, J.-C., S. Kitao, and D. Krueger (2009). Taxing capital: Not a bad idea after all! *American Economic Review 99*(1), 25–48.

Conesa, J.-C. and D. Krueger (2006). On the optimal progressivity of the income tax code. *Journal of Monetary Economics 53*(7), 1425–1450.

Coronado, J. L., D. Fullerton, and T. Glass (2011). The progressivity of social security. *The BE Journal of Economic Analysis & Policy 11*(1), Article 70.

de Nardi, C., G. Fella, and G. Paz-Pardo (2020). Nonlinear household earnings dynamics, self-insurance, and welfare. *Journal of the European Economics Association 18*(2), 890–926.

Deutsche Bundesbank (2022). Zeitreihen-Datenbanken. Online Database available at https://www.bundesbank.de/de/statistiken/zeitreihen-datenbanken.

Deutsche Rentenversicherung Bund (2019). Rentenversicherung in Zahlen 2019. Berlin.

Deutsche Rentenversicherung Bund (2020). Rentenversicherung in Zeitreihen 2020. Available at https://statistik-rente.de/drv/, Berlin.

Domeij, D. and J. Heathcote (2004). On the distributional effects of reducing capital taxes. *International Economic Review 45*, 523–554.

Du, K.-L. and M. N. S. Swamy (2016). *Search and optimization by metaheuristics: Techniques and Algorithms Inspired by Nature*. Basel: Birkhauser.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*, 937–969.

Fehr, H. and C. Habermann (2008). Risk sharing and efficiency implications of progressive pension arrangements. *Scandinavian Journal of Economics 110*(2), 419–443.

Fehr, H., M. Kallweit, and F. Kindermann (2013). Should pensions be progressive? *European Economic Review 91*, 30–56.

Fehr, H. and F. Kindermann (2015). Taxing capital along the transition – not a bad idea after all? *Journal of Economic Dynamics and Control 51*, 64–77.

Fehr, H. and F. Kindermann (2018). *ntroduction to Computational Economics using Fortran*. Oxford: Oxford University Press.

French, E., A. Lindner, C. O'Dea, and T. Zawisza (2021). Labor supply and the pension contribution-benefit link. Working Paper, Yale University.

German Statistical Office (2020). Genesis database. Online Database available at https://www-genesis.destatis.de/.

Goda, G. S., J. B. Shoven, and S. N. Slavov (2011). Differential mortality by income and social security progressivity. In D. A. Wise (Ed.), *Explorations in the Economics of Aging*, pp. 189–208. University of Chicago Press.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015). What do data on millions of u.s. workers reveal about life-cycle earnings dynamics? NBER Working Paper No. 20913.

Haan, P., D. Kemptner, and H. Lüthen (2020). The rising longevity gap by lifetime earnings – distributional implications for the pension system. *Journal of the Economics of Ageing 117*, 100–199.

Hall, R. E. (1982). The importance of lifetime jobs in the us economy. *American Economics Review 72*, 716–724.

Halvorson, E., H. Holter, S. Ozkan, and K. Storesletten (2020). Dissecting idiosyncratic earnings risk. Working Paper.

Hansen, E. (2021). Optimal income taxation with labor supply responses at two margins: When is an earned income tax credit optimal? *Journal of Public Economics 195*, 1043–1065.

Heckman, J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics 1*, 1–58.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control 17*(5), 953–969.

Huggett, M. and G. Ventura (1999). On the distributional effects of social security reform. *Review of Economic Dynamics 2*(3), 498–531.

Human Mortality Database (2020). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.

Imrohoruglu, A. (1989). Cost of business cycles with indivisibilities and liquidity constraints. *Journal of Political Economy 97*(6), 1364–83.

Jones, J. B. and Y. Li (2022). Social security reform with heterogeneous mortality. *forthcoming: Review of Economic Dynamics.*

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

Keane, M. P. and R. Rogerson (2011). Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literatur 50*(2), 464–476.

Kindermann, F. and D. Krueger (2022). High marginal tax rates on the top 1%? lessons from a life cycle model with idiosyncratic income risk. *American Economic Journal: Macroeconomics 14*, 319–366.

Kindermann, F., L. Mayr, and D. Sachs (2020). Inheritance taxation and wealth effects on the labor supply of heirs. *Journal of Public Economics 191*, 104–127.

Kopecky, K. A. and R. M. H. Suen (2010). Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics 13*(3), 701–714.

Kuhn, M. and G. Ploj (2020). Job stability, earnings dynamics, and life-cycle savings. Working Paper.

Lalive, R., A. Magesan, and S. Staubli (2023). How social security reform affects retirement and pension claiming. *forthcoming: American Economic Journal: Economic Policy*.

Ludwig, A. (2007). The gauss–seidel–quasi-newton method: A hybrid algorithm for solving dynamic economic models. *Journal of Economic Dynamics and Control 31*(5), 1610–1632.

Luy, M., C. Wegner-Siegmundt, A. Wiedemann, and J. Spijker (2015). Life expectancy by education, income and occupation in germany: Estimations using the longitudinal survival method. *Comparative Population Studies 40*(4), 399–436.

Manjunath, B. G. and S. Wilhelm (2012). Moments calculation for the doubly truncated multivariate normal density. arXiv eprint: 1206.5387.

Meyer, B. D. (2002, May). Labor supply at the extensive and intensive margins: The eitc, welfare, and hours worked. *American Economic Review 92*(2), 373–379.

Nam, L. (2022). Optimal progressive pension systems in a life-cycle model with heterogeneity in job stability. Working Paper, University of Bonn.

Nishiyama, S. (2019). The joint labor supply decision of married couples and the us social security pension system. *Review of Economic Dynamics 31*, 277–304.

Nishiyama, S. and K. Smetters (2008). The optimal design of social security benefits. Michigan Retirement Research Center Research Paper No. 2008-197.

O'Dea, C. (2018). Insurance, efficiency and the design of public pensions. Working Paper, Yale University.

OECD (2016). Education at a Glance 2016: OECD Indicators. OECD Publishing, Paris.

Ortigueira, S. and N. Siassi (2022). The u.s. tax-transfer system and low-income households: Savings, labor supply, and household formation. *Review of Economic Dynamics 44*, 184–210.

Peterman, W. B. (2016). Reconciling micro and macro estimates of the frisch labor supply elasticity. *Economic Inquiry 54*(1), 100–120.

RDC of the Federal Statistical Office and Statistical Offices of the Federal State (2017). Microcensus 2017, Scientific Use Files. doi=10.21242/12211.2017.00.00.3.1.0, own calculations.

Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. *The Quarterly Journal of Economics 117*(3), 1039–1073.

Swanson, E. T. (2018). Risk aversion, risk premia, and the labor margin with generalized recursive preferences. *Review of Economic Dynamics 28*, 290–321.

Wallenius, J. (2013). Social security and cross-country differences in hours: A general equilibrium analysis. *Journal of Economic Dynamics and Control 37*(12), 2466–2482.

Progressive Pensions as an Incentive for Labor Force Participation

Appendix for Online Publication

Fabian Kindermann and Veronika Püschel

A Datawork

The productivity profiles in this paper are based on administrative data from the German Pension Insurance. In particular we use the 2017 wave of the scientific use-file of the Versichertenkontenstichprobe that contains monthly earnings data of 69,520 insured individuals. This is about 0.18% of the actively insured population.[19] We restrict our attention to the male sample population aged between 25 and 60 of which we have information on the education level. Our measure of monthly labor earnings comprises income from regular work, marginal employment and short-term unemployment (up to one year). We count all other source of pension accumulation (like times of care for children or sickness) as zero earnings months. We sum up monthly earnings observations to construct an annual earnings measure for each individual. This appendix explains the data selection and estimation process in detail.

A.1 The Administrative Dataset

The data set consists of two parts: One provides demographic characteristics such as age, gender and education for the year 2017. The other one records the entire history of an individual's accumulated pension claims and employment status on a monthly basis. The sample covers worker who were born between 1950 and 1987 and who were not permanently retired in 2017. The historical record starts in the year an individual turns 14 and ends when she turns 65. Hence, the maximum length of an employment history is 624 month. Overall, the data set includes more than 28 million worker-month observations for the years 1964 to 2017. As the sample ends in Dezember 2017, individuals who were born in 1953 or later have shorter histories (e.g. 612 month for the 1953 cohort). Those who have never been employed are not represented, as they never were registered with the insurance.

[19]The German pension scheme covered of 38 million actively insured individuals in 2017.

A.1.1 Earnings measurement

Earnings y_{isjt} of an individual i of education s and age j at time t are subject to social security contribution. There is a contribution threshold $y_{max,t}$ and any earnings beyond that value are non-contributory. Contributory earnings hence amount to $\min(y_{isjt},\ y_{max,t})$. They are converted into pension claims y^p_{isjt} by diving them through average earnings $\bar{y}_t$. We account for the fact that pension claims from so-called mini-and midi jobs[20] are subject to a reduced pension contribution rate. Both, the contribution threshold $y_{max,t}$ and average earnings $\bar{y}_t$ are adjusted annually to account for wage growth. The contribution threshold $y_{max,t}$ currently amounts to about twice the average earnings $\bar{y}_t$.[21]

For our analysis, it is most convenient to use pension claims y^p_{isjt} as an earnings measure, as they are stationary over time. In particular, we define

$$y^p_{isjt} = \frac{\min(y_{isjt},\ y_{max,t})}{\bar{y}_t}. \tag{17}$$

Obviously, the data are right-censored at $y_{max,t}$, see also Figure 1.

A.1.2 Data Selection

Although the monthly records start in 1964, we only consider observations for the years 2000 to 2016. This has certain advantages: First, our estimates are based on recent data; second, we avoid structural breaks arising from German reunification and policy-changes in the 1990s and third, different age cohorts are represented in the sample at similar shares in each year (early sample years cover only young individuals). The data-selection process is summarized in Table 6.

We restrict the sample such that it targets workers who are attached to the labor market. We therefore limit our attention to men aged between 25 and 60 who are likely to already have finished education and military service and are not in the process of retiring. We drop all individuals who already received pensions such as disability pensions or early-retirement pensions.

We divide the sample into two educational groups. We adapt the scheme to the International Standard Classification of Education of the UNESCO (ISCED 2011) to allow for international comparison. A person is defined to be college-educated[22] if she is classified ISCED 6 (Bachelor's or equivalent level) or above, excluding ISCED 65 (trade and technical schools, including master craftsman training). She is non-college-educated[23] if she is classified ISCED 5 and below or ISCED 65. We drop those with unknown education status.

[20] In a mini-job, an individual can earn a maximum of EUR 450. Midi-jobs cover earnings from 451 to 850 Euros.

[21] See Section 11 in Deutsche Rentenversicherung Bund (2020) for a full history of reference values.

[22] corresponds to KldB 2010 4-6

[23] corresponds to KldB 2010 1-3

Table 6: Data Selection

	Individuals	Observations
Initial data set (1975 - 2017)	69,520	28,166,952
Initial data set (2000 - 2016)	69,520	14,139,972
- Women	−36,634	−7,451,736
- Ages < 25		−1,014,120
- Ages > 60		−152,976
	32,886	5,521,140
- Ind. that receive pensions	−3,606	−605,208
	29,280	4,915,932
- Ind. with unknown education	−13,677	−2,346,840
	15,603	2,569,092
Annualized data (2000 - 2016)	15,603	214,091
No contributory earnings in 2000 - 2016	−361	−6,137
No contributory earnings in entire year		−18,770
Final data set	15,242	189,184
Non-college education	11,821	149,929
College education	3,421	39,255
Observations on regular workers		181,469
Observations on low earners		7,715

For estimating earnings profiles we use all pension claims y^p_{isjt} that stem from (1) regular-employment, (2) mini-jobs or (3) unemployment benefits (short-term, max. 12 month) according to the variable `SES`. Since individuals are productive when searching for a new job, we consider short-term unemployment as an employment type. Table 7 shows the distribution of employment states across monthly observations. About 13 percent of all observations are on months with no contributory earnings. Such observations emerge when individuals become self-employed or civil servants, when they take care leave, face a longer spell of unemployment or just decide to drop out of the workforce. We code non-contributory months as periods of zero earnings.

To make the data comparable with our simulation model, we have to change the time-dimension of the panel from monthly to annual. We do so by computing the sum of acquired pensions claims for each calendar year. Finally, we drop all sample individuals who had no contributory earnings at all in the period from 2000 to 2016. In addition, we exclude observations with no contributory earnings in an entire calendar year, see Table 6. Our final data set is an unbalanced annual

Table 7: Distribution of Employment States (across monthly observations)

Employment Status	Observations	Percent
Regular employment	2,139,302	83.27
Mini-job	44,113	1.72
Unemployment (short-term)	55,138	2.15
No contributory earnings	330,539	12.86
Total	2,569,092	100.00

panel for the years 2000 to 2016 with 15,242 individuals – of which 22.4 percent are college-educated – and a total of 189,184 observations.

In order to take account of the substantial mass of individuals at the lower end of the earnings distribution, see the discussion in Section 2 and Figure 1, we split the sample into two sub-samples. The first one contains individuals with normal labor earnings and the second one those with extraordinarily low earnings. An individual i is defined as a low earner in year t if she acquires pension claims y^p_{isjt} that corresponds to somebody working full-time for six month at minimum wage. With 250 annual working days, 8 hours of work per day, a minimum wage of 8.50 Euros and an average income of 36,187 Euros in 2016, the threshold below which an individual counts as low earner is

$$\frac{125 \times 8 \times 8.5}{36,187} = 0.23. \tag{18}$$

Within our sample, 95.9% of observations are regular earnings and 4.1% are low earnings. We use observations from regular workers to estimate earnings profiles as shown in the left panel of Figure 14. Earnings estimates for low earners are shown in the right panel of Figure 2.

A.2 Earnings estimates for Regular Workers

In the following, we describe the estimation process for the life-cycle earnings profiles and labor earnings risk of regular workers in detail.

A.2.1 Identifying the top censoring threshold

Our starting point is the data set of regular workers with 181,469 observations as summarized in Table 6. While we fixed the bottom threshold that marks the difference between a regular worker and a low earner at a constant value of 0.23, see equation (18), identifying the top censoring threshold is not as straightforward. Although the German public pension insurance provides an official contribution ceiling $\tilde{y}_{max,t}$ for contributory earnings in every year, see Deutsche Rentenversicherung Bund (2020), we cannot take this value directly. The reason is that the ceiling is applied on a monthly basis while we are working with annual data. Hence, an observation could be subject to censoring, although the observed annual earnings y^p_{isjt} are below the official cut-off value. This is the case if the contribution threshold is reached in some months of the year, but not in others (for instance because of salary changes). In addition, we observe a few outliers where annual pension claims y^p_{isjt} are beyond the corresponding official threshold, which might be due to value adjustments.

To overcome these problems, we use the following strategy to identify a threshold $y_{max,t}$ for every year that allows us to capture most observations that have been top-coded at least in one month:

1. First, we find the value of pension claims $mode_{y,t}$ at the upper end of the distribution where most of the observations pile up and compare it to the official threshold $\frac{\tilde{y}_{max,t}}{\bar{y}_t}$. $mode_{y,t}$ typically is in the order of 0.0002 smaller than $\frac{\tilde{y}_{max,t}}{\bar{y}_t}$, which corresponds to about 7 Euros in 2016 compared to an average income of 36,000 Euros.

2. We then define our censoring threshold as

$$\frac{y_{max,t}}{\bar{y}_t} = mode_{y,t} - 0.0003.$$

   This guarantees that (i) $y_{max,t}$ is always smaller than $\tilde{y}_{max,t}$ and (ii) as little information as possible is cut off.

3. Next, we identify outliers as observations with

$$y^p_{isjt} > 1.05 \times \frac{y_{max,t}}{\bar{y}_t},$$

   that is those that exceed the contribution ceiling by more than 5 percent. These outliers are treated as observations with no contributory earnings and therefore deleted from the data set (285 observations).

4. Finally, we recalculate pension claims for all individuals that exceed the contribution ceiling by less than the outlier threshold. Specifically, we set

$$y^p_{isjt} = \frac{y_{max,t}}{\bar{y}_t} \text{ for all } i \text{ with } y^p_{isjt} > \frac{y_{max,t}}{\bar{y}_t}.$$

   We therefore have to modify 16,597 observations.

After these steps, the data is subject to a sharp annual censoring threshold $y_{max,t}$, which is required for the estimation. Table 8 shows the exact values of $\tilde{y}_{max,t}$, $y_{max,t}$, and the share of observation at both thresholds for each year. About 7 to 12 percent of the annual observations are on the threshold value $y_{max,t}$.

Table 8: Identification of $y^*_{max,t}$

Year t	$\tilde{y}_{max,t}$	% at $\tilde{y}_{max,t}$	$y_{max,t}$	% at $y_{max,t}$	Observations n
2000	1.9021	0.9140	1.9017	9.0382	6,893
2001	1.8908	8.4678	1.8905	9.5849	7,251
2002	1.8864	1.2084	1.8858	10.0832	7,696
2003	2.1149	0.2959	2.1143	7.2115	8,112
2004	2.1266	0.6251	2.1261	7.6197	8,478
2005	2.1368	7.4983	2.1365	7.6889	8,922
2006	2.1360	7.3366	2.1358	7.4732	9,514
2007	2.1034	0.9538	2.1029	8.5742	10,170
2008	2.0767	1.0249	2.0763	9.1874	10,830
2009	2.1242	0.4134	2.1239	8.4528	11,369
2010	2.1192	8.6243	2.1191	8.6578	11,943
2011	2.0561	0.6724	2.0556	9.6590	12,641
2012	2.0362	9.4922	2.0361	9.6429	13,274
2013	2.0678	9.6261	2.0675	10.0647	13,453
2014	2.0687	0.7156	2.0683	10.2464	13,556
2015	2.0530	10.6598	2.0528	10.7109	13,687
2016	2.0560	0.7675	2.0553	11.6082	13,680
					181,469

* Values for $\tilde{y}_{max,t}$ and $y_{max,t}$ are expressed relative to average earnings $\bar{y}_t$.

A.2.2 Statistical Model and Moments

We describe the earnings dynamics of the normal earner sample by a standard AR(1) process in logs. We therefore split the normal labor earnings sample according to an individuals' education level $s \in \{0, 1\}$. $s = 0$ summarizes all individuals with high school education, while $s = 1$ indicates the college educated workforce. For each education group, we estimate the statistical model

$$\log(y_{isjt}) = \kappa_{t,s} + \theta_{j,s} + \eta_{isjt} \quad \text{with} \quad \eta_{isjt} = \rho_s \eta_{isj-1,t-1} + \varepsilon_{isjt}, \tag{19}$$

for labor earnings y_{isjt} of an individual i with education s at age j in year t. $\kappa_{t,s}$ is a year fixed effect that controls for earnings changes along the business cycle. $\theta_{j,s}$ is an age fixed effect that informs us about the age-earnings relationship. The noise term ε_{isjt} is assumed to follow a normal distribution with mean 0. Furthermore, we let the stochastic process start from its long-run variance σ^2_s. This means that

$$\varepsilon_{isjt} \sim N(0, \sigma^2_{\varepsilon,s}) \quad \text{and} \quad \eta_{is20t} \sim N\left(0, \sigma^2_s\right) \text{ with } \sigma^2_s = \frac{\sigma^2_{\varepsilon,s}}{1 - \rho^2_s}.$$

We use a generalized method of moments estimator to determine the parameters of this model. We thereby control for the fact that the data are top-coded at the threshold $y_{max,t}$ and that we truncated them at the low earner threshold $y_{min} = 0.23$. Using

$$x_{sjt} = \frac{\log(y_{min}) - \kappa_{t,s} - \theta_{j,s}}{\sigma_s} \quad \text{and} \quad z_{sjt} = \frac{\log(y_{max,t}) - \kappa_{t,s} - \theta_{j,s}}{\sigma_s}$$

as notation for the standardized truncation and censoring thresholds, the education-, age-, and year-specific mean of the left-truncated and right-censored distribution of earnings is

$$\begin{aligned} E_{sjt} &= E\Big[\log(y_{isjt})\Big|\ y_{min} \le y_{isjt} \le y_{max,t}\Big] = \\ &= [1 - P_{sjt}] \times \left[\kappa_{t,s} + \theta_{j,s} + \sigma_s \frac{\phi(x_{sjt}) - \phi(z_{sjt})}{\Phi(z_{sjt}) - \Phi(x_{sjt})}\right] + P_{sjt} \times \log(y_{max,t}) \end{aligned}$$

with

$$P_{sjt} = P\Big(\{y_{isjt} = y_{max,t}\}\Big) = \frac{1 - \Phi(z_{sjt})}{1 - \Phi(x_{sjt})}.$$

When calculating the variance, we exclude the censored data, i.e. all observations with $y_{isjt} = y_{max,t}$. The variance of the double-truncated distribution of earnings then reads

$$\begin{aligned} \mathrm{Var}_{sjt} &= \mathrm{Var}\Big[\log(y_{isjt})\Big|\ y_{min} \le y_{isjt} < y_{max,t}\Big] = \\ &= \sigma_s^2 \times \left[1 + \frac{x_{sjt}\phi(x_{sjt}) - z_{sjt}\phi(z_{sjt})}{\Phi(z_{sjt}) - \Phi(x_{sjt})} - \left(\frac{\phi(x_{sjt}) - \phi(z_{sjt})}{\Phi(z_{sjt}) - \Phi(x_{sjt})}\right)^2\right]. \end{aligned}$$

Following Manjunath and Wilhelm (2012), we derive the intertemporal covariance of the double-truncated distribution of earnings as

$$
\begin{aligned}
\mathrm{Cov}_{sjt} = \mathrm{Cov}\Big[& \log(y_{isjt}), \log(y_{isj+1,t+1}) \\
& \Big| \; y_{min} \leq y_{isjt} < y_{max,t} \wedge y_{min,t+1} \leq y_{isj+1,t+1} < y_{max,t+1} \Big] \\
= \rho\sigma_s^2 \Bigg\{ & 1+ \\
& + M x_{sjt} \phi(x_{sjt}) \left[\Phi\left(\frac{z_{sj+1,t+1} - \rho x_{sjt}}{\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{x_{sj+1,t+1} - \rho x_{sjt}}{\sqrt{1-\rho^2}} \right) \right] \\
& - M z_{sjt} \phi(x_{sjt}) \left[\Phi\left(\frac{z_{sj+1,t+1} - \rho z_{sjt}}{\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{x_{sj+1,t+1} - \rho z_{sjt}}{\sqrt{1-\rho^2}} \right) \right] \\
& + M x_{sj+1,t+1} \phi(x_{sj+1,t+1}) \left[\Phi\left(\frac{z_{sjt} - \rho x_{sj+1,t+1}}{\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{x_{sjt} - \rho x_{sj+1,t+1}}{\sqrt{1-\rho^2}} \right) \right] \\
& - M z_{sj+1,t+1} \phi(x_{sj+1,t+1}) \left[\Phi\left(\frac{z_{sjt} - \rho z_{sj+1,t+1}}{\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{x_{sjt} - \rho z_{sj+1,t+1}}{\sqrt{1-\rho^2}} \right) \right] \\
& + M \frac{\sigma_\varepsilon^2}{\rho} \left[\phi_{0,\Sigma}\begin{pmatrix} x_{sjt} \\ x_{sj+1,t+1} \end{pmatrix} - \phi_{0,\Sigma}\begin{pmatrix} x_{sjt} \\ z_{sj+1,t+1} \end{pmatrix} \right] \\
& - M \frac{\sigma_\varepsilon^2}{\rho} \left[\phi_{0,\Sigma}\begin{pmatrix} z_{sjt} \\ x_{sj+1,t+1} \end{pmatrix} - \phi_{0,\Sigma}\begin{pmatrix} z_{sjt} \\ z_{sj+1,t+1} \end{pmatrix} \right] \Bigg\} \\
& - \sigma_s^2 \left[\frac{\phi(x_{sjt}) - \phi(z_{sjt})}{\Phi(z_{sj+1,t+1}) - \Phi(x_{sj+1,t+1})} \right] \left[\frac{\phi(x_{sj+1,t+1}) - \phi(z_{sj+1,t+1})}{\Phi(z_{sj+1,t+1}) - \Phi(x_{sj+1,t+1})} \right],
\end{aligned}
$$

where

$$
M = \left[\Phi_{0,\Sigma}\begin{pmatrix} z_{sjt} \\ z_{sj+1,t+1} \end{pmatrix} - \Phi_{0,\Sigma}\begin{pmatrix} x_{sjt} \\ x_{sj+1,t+1} \end{pmatrix} \right]^{-1} \quad \text{and} \quad \Sigma = \begin{bmatrix} 1 & \rho^2 \\ \rho^2 & 1 \end{bmatrix}.
$$

A.2.3 Moment Conditions and Estimation

To estimate the statistical model in (19) with our data, we have to determine a total of 110 parameters:

1. 34 year fixed effects $\kappa_{t,s}$ for the years 2000 to 2016 and the education levels $s \in \{0, 1\}$;
2. 72 age fixed effects $\theta_{j,s}$ for the ages 25 to 60 for each education level s;
3. the two unconditional variances σ_s^2;
4. the two autocorrelation parameters ρ_s.

In order to estimate these parameters, we use the labor earnings data y_{isjt}^p to calculate the empirical moments that correspond to the means E_{sjt}, censoring shares

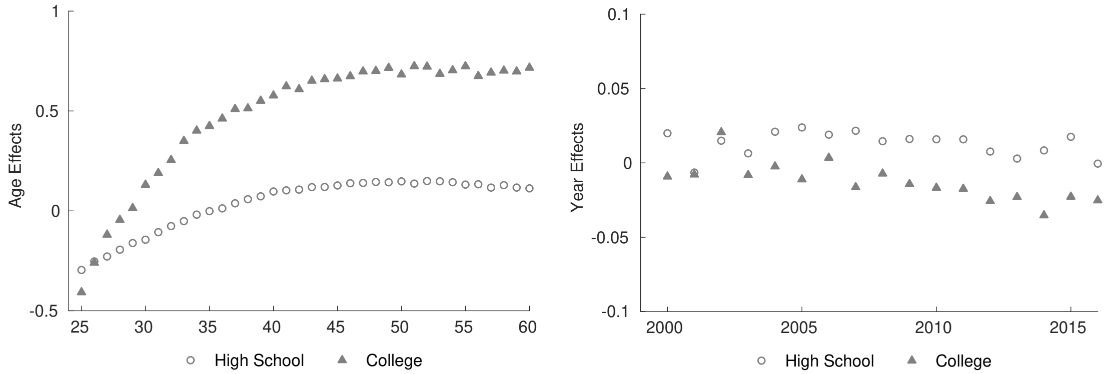
P_{sjt}, variances Var_{sjt} and covariances Cov_{sjt} discussed above for each education level s, age j and year t. We exclude moments when the number of individuals in the corresponding education-age-year bin is smaller than 30, or when the empirical standard error of the moment is equal to zero. This gives us the following moments:

- **sample means:** we estimate 974 means $\hat{\mu}_{sjt}$ of $\log(y^p_{isjt})$ including the censored observations $y_{isjt} = y_{max,t}$ and the corresponding standard errors $\frac{\hat{\sigma}_{sjt}}{\sqrt{n_{sjt}}}$;
- **share of observations at threshold $y_{max,t}$:** we compute 930 shares $\widehat{shr}_{sjt}$ of the observations that sit exactly on the threshold $y_{max,t}$ and the corresponding standard errors $\sqrt{\frac{shr_{sjt}(1-shr_{sjt})}{n_{sjt}}}$;
- **sample variances:** we estimate 943 variances $\hat{\sigma}^2_{sjt}$ of $\log(y^p_{isjt})$ excluding the censored observations as well as the corresponding standard errors of the variance $\hat{\sigma}^2_{sjt}\frac{\sqrt{2}}{n_{sjt}-1}$;
- **sample covariances:** we compute 877 covariances $\hat{\sigma}_{sjt,t+1}$ of $\log(y_{isjt})$ excluding the censored observations as well as the corresponding standard errors of the covariance $\sqrt{\frac{(\hat{\sigma}_{sjt,t+1})^2+\hat{\sigma}^2_{sjt}\hat{\sigma}^2_{sj+1,t+1}}{n_{sjt}-1}}$.

We use these 3724 empirical moments to calculate a residual sum of squares measure. We use a diagonal weighting matrix that has the inverse of the squared standard errors of the empirical moments on the diagonal. To minimize the residual sum of squares and account for multiple local minima, we use the method of simulated annealing, see Du and Swamy (2016). We estimate parameters separately for each education level s.

The results of this estimation process are quite standard in the sense that the estimates exhibit typical life-cycle labor earnings profiles, a significant college wage premium as well as a high auto-correlation of earnings, see Figure 14. We will use these estimates as prime inputs into the calibration of our quantitative model. Yet, as the statistical model describes labor earnings and not labor productivity, we can not use the estimated parameters as direct inputs, see the discussion Section 5.3.2. The left panel of Figure 14 visualizes the point estimates of the age fixed effects by education level. Up to the age of 45, earnings steeply increase for both education groups, especially so for the college educated. Afterwards, they stagnate or decline slightly for the rest of an individual's working life. This shape of life-cycle earnings is quite common in the empirical literature and has been found for other countries as well, see for example Heckman et al. (1998) or Casanova (2013). The college-wage premium implied by these profiles is equal to 60 percent, which is in line with empirical findings (OECD, 2016). The right panel of the figure shows the year fixed effects. These are generally small relative to the age effects and exhibit some cyclical dynamics. Table 9 summarizes the

Figure 14: Age fixed-effects and year fixed-effects



estimation results for the residual earnings process. The parameter estimates are

Table 9: Estimates of residual log-earnings process

	High School $s = 0$	College $s = 1$
Autocorrelation $\hat{\rho}_s$	0.9881	0.9900
Innovation Variance $\hat{\sigma}^2_{\varepsilon,s}$	0.0042	0.0040
Unconditional Variance $\frac{\hat{\sigma}^2_{\varepsilon,s}}{1-\hat{\rho}^2_s}$	0.1787	0.2016

fairly standard. Both high school and college educated workers exhibit a high persistence in labor earnings with an unconditional earnings variance of around 15 to 20 percent. This is in line with what has been found in Bayer and Juessen (2012), for example.

A.2.4 The low labor earnings group

In a second step, we examine the statistical properties of the low labor earnings sample. The left hand side of Figure 2 shows – for each age between 25 and 60 – the fraction of individuals in an age cohort that is a member of the low earnings group (circles for high school and triangles for college educated workers). This fraction declines over time, which indicates that individuals transition between the states of low and normal labor earnings while moving through their life cycle. College educated workers predominantly experience low labor earnings early in their career, for example when doing internships or while working in addition to studying in college, while for high school workers, experiencing a low earnings episode is a phenomenon that is more equally distributed across ages. Labor earnings of individuals in the low earnings group are by and large independent of age and education, see the right panel of Figure 2.[24] For all ages and education types, average earnings of the low earnings group is approximately equal to 10 percent of average labor earnings. The typical low earnings individual consequently makes about 3,700 Euros a year, or 308 Euros a month.

We interpret the findings in Figure 2 in accordance with empirical evidence from the labor literature that starts with Hall (1982). In particular, we assume that individuals face different degrees of career stability.[25] We model career stability as a one-time discrete shock $m \in \{0, 1\}$ that an individual draws at the beginning of working life. While individuals with $m = 0$ face a stable career path and never experience a low earnings episode, those with $m = 1$ may transition into and out of low earnings throughout their entire working life. We denote by ϕ_m the probability to draw $m = 1$, and set it to $\phi_m = 0.5$ in our benchmark simulations. We provide sensitivity checks in Section 6.7.2.

A.3 The Transition Process for Low Earnings Episodes

We model the transition into and out of low earnings as a first-order discrete Markov process with a transition matrix as shown in equation (20). In particular, we assume that households with unstable careers ($m = 1$) face the education-specific transition matrix

$$\Pi^s_{low} = \begin{bmatrix} 1 - \pi^s_{low,0} & \pi^s_{low,0} \\ 1 - \pi^s_{low,1} & \pi^s_{low,1} \end{bmatrix}. \tag{20}$$

The probability $\pi^s_{low,0}$ indicates the likelihood of a normal earner to transition into the low earnings state in the next period, while $\pi^s_{low,1}$ is the probability to remain in the low earnings state. We assume that at age 25, a fraction

$$\Omega_{25,s} = \omega^s_{low}$$

of all individuals with an unstable career path ($m = 1$) start out in the low earnings state. Over time, the share of low earnings individuals evolves according to

$$\Omega_{j+1,s} = \Omega_{j,s} \times \pi^s_{low,1} + (1 - \Omega_{j,s}) \times \pi^s_{low,0}.$$

Knowing that only a share ϕ_m of the population of education level s is exposed to low earnings shocks at all, we can calculate the fraction of individuals in each education-age bin that currently experiences a low earnings episode as

$$\Phi_{j,s} = \phi_m \times \Omega_{j,s}.$$

[24] Partly this may be owing to our choice of the earnings threshold that separates normal and low earners, which is independent of age and education as well.

[25] See Kuhn and Ploj (2020) for a recent investigation of the importance of career stability for heterogeneity in household wealth.

We use the empirical counterparts to these shares $\hat{\Phi}_{j,s}$ shown in the left panel of Figure 2 to estimate the six free parameters ω^s_{low}, $\pi^s_{low,0}$ and $\pi^s_{low,1}$ for $s \in \{0,1\}$ of this statistical model. Our choices of parameters minimizes a simple residual sum of squares between the empirical and the model based moments $\Phi_{j,s}$. Table 10 summarizes the point estimates that provide the best fit to the data in a least squares sense. The solid and dashed lines in the left panel of Figure 2 indicate the model's predicted share of households in the low earnings group. As noted above,

Table 10: Estimates of low-earnings transition process

	High School $s=0$	College $s=1$
Initial share of low income earners ω^s_{low}	0.2022	0.8004
Probability to transition to low earnings $\pi^s_{low,0}$	0.0063	0.0051
Probability to stay low income earner $\pi^s_{low,1}$	0.8374	0.7282

college educated workers experience low earnings episodes predominantly early in their life, while for high-school workers the risk of drawing a low income shock is more equally distributed over the life cycle. This is reflected in the estimates of ω^s_{low}, i.e. the share of low earners at age 25. Throughout her working life, the chance for a regular worker to transition into a low earnings episode is very small (less than 1 percent for both education groups). Being in the low income state however has quite some persistence. With a persistence of 0.84 and 0.73, the average duration of a low earnings episode is 6.15 years for high school workers and 3.74 years for the college educated, respectively.
Summing up, the investigation of the labor earnings process of individuals in our administrative data set has shown that a simple log-normal AR(1) process is not rich enough to describe the earnings dynamics of households. While it might be a fair description of what happens in "normal" times, individuals can also experience very low earnings episodes. We provide a statistical model that can fit the data on low earners by age and education. Note that the recent literature on fiscal redistribution has highlighted the importance of generating a realistic earnings distribution, see for example Castaneda et al. (2003) or Kindermann and Krueger (2022), which can not simply be captured by a single AR(1) labor productivity component. While the aforementioned papers concentrate on income at the top end of the distribution, we use a similar methodology to more realistically characterize households at the bottom, who might be more loosely attached to the labor force and therefore responsive to employment incentives.

B Building Intuition: Solutions

In this Appendix, we present a (partial) equilibrium version of the simple model discussed in Section 3. Households in this framework live for two periods $j = 1, 2$. At each date t, a new generation of mass N_t is born. At the moment they enter the economy, households draw two different shocks: (i) a labor productivity z according to the cumulative distribution function $\Phi_z(\cdot)$ and (ii) a utility cost of employment ξ according to the cumulative distribution function $\Phi_\xi(\cdot)$. We assume both shocks to be independent and identically distributed across households. The interest rate r as well as the wage rate w for effective labor are exogenous. We consider steady state allocations only.[26]

B.1 The Household Decision Problem

As in Section 3 households maximize utility

$$U(c_1, c_2, \ell, e) = c_1 + \frac{c_2}{1+r} - \frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e. \tag{21}$$

subject to the budget constraint

$$c_1 + \frac{c_2}{1+r} = (1-\tau_p) wze\ell + \frac{p}{1+r}. \tag{22}$$

The government operates an employment-linked pension system (ELS), such that

$$p = \kappa \times \Big[\lambda \bar{y} e + (1-\lambda) wze\ell\Big]. \tag{23}$$

Plugging the pension formula into the household's budget constraint, we can write

$$\begin{aligned} c_1 + \frac{c_2}{1+r} &= (1-\tau_p) wze\ell + \frac{\kappa \times \Big[\lambda \bar{y} e + (1-\lambda) wze\ell\Big]}{1+r} \\ &= \left[1 - \tau_p + \frac{\kappa}{1+r} \times (1-\lambda)\right] wze\ell + \frac{\kappa}{1+r} \times \lambda \bar{y} e. \end{aligned}$$

B.2 The equilibrium pension system

For an equilibrium in this economy to exist, we require $r, n \geq -1$, which is not restrictive. Recall that labor productivity z is distributed in this economy according to the distribution function Φ_z. Further, denote by $e(z)$ and $\ell(z)$ the optimal household choices as functions of labor productivity, which we discuss in more detail below. Average labor earnings of the employed then are given by

$$\bar{y} = \frac{\int wze(z)\ell(z)\ \Phi_z(dz)}{\int e(z)\ \Phi_z(dz)}.$$

[26]We hence drop the time index t wherever possible.

The pension system collects pension contributions $\tau_p wze(z)\ell(z)$ from each employed household and pays pensions according to the pension formula discussed above. Let population growth be constant over time and let n denote the population growth rate. In a balanced-budget pay-as-you-go pension system the sum of pension contributions needs to be equal to the sum of pension payments, i.e.

$$\int \tau_p wze(z)\ell(z)\ \Phi_z(dz) = \frac{\int \kappa \times [\lambda \bar{y} e + (1-\lambda) wze\ell]\ \Phi_z(dz)}{1+n}.$$

Dividing this equation by the measure of employed households, we immediately obtain

$$\tau_p \times \bar{y} = \frac{\kappa}{1+n} \times \left[\lambda \bar{y} + (1-\lambda)\bar{y}\right].$$

The equilibrium replacement rate of the pension system hence is

$$\kappa = (1+n)\tau_p. \tag{24}$$

B.3 Implicit taxes and employment subsidies

Let us denote by $\varrho = \frac{1+n}{1+r}$ the ratio between population growth and the economy's interest rate. ϱ is an indicator for the rate-of-return difference between the pension system and the capital market. The smaller is ϱ, the higher is the return to financial investments relative to investments into public pensions. In our benchmark case in Section 3, we assumes that $r = n$ and therefore $\varrho = 1$. However, we now want to prove our results more generally.

Using the relationship in (24), the household budget constraint becomes

$$c_1 + \frac{c_2}{1+r} = \Big[1 - \underbrace{\Big(1 - \varrho(1-\lambda)\Big)\tau_p}_{=:\tau_p^{\text{imp}}}\Big] wze\ell + \underbrace{\lambda \varrho \tau_p \bar{y}}_{=:\tau_p^{\text{sub}}} e. \tag{25}$$

τ_p^{imp} is the implicit tax rate. Note that we have

$$\tau_p^{\text{imp}} \geq 0 \quad \text{whenever} \quad n \leq r + \frac{\lambda}{1-\lambda}(1+r).$$

In a proportional pension system with $\lambda = 0$, the implicit tax rate on labor earnings is hence non-negative if $n \leq r$, and it is zero in case of $n = r$. In a dynamically efficient economy ($n \leq r$), the implicit tax rate is always positive for any $\lambda > 0$. τ_p^{sub} is an employment subsidy. This subsidy is positive whenever $\lambda > 0$.

B.4 Optimal choices

Using the budget constraint in (25), the household optimization problem becomes

$$\max_{c_1,c_2,\ell,e} \quad u(c_1,c_2,\ell,e) = c_1 + \frac{c_2}{1+r} - \frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e$$
$$\text{s.t. } c_1 + \frac{c_2}{1+r} = \left[1-\tau_p^{\text{imp}}\right] wze\ell + \tau_p^{\text{sub}} e.$$

The first-order condition with respect to intensive margin labor supply is

$$\begin{aligned} -\ell(z|e=1)^{\frac{1}{\chi}} + \left[(1-\tau_p^{\text{imp}})wz\right] &= 0 \\ \Leftrightarrow \qquad \ell(z|e=1) &= \left[(1-\tau_p^{\text{imp}})wz\right]^{\chi}. \end{aligned} \tag{26}$$

Plugging $\ell(z|e=1)$ into the household utility function, we immediatley obtain

$$\begin{aligned} U(z|e=1) &= [1-\tau_p^{\text{imp}}]wz[(1-\tau_p^{\text{imp}})wz]^{\chi} + \tau_p^{\text{sub}} - \frac{[(1-\tau_p^{\text{imp}})wz]^{1+\chi}}{1+\frac{1}{\chi}} - \xi \\ &= \frac{\left[(1-\tau_p^{\text{imp}})wz\right]^{1+\chi}}{1+\chi} + \tau_p^{\text{sub}} - \xi. \end{aligned}$$

As $\ell(z|e=0)=0$, we have $U(z|e=0)=0$ and hence the utility difference between being employed and not is

$$U(z|e=1) - U(z|e=0) = \frac{\left[(1-\tau_p^{\text{imp}})wz\right]^{1+\chi}}{1+\chi} + \tau_p^{\text{sub}} - \xi.$$

Given the distribution Φ_ξ of the utility costs of employment, the probability that an individual with labor productivity z is employed is given by

$$\begin{aligned} P(e=1|z) &= P\left(\left\{U(z|e=1) - U(z|e=0) \geq 0\right\}\right) \\ &= \Phi_\xi\left(\frac{\left[(1-\tau_p^{\text{imp}})wz\right]^{1+\chi}}{1+\chi} + \tau_p^{\text{sub}}\right). \end{aligned} \tag{27}$$

B.5 Incentive effects of progressive pensions

To study the incentive effects of employment-linked progressive pensions on labor supply, we take the derivative of a household's employment decision with respect to λ. For the intensive hours choice in (26) this derivative is

$$\frac{\partial \ell(z|e=1)}{\partial \lambda} = -\tau_p \times \varrho \times \chi \times \frac{\ell(z|e=1)}{1-\tau_p^{\text{imp}}} < 0.$$

The probability of being employed in (27) changes with λ according to

$$\begin{aligned}\frac{\partial P(e=1|z)}{\partial\lambda} &= \phi_\xi(\cdot)\cdot\left[\left[(1-\tau_p^{\text{imp}})wz\right]^{\chi}(-wz)\cdot\frac{\partial\tau_p^{\text{imp}}}{\partial\lambda}+\frac{\partial\tau_p^{\text{sub}}}{\partial\lambda}\right] \\ &= \phi_\xi(\cdot)\cdot\left[-wz\ell(z|e=1)\cdot\frac{\partial\tau_p^{\text{imp}}}{\partial\lambda}+\frac{\partial\tau_p^{\text{sub}}}{\partial\lambda}\right]\end{aligned}$$

With $\frac{\partial\tau_p^{\text{imp}}}{\partial\lambda}=\varrho\tau_p$ and $\frac{\partial\tau_p^{\text{sub}}}{\partial\lambda}=\varrho\tau_p\bar{y}$, we get

$$\frac{\partial P(e=1|z)}{\partial\lambda}=\tau_p\times\varrho\times\phi_\xi(\cdot)\times\left[\bar{y}-wz\ell(z|e=1)\right],$$

where the sign of the effect depends on the relative income position of the household. It is positive for all individuals with earnings less than the average earnings of the workforce, and negative otherwise.

C Simulation Model: Computational Details

C.1 First-order conditions for the ELS

In the following, we describe the first-order conditions of the household problem under an employment-linked pension system.

The dynamic household optimization problem reads

$$v(\mathbf{x}) = \max_{c,\ell,e,a^+,ep^+} \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \nu \frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e \\ + \frac{\beta\psi_{j+1,h}}{1-\frac{1}{\sigma}} E\left[\left[\left(1-\frac{1}{\sigma}\right) v(\mathbf{x}^+)\right]^{1+\gamma} \;\middle|\; j,s,m,\eta,h\right]^{\frac{1}{1+\gamma}}$$

with $\mathbf{x} = (j,s,m,\eta,h,a,ep)$ and $\mathbf{x}^+ = (j+1,s,m,\eta^+,h^+,a^+,ep^+)$. Households maximize their utility with respect to the budget constraint

$$(1+\tau_c)c + a^+ + T_p(y) + T\Big(y - T_p(y) + p\Big) = (1+r)a + y + p + b \\ \text{with } \; y = wz(j,s,m,\eta)e\ell$$

and the accumulation equation for pension claims

$$ep^+ = ep + \Big[\lambda\bar{y}e + (1-\lambda)\min\Big(wz(j,s,m,\eta)e\ell,\; 2\bar{y}\Big)\Big].$$

In the following, we assume that $y < 2\bar{y}$, meaning that the household is below the contribution ceiling of the pension system. Let us denote by μ_1 and μ_2 the multipliers on the budget constraint and the pension accumulation equation in the Lagrangian $\mathcal{L}$, respectively. The first-order conditions of the household then read

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial c} &= c^{-\frac{1}{\sigma}} - \mu_1(1+\tau_c) = 0\\
\frac{\partial\mathcal{L}}{\partial \ell} &= -\nu\ell^{\frac{1}{\xi}} + \Big[(1-\tau_p)\Big(1-T'(y_{tax})\Big)\mu_1 + (1-\lambda)\mu_2\Big] wz(j,s,m,\eta)e = 0\\
\frac{\partial\mathcal{L}}{\partial a^+} &= -\mu_1 + \beta\psi_{j+1,h}E\Big[M(\mathbf{x}^+)V_a(\mathbf{x}^+)\;\Big|\; j,s,m,\eta,h\Big] = 0\\
\frac{\partial\mathcal{L}}{\partial ep^+} &= -\mu_2 + \beta\psi_{j+1,h}E\Big[M(\mathbf{x}^+)V_{ep}(\mathbf{x}^+)\;\Big|\; j,s,m,\eta,h\Big] = 0
\end{aligned}$$

where $y_{tax} = y - T_p(y) + p$ and

$$M(\mathbf{x}^+) = E\left[\left[\left(1-\frac{1}{\sigma}\right)v(\mathbf{x}^+)\right]^{1+\gamma} \;\middle|\; j,s,m,\eta,h\right]^{\frac{-\gamma}{1+\gamma}} \times \left[\left(1-\frac{1}{\sigma}\right)v(\mathbf{x}^+)\right]^{\gamma}. \tag{28}$$

Note that the state-specific discount factor $M(\mathbf{x}^+)$ determines the weight a household attaches to different future events. In the case of standard CRRA preferences, i.e. when $\gamma = 0$, we have $M(\mathbf{x}^+) = 1$ and risk aversion solely emerges from the curvature of the household's utility functions. In case of $\gamma > 0$, the household attaches a higher weight to negative future events and therefore risk aversion increases.

Using the envelope theorem, we immediately obtain

$$\begin{aligned} V_a(\mathbf{x}) &= (1+r)\mu_1 \quad \text{and} \\ V_{ep}(\mathbf{x}) &= \begin{cases} \mu_2 & \text{if } j < j_R \text{ and} \\ (1 - T'(y_{tax}))\frac{\kappa}{j_R - 20}\mu_1 + \mu_2 & \text{otherwise.} \end{cases} \end{aligned}$$

Under the assumption of a time-invariant consumption tax rate, the Euler equation then reads

$$c^{-\frac{1}{\sigma}} = (1+r)\beta\psi_{j+1,h}E\left[M(\mathbf{x}^+)V_a(\mathbf{x}^+) \;\middle|\; j, s, m, \eta, h\right]. \tag{29}$$

The first order condition for labor supply is

$$\begin{aligned} \nu\ell^{\frac{1}{\xi}} = \Bigg[&(1-\tau_p)\Big(1 - T'(y_{tax})\Big)\frac{c^{-\frac{1}{\sigma}}}{1+\tau_c} \\ &+ (1-\lambda)\beta\psi_{j+1,h}E\left[M(\mathbf{x}^+)V_{ep}(\mathbf{x}^+) \;\middle|\; j, s, m, \eta, h\right]\Bigg]wz(j, s, m, \eta)e. \end{aligned} \tag{30}$$

C.2 Stationary Recursive Competitive Equilibrium

Definition 1. *Given an international interest rate $\bar{r}$, government expenditures G, a level of government debt B, a consumption tax rate τ_c, a progressive tax system $T(\cdot)$ as well as a characterization of the pension system $\{\tau_p, \kappa\}$, a stationary recursive equilibrium with population growth n is a collection of value and policy functions $\{v, c, \ell, e, a^+, ep^+\}$ for the household, optimal production inputs $\{K, L\}$, accidental bequests $\{b_j\}_{j=1}^J$, a net foreign asset position and a trade balance $\{Q, TB\}$ as well as factor prices $\{r, w\}$ that satisfy*

1. *(Household Optimization) Given prices and characteristics of the tax and pension system, the value function v satisfies the Bellman equation* (9) *together with the budget constraint, the accumulation equation for pension claims, the borrowing constraint and the laws of motion for productivity risk and health. c, ℓ, e, a^+, and ep^+ are the associated policy functions.*

2. *(Firm Optimization) Given the international interest rate $\bar{r}$ as well as the wage rate w, firms employ capital and labor according to the demand functions*

$$\bar{r} = \Omega\alpha \left(\frac{L}{K}\right)^{1-\alpha} - \delta \ \text{ and } \ w = \Omega(1-\alpha)\left(\frac{K}{L}\right)^{\alpha}.$$

3. *(Government Constraints) The budget constraints of the pension system* (11) *and the tax system* (12) *hold, and accidental bequests are calculated from* (14)*.*

4. *(Market Clearing:)*

   (a) *The labor market clears:*

$$L = \int z(j, s, m, \eta) e(\boldsymbol{x}) l(\boldsymbol{x}) \, d\Phi$$

   (b) *The capital market clears:*

$$K + B + Q = \int a \, d\Phi$$

   (c) *The balance of payments identity is satisfied:*

$$TB = (n - \bar{r})Q$$

   (d) *The goods market clears:*

$$Y = \int c(\boldsymbol{x}) \, d\Phi + (n + \delta)K + G + TB.$$

5. *(Consistency of Probability Measure Φ) The invariant probability measure is consistent with the population structure of the economy, with the exogenous processes of labor productivity η and health h, and the household policy functions a^+ and ep^+. A formal definition is provided in Appendix C.3.*

C.3 The Measure of Households

First, we construct the measure of households at age 20 across the characteristics (s, m, η, h, a, ep). Households draw one of two possible education levels $s \in \{0, 1\}$, where $s = 1$ occurs with probability ϕ_s. They are also assigned a career-path characteristic $m \in \{0, 1\}$, where $m = 1$ occurs with probability ϕ_m. Conditional on their career path m, households draw an initial labor productivity η at age 20 from the distribution $\pi_{\eta,20}(\eta \,|\, m)$, see equation (36). Finally, households enter the economy with average health $\bar{h}$, zero assets and zero pension claims. Thus,

$$
\begin{aligned}
&\Phi(\{20\}, \{s\}, \{m\}, \{\eta\}, \{\bar{h}\}, \{0\}, \{0\}) = \\
&\qquad = \Big[s\phi_s + (1-s)(1-\phi_s)\Big] \times \Big[m\phi_m + (1-m)(1-\phi_m)\Big] \times \pi_{\eta,20}(\eta \,|\, m)
\end{aligned}
$$

and zero otherwise.

We can then construct the probability measure for all ages $j > 1$. For all Borel sets of assets $\mathcal{A}$ and pension claims $\mathcal{EP}$ we have

$$
\begin{aligned}
&\Phi(\{j+1\}, \{s\}, \{m\}, \{\eta^+\}, \{h^+\}, \mathcal{EP}, \mathcal{A}) = \\
&= \frac{\psi_{j+1,h} \times \pi_\eta(\eta^+ \mid \eta, j, s, m) \times \pi_h(h^+ \mid h, j, s, \eta)}{1+n} \\
&\quad \times \int \mathbb{1}_{\{a'(j,s,m,\eta,h,a,ep) \in \mathcal{A}\}} \times \mathbb{1}_{\{ep'(j,s,m,\eta,h,a,ep) \in \mathcal{EP}\}} \; \Phi(\{j\}, \{s\}, \{m\}, \{\eta\}, \{\bar{h}\}, dep, da)
\end{aligned}
$$

where

$$
\int \mathbb{1}_{\{a'(j,s,m,\eta,h,a,ep) \in \mathcal{A}\}} \times \mathbb{1}_{\{ep'(j,s,m,\eta,h,a,ep) \in \mathcal{EP}\}} \; \Phi(\{j\}, \{s\}, \{m\}, \{\eta\}, \{\bar{h}\}, dep, da)
$$

is the measure of assets a and pension claims ep today such that, for fixed (j, s, m, η, h), the optimal choice today of assets for tomorrow $a^+(j, s, m, \eta, h, a, ep)$ lies in $\mathcal{A}$ and the optimal choice today of pension claims for tomorrow $ep^+(j, s, m, \eta, h, a)$ lies in $\mathcal{EP}$.

C.4 Computational Algorithm

Following Kindermann et al. (2020), we solve the model in three steps. We apply the method of endogenous grid points to solve the household problem. We can than compute policy functions $c(x), \ell(x)$ and $a^+(x)$, as well as the value function $v(x)$. Second, we determine equilibrium quantities and prices following closely the Gauss-Seidel-Quasi-Newton procedure proposed in Ludwig (2007). Finally, we calculate compensating transfers using a standard rootfinding method.

C.4.1 Computation of Policy and Value Functions

We use the method of endogenous gridpoints to compute the policy and value functions. The state space of the quantitative model is $\mathbf{x} = (j, s, m, \eta, h, a, ep)$.

To solve the model on a computer, we start with discretizing the continuous elements a, ep and η. We use routines provided by the toolbox that accompanies Fehr and Kindermann (2018).

- We specify the asset grid $\hat{\mathcal{A}} = \{\hat{a}_0, \dots, \hat{a}_{100}\}$ as nodes with growing distance on the interval $[\bar{a}_l, \bar{a}_u]$. In particular, we let

$$\hat{a}_i = \bar{a}_l + \frac{\bar{a}_u - \bar{a}_l}{(1+g_a)^{100} - 1} \times [(1+g_a)^i - 1] \text{ for } i = 0, 1, \dots, 100.$$

  The lower limit of the asset grid is $\bar{a}_l = 0$, the upper limit of the asset grid is $\bar{a}_u = 70$, the growth rate of gridpoints $g_a = 0.08$.

- We specify the earnings points grid $\widehat{\mathcal{EP}} = \{\hat{ep}_0, \dots, \hat{ep}_{30}\}$ as a grid with $\bar{ep}_l = 0$, $\bar{ep}_u = 2$ and equally spaced nodes.

- We approximate the stochastic process of the AR(1) labor productivity process of normal labor earnings by a Markov chain. We use the Rouwenhorst method to discretize the stochastic process of the innovations[27] $\hat{\mathcal{E}} = \{\hat{\eta}_1, \dots, \hat{\eta}_7\}$ and to determine a transition matrix

$$\pi_\eta(\eta^+|\eta) = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{17} \\ \pi_{21} & \pi_{22} & \dots & \pi_{27} \\ \vdots & \vdots & \ddots & \dots \\ \pi_{71} & \pi_{72} & \dots & \pi_{77} \end{bmatrix}. \tag{31}$$

- In order to account for the low productivity shocks, we extend the stochastic process $\hat{\mathcal{E}}$ to $\hat{\mathcal{E}} = \{\hat{\eta}_0, \hat{\eta}_1, \dots, \hat{\eta}_7\}$ and augment the 7×7 Markov transition matrix as outlined in Appendix D.4.

- We determine the health shocks $h \in \{0, \dots, 7\}$ and the transitions matrix $\pi_h(h^+|h, j, s, \eta)$ as outlined in Appendix D.1.

The policy and value functions can now be solved via backward induction. In the last possible age J, the household will not work[28] and not save, but will consume all remaining resources. This determines the policy functions

$$\begin{aligned} c(J, s, m, \hat{\eta}_g, h, \hat{a}_i, \hat{ep}_k) &= \frac{(1+r) \times \hat{a}^i + p - T(p) + b}{1 + \tau_c}, \\ l(J, s, m, \hat{\eta}_g, h, \hat{a}_i, \hat{ep}_k) &= 0, \\ a^+(J, s, m, \hat{\eta}_g, h, \hat{a}_i, \hat{ep}_k) &= 0 \end{aligned}$$

[27] Where ρ_s and $\sigma^2_{\varepsilon,s}$ are as specified in Table 1 and $\mu = 0$.

[28] Remember, the compulsory retirement age is J_R.

and the value function

$$v(J,s,m,\hat{\eta}_g,h,\hat{a}_i,\hat{ep}_k) = \frac{\Big[c(J,s,m,\hat{\eta}_g,h,\hat{a}_i,\hat{ep}_k)\Big]^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}$$

for all $i = 0,\dots,100,\ k = 0,\dots,30,\ g = 0,\dots,7$.
With the final period policy functions and value function at hand, we can iterate backwards over ages to determine the full history of household decisions. We describe the procedure for working-age households. Assume the problem is solved for age $j+1$, then the problem for an individual at state $\mathbf{x} = (j,s,m,\eta,h,a,ep)$ reads

$$\begin{aligned} v(\mathbf{x}) = \max_{c,\ell,e,a^+,ep^+} \ & \frac{c^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \nu\frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e \\ & - \beta\psi_{j+1,h}\left[\sum_{\eta^+}\pi_\eta(\eta^+|\eta)\sum_{h^+}\pi_h(h^+|h)\cdot[-v(\mathbf{x}^+)]^{1+\gamma}\right]^{\frac{1}{1+\gamma}} \end{aligned} \tag{32}$$

with $\mathbf{x} = (j,s,m,\eta,h,a,ep)$ and $\mathbf{x}^+ = (j+1,s,m,\eta^+,h^+,a^+,ep^+)$. Households maximize their utility with respect to the budget constraint

$$\begin{aligned} (1+\tau_c)c + a^+ + T_p(y) + T\Big(y - T_p(y) + p\Big) = (1+r)a + y + p + b \\ \text{with } \ y = wz(j,s,m,\eta)e\ell, \end{aligned}$$

the accumulation equation for pension claims

$$ep^+ = ep + \Big[\lambda\bar{y}e + (1-\lambda)\min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big)\Big].$$

and the positive asset restriction $a^+ \geq 0$ and the time restriction $0 \leq \ell \leq 1$. The first order conditions are outlined in Appendix C.1.
We now apply the method of endogenous gridpoints. We first define an exogenous grid on the state space $\{\hat{a}_v\}_{v=0}^{100}$, which denotes the remainder of assets in the next period, i.e. $a^+ = \hat{a}_v$. For each state $\tilde{x} = (j,s,m,\eta,h,\mathbf{a}^+,ep)$, we

- search for the optimal $\ell(\tilde{x})$ according to the first order condition (30) using a quasi-Newton rootfinding method

  1. given $\ell(\tilde{x})$[29] we determine

  $$ep^+ = \frac{(j-1)ep}{j} + \frac{\lambda\bar{y}e + (1-\lambda)\min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big)}{j}$$

  2. given a^+ and ep^+ we determine $c(\tilde{x})$ from the Euler Equation (30)
  3. with $l(\tilde{x})$ and $c(\tilde{x})$, we use the budget constraint (33) to get $a(\tilde{x})$

[29] we guess $\ell = \ell^+$ in the first iteration

- once $l(\tilde{x})$, $c(\tilde{x})$ and $a(\tilde{x})$ are solved, we can interpolate along a to obtain the policy functions $l(x)$, $c(x)$ and $a^+(x)$ as well as the value function

$$
\begin{aligned}
v(\mathbf{x}) = & \frac{c(x)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \nu \frac{\ell(x)^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e(x) \\
& - \beta \psi_{j+1,h} \left[\sum_{\eta^+} \pi_\eta(\eta^+|\eta) \sum_{h^+} \pi_h(h^+|h) \cdot [-v(\mathbf{x}^+)]^{1+\gamma} \right]^{\frac{1}{1+\gamma}}, \quad (33)
\end{aligned}
$$

  for each today's asset value $\hat{a}_i,\ i = 0, \ldots, 100$ and earnings points amount $e\hat{p}_k, k = 0, \ldots, 30$ by piecewise linear interpolation[30]

In case the asset restriction $a^+ \geq 0$ is binding, we extend the interpolation data by another point of value 0 on the left and determine the policy and value functions at this point. We assume the household consumes all available resources and has no savings left over for tomorrow.

C.5 The initial equilibrium of the macroeconomy

We model a small open economy, hence prices r and w are fixed. In order to determine aggregate quantities and policy parameters in the initial equilibrium ($t = 0$) we need to determine the following four variables numerically:

- the government budget balancing consumption tax rate τ_c as outlined in equation (12)
- the pension replacement rate κ that balances pension contributions and pension payments as outlined in equation (11)
- average earnings $\bar{y}_t$[31]
- aggregate bequests $\bar{B}$, which immediately allows us to compute cohort bequests $\{b_j\}_{j=1}^{J}$ (equally shared between all working-age individuals)

Once a guess of these four variables is available, we can use the following algorithm to compute the remainder individual and aggregate variables of the economy:

1. We solve the household optimization problem using the guesses for $\tau_c, \kappa, \bar{y}, \bar{B}$ and determine the measure of households.

[30]Note, we interpolate $\left[(1-\frac{1}{\sigma})v(x)\right]^{\frac{1}{1-\frac{1}{\sigma}}}$ rather than $v(x)$ directly and than transform it back to the original shape. This leads to more accurate results for discretized functions with high curvature.

[31]This is an important parameter, as it determines the pension contribution cap, pension payments and earnings of the low-earning group.

2. We compute aggregate quantities $\{L, K, TB, Y, C, G, I, \Omega, B\}$ from individual decisions and the measure of household and determine the gap $D = Y - C - I - G$ between demand and supply.

We determine the four central parameters $(\tau_c, \kappa, \bar{y}_t, \bar{B})$ by means of a quasi-Newton rootfinding method. The method receives an initial guess of these variables and updates them in each iteration step using the Jacobian of the determining equation system. The iteration process stops when the government and the pension budget are in equilibrium and the model implied average earning and aggregate bequest equal the guess provided by the method. After the iteration procedure has finished, we extract the Jacobian which is essential for running a Gauss-Seidel-Quasi-Newton method as proposed in Ludwig (2007) to compute the transitional dynamics.

C.6 The transition path of the macroeconemy

To quantify the intergenerational effects of the pension reforms, we simulate the economy along the transition path. We distinguish between different simulation periods $t = 0, 1, \ldots, T$, where period $t = 0$ is the initial equilibrium as determined before. In period $t = 1$, the pension reform is introduced such that households adopt their decisions and the macroeconomy adjusts. The economy slowly converges to a new long-run equilibrium, which is reached after $T = 300$ periods in the numerical model.

While most variables and parameters are time dependent, government consumption G is fixed. The measure for average earnings $\bar{y}_t$ adjusts along the transition. However, we use the value from the initial equilibrium $\bar{y}_0$ as reference value for computing policy parameters such as pension payments. As a result, we only have to determine $\tau_{c,t}, \kappa_t$ and aggregate bequests$\bar{B}_t$ in each period $t = 1, 2, \ldots, T$ of the transition.

We use the Gauss-Seidel-Quasi-Newton procedure proposed by Ludwig (2007) to solve for our variables of interest. This procedure works like a standard rootfinding method with the difference that the Jacobian is not computed numercially but initialized using the initial equilibrium Jacobian. To speed up the computational process, we use openMP to parallelize the computation of household decisions and invariant household measure across different cohorts.

C.7 Welfare Calculations

C.7.1 Consumption Equivalent Variation

We calculate utility of each household recursively according to

$$v_t(\mathbf{x}) = u\Big(c_t(\mathbf{x}), \ell_t(\mathbf{x}), e_t(\mathbf{x})\Big) - \beta\psi_{j+1,h}E_t\left[\Big(-v_{t+1}(\mathbf{x}^+)\Big)^{1+\gamma}\right]^{\frac{1}{1+\gamma}}.$$

In addition, we can calculate the discounted marginal utility of consumption as

$$PVC_t(\mathbf{x}) = u_c\Big(c_t(\mathbf{x}), \ell_t(\mathbf{x}), e_t(\mathbf{x})\Big) - \beta\psi_{j+1,h}E_t\left[M(\mathbf{x}^+)PVC_{t+1}(\mathbf{x}^+)\right],$$

with the stochastic discount factor defined as in (28).

Ex-ante expected utility of a cohort born at some date t then is given by

$$EV_t = -E_0\left[(-v_t(x))^{1+\gamma}\right]^{\frac{1}{1+\gamma}}.$$

where $\mathbf{x} = \{1, s, m, \eta, \bar{h}, 0, 0\}$ and E_0 uses the invariant distribution of this cohort at age $j = 1$. Similarly, we can calculate the ex-ante discounted marginal utility of consumption as

$$PVC_t = E_0\left[M(\mathbf{x})PVC_t(x)\right].$$

Let us denote by $\overline{EV}$ and $\overline{PVC}$ the ex-ante welfare measure and the discounted marginal utility of consumption of a cohort that was born and has lived entirely in the initial equilibrium with a proportional pension system. We compute the consumption equivalent variation welfare measure between this initial cohort and any other cohort that was born at some date t and has experienced the pension reform as

$$CEV_t = \frac{EV_t - \overline{EV}}{\overline{PVC}}.$$

C.7.2 Efficiency Measure

To derive our measure of aggregate efficiency, we numerically compute the transfer payment Ψ_t that we have to give to each cohort affected by the pension reform so as to make this cohort as well of as the initial equilibrium cohort. Technically, we use a quasi-Newton method and determine the payment Ψ_t such that

$$EV_t(\Psi_t) = \overline{EV}.$$

The negative of Ψ_t is a monetary measure of the welfare increase the cohort t experiences from the pension reform.

We then derive the present value of all transfers, which gives us a wealth-based measure W of the economic efficiency effect

$$W = \sum_{t=-J+1}^{\infty} \left[\frac{1+n}{1+\bar{r}}\right]^t \Psi_t.$$

We convert the wealth-based measure into an annuity that pays out a constant stream along the transition path and in the new long-run equilibrium:

$$C = W \times \left[\sum_{t=1}^{\infty} \left[\frac{1+n}{1+\bar{r}}\right]^t\right]^{-1}.$$

Our final measure of economic efficiency relates this annuity stream to the initial equilibrium aggregate consumption level, i.e., we compute

$$\varphi = -\frac{C}{C_0} \cdot 100.$$

Note that we have to use $-C$ in this computation, as we want our measure to be positive when the economy experiences aggregate efficiency gains.

D Further Information on the Calibration Process

D.1 Determining the health shock

This section provides details on the calibration process of the probabilities $P(h|s,\eta)$ to draw a certain health shock upon entering retirement. We assume $P(h|s,\eta)$ to be the probability mass function of a binomial distribution with success probabilities $p_{s,\eta}$ depending on education and labor productivity. In particular, we let

$$p_{s,\eta} = \Phi\left(\iota_0 + \iota_1 \times \mathbb{1}_{s=\text{college}} + \iota_2 \times \eta\right), \tag{34}$$

where Φ is the probability distribution function of the standard normal distribution and $\mathbb{1}_{s=\text{college}}$ is an indicator function that takes a value of one for households with college education. We set the parameters $\iota_1 = 0.32$ and $\iota_2 = 0.64$ to target the reported life expectancy gaps by education level and life-time labor earnings. Finally, we choose $\iota_0 = -0.05$ such that the average life expectancy of the total population amounts to 79.5 years, the value we obtain from the Human Mortality Database (2020) life tables. The right panel of Figure 15 shows the relation between life-time labor earnings and life expectancy. While individuals in the bottom decile expect their life to be about three years shorter than that of the population average, the average life of a top decile earner is four years longer.

Incorporating these probabilities into model notation, we have

$$\pi_h(h^+|h,j,s,\eta) = \begin{cases} P(h|s,\eta) & \text{if } j = j_r - 1 \text{ and} \\ \mathbf{I} & \text{otherwise,} \end{cases}$$

with $\mathbf{I}$ being the identity matrix. Consequently, our model features one single health shock that individuals are exposed to right before entering retirement. After the individual health status is revealed, households retain their health level for the rest of their life.

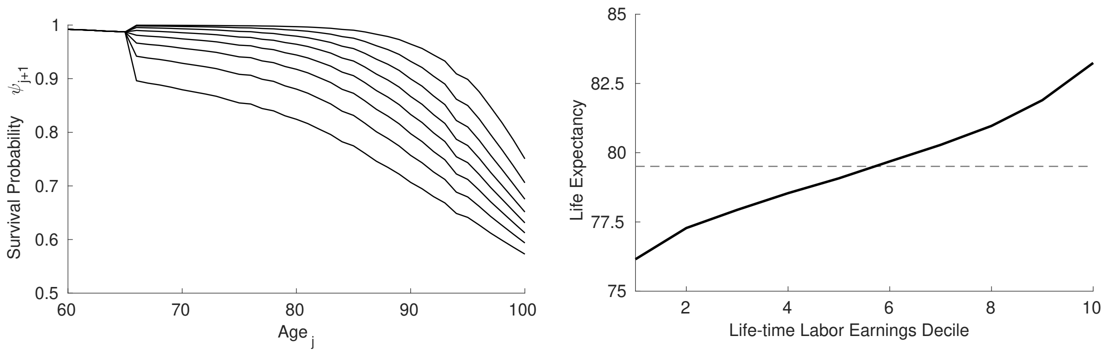
D.2 Determining survival probability profiles

We calculate average survival probabilities $\bar{\psi}_j$ from the 2017 annual life tables for men from the Human Mortality Database (2020). $\bar{\psi}_j$ is hence the average probability of an individual of age j to survive to age $j+1$. During working life ($j < j_R$) we set the individual survival probabilities $\psi_{j,h}$ equal to $\bar{\psi}_j$. When entering retirement, each individual draws one out of eight different health shocks $h \in \{0,\dots,7\}$ according to a probability distribution $P(h|s,\eta)$. A health shock is associated with a set of survival probabilities $\psi_{j,h}$ that we calculate from a logistic model

$$\psi_{j,h} = \frac{1}{1+\exp(-\iota_h \times \bar{x}_j)} \quad \text{with} \quad \bar{x}_j = \log\left(\frac{1}{\bar{\psi}_j} - 1\right). \tag{35}$$

We choose the multipliers ι_h such that (i) life expectancy at the lowest health shock $h = 0$ is ten years below average, (ii) life expectancy at the highest health shock $h = 7$ is ten years above average and (iii) life expectancy evolves linearly with health shocks h.[32] The left panel of Figure 15 in the main text shows the resulting survival probability profiles.

Figure 15: Survival probabilities and life expectancy



D.3 Estimating model-implied participation elasticities

For estimating participation elasticities we follow the evidence from Table 2(2) in Bartels and Pestel (2016). They empirically test to what extent a lower participation tax rate PTR is associated with an increased probability of taking up work. They define a household's participation tax rate as

$$PTR_{ih} = \frac{T(y_h^E) - T(y_h^U)}{y_i^{E,w}},$$

where y_h^E is gross household income (i.e. the sum of labor earnings, asset income and transfers of all household members), $T(y_h^E)$ is a household's net taxes and $y_i^{E,w}$ are labor earnings of individual i when being employed E. $T(y_h^U)$ denotes a household's net taxes if individual i is unemployed U. The binary outcome variable *switch* takes a value of one if individual i switches from non-participation in period $t-1$ to participation in period t. Bartels and Pestel (2016) estimate the effect of changes in the short-term participation tax rate Δ_{PTR} on male labor force participation in Germany, evaluated at 40 h, using the following statistical model:

$$\begin{aligned} switch = b_1\Delta_{PTR} + b_2 Age_{35-44} + b_3 Age_{45-54} + b_4\Delta_{U-rate} + b_5 East \\ + b_6 \, Year_{FE} + b_7 HH_{FE} + b_8 Skill_{FE} + \epsilon. \end{aligned}$$

[32]Note that for $\iota_h = 1$, we recover the average survival probability $\psi_{j,h} = \bar{\psi}_j$.

b_1 is the coefficient of interest, which takes a value of -0.106 and is significant at the 1% level. The impact of changes in the the short-term participation tax rate on the probability to take up work is substantial. Reducing the participation tax rate by 10 percentage points increases the probability of taking up work by 1.06 percentage points. Coefficients on age-group dummies, changes in the unemployment rate and on whether a household is located in East Germany b_2, b_3, b_4 and b_5 are all insignificant.

We adopt this method to estimate the participation elasticity implied by our model using simulated data. We restrict the simulated data such that it corresponds to the data selection of Bartels and Pestel (2016). We meet most of the specifications by construction as, for instance, self-employed, civil servants and disabled individuals are not represented in our model anyway. We limit the analysis to individuals of ages 25 to 54.

Our measure for PTR is constructed as follows: We estimate participation taxes in the benchmark equilibrium of our model that most closely resembles the German economy. For each potential household characterized by the state vector $\mathbf{x} = (j, s, m, \eta, h, a, ep)$ with $j \in \{25, \ldots, 54\}$, we compute the initial share of employed individuals $e(\mathbf{x})$, the initial taxable income

$$y_{tax}(\mathbf{x}) = y(\mathbf{x}) - \tau_p \min\Big(y(\mathbf{x}), 2\bar{y}\Big) \quad \text{with} \quad y(\mathbf{x}) = wz(j, s, m, \eta)\ell(\mathbf{x})$$

and the initial participation tax rate as

$$PTR(\mathbf{x}) = \frac{T_p\Big(y(\mathbf{x})\Big) + T\Big(y_{tax}(\mathbf{x})\Big)}{y(\mathbf{x})}.$$

Next, we reduce the contribution rate to the pension system τ_p by 10 percentage points without recalculating equilbrium prices. Under this new contribution rate, we compute a new share of employed households $e_{new}(\mathbf{x})$ and a new participation tax rate $PTR_{new}(\mathbf{x})$.

Under the benchmark equilibrium, a fraction $1-e(\mathbf{x})$ of households was not in employment. Under the system with a lower pension contribution rate, the fraction of non-employed changed to $1 - e_{new}(\mathbf{x})$. We split the sample of $1 - e(\mathbf{x})$ non-employed individuals into those $e_{new}(\mathbf{x}) - e(\mathbf{x})$ that switched from non-employment to employment and assign to them a value of 1 for the variable *switch*. For the other $1 - e_{new}(\mathbf{x})$ that remained in non-employment, *switch* takes a value of 0. The change in the participation tax rate of these individuals is equal to

$$\Delta_{PTR} = PTR_{new}(\mathbf{x}) - PTR(\mathbf{x}).$$

To account for the distribution of households over the state-space, we create a weighted data set using the distribution $\Phi(\cdot)$ as individual weights. In addition, we collect households' age and education level.

Employing this simulated data and the empirical evidence of Bartels and Pestel (2016), we use the method of indirect inference to calibrate the variance σ_ξ^2 of participation costs ξ. In particular, we run the following regression on our simulated data

$$switch = b_0 + b_1 \Delta_{PTR} + b_2 Age_{35-44} + b_3 Age_{45-54} + b_8 College + \epsilon$$

and target a participation elasticity b_1 of -0.106. Stetting σ_ξ^2 to 0.138 delivers a similar value. This means that the probability of switching from non-employment to employment after reducing the pension contribution rate τ_p by 10 percentage points (from 0.1870 to 0.0870) increases by 1 percentage point. This change is substantial given a benchmark participation rate of 87% for the age group 24-54. Unlike in Bartels and Pestel (2016), coefficients on the age and college dummies are significant. However, this is not surprising given that the simulated data set features more than 1.6 million observations. Table 11 provides details on the estimation results from our simulated data.

Table 11: Effect of Δ_{PTR} on the probability of taking up work

Switch (U $\rightarrow$ E)	
Δ_{PTR}	-0.099
	(0.0175)
Age_{35-44}	-0.0055
	(0.0005)
Age_{45-54}	0.0223
	(0.0005)
$College$	-0.0023
	(0.0007)

Observations: 1,639,696, standard errors in parenthesis

D.4 Parameterizing Labor Productivity

This section provides further details on the calibration of labor productivity profiles and productivity risk as outlined in Section 5.3.2.

Normal labor productivity We first concentrate on normal labor productivity, meaning the labor productivity process of individuals with permanent state $m = 0$. Labor earnings and labor productivity are not identical when individual labor hours vary across ages and states, as they do in our quantitative model. Hence, we can not simply take the labor earnings estimates one for one. Instead, to calibrate the process of normal labor productivity, we proceed as follows: We assume the average labor productivity profile to evolve according to

$$\theta_{j,s} = b_{0,s} + b_{1,s}\frac{\min(j, j_{M,s})}{10} + b_{2,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^2 + b_{3,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^3.$$

This functional form is flexible enough to capture both a hump-shaped ($j_{M,s} = \infty$) and a stagnating ($j_{M,s} < j_R$) life-cycle labor productivity profile. Note that in the case of a stagnating profile, labor productivity is constant from age $j_{M,s}$ onward. We calibrate the coefficients of this polynomial such that our model implied average labor earnings profile for each education type matches its empirical counterpart. Figure 4 compares the empirical and model implied average earnings profiles.[33] The top panel of Table 1 in the main text shows the calibrated values for the polynomial coefficients $b_{i,s}$ and the stagnation thresholds $j_{M,s}$.

Next, we model residual labor productivity as an AR(1) process. In particular, we discretize the AR(1) process by a seven state Markov chain using a Rouwenhorst method, see Kopecky and Suen (2010). As autocorrelation parameter ρ_s we directly use the estimates from Table 9. We then calibrate the innovation variance $\sigma^2_{\varepsilon,s}$ such that the model implied variance of residual labor earnings equals its empirical counterpart, see Table 9. In doing so, we obtain a set of seven productivity realizations $\{\eta_{1,s}, \ldots, \eta_{7,s}\}$ as well as a transition matrix π^s that governs the transition between these seven normal productivity states.

Low labor productivity shocks The shock process for low labor productivity shocks follows the structure discussed in Appendix A.2.4. In particular, we assume that at the beginning of life ($j = 1$) a fraction ω^s_{low} of households with permanent state $m = 1$ starts in the low productivity state. The share $1 - \omega^s_{low}$ has normal labor productivity. Individuals transition between the state of normal productivity and a low productivity shock according to the transition matrix specified in (16). We take the estimates of the initial share of households as well as the transition matrix directly from our empirical findings as summarized in Table 10. When individuals draw the low labor productivity shock, they get assigned a labor productivity level of $\exp(\eta_0) = 0.0675$. This productivity level ensures that the average earnings of low productivity workers are equal to 10 percent of the average labor earnings of the total population, see the right panel of Figure 2.

Bringing the two processes together At the beginning of life, a fraction ϕ^s_m of households of education level s draws a permanent shock $m = 1$. These households face a labor productivity process that combines normal labor productivity with low productivity shocks. Households with $m = 0$, on the other hand, only experience a normal labor productivity process. We set the transition matrix between potential labor productivity states $\{\eta_0, \eta_{1,s}, \ldots, \eta_{7,s}\}$ to

[33] Note that, owing to the log-normal nature of labor productivity shocks, the model-implied average life-cycle wage profile is equal to

$$\exp\left(\theta_{j,s} + \frac{\sigma^2_s}{2}\right).$$

$$\pi_\eta(\eta^+|\eta,j,s,m)=\begin{bmatrix} m\pi^s_{low,1} & (1-m\pi^s_{low,1})\phi^s_\eta(1) & \dots & (1-m\pi^s_{low,1})\phi^s_\eta(7) \\ m\pi_{low,0} & (1-m\pi^s_{low,0})\pi^s_{11} & \dots & (1-m\pi^s_{low,0})\pi^s_{17} \\ m\pi_{low,0} & (1-m\pi^s_{low,0})\pi^s_{21} & \dots & (1-m\pi^s_{low,0})\pi^s_{27} \\ \vdots & \vdots & \dots & \vdots \\ m\pi_{low,0} & (1-m\pi^s_{low,0})\pi^s_{71} & \dots & (1-m\pi^s_{low,0})\pi^s_{77} \end{bmatrix}.$$

Hence, when being in the normal productivity state, households transition into the low productivity state η_0 with a constant probability $m\pi_{low,0}$, meaning 0 when $m = 0$ and $\pi_{low,0}$ when $m = 1$. Once they are facing low productivity, they stay in the low productivity state with probability $m\pi^s_{low,1}$. If they revert to normal productivity, they draw a regular productivity shock from the unconditional distribution $\phi^s_\eta(i)$.

At the beginning of life, individuals are distributed over the potential productivity levels $\{\eta_0, \eta_{1,s}, \dots, \eta_{7,s}\}$ according to the distribution

$$\pi_{\eta,20}(\eta\,|\,m,s)=\Big[m\omega^s_{low} \quad (1-m\omega^s_{low})\phi^s_\eta(1) \quad \dots \quad (1-m\omega^s_{low})\phi^s_\eta(7)\Big]. \tag{36}$$

Hence, those individuals who do not experience low productivity from the outset of their life draw an initial labor productivity from the unconditional distribution of the normal productivity process. Finally, individual labor productivity is given by

$$z(j,s,m,\eta_{i,s})=\begin{cases}\exp\left(\theta_{j,s}+\eta_{i,s}\right) & \text{if } i>0 \text{ and} \\ \exp\left(\eta_0\right) & \text{otherwise.}\end{cases}$$

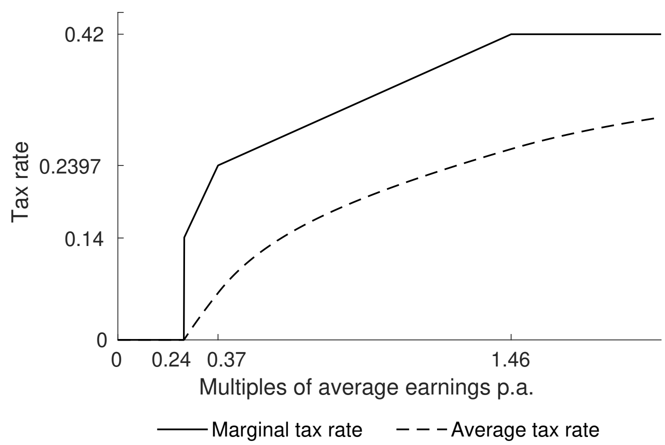
Agents with a low productivity shock consequently have a productivity level that is independent of age.

D.5 The progressive income tax code

We apply the German income tax code for the year 2017 to labor earnings and pension income. Individuals with earnings less than 0.24 times the average earnings are exempt from taxes. For earnings between 0.24 and 1.46 times the average, the marginal tax rate increases from 14 to 42 percent. For earnings exceeding 6.93 times the average, the top marginal tax rate of 45 percent is applied. Figure 16 shows the tax code $T_{I,t}$ in the case of individual taxation. However, about two-thirds of working-age German households consist of couples, as reported by RDC 2017. They enjoy a tax advantage in the form of income splitting. Hence, we set the splitting factor to 1.65. This results in

$$T_t\Big(y-T_{p,t}(y)+p\Big)=1.65\times T_{I,t}\left(\frac{y-T_{p,t}(y)+p}{1.65}\right).$$

Figure 16: Marginal and average tax rates for labor earnings and pension income



E Further Simulation Results

Figure 17: Employment changes and labor productivity

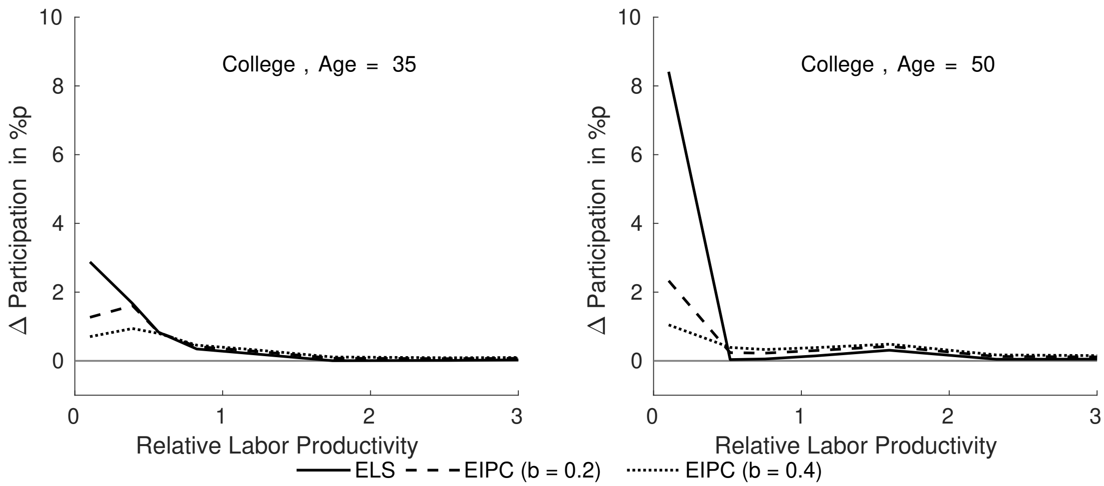


Figure 18: Intensive margin labor supply changes and labor productivity