DISCUSSION PAPER SERIES

DP16380

Progressive Pensions as an Incentive for Labor Force Participation

Fabian Kindermann and Veronika Pueschel

MACROECONOMICS AND GROWTH

PUBLIC ECONOMICS

ISSN 0265-8003

Progressive Pensions as an Incentive for Labor Force Participation

Fabian Kindermann and Veronika Pueschel

Discussion Paper DP16380
Published 21 July 2021
Submitted 16 July 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Progressive Pensions as an Incentive for Labor Force Participation

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. So far, the literature has argued that higher pension progressivity leads to more redistribution and insurance provision on the one hand, but increases implicit taxes and therefore distorts labor supply choices on the other. In contrast, we show that a well-designed reform of the pension system has the potential to encourage labor force participation. We propose a progressive pension component linked to the employment decision of households, which implicitly subsidizes employment of the productivity poor. A simulation analysis in a quantitative stochastic overlapping generations model with productivity and longevity risk indicates that this positive employment effect can be sizable and welfare enhancing.

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

Keywords: N/A

Fabian Kindermann - fabian.kindermann@ur.de
University of Regensburg and CEPR

Veronika Pueschel - veronika.pueschel@ur.de
University of Regensburg

Progressive Pensions as an Incentive for Labor Force Participation

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

Veronika Püschel
University of Regensburg

July 16, 2021

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. So far, the literature has argued that higher pension progressivity leads to more redistribution and insurance provision on the one hand, but increases implicit taxes and therefore distorts labor supply choices on the other. In contrast, we show that a well-designed reform of the pension system has the potential to encourage labor force participation. We propose a progressive pension component linked to the employment decision of households, which implicitly subsidizes employment of the productivity poor. A simulation analysis in a quantitative stochastic overlapping generations model with productivity and longevity risk indicates that this positive employment effect can be sizable and welfare enhancing.

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 as well as Hans Fehr 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. A progressive pension consequently redistributes between individuals of different permanent types and provides insurance against wage shocks. However, the literature has also highlighted that it comes at the expense of higher labor supply distortions.

In this paper, we challenge the conventional idea that progressive old-age pensions unanimously induce negative incentives on individual labor supply. In particular, we argue that a well-designed progressive pension system can actually encourage labor force participation and may therefore lead to more instead of less employment. The starting point of our analysis is a *proportional system*, in which old-age pension benefits are directly proportional to lifetime earnings. We modify this proportional system by adding an employment-linked component. Through this component, households acquire pension claims for every year they were employed, irrespective of how much they earned. Such an employment-linked component increases the progressivity of the pension system by allowing for an overproportional accumulation of pension claims for working households who are income poor. We refer to such a system as *employment-linked progressive system*.

We first explore the incentive effects of such a progressive pension system in a simple and tractable two-period model. In this model, young individuals have to make a binary employment decision (being employed or not) and, conditional on being employed, they decide about how many labor hours to provide. A government operates a pay-as-you-go pension system that collects payroll taxes in order to finance pension payments to retired households. The pension formula contains both a proportional and an employment-linked component. We show that increasing the weight on the employment-linked component at the expense of the proportional one decreases work incentives at the intensive margin. It weakens the link between earned income and individual pensions, and consequently a larger share of the household's pension contribution is perceived as a tax. However, along the extensive margin the employment-linked component implicitly subsidizes employment of productivity poorer households. As a result, the participation rate increases for all households with labor earnings less than average earnings.

The positive employment effect of an employment-linked progressive pension system has the potential to mitigate the negative labor supply distortions induced at the intensive margin. Conventional studies of the optimal progressivity of the pension system that solely focus on intensive margin labor supply, amongst them Huggett and Ventura (1999), Nishiyama and Smetters (2008), Fehr and Habermann (2008), and Fehr et al. (2013), might therefore overstate the negative labor market consequences of fiscal redistribution. Our ultimate goal is to evaluate the importance of the intensive and the extensive margin channel in a quantitative stochastic overlapping model. Yet, such a quantitative study also has to adequately describe the major dimensions of life-time inequality such as productivity and health.

To this end, we use administrative data from the German pension insurance system to investigate the properties of individual labor earnings dynamics over the life cycle. Concentrating on the male sample population, we find that individuals are exposed to a significant amount of earnings 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. The dynamics of low earnings episodes over the individual life cycle are quite distinct across education groups. 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. 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. We estimate a first-order Markov process that captures the salient features of low income shocks over a household's working life and use our estimation results to inform our quantitative model.

A second major dimension of inequality relates to individual life expectancy. While individual life expectancy has increased substantially for younger cohorts, a recent literature documents that the increase in life expectancy is not equally distributed within cohorts, see for example Meara et al. (2008), Mackenbach et al. (2015), de Gelder et al. (2017), Waldron (2007) or Cristia (2011). Differences in life expectancy along the education and income distribution alter the implicit rate of return an individual can expect from its pension contributions. The size of pension payments typically is related to the average life expectancy of all pensioners. Hence, an individual with an unusually low life expectancy typically makes a low return on pension contributions, and vice versa for those with a high individual life expectancy. When life expectancy correlates positively with labor earnings, a progressive pension system can undo such inequality and leads to a more balanced distribution of internal rates of return across pensioners.

The above theoretical and empirical considerations eventually inform our quantitative analysis. We provide an assessment of the strength of labor supply distortions, the positive employment effect as well as the redistribution and insurance benefits of an employment-linked progressive pension system. In our simulated stochastic overlapping generations model, households decide about consumption and savings, about whether to participate in the labor market and, if so, about how many hours to work. They do so under the presence of persistent shocks to labor productivity, additional low productivity shocks, longevity risk as well as shocks to individual life expectancy that correlate with education and labor productivity. A government collects progressive taxes on labor earnings to finance government expenditure, and operates a pay-as-you-go pension system that is financed by payroll taxes. We calibrate this model to the German economy which, in its status quo, operates a purely proportional system. This economy works as a nice benchmark for introducing progressive pensions, as it allows us to start from a pension system that minimizes labor supply distortions. When calibrating the benchmark equilibrium of our model, we pay particular attention to the calibration of the labor productivity process and the labor supply decision, both at the intensive and the extensive margin. In particular, we choose the relevant model parameters such that our model is able to replicate empirical evidence on the participation reaction to changes in wages.

With this calibrated simulation model, we explore the effects of introducing an employment-linked component into the pension formula. We find that the positive effects on employment can be sizable. Particularly older workers are encouraged to participate more often in the labor market, and especially so when they currently experience adverse shocks to their labor productivity. In fact, for the productivity poorest 50-year old in the economy, the employment rate increases by up to 6 percentage points. However, along the intensive margin, the labor supply decision of households is mostly distorted downwards, leading to a decline of about 0.5 to 1.6 hours per week. The introduction of an employment component compresses the distribution of pension entitlements for the retired population. It therefore provides more redistribution – between ex ante identical but ex post heterogeneous households – and more insurance against labor productivity shocks that individuals receive over their working life. Furthermore, it levels the internal rates of return between individuals with different life expectancies.

We compare the employment-linked progressive system with an alternative *basic progressive pension system*. While the employment-linked system grants pension payments based on a worker's employment status, the basic pension system provides an unconditional minimum income at retirement. We find that the two systems have quite distinct effects. While the employment-linked progressive system encourages employment of the productivity poor, the basic pension system discourages it. In addition, the basic pension system inflates pension payments, as it also pays benefits to the non-employed population, a group otherwise excluded from the system. This results in higher payroll taxes which further distort economic activity.

We finally conduct a welfare analysis. We therefore compare the consumption equivalent variation between an initial equilibrium with a purely proportional pension system and a new long-run steady state with a pension system that awards half of the benefits proportionally and the other half based on either an employment-linked or a basic pension component. In our benchmark simulations, an employment-linked system increases long-run welfare by 0.22 percent of aggregate consumption, about 4 billion Euros per year. This system strikes a balance between redistribution, insurance, the equalization of rates of return, positive employment incentives and negative distortions at the intensive margin. A basic pension, on the other hand, deteriorates long-run welfare, and the welfare losses are substantial. Under reasonable degrees of individual risk aversion, the simulation model predicts that the welfare gains from employment-linked progressive pensions can be even more sizeable, in the order of at least 1 percent of aggregate consumption.

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 (2021).

Huggett and Ventura (1999) were among the first to evaluate the welfare implications of redistributive social security. They compare the current US social security system with a two-tier pay-as-you-go pension. The first tier consists of a perfectly earnings related pension, while the second tier provides some minimum income at retirement. They find that the current US system is to be favored of the two-tier system. Nishiyama and Smetters (2008), Fehr and Habermann (2008), and Fehr et al. (2013) study the optimal design of the social security benefit formula. Similar to Huggett and Ventura (1999), they allow for a policy mix between earnings related and redistributive pension components. All of them find that a social welfare maximizing policy maker should implement a redistributive pension system. However, they have in common that they only look at the intensive margin labor supply decision of households. Hence, they stress the trade off between redistribution and insurance provision on the one hand, and negative labor supply incentives on the other hand.

Studies like Sánchez-Martín (2010) or Wallenius (2013) point to the fact that social security reforms can also have an impact on extensive margin labor supply choices. Yet, they only look at a households' optimal decision to retire. In allowing for employment decisions over the entire working life, our paper also speaks to the literature that studies employment incentives, often through the EITC or other fiscal policy measures targeted to the poor. Examples include Chan (2013), Athreya et al. (2014), and Ortigueira and Siassi (2019). For the German case, Bartels and Pestel (2016) identify variations in participation taxes over time and use them to empirically quantify the employment reaction to such incentives.

Finally, our paper is also 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), or Bagchi (2019) 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: In Section 2, we investigate the basic economic mechanisms at work in a tractable, two-period analytical model. Section 3 points to important dimensions of life-time inequality using administrative data from the German pension system. In Section 4, we present our full quantitative simulation model, and discuss its calibration in Section 5. In Section 6, we present simulation results for life-cycle choices, macroeconomic performance and long-run welfare. The last section concludes.

2 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 $j = 1, 2$. 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.[1]

2.1 The Household Decision Problem

Households can supply labor only in the first period of life, in the second period they are retired. 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.

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

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 mainly governed by the Frisch elasticity of labor supply χ. The latter kicks in through utility costs of employment ξ, which emerge whenever an individual is employed ($e = 1$).

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

$$c_1 + \frac{c_2}{1+r} = (1-\tau_p)wze\ell + \frac{p}{1+r} + R, \tag{2}$$

where R denotes some unearned income. Households pay contributions to the pension system in the form of a payroll tax τ_p on their total labor earnings $wze\ell$. As a reward, they receive a pension payment p when retired. Note that households only have the capacity to earn income by providing hours ℓ if they formally joined the labor force ($e = 1$).

2.2 Implicit Taxes and Participation Subsidies

Pension payments p are due to two components: First, the household's employment status in period 1 and second, her individual labor earnings. Specifically, we let

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

κ denotes the replacement rate of the pension system and $\bar{y}$ average labor earnings of the employed. When being employed, households receive a fixed pension reward for employment, which is indexed to average earnings and independent of the households own income position, plus an earnings-tied pension. 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 $\lambda = 0$, pension claims are purely earnings related. We call such a system a *proportional pension*. If $\lambda > 0$, we call the pension system *progressive*. Note, however, 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 ($e = 0$). We therefore also call the system *employment-related*.

In the following, we deliberately assume that the population growth rate of the economy, which defines the implicit return on pension contributions, is equal to the interest rate on financial investments, i.e. $r = n$. In the context of our model, this means that $\tau_p = \frac{\kappa}{(1+r)}$.[2] 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

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

The pension system influences the household budget constraint in two ways. On the one hand, it imposes an *implicit tax* on intensive labor supply $\tau_p^{\text{imp}} = \lambda\tau_p$. This implicit tax is equal to zero when the pension system is fully earnings related ($\lambda = 0$). In this case, any additional Euro a household contributes to the system pays the same return as a financial investment. Hence, contributing to the system is as valuable to the household as not contributing and saving the money in a private financial account. Yet, when we weaken the link between pension contributions and pension payments by setting $\lambda > 0$, the implicit tax rate of the pension system increases. In the extreme case where $\lambda = 1$, an increase in intensive margin labor supply has no effect on the size of the pension a household receives. 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 independent of individual income, but are linked only to the employment status of a household. 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.

2.3 Intensive and Extensive Margin Choices

We now take a deeper look at the household's labor supply problem and determine the incentive effects of an increase in pension progressivity more formally.[3] We start with the intensive margin labor supply decision of an employed household with productivity z. Maximizing utility in (1) subject to the household budget constraint (4) yields

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

In the absence of income effects, the intensive margin labor supply choice is immediately determined by individual productivity z as well as the implicit tax rate τ_p^{imp} of the pension system.

[2]In Appendix A, we show that our results also hold in a more general framework where $r \neq n$. The intuition is exactly the same, with the only difference that formulas get more complicated.

[3]All formal derivations can be found in Appendix A.

To make an employment choice at the extensive margin, the household has to compare her utility from working to the utility from not working. This utility difference is

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

Consequently, given the distribution of the utility costs of employment ξ, the probability that an individual with labor productivity z chooses to be employed is

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

The term in parentheses denotes the utility gain from working and marks the indifference point of households. Any individual with ξ below this utility gain chooses to be employed, anyone with ξ larger than the respective utility gain chooses to not be employed. Total labor supply of all households with labor productivity z consequently is

$$h(z) = P(e=1|z) \times \ell(z|e=1).$$

2.4 The Incentive Effects of Progressive Pensions

Equipped with the solution to the household's labor supply choice problem, we want to study how a change in pension progressivity λ impacts on the intensive and the extensive labor supply decision of a household. Taking the derivative with respect to λ, we immediately obtain

$$\frac{\partial \ell(z|e=1)}{\partial \lambda} = -\tau_p \times \chi \times \frac{\ell(z|e=1)}{1-\tau_p^{\text{imp}}} < 0. \tag{5}$$

As already argued above, an increase in pension progressivity leads to an increase in the implicit tax rate and therefore directly distorts labor supply on the intensive margin. The extent of this distortion is due to two factors: first, the size of the pension system as indicated by its contribution rate τ_p; second, the elasticity χ that governs the reaction to changes in the price of labor.

Regarding the employment decision, we find that

$$\frac{\partial P(e=1|z)}{\partial \lambda} = \tau_p \times \phi_\xi(\cdot) \times \left[\bar{y} - wz\ell(z|e=1)\right]. \tag{6}$$

This derivative again depends on the size of the pension system τ_p and on the extent to which individuals react to changes in participation incentives. The latter is determined by the density $\phi_\xi(\cdot)$ of households located exactly at the indifference point between employment and non-employment. Most importantly, however, the degree to which pension progressivity incentivizes labor force participation depends on the relative labor market position of a household. All households who would earn an income below average labor earnings $\bar{y}$ are encouraged to be employed, while households earning more than $\bar{y}$ are discouraged. This is owing to the progressive nature of the employment component that pays a relatively high subsidy to the earnings poor, but a relatively low subsidy to the earnings rich.

Summing up, the proposed employment-linked progressive pension system affects aggregate labor in two ways. While it distorts labor supply along the intensive margin, it also provides incentives for employment along the extensive margin, especially 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 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 life-expectancies and by providing insurance against productivity fluctuations over the life cycle. In Section 4 we quantify the importance of labor supply distortions, redistribution and insurance for aggregate welfare in a quantitative simulation model. In order for this exercise to deliver credible results, we however have to provide a proper model of the important dimensions of life-time inequality. Before diving into the quantitative setup, we therefore document some facts on inequality along the life cycle.

3 Facts about life-time inequality

In this section, we document salient facts on different dimensions of inequality and risk that households face over their life cycle. We restrict our attention to mechanisms that we think have a first-order impact on old-age income and that could shape the need for redistribution through the pension system. These facts will guide us in constructing our simulation model. Specifically, we first examine the statistical properties of the household's labor earnings process using adminstrative data. In a next step, we discuss the relationship between life-time income and individual life expectancy.

We use data from Germany in this discussion, 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). In addition, we document a strong positive correlation between individual life-time income and life expectancy. Meara et al. (2008), Mackenbach et al. (2015), de

Gelder et al. (2017), Waldron (2007) and Cristia (2011) find a similar relationship for various other countries.

3.1 Inequality in labor earnings

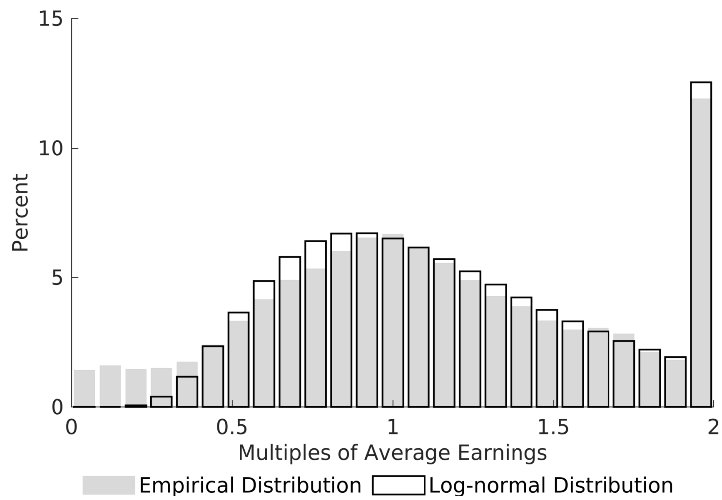
Providing a proper model for the household's life-cycle labor earnings process is crucial if one wants to assess the benefits of fiscal redistribution and insurance. To this end, we use administrative data from the German public pension insurance system to estimate life-cycle labor earnings profiles and earnings risk. Our 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.[4] 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. Hence, they are a good indicator for estimating earnings processes.

We restrict our attention to the male sample population aged between 25 and 60 of which we have information on the education level. To avoid confounding factors from German reunification, we only include observations starting from the year 2000 up to 2016. 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. Observations are excluded from the sample if individuals had a full year of zero earnings months. This selection procedure leaves us with a total of 15,242 individuals and 189,184 annual earnings observations. Appendix B.1 provides more information on the data set and sample selection.

Figure 1 shows a histogram of raw individual annual earnings (gray bars) expressed as multiples of average labor earnings of the total population. The figure reveals two salient features of the data: First, the data are top-coded at about two times 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

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

Figure 1: Histogram of pension claim distribution



earnings seems invalid.

To deal with this feature of the labor earnings data, we split the dataset in two parts for the remainder of the 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*. All those with earnings below the threshold are *low earnings individuals*. Low earnings individuals can be thought of as having some months of temporary unemployment or non-employment throughout a year or as being marginally employed (i.e. having a so-called mini-job).

3.1.1 The dynamics of normal earnings

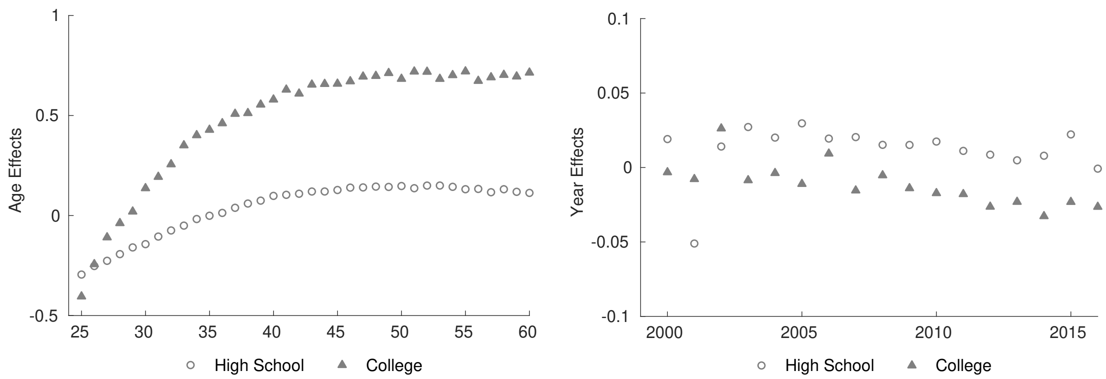
We describe the earnings dynamics of the normal earner sample by a standard AR(1) process in logs. We therefore split the normal labor earnings sample according to an individuals' education level $s \in \{0, 1\}$. $s = 0$ summarizes all individuals with high school education, while $s = 1$ indicates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

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

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. We use a method of moments estimator to determine the parameters of this model. We thereby control for the fact that the data are top-coded at two times the average earnings and truncated by the lower earnings threshold, see Appendix B.2 for more detail.

The left panel of Figure 2 visualizes the point estimates of the age fixed effects by education level. Up to the age of 45, earnings steeply increase for both education

Figure 2: Age fixed-effects and year fixed-effects



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 1 summarizes the estimation results for the residual earnings process. The parameter estimates are fairly standard. Both high school

Table 1: Estimates of residual log-earnings process

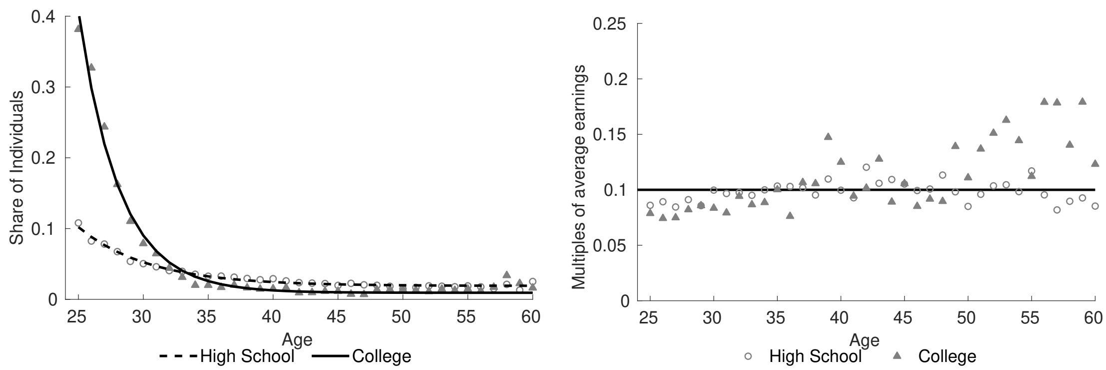
	High School $s = 0$	College $s = 1$
Autocorrelation $\hat{\rho}_s$	0.9869	0.9900
Innovation Variance $\hat{\sigma}^2_{\varepsilon,s}$	0.0046	0.0039
Unconditional Variance $\frac{\hat{\sigma}^2_{\varepsilon,s}}{1-\hat{\rho}^2_s}$	0.1780	0.1982

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.

3.1.2 The low labor earnings group

As discussed before, a simple log-normal distribution is not enough to capture the bimodal distribution of the earnings data. In a second step, we therefore examine the statistical properties of the low labor earnings sample. The left hand side of Figure 3 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

Figure 3: Life-cycle dynamics of low labor earnings



predominantly experience low 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. Labor earnings of individuals in the low earnings group are by and large independent of age and education.[5] The right panel of Figure 3 shows the age-education-earnings relationship of the low earnings segment of the population. For all ages and education types, average earnings of the low earnings group is approximately equal to 10 percent of average labor earnings. With average earnings amounting to roughly 37,000 Euros a year, the typical low earnings individual makes 3,700 Euros a year, or 308 Euros a month.

We interpret the findings in Figure 3 in the following way: Following empirical evidence from the labor literature that starts with Hall (1982), we assume that individuals face different degrees of career stability. While some exhibit stable career paths, others frequently transition into and out of employment.[6] We model

[5] 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.

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

career stability as a one-time discrete shock $m \in \{0, 1\}$ that an individual draws at the beginning of working life. The probability of drawing $m = 1$ is denoted by ϕ_m. While individuals with $m = 0$ face a stable career path and consequently never experience a low earnings episode, those with $m = 1$ may transition into and out of low earnings throughout their entire working life. In our benchmark estimation we set $\phi_m = 0.5$, meaning that 50 percent of the population experiences unstable careers. We provide sensitivity checks in Section 6.7.2.

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

$$\Pi^s_{low} = \begin{bmatrix} 1 - \pi^s_{low,0} & \pi^s_{low,0} \\ 1 - \pi^s_{low,1} & \pi^s_{low,1} \end{bmatrix}. \tag{7}$$

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 ω^s_{low} of individuals of education s starts as a low earnings individual. For each particular choice of the parameters ω^s_{low} , $\pi^s_{low,0}$, $\pi^s_{low,1}$, we can then calculate the model-implied share of individuals in each age-education bin that is a member of the low earnings group (see the left panel of Figure 3). This gives us a total of 72 moments which we use to estimate the six free parameters of our statistical model: ω^s_{low} , $\pi^s_{low,0}$, $\pi^s_{low,1}$ for $s = 1, 2$, see Appendix B.3 for details.

Table 2 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 3 indicate the model's predicted share of households in the low earnings group. As noted

Table 2: Estimates of low-earnings transition process

	High School $s = 0$	College $s = 1$
Initial share of low income earners ω^s_{low}	0.2040	0.8136
Probability to transition to low earnings $\pi^s_{low,0}$	0.0063	0.0051
Probability to stay low income earner $\pi^s_{low,1}$	0.8399	0.7324

above, 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 6.24 years for high school workers and 3.70 years for the college educated, respectively.

Summing up, the investigation of the labor earnings process of individuals in our administrative dataset 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 provided 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 (2021), 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.

3.2 Inequality in life expectancy

From the perspective of the pension system, inequality in earnings and earnings risk is not the only factor that can justify redistributive elements. While individual life expectancy has increased substantially for younger cohorts, a recent literature also documents that the increase in life expectancy is not equally distributed within cohorts. Meara et al. (2008) show that the decline in mortality rates at older ages in the US in between 1980 and 2000 can almost exclusively be attributed to a rising life expectancy of highly educated individuals. For the lower skilled, life expectancy has stagnated in the same time period, leading to a 30 percent increase in the longevity-education gap. Mackenbach et al. (2015) and de Gelder et al. (2017) find similar dynamics in individual life expectancy for selected European countries. Yet, it is not only education that correlates with life expectancy. Waldron (2007) uses data from the US social security system to calculate life expectancy at age 65 for the cohorts born in 1912 and in 1941. While for the lowest income group life expectancy of the 1941 cohort is only about half a year greater than that of the 1912 cohort, this difference amounts to 5.6 years for the highest income group. Cristia (2011) supports these findings.

Life expectancy is a major determinant of the internal rate of return an individual obtains from a public pension system, as those systems pay out annuity streams of income. The amount of payment an individual gets typically is related to the average life expectancy of all pensioner. Hence, an individual with an unusually low life expectancy makes a low return on pension contributions, and vice versa for those with a high individual life expectancy. Liebman (2002), Goda et al. (2011) and Bosworth et al. (2016) calculate the internal rates of return for individuals of different income groups in the US paying particular attention to the group specific life expectancy. All these studies find that the progressivity of the US system – that leads low income earners to get a higher replacement rate than high income individuals – is undone by the differences in life expectancy across income groups. In some cases, the internal rate of return is even lower for low income earners than for higher income groups.

In Germany, the relation between education or income and longevity is comparable to the international evidence. Luy et al. (2015) find individuals with college education to live on average 2.5 years longer than those with lower education levels. Haan et al. (2020) report a life expectancy gap of around 7 years between individuals in the top and the bottom life-time labor earnings decile using administrative data from the German pension insurance system. As the German pension system is fully earnings related, the differences in life expectancy along the income distribution lead the internal rate of return to be particularly low for low income individuals. Taking this into account, the German statutory pension system is in fact regressive, redistributing income from the income poor to the rich through the life expectancy channel. Consequently, Breyer and Hupfeld (2010) argue in favor of a more progressive pension formula that explicitly takes the earnings-longevity relationship into account and guarantees a constant internal rate of return along the income distribution.

Overall, this section has shown that there can be multiple reasons for having a progressive pension system. On the one hand, individuals are exposed to a significant amount of earnings risk, much richer than the typical AR(1) process for log-labor earnings would predict. Most importantly, individuals face a serious portion of low income episodes, which not only lowers their life-time earnings, but also makes them marginally attached to the labor force. On the other hand, differences in life expectancy along the education and income distribution alter the implicit rate of return an individual can expect from its pension contributions. Whether the potential benefits of redistribution, insurance and of equalizing individuals' rates of return outweigh the labor supply distortions inherent in any progressive pension system is an open question which we now address using a quantitative simulation model.

4 The Quantitative Simulation Model

In this section, we present our full quantitative simulation model, which is based on the previous theoretical considerations and informed by the empirical facts regarding income risk and longevity. We employ a general equilibrium overlapping generations model with population growth and 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 the progressive taxation of 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. Since we only consider long-run equilibria, we omit the time index t in the following wherever possible.

4.1 Demographics

The economy is populated by overlapping generations of heterogeneous individuals.[7] 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 under perfect competition. They hire both capital K at price r and labor L at price w on competitive spot markets. Firms operate a constant returns to scale technology

$$Y = \Omega K^{\alpha} L^{1-\alpha}. \tag{8}$$

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 (8). 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 $c_j \geq 0$, labor supply $\ell_j \geq 0$ and employment $e_j \in \{0, 1\}$. They maximize a

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

discounted expected utility

$$U_0 = E_0 \left[\sum_{j=1}^{J} \beta^{j-1} u(c_j, \ell_j, e_j) \right],$$

where expectations are formed with respect to both survival risk and idiosyncratic wage risk. Individuals discount the future with the constant time discount factor β and incur a utility loss from being employed.

Labor productivity risk Households are ex ante homogeneous, but differ ex post in their labor productivity $z(j, s, m, \eta)$. 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 $\theta_{j,s}$. Knowing their education level, households again 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 ($m = 1$), see Section 3.1.2. The probability to draw the state $m = 1$ is denoted by ϕ_m.

Throughout their working life, households' labor productivity is due to idiosyncratic shocks η. For individuals with a stable career ($m = 0$), we assume that their productivity follows a standard AR(1) process in logs

$$\eta^+ = \rho\eta + \varepsilon^+ \quad \text{with} \quad \varepsilon^+ \sim N(0, \sigma_\varepsilon^2), \tag{9}$$

where innovations ε^+ are iid across households. Following the evidence presented in Section 3.1, we augment this standard shock process by a persistent (but not permanent) *low productivity shock* for all individuals with unstable careers ($m = 1$). Those agents can be hit by a 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. 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 $w \times z(j, s, m, \eta)$.

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. 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 $y = wz(j, s, m, \eta)e\ell$, intergenerational transfers b,[8] as well as pension payments p. They use these resources to finance consumption expenditure $(1+\tau_c)c$ (including consumption taxes) and savings into the next period a^+, contributions to social security $T_p(y)$ as well as progressive income taxes $T\Big(y - T_p(y) + p\Big)$. 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

$$v(\mathbf{x}) = \max_{c,\ell,e,a^+,ep^+} \; u(c,\ell,e) + \beta\psi_{j+1,h} E\Big[v(\mathbf{x}^+) \;\Big|\; j,s,m,\eta,h\Big] \tag{10}$$

with $\mathbf{x}^+ = (j+1,s,m,\eta^+,h^+,a^+,ep^+)$. Households maximize (10) subject to the borrowing constraint $a^+ \geq 0$, the budget constraint

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

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.

4.4 The Pension System

In the benchmark economy, we consider a purely proportional pension system. The pension system collects payroll taxes at rate τ_p on all income below two

[8] 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.

times average labor earnings of the employed $2\bar{y}$. The pension contribution $T_p(y)$ of a household with labor earnings y hence reads

$$T_p(y) = \tau_p \times \min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big).$$

In reward for contributing to the system, households earn pension claims ep according to

$$ep^+ = ep + \min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big). \tag{11}$$

In the proportional system, earned pension claims are solely determined by earned income and, like pension contributions, capped at twice the average earnings. Finally, individual pension benefits $p(ep)$ are calculated from earned pension claims as

$$p(ep) = \kappa \times \frac{ep}{j_R - 20}.$$

The second factor in this equation reflects the average pensionable earnings of an individual and κ is the replacement rate.

The pension system operates on a pay-as-you-go basis. In equilibrium, total pension contributions hence need to be equal to the total amount of pension payments. Letting Φ denote the cross-sectional measure of households over the state space,[9] we require

$$\kappa \times \underbrace{\int ep \times \mathbb{1}_{j \geq j_R}\, d\Phi}_{\text{total pension claims}} = \tau_p \times \underbrace{\int \min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big)\, d\Phi}_{\text{contribution base}}. \tag{12}$$

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. Tax revenue is used to finance (wasteful) government spending. As we abstract from any government debt, the tax system is balanced whenever

$$\tau_c \times C + \int T\Big(y - T_p(y) + p\Big)\, d\Phi = G \ \text{ with } \ y = wz(j,s,m,\eta)e\ell. \tag{13}$$

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

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

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 employed by the domestic production sector are invested abroad at the international interest rate $\bar{r}$. The capital market equilibrium reads

$$K + Q = A,$$

where A are aggregate private savings and Q 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 $-nQ_{t+1}$. 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

$$TB = (n - \bar{r})Q. \tag{14}$$

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

$$b_j = \frac{\int \frac{1-\psi_{j,h}}{\psi_{j,h}} \times (1+r)a \; d\Phi}{\int \mathbb{1}_{j<j_R} \; d\Phi} \quad \text{if } j < j_R. \tag{15}$$

Given an international interest rate and the exogenous 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

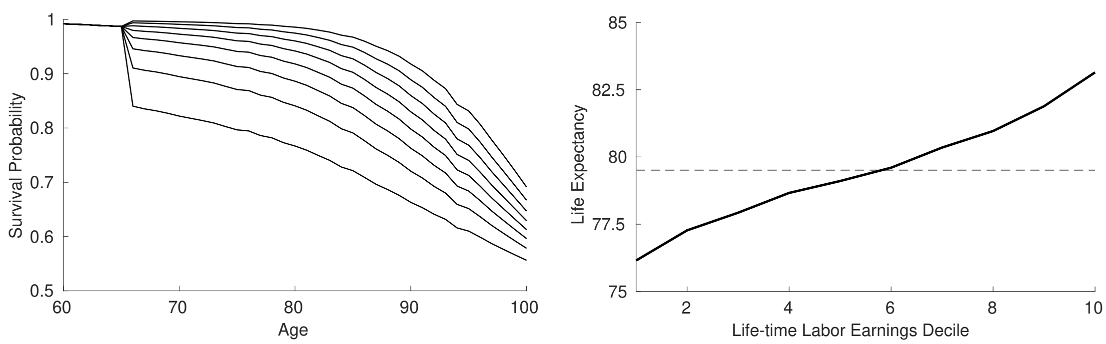
This section discusses our choices of functional forms and parameters in detail. We pay particular attention to the labor supply decision of households along the extensive and the intensive margin and to the relationship between life-time income and life expectancy. We calibrate our model to the German 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).[10] 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 different 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.[11] The left panel of Figure 4 shows the respective survival probability profiles.

Figure 4: Survival probabilities and life expectancy



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 labor productivity shock η at the date directly prior to retirement. This modeling choice is grounded on two pieces of empirical evidence: First, Luy et al. (2015) find that in Germany individuals with college education live on average 2.5 years longer than those with lower education levels. Second, Haan et al. (2020) report a life expectancy gap of around 7 years between individuals in the top and the bottom life-time labor earnings decile. In accordance with these findings, 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

$$p_{s,\eta} = \Phi\left(\iota_0 + \iota_1 \times \mathbb{1}_{s=\text{college}} + \iota_2 \times \eta\right), \tag{16}$$

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

[11] See Appendix D for more details on how we derive these profiles from the average survival probabilities.

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 $\iota_2 = 0.61$ to target the reported life expectancy gaps by education level and life-time labor earnings. Finally, we choose $\iota_0 = -0.06$ 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 4 shows the relation between life-time labor earnings and life expectancy. While individuals in the bottom decile expect their life to be about four years shorter than that of the population average, the average life of a top decile earner is three 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. While agents share a common set of survival probabilities during their entire working life, they still form expectations with respect to their potential health shocks 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. 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 is equal to 1. Finally, we assume an international interest rate of $\bar{r} = 0.03$, which constitutes as mix between the (currently) 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

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

We choose an intertemporal elasticity of substitution σ of 0.67.[12] The choice of σ has important implications for the size of the income effect of wage changes on

[12] In a model with inelastic labor supply, the implied risk aversion would then be equal to 1.5.

labor supply. Heathcote et al. (2014) estimate a similar value for this parameter in a life-cycle model using cross-sectional data on earnings and consumption from PSID and CEX. Our preferred value for the Frisch elasticity is $\chi = 0.6$, like in Kindermann and Krueger (2021). This medium range value represents a mix between the very low empirical estimates of the labor supply elasticity of men and the much higher elasticities of women, see for example Keane (2011). Finally, we set the time discount factor to $\beta = 0.98$ and chose the level parameter of intensive labor supply $\nu = 20.33$ so as to target a 40 hour work week for the employed.

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). As already indicated in equation (6), 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. If a large fraction of households is located directly at the threshold between not working and working, an increase in wages causes a greater fraction of households to switch from non-employment to employment.

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 $z(j, s, m, \eta)$. We let ξ follow a log-normal distribution with mean μ_ξ and variance σ_ξ^2. The mean $\mu_\xi = 0.77$ is set so as to target an employment-to-population ratio for the 25 to 54 year old of 88 percent, as reported in the Labor Force Statistics of the OECD (2020) for Germany. The variance σ_ξ^2 is chosen to target the evidence on participation choices in Bartels and Pestel (2016). In this study, the authors estimate the impact of changes in participation tax rates on the probability of moving from non-employment to employment. In a regression of these transition probabilities on changes in participation taxes, they find a coefficient of -0.106 (Table 2(2) in their paper). Setting $\sigma_\xi^2 = 5.75$ and estimating the effect of a change in the pension contribution rate[13] on the likelihood of switching from non-employment to employment, we obtain the same coefficient. Appendix D provides more details of the estimation procedure.

5.3.2 Labor Productivity

In Section 3.1, we already analyzed 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.

To provide a suitable calibration for the labor productivity process, we first set the share of college educated workers to $\phi_s = 0.2373$ and assume $\phi_m = 0.5$. We then take a subset of parameters directly from Tables 1 and 2. This includes the

[13]The pension contribution rate is the most important tax for low income households as it, unlike the income tax, has no exemption levels.

autocorrelation ρ_s of normal labor productivity risk as well as 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.

Next, we parameterize the age-productivity relationship using the functional form

$$\theta_{j,s} = b_{0,s} + b_{1,s}\frac{\min(j, j_{M,s})}{10} + b_{2,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^2 + b_{3,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^3. \quad (17)$$

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. 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 (28) 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 age fixed-effects for labor earnings in Figure 2;
2. the unconditional variance of normal labor earnings in Table 1;
3. average labor earnings of low productivity individuals as shown in the right panel of Figure 3.

Table 3 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 Appendix D.

5.4 Government Policies

We fix the pension contribution rate at $\tau_p = 0.186$, the current statutory rate of the German pension system. In equilibrium, our choice of τ_p results in a value of $\kappa = 0.5512$, the gross replacement rate of the system. This is higher than the current replacement rate in Germany (around 45 percent in 2017) as we (i) do not consider other pensions than old-age pensions (like for example disability pensions), (ii) assume that all individuals claim their pensions at age 64 and not before and (iii) use the survival probabilities of men who expect a shorter life than women.

In our benchmark economy, we fix government consumption at 19 percent of GDP. We set the consumption tax rate at $\tau_c = 0.18$, thereby acknowledging the fact that some goods are taxed at rates smaller than the regular consumption tax rate of

Table 3: Parameter values of labor productivity profiles and risk

	High School $s = 0$	College $s = 1$
Normal labor productivity		
Intercept $b_{0,s}$	-2.0732	-7.3497
Linear age term $b_{1,s}$	0.5981	4.3161
Quadratic age term $b_{2,s}$	-0.0570	-0.8465
Cubic age term $b_{3,s}$	0.0000	0.0562
Stagnation threshold $j_{M,s}$	∞	50
Autocorrelation ρ_s	0.9869	0.9900
Innovation Variance $\sigma^2_{\varepsilon,s}$	0.0054	0.0047
Low labor productivity		
Productivity level $\exp(\eta_0)$	0.1700	0.1700
Initial share of low productivity earners ω^s_{low}	0.2040	0.8136
Probability to transition to low productivity $\pi^s_{low,0}$	0.0063	0.0051
Probability to stay low productivity earner $\pi^s_{low,1}$	0.8399	0.7324

19% in Germany. Modeling the progressive labor tax is important, as progressive income taxation already implies some redistribution and insurance. Following Benabou (2002), we assume a progressive labor income tax function of the form

$$T(y) = y - (1 - \tau_0)y^{1-\tau_1}.$$

In this specification, τ_0 roughly resembles the average tax rate and τ_1 is a measure for progressivity. If $\tau_1 = 0$, the tax function collapses to a purely proportional one. A larger τ_1 means more redistribution across households of different income levels. As in Kindermann et al. (2020), we set $\tau_1 = 0.128$. We then choose $\tau_0 = 0.1435$ such that the government collects enough tax revenue to finance its expenditures. Tables 4 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 benchmark economy. We then turn to counterfactual policy simulations, in which we introduce progressive components into the pension formula.

6.1 The Benchmark Economy

Table 5 summarizes central macroeconomic aggregates of our benchmark economy with a proportional pension system as outlined in Section 4.4. Private savings are

Table 4: Summary of model parameters

Parameter	Value	Data/Target
Exogenous parameters		
Share college educated ϕ_s	0.237	German Statistical Office (2020)
Share unstable careers ϕ_m	0.500	Assumption
Average survival probabilities $\bar{\psi}_j$		HMD (2020)
Population growth rate n	0.000	German Statistical Office (2020)
Retirement age	64	DRV (2019)
Pension contribution rate τ_p	0.186	DRV (2019)
International interest rate $\bar{r}$	0.030	
Capital share in production α	0.300	Labor share of 0.70
Intert. elasticity of substitution γ	0.667	Heathcote et al. (2014)
Frisch elasticity of labor supply σ	0.600	Kindermann and Krueger (2021)
Tax progressivity τ_1	0.128	Kindermann et al. (2020)
Consumption tax rate τ_c	0.180	German Statistical Office (2020)
Government consumption G/Y	0.190	German Statistical Office (2020)
Endogenous parameters		
Depreciation rate δ	0.070	I/Y: 0.21
Technology level Ω	0.923	Wage rate: 1
Disutility of intensive labor ν	20.33	Working hours per week: 40.0
Mean disutility employment μ_ξ	0.77	Participation rate: 0.88
Variance disutility employment σ_ξ^2	5.75	Participation elasticity: -0.10
Health shock probabilities $P(h\|s, \eta)$		see Section 3.2
Replacement rate κ	0.551	Budget balancing pension system
Average labor tax rate τ_0	0.144	Budget balancing tax system

not enough to cover total capital demand. As a result, the economy exhibits a negative net foreign asset position of around 1.24 times GDP. On the goods market, this implies exports amounting to 3.67 percent of GDP to foreign countries. The government consumes 19 percent of GDP and 21 percent are invested into the future capital stock. The remainder is consumed by private households. The average work week of employed individuals between ages 25 and 64 amounts to 40 hours. This is above the German economy-wide average. However, in our model we consider single-earner households only. Hence, both population-wide participation and hours worked are greater than in an economy with two-earner couples, in which one partner might work part-time. The employment to population ratio (of the cohorts aged 25-54) is at 88.25 percent, and therefore similar to what we observe for male earners, see OECD (2020).

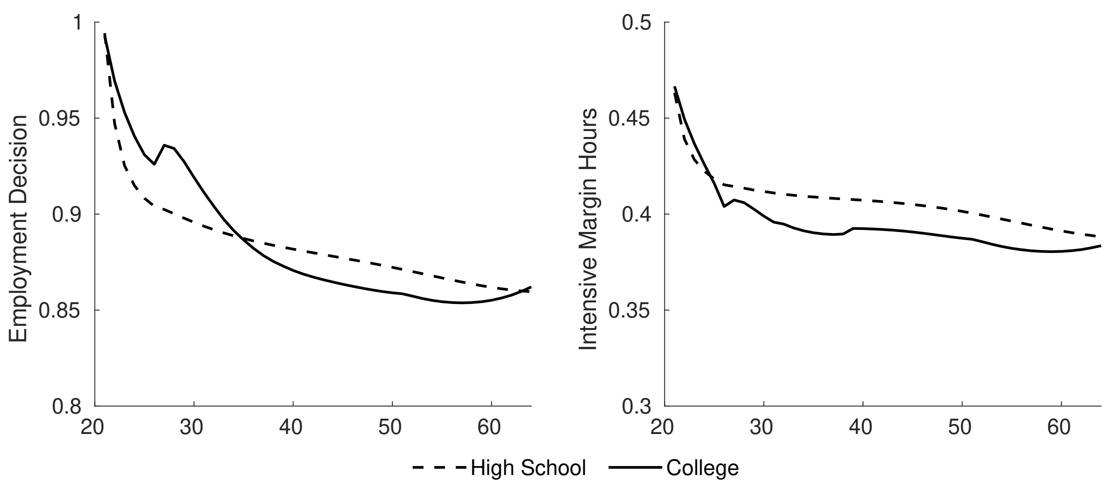
The left panel of Figure 5 shows the labor force participation profiles over the active working life of households. Households start their life with zero assets and only receive some intergenerational transfers in the form of accidental bequests. In addition, they expect their productivity to rise with age, which enforces liquid-

Table 5: Macroeconomic aggregates

Variable	Value
Private Savings	175.80
Capital Stock	300.00
Net Foreign Assets	−124.20
Private Consumption	56.33
Government Consumption	19.00
Investment	21.00
Trade Balance	3.67
Labor Tax Revnue	8.86
Consumption Tax Revnue	10.14
Average Work Week of Employed 25-64 (in hrs)	40.00
Employment-to-Population Ratio 25-54 (in %)	88.25

Variables in percent of GDP if not indicated otherwise.

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



ity constraints. Consequently, the employment share is almost equal to one 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 10 in Appendix D. As a result, labor force participation of the former drops faster than that of the latter at young ages. However, since college workers have a greater labor productivity on average, they have the chance to accumulate more wealth in their early working years. Consequently, the participation rate of college graduates is even below that of high school workers at higher ages.

The right panel of Figure 5 shows the labor hours a household provides conditional on being employed. After a phase of constrained liquidity at early ages, changes in productivity and wealth hardly have an impact on the average labor supply profile anymore. From age 30 onwards, the hours profile is almost flat. This is a consequence of our choice of preference parameters and the fact that households can completely drop out of the labor force. As a result, a substantial part of the variation in life cycle labor supply stems from the decision to be employed or not, and not from intensive hours changes.

6.2 The Thought Experiment and Incentive Effects

We now present results from counterfactual policy analyses that emerge from introducing a progressive component into the accumulation formula of the pension system. In particular, we calculate alternative economies that feature different degrees of pension progressivity as well as different styles of progressive pension systems. To ensure comparability between simulations, we use the same set of structural parameters, but fix per-capita government consumption as well as the replacement rate of the pension system at the benchmark economy's levels. The average tax rate in the labor tax system and the contribution rate of the pension system serve to balance the government's budgets.

The employment-linked progressive system The pension system we propose in this paper is an *employment-linked progressive system*, which is closely related to the system discussed in Section 2. Compared to the benchmark model, we modify the pension accumulation formula by adding an *employment component*. Pension claims ep^+ then evolve according to

$$ep^+ = ep + \Big[\lambda \bar{y} e + (1-\lambda)\min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big)\Big], \tag{18}$$

where $\bar{y}$ denotes the average labor earnings of the employed. For each year in which they are employed ($e = 1$), households receive pension claims of size $\lambda\bar{y}$ through the employment component, which is explicitely indexed to average labor earnings and not to individual income. In addition to employment, households are rewarded for higher contributions to the system through the *earnings component*, which is scaled with $1-\lambda$. The factor λ governs the weight on the two different components and defines the degree of progressivity of the pension system. A more progressive system incentivizes employment especially for the income poor population. It does so, however, at the expense of the earnings component.[14] Note that the benchmark system can be restored by simply setting $\lambda = 0.0$.

Increasing λ encourages employment, but it distorts intensive margin labor supply ℓ. This can readily be seen from the first-order condition for labor supply, which

[14]We assume that λ only changes the weight of the employment and the earnings component, and do not allow the pension system to increase or decrease in its overall size.

reads

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

The marginal disutility of providing an additional hour of work has to equal the marginal benefits. Providing an additional hour of work increases gross income by an amount $wz(j, s, m, \eta)$. This has an instantaneous benefit, as it allows the household to increase consumption, yet, only after paying contributions to the pension system and taxes. In addition, earning more has an impact on future pension income. The term $E\Big[v_{ep}(\mathbf{x}^+)\;\Big|j,s,m,\eta,h\Big]$ measures the utility value of accumulating additional pension claims. When λ increases and the pension system becomes more progressive, the link between earning more income and accumulating more pension claims is weakened and the return to providing additional working hours declines.

The basic progressive system To demonstrate the importance of linking the progressive part of the pension formula to the individual employment decision, we contrast our preferred pension reform to a *basic progressive system*. Unlike the employment-linked system, the basic progressive pension system provides a consumption floor to the entire population at old age. It hence allows individuals to accumulate pension claims even if they were not employed. In this counterfactual policy scenario, we assume that the pension accumulation formula reads

$$ep^+ = ep + \Big[\lambda\bar{y} + (1-\lambda)\min\Big(wz(j,s,m,\eta)e\ell,\ 2\bar{y}\Big)\Big]. \tag{19}$$

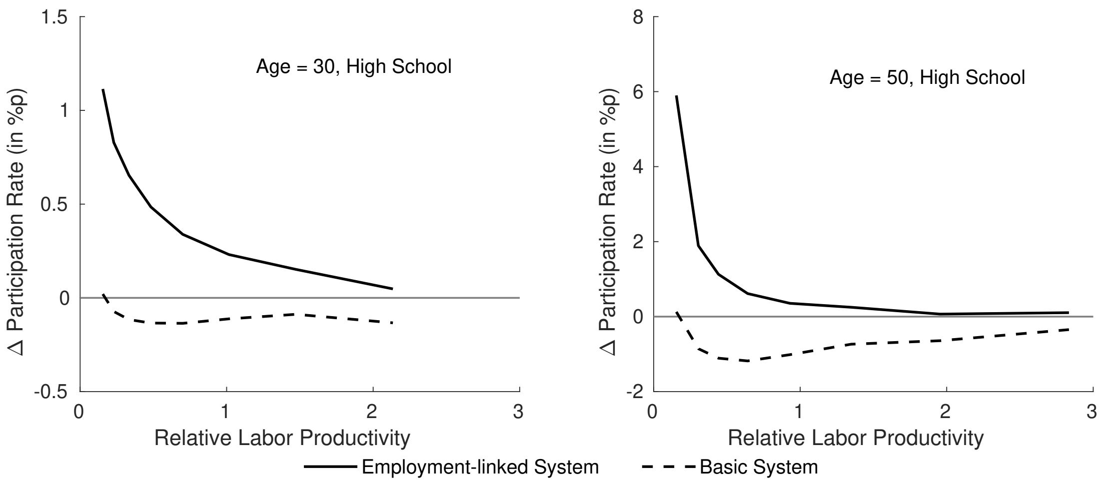
This formula is almost identical to the one in (18). The only difference is that progressive pension claims $\lambda\bar{y}$ are given to every individual, regardless of her employment status. It is quite easy to see that the introduction of such a basic pension system does not encourage employment for the productivity poor. Instead, it rather discourages employment unanimously by also awarding pension claims to individuals who are not active in the labor force.

In the next sections, we illustrate the effects of introducing these different progressive pension components into the statutory German economy. Throughout the entire analysis, we use the case of $\lambda = 0.5$, a medium range value for pension progressivity. We focus on the impact such progressive systems have on the life-cycle labor supply decision of households, the distribution of pension claims, the macroeconomy as a whole, and long-run welfare.

6.3 Labor Supply Effects of Pension Progressivity

Figure 6 summarizes the employment effects of progressive pension systems with $\lambda = 0.5$. 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 benchmark proportional system and the new progressive pension system in percentage points. We show the employment effects for 30- and 50-year old high-school workers. Results for the college-educated workforce are qualitatively identical and can be found in Figure 11 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



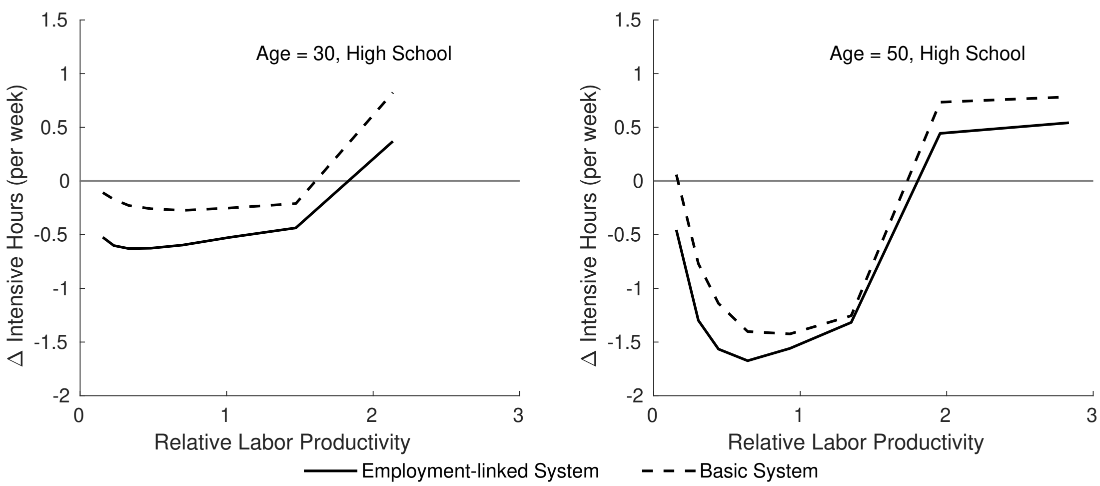
We first focus on the solid lines, which represent the results for the employment-linked progressive system. Introducing progressivity into the pension formula comes at sizable employment effects for the productivity poor. Regardless of age, all households with a less-than-average productivity experience an increase in labor force participation. At young ages, where employment is already very high and individuals do not have a lot of wealth, the employment effect is rather moderate. It is more pronounced though for the older working-age generations. For the productivity poorest 50-year-old individuals, employment even increases by a remarkable 5.9 percentage points. Around the age of 50, individuals have already accumulated some assets and, in the absence of an employment subsidy, would drop out of the labor force when their productivity is low. The potential for increasing labor force participation is therefore larger for these older workers.

Households with an above-than-average productivity hardly react to the reform in terms of employment. This is owning to the interplay between income and substitution effects. On the one hand, an increase in pension progressivity lowers the returns to working for productivity rich households. The amount of pension claims they receive for being employed under a progressive system ($\lambda = 0.5$) are typically smaller than what they would get under a proportional system ($\lambda = 0.0$). On the other hand, this cut in pension benefits induces a negative income effect and therefore fosters employment. In net terms, the two effects approximately cancel out.

The dashed lines in Figure 6 display the simulation results for a basic progressive system. The basic progressive system assigns pension claims to the entire working age population and therefore guarantees a comprehensive consumption floor at old age. As argued in the previous section, such a progressive system does not encourage employment for the productivity poor. Instead, it rather discourages employment unanimously as pension claims are also given to the non-employed. Consequently, employment of the productivity poor population drops by as much as 1.2 percentage points.

Figure 7 shows the intensive margin labor supply reactions to increased pension progressivity.[15] 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 productivity-poorer population, pension progressivity comes with negative labor supply incentives at the intensive margin. As argued in Sections 2 and 6.2, this is because a larger λ weakens the link between individual earnings and accumulated pension claims. As a result, a larger fraction of the contribution to the pension system is perceived as a tax, which distorts intensive margin hours. This effect is, however, only present for individuals with incomes 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. For individuals above the contribution ceiling, there consequently are no negative labor supply distortions. Instead, these agents face a decline in expected old-age income as a result of increased pension progressivity. This constitutes a negative income effect, which in turn leads to an increase in intensive margin hours. Note that the

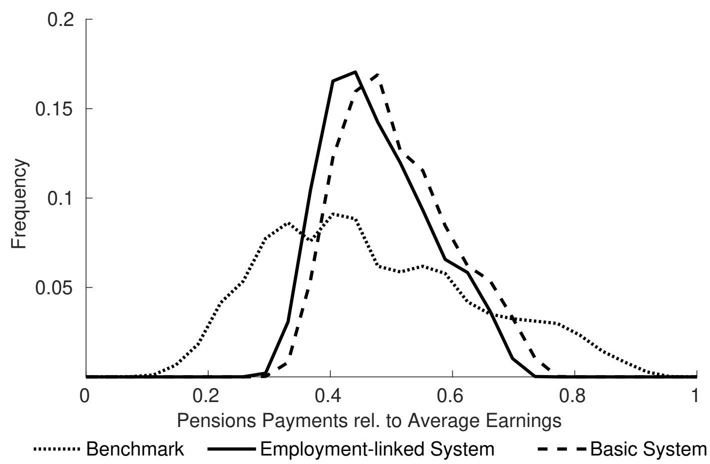
[15] Results for the college educated can be found in Figure 12 in Appendix E.

intensive labor supply effects are quite similar for the employment-linked and the basic progressive system, as they both distort the first-order condition for labor supply in the same way.

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

Figure 8: Distribution of pension claims



replacement rate is $\kappa = 0.551$ in all three scenarios. However, the mode of the pension payment distribution is somewhat lower at around 0.44. 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 dotted line in Figure 8 displays the distribution of pension payments in the benchmark equilibrium. As pension claims are perfectly earnings related, this distribution is closely linked to the lifetime earnings distribution of households. Under an employment-linked progressive system, the distribution of pension claims is much more concentrated, see the solid line in Figure 8. In fact, the mass of individuals with a pension of less than 30 percent of average earnings shrinks to almost zero. The dashed line finally indicates the distribution of pension payments under the basic progressive system. Under such a system, the fraction of "pension poor" households decreases even further, as interruptions in the individual earnings history play only a minor role for accumulating pension claims.

6.5 A Macroeconomic Evaluation

The long-run macroeconomic effects that result from introducing progressive components into the pension formula are summarized in Table 6. The first column shows the changes in macroeconomic quantities and tax rates under an employment-linked progressive system. Overall, the effects are moderate. Private savings increase by about 2 percent, which is a result of reduced pension payments to the income rich and hence more need for old-age savings. Note that there is an opposing effect on private savings, though. By providing insurance against unlucky labor productivity draws, a progressive pension reduces the need for precautionary savings. Yet, the second effect is quantitatively smaller. As total labor supply falls, the capital stock has to decline as well to ensure a constant capital to labor ratio. The result is a decline in the net foreign asset position of −6.78 percent.

Table 6: Macroeconomic effects of increased pension progressivity

	Progressive Component	
Variable	employment-linked	basic
Private Savings	2.03	−4.43
Capital Stock	−1.62	−2.17
Net Foreign Assets	−6.78	0.92
Total Intensive Labor Hours	−2.55	−2.025
Employment	0.41	−0.48
GDP	−1.61	−2.17
Private Consumption	−1.62	−3.12
Average labor tax rate (in %p)	0.48	0.74
Pension contribution rate (in %p)	−0.04	1.25
Aggregate pension payments	−1.82	4.41

Table reports percentage changes over initial equilibrium values if not indicated otherwise.

On the labor market side, the introduction of an employment-linked progressive pension comes with a decline in total intensive labor hours of the employed of around 2.5 percent. However, owing to the positive labor supply incentives on the productivity poor, total employment increases by 0.4 percent. The reduction in aggregate labor and capital causes a drop in GDP and private consumption. On the fiscal side, the decline in total labor earnings needs to be compensated by a higher average labor tax rate. However, the pension contribution rate slightly decreases because of the employment gains at the lower end of the earnings distribution.

The picture looks quite different if we instead introduce a basic progressive system which awards pension claims to the non-employed. A basic progressive pension induces a much larger drop in private savings. The system has a greater insurance value, as it also provides insurance against not being employed. This reduces the need for precautionary savings even further, so that aggregate savings decline by 4.4 percent. Since the capital stock only falls by 2.2 percent, the net foreign asset position increases by 0.9 percent.

Along the intensive margin, the basic progressive pension shows a similar impact than the employment-linked system. Along the extensive margin, however, there are remarkable differences. While an employment-linked progressive system increases work incentives especially for the poor, employment falls by 0.5 percent under the basic pension system. As already discussed in Figure 6, a basic pension system that allows the non-employed to accumulate pension claims sets purely negative employment incentives. The drop in aggregate employment leads to a weaker long-run macroeconomic performance. GDP drops by 2.2 percent and private consumption by even 3.1 percent. Finally, with respect to the fiscal sector, the decline in aggregate labor supply has to be compensated by a higher average labor earnings tax. The pension contribution rate increases by 1.25 percentage points. This is not only owing to the decline in labor earnings, but also to a substantial expansion in aggregate pension payments of 4.4 percent, see again Figure 8.

Summing up, our simulation results indicate that the macroeconomic consequences of a progressive pension system that features an employment-linked component are much more modest than those of a basic progressive pension system. Since under the latter the non-employed are eligible for pensions claims as well, the budget of the pension system would have to expand substantially, leading to a further distortion of economic activity through higher payroll taxes. On the side of the labor market, the positive incentive effects of the employment-linked component can even increase aggregate labor force participation, therefore dampening the reaction in aggregate labor supply. Of course, we expect these effects to also impact on aggregate welfare, which we illustrate next.

6.6 Welfare Analysis

We now evaluate the welfare effects of progressive pensions. To this end, we calculate ex-ante expected life-time utility EV before any information about the household's education level or labor productivity has been revealed. We then compare two steady state allocations: the benchmark scenario with a proportional pension system and utility level EV_0, and a scenario with a progressive pension system with an associated utility level EV_∞. To give the welfare numbers a meaningful interpretation, we calculate the consumption equivalent variation CEV between the two utility levels. 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 benchmark equilibrium in order to make them as well off as in a reform scenario with progressive pensions.

A negative value for CEV indicates that a reform of the pension system deteriorates long-run welfare and that households would be willing to pay a positive amount of resources in order to stay in the benchmark equilibrium.

The first row of the first panel in Table 7 shows the ex-ante welfare effects of an employment-linked and a basic progressive pension system. There are clear differences between the two: The employment-linked system sets employment incentives for poorer households who would otherwise not work. This limits labor supply distortions and substantially lowers the welfare costs of increased redistribution. As a result, introducing such a system into the economy comes at long-run welfare gains of 0.22 percent of aggregate consumption, about 4 billion Euros annually. As for a basic progressive pension, the negative distortionary effects on

Table 7: Welfare effects of increased pension progressivity

	Progressive Component	
Variable	employment-linked	basic
Benchmark Simulations		
Change in ex-ante long-run welfare	0.22	−1.45
– for high school educated households	0.55	−1.07
– for college educated households	−0.88	−2.73
Decomposition with respect to life expectancy		
– correlated with labor productivity (benchmark)	0.22	−1.45
– uncorrelated with labor productivity	−0.15	−1.79
Welfare benefit of equalizing returns	0.37	0.34
Sensitivity analysis		
– Career stability $\phi_m = 1$	0.21	−1.45

Table reports CEV over initial equilibrium in percent.

long-run welfare weigh much stronger. While such a system provides even more redistribution, it sets negative incentives on labor supply both at the intensive and at the extensive margin and it hinders wealth accumulation. In addition, the size of the pension system expands and the pension contribution rate increases, which further distorts economic activity. The welfare effects from a basic progressive pension consequently are negative and amount to a substantial -1.45 percent of aggregate consumption.

The second and third row of Table 7 compare the welfare consequences of progressive pension systems for individuals after their education level s has been revealed. In this figure, the redistributive properties of progressive pensions become immediately apparent. High-school workers, who face lower earnings and shorter lives, are the beneficiaries of progressive pensions. Under an employment-linked system, this group of households experiences welfare gains, because they receive both higher pensions and an employment subsidy. The college educated, on the other hand, are net-payers of a progressive pension reform and therefore experience welfare losses.

The second panel of Table 7 decomposes the long-run welfare effects with respect to the central assumptions about life expectancy. In the first row, we repeat our benchmark reform numbers. In the benchmark simulations, individuals receive a shock to life expectancy upon retirement, and this shock correlates with both education and labor productivity. As a result, individuals with lower education and life-time earnings systematically make a lower financial return on their pension contributions and a proportional pension system actually turns out to be regressive. By granting over-proportional pension claims to the education- and productivity poor, a progressive pension system equalizes the internal rates of return across the population and therefore generates welfare gains. To elaborate on the strength of this effect, the second row of the second panel shows the welfare results from a variant of our simulation model in which we assume away the correlation between labor earnings and life expectancy. Under this assumption, the welfare effects of both pension reforms are more negative. Taking the difference between the welfare numbers in the first and second row, we quantify the welfare benefits of equalizing rates of return to roughly 0.35 percent of aggregate consumption.

6.7 Sensitivity Analysis

In this section, we provide sensitivity checks with respect to two central elements of our quantitative model: individual risk aversion and the share of households with unstable career paths.

6.7.1 Risk Aversion

In our benchmark calibration, we use an intertemporal elasiticity of substitution of $\sigma = 0.67$. Following Swanson (2018), this implies a *relative consumption risk aversion* of $R_c = \frac{1}{\sigma+\chi} = 0.79$. While our choice of σ is quite standard for the macroeconomic literature and delivers plausible results for the uncompensated labor supply elasticity, the resulting degree of risk aversion is obviously very small. Hence, households' appreciation for insurance provision through the pension system is not very strong. We therefore see the welfare results of the previous section rather as a lower bound.

To check how sensitive the welfare results are with respect to assumptions about individual risk aversion, we use the generalized Epstein-Zin preference formulation

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

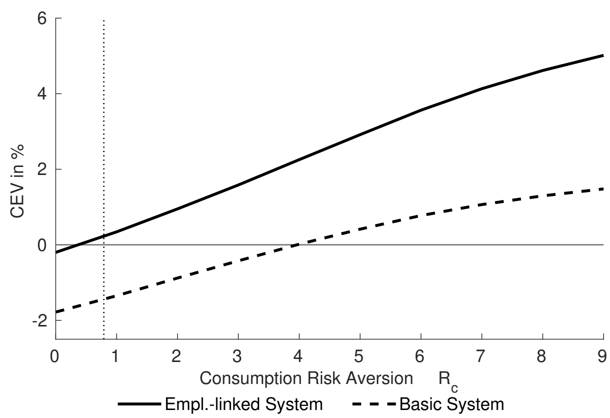
proposed by Swanson (2018). The advantage of this formulation is that it allows for a separate variation of the intertemporal elasticity of substitution σ and a risk aversion parameter γ. The relative consumption risk aversion for these preferences is approximately equal to

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

Setting $\gamma = 0$ brings us back to our benchmark case. A higher γ implies a higher degree of risk aversion.

Figure 9 shows the welfare effects of progressive pensions with a scale factor $\lambda = 0.5$ for different degrees of consumption risk aversion. The horizontal dashed

Figure 9: Sensitivity analysis



line marks our benchmark case. A higher risk aversion leads to additional welfare gains from insurance. While the welfare gains of an employment-linked progressive pension are quite small in our benchmark scenario, they are already sizeable (in the order of 1 percent) with a moderate value for risk aversion of 2. The picture looks quite different for the basic progressive pension system. There, we need a consumption risk aversion of 4 to turn the welfare effects positive, and even then they remain small.

While moving from our benchmark scenario to the right is indicative of the potential size of the welfare gains from redistribution, there is another point of interest in Figure 9. When we choose a value of $\gamma = -1.875$, individual risk aversion drops to zero. In this case, gains from redistribution are absent and the welfare effects from progressive pensions emerge solely from distortions and intergenerational redistribution. The employment-linked progressive pension comes with much less negative economic consequences than the basic pension. A major driver of this result is the fact that the former encourages employment of the productivity-poor population and therefore limits the distortionary impact of increased redistribution on economic performance.

6.7.2 Career Stability

The last row in Table 7 displays the welfare consequences of introducing progressive pensions for a different assumption about career stability. 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 (see Table 2) to guarantee consistency with the data. As the results in Table 7 reveal, the consequences for the welfare effects of progressive pensions are only minor.

7 Conclusion

When thinking about the incentive effects of a progressive component in the pension system, the details matter. Our analysis has shown that, in contrast to the traditional view of pension progressivity, a well-designed pension reform has the potential to increase employment, especially of the productivity poor.

Starting from a purely proportional pension system, we introduce progressivity into the pension formula by adding an employment-linked component. Pension claims are then rewarded (i) for being employed and (ii) for individual earnings. The progressive pension system distorts labor supply along the intensive margin by weakening the link between individual earnings and accumulated pension claims. However, it implicitly subsidizes employment of productivity-poorer households through the employment component. In addition, it compresses the distribution of pension entitlements when individuals enter retirement. As a results, an employment-linked progressive pension provides more redistribution – between ex ante identical but ex post heterogeneous households – and more insurance against labor productivity shocks with limited economic distortions.

We then compare an employment-linked progressive pension system to a basic progressive system that guarantees a consumption floor at old age to the entire population, and therefore also attributes pension claims to the non-employed. We find the two systems to have quite distinct effects. While the former system encourages employment of the productivity poor, the latter discourages it. In addition, a basic progressive pension system inflates the overall size of the pension system by also paying benefits to the non-employed population. This results in higher payroll taxes. Under a basic progressive system, the aggregate labor supply effects are much more negative and the economic performance is weaker overall.

A welfare analysis reveals that an employment-linked progressive pension system has the potential to increase long-run welfare, while a basic pension deteriorates welfare. When we calibrate our model to feature realistic degrees of risk aversion, the welfare consequences of an employment-linked progressive system can be sizeable.

8 Bibliography

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

Athreya, K., D. Reilly, and N. Simpson (2014). Uninsured idiosyncratic risk and aggregate saving. Richmond Fed Working Paper No. 14-11R.

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.

Benabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica 70*(2), 481–517.

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.

Bosworth, B., G. Burtless, and K. Zhang (2016). *Later Retirement, Inequality in Old Age, and the Growing Gap in Longevity Between Rich and Poor*. Washington, D.C.: Economic Studies at Brookings.

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.

Cristia, J. (2011). Rising mortality and life expectancy differentials by lifetime earnings in the united states. *Journal of Health Economics 28*, 984–995.

de Gelder, R., G. Menvielle, G. Costa, K. Kovács, P. Martikainen, B. H. Strand, and J. P. Mackenbach (2017). Long-term trends of inequalities in mortality in 6 european countries. *International Journal of Public Health 62*, 127–141.

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

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

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.

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.

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. *forthcoming: Journal of the Economics of Ageing*.

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.

Heathcote, J., K. Storesletten, and G. Violante (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review 104*(7), 1–52.

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.

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 (2021). High marginal tax rates on the top 1%? lessons from a life cycle model with idiosyncratic income risk. *forthcoming: American Economic Journal: Macroeconomics*.

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.

Liebman, J. B. (2002). Redistribution in the current u.s. social security system. In M. Feldstein and J. B. Liebman (Eds.), *The Distributional Aspects of Social Security and Social Security Reform*, pp. 11–48. University of Chicago Press.

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.

Mackenbach, J. P., I. Kulhánová, G. Menvielle, M. Bopp, C. Borrell, G. Costa, P. Deboosere, S. Esnaola, R. Kalediene, K. Kovacs, M. Leinsalu, P. Martikainen, E. Regidor, M. Rodriguez-Sanz, B. H. Strand, R. Hoffmann, T. A. Eikemo, O. Östergren, and O. Lundberg (2015). Trends in inequalities in premature mortality: a study of 3.2 million deaths in 13 european countries. *Journal of Epidemiology & Community Health 69*(3), 207–217.

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

Meara, E., S. Richards, and D. M. Cutler (2008). The gap gets bigger: Changes in mortality and life expectancy by education, 1981-2000. *Health Affairs 27*, 350–360.

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.

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

OECD (2020). OECD iLibrary. Available at https://www.oecd-ilibrary.org/.

Ortigueira, S. and N. Siassi (2019). The u.s. tax-transfer system and low-income households: Savings, labor supply, and household formation. Working Paper.

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

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

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

Waldron, H. (2007). Trends in mortality differentials and life expectancy for male social security-covered workers by socioeconomic status. *Social Security Bulletin 67*, 1–28.

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 Building Intuition: Solutions

Households solve the maximization problem

$$\max_{c_1,c_2,\ell,e} \; u(c_1,c_2,\ell,e) = c_1 + \frac{c_2}{1+r} - \frac{\ell^{1+\frac{1}{\chi}}}{1+\frac{1}{\chi}} - \xi e$$
$$\text{s.t. } c_1 + \frac{c_2}{1+r} = (1-\tau_p)wze\ell + \frac{p}{1+r} + R,$$
$$\text{and } p = \kappa \times \Big[\lambda \bar{y} e + (1-\lambda) wze\ell\Big].$$

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

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

A.1 The equilibrium pension system

For an equilibrium in this economy to exist, we require $r, n \geq -1$, which is not restrictive. Let us assume that labor productivity z was 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

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

The pension system collects pension contributions $\tau_p wze(z)\ell(z)$ from each employed households 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

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

The equilibrium replacement rate of the pension system hence is

$$\kappa = (1+n)\tau_p. \tag{20}$$

A.2 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 2, we assumes that $r = n$ and therefore $\varrho = 1$. However, we now want to prove our results more generally.

Using the relationship in (20), the household budget constraint becomes

$$c_1 + \frac{c_2}{1+r} = \Big[1 - \underbrace{\Big(1-\varrho(1-\lambda)\Big)\tau_p}_{=:\tau_p^{\text{imp}}}\Big]wze\ell + \underbrace{\lambda\varrho\tau_p\bar{y}}_{=:\tau_p^{\text{sub}}}e + R. \tag{21}$$

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 $\lambda > 0$.

A.3 Optimal choices

Using the budget constraint in (21), the household optimization problem becomes

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

$$\text{s.t. } c_1 + \frac{c_2}{1+r} = \Big[1-\tau_p^{\text{imp}}\Big]wze\ell + \tau_p^{\text{sub}}e + R.$$

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

$$
\begin{aligned}
& -\ell(z|e=1)^{\frac{1}{\chi}} + \left[(1-\tau_p^{\text{imp}})wz\right] = 0 \\
\Leftrightarrow \qquad & \ell(z|e=1) = \left[(1-\tau_p^{\text{imp}})wz\right]^{\chi}.
\end{aligned} \tag{22}
$$

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

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

As $\ell(z|e=0)=0$, we have $U(z|e=0)=R$ 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

$$
\begin{aligned}
P(e=1|z) &= P\left(\left\{U(z|e=1) - U(z|e=0) \geq 0\right\}\right) \\
&= \Phi_\xi\left(\frac{\left[(1-\tau_p^{\text{imp}})wz\right]^{1+\chi}}{1+\chi} + \tau_p^{\text{sub}}\right).
\end{aligned} \tag{23}
$$

A.4 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 (22) 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 (23) changes with λ according to

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

where the sign of the effect depends on the relative income position of the household. Recall that we can obtain the results in Section 2 by simply setting $\varrho = 1$.

B Datawork

The parameter estimates in our 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 information from the insurance accounts of 69,520 insured individuals. This is about 0.18% of the actively insured population. While we already provided some information on the estimation strategy in the main text of the paper, this appendix explains the data selection and estimation process in detail.

B.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. A 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.

B.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 $\min\left(y_{isjt},\ y_{max,t}\right)$. They are converted into pension claims y^p_{isjt} by diving them through average earnings $\bar{y}_t$. 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 $\bar{y}_t$.[16]

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

$$y^p_{isjt} = \frac{\min\left(y_{isjt},\ y_{max,t}\right)}{\bar{y}_t}. \tag{24}$$

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

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

B.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). We exclude data from 2017 because of a statistical break. The data-selection process is summarized in Table 8.

Table 8: 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		181,298
Observations on low earners		7,886

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 received pensions such as disability pensions or early-retirement pensions.

We divide the sample into two educational groups. The data set provides the variable `TTSC3_KLDB2010` which indicates an individual's highest degree in 2017 according to the classification of education scheme of the Federal Statistical Office of Germany (Klassifikation der Berufe 2010 - KldB 2010). 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[17] if he is classified ISCED 6 (Bachelor's or equivalent level) or above, excluding ISCED 65 (trade and technical schools, including master craftsman training). He is non-college-educated[18] if he is classified ISCED 5 and below or ISCED 65. We drop those with unknown education status.

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 9 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.

Table 9: 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

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 8. Our final data set is an unbalanced annual 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 3.1 and Figure 1, we split the sample into two sub-samples. The first one contains individuals with normal

[17] corresponds to KldB 2010 4-6

[18] corresponds to KldB 2010 1-3

labor earnings and the second one those with extraordinarily low earnings. An individual i is defined as a low earner in year t if he 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

$$\frac{125 \times 8 \times 8.5}{36,187} = 0.23. \tag{25}$$

Within our sample, 95.8% of observations are regular earnings and 4.2% are low earnings. We use observations from regular workers to estimate earnings profiles as shown in the left panel of Figure 2. Earnings estimates for low earners are shown in the right panel of Figure 3. A detailed description of the estimation strategy for both profiles follows in sections B.2 and B.3.

B.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.

B.2.1 Identifying the top censoring threshold

Our starting point is the data set of regular workers with 181,298 observations as summarized in Table 8. 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 (25), 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, our observations 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 $y_{max,t}$. Table 10 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 $y_{max,t}$.

B.2.2 Statistical Model and Moments

We describe the earnings dynamics for each education group s of the normal earner sample using the following statistical model

$$\log(y_{isjt}) = \kappa_{t,s} + \theta_{j,s} + \eta_{isjt} \quad \text{with} \quad \eta_{isjt} = \rho_s \eta_{isj-1,t-1} + \varepsilon_{isjt}. \tag{26}$$

y_{isjt} denotes labor earnings 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 and variance $\sigma^2_{\varepsilon,s}$. Furthermore, we let the stochastic process start from its long-run variance σ^2_s. This means that

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

Table 10: 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	0.9141	1.9017	9.0395	6,892
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	0.2962	2.1143	7.2195	8,103
2004	2.1266	0.6262	2.1261	7.6323	8,464
2005	2.1368	7.5118	2.1365	7.7027	8,906
2006	2.1360	7.3505	2.1358	7.4874	9,496
2007	2.1034	0.9543	2.1029	8.5793	10,164
2008	2.0767	1.0261	2.0763	9.1976	10,818
2009	2.1242	0.4140	2.1239	8.4647	11,353
2010	2.1192	8.6330	2.1191	8.6665	11,931
2011	2.0561	0.6732	2.0556	9.6705	12,626
2012	2.0362	9.5023	2.0361	9.6531	13,260
2013	2.0678	9.6354	2.0675	10.0744	13,440
2014	2.0687	0.7160	2.0683	10.2524	13,548
2015	2.0530	10.6683	2.0528	10.7195	13,676
2016	2.0560	0.7679	2.0553	11.6133	13,674
					181,298

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

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

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

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

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

where

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

B.2.3 Moment Conditions and Estimation

To estimate the statistical model in (26) 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^p_{isjt} 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 $\hat{shr}_{sjt}$ 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.

B.3 The Transition Process for Low Earnings Episodes

We model the transition out of and into low earnings episodes as a first order Markov process with a transition matrix as shown in equation (7). 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

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

We use the empirical counterparts to these shares $\hat{\Phi}_{j,s}$ shown in the left panel of Figure 3 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 parameter minimizes a simple residual sum of squares between the empirical and the model based moments $\Phi_{j,s}$.

C Simulation Model: Further Information

C.1 First-order conditions of the Household Problem

In the following, we describe the first-order conditions of the household problem. We use the generalized Epstein-Zin specification from Section 6.7.1. The results for the standard model can easily be recovered by setting $\gamma = 0$.

The dynamic household optimization problem reads

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

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

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

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

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

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

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

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

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

The first order condition for labor supply is

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

C.2 Recursive Competitive Equilibrium

Definition 1. *Given an international interest rate $\bar{r}$, government expenditures G, 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* (10) *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* (12) *and the tax system* (13) *hold, and accidental bequests are calculated from* (15).

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:*

$$K + Q = \int a \; d\Phi$$

   (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 (29). 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 $\mathcal{EP}$.

D Further Information on the Calibration Process

D.1 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 $h \in \{0, \ldots, 7\}$ 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

$$\psi_{j,h} = \frac{1}{1 + \exp(-\iota_h \times \bar{x}_j)} \quad \text{with} \quad \bar{x}_j = \log\left(\frac{1}{\bar{\psi}_j} - 1\right). \tag{27}$$

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.[19] The left panel of Figure 4 in the main text shows the resulting survival probability profiles.

D.2 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:

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

[19] 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 and exclude individuals with earnings below 33% of the marginal employment threshold of EUR 4,800.

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 5.75 delivers exactly this 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 0.1860 to 0.0860) increases by 1 percentage point. This change is substantial given a benchmark participation rate of 88.25% 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 800,000 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}	-0.1063
	(0.0126)
Age_{35-44}	-0.0062
	(0.0003)
Age_{45-54}	0.0019
	(0.0003)
$College$	0.1064
	(0.0126)

Observations: 817,061, standard errors in parenthesis

D.3 Parameterizing Labor Productivity

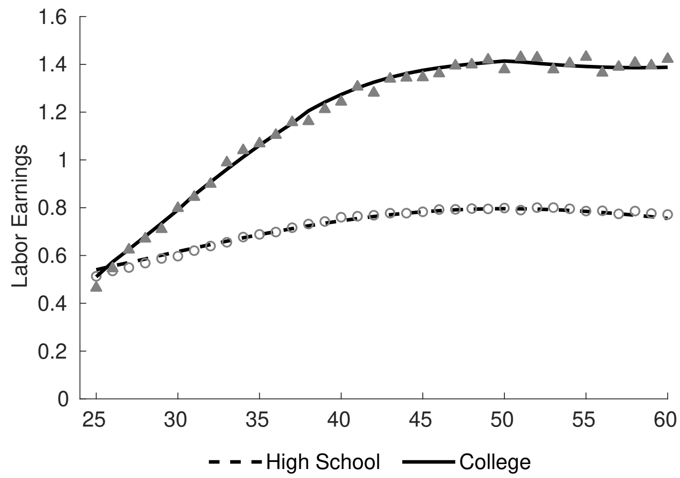
This section provides further details on the calibration of labor productivity profiles and productivity risk as outlined in Section 5.3.2.

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

$$\theta_{j,s} = b_{0,s} + b_{1,s}\frac{\min(j, j_{M,s})}{10} + b_{2,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^2 + b_{3,s}\left[\frac{\min(j, j_{M,s})}{10}\right]^3. \quad (28)$$

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 10 compares the empirical and model implied average earnings

Figure 10: Empirical and model implied average life-cycle earnings profiles



profiles.[20] The top panel of Table 3 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 1. 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 1. In doing so, we obtain a set of seven productivity realizations $\{\eta_{1,s}, \dots, \eta_{7,s}\}$ as well as a transition matrix π^s that governs the transition between these seven normal productivity states.

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

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

Low labor productivity shocks The shock process for low labor productivity shocks follows the structure discussed in Section 3.1.2. 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 (7). 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 2. When individuals draw the low labor productivity shock, they get assigned a labor productivity level of $\exp(\eta_0) = 0.17$. 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 3.

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

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

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 $\{\eta_0, \eta_{1,s}, \ldots, \eta_{7,s}\}$ according to the distribution

$$\pi_{\eta,20}(\eta \mid m, s) = \begin{bmatrix} m\omega^s_{low} & (1 - m\omega^s_{low})\phi^s_\eta(1) & \ldots & (1 - m\omega^s_{low})\phi^s_\eta(7) \end{bmatrix}. \tag{29}$$

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

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

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

E Further Simulation Results

Figure 11: Employment changes for the college educated ($\lambda = 0.5$)

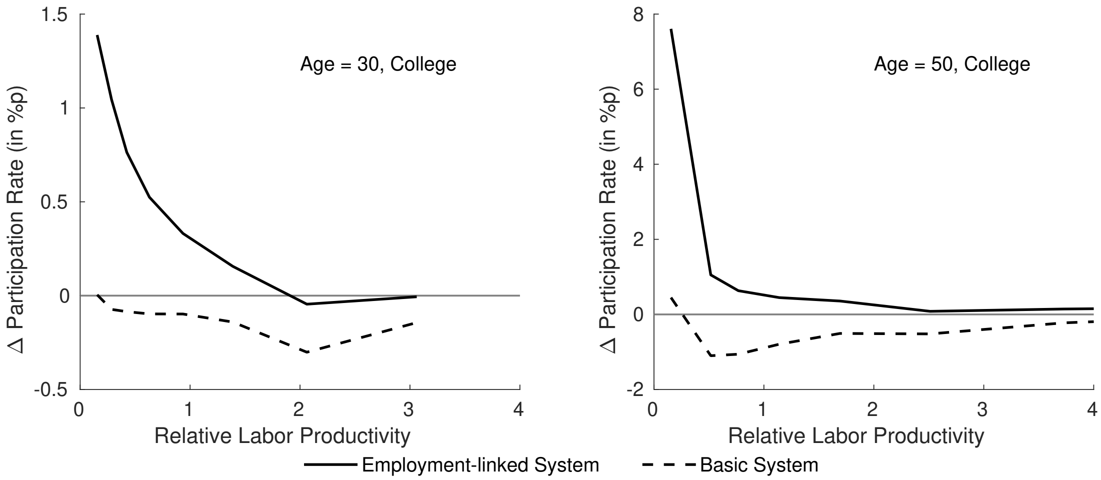


Figure 12: Intensive margin labor supply changes for the college educated ($\lambda = 0.5$)