Freie Universität Berlin

The Private and Fiscal Returns to Higher Education
A Simulation Approach for a Young German Cohort

Benjamin Fischer
Dominik Hügle

School of Business & Economics
Discussion Paper

Economics

2020/21

The Private and Fiscal Returns to Higher Education: A Simulation Approach for a Young German Cohort*

Fischer, Benjamin† Hügle, Dominik‡

December 15, 2020

Abstract

We quantify the private and fiscal lifetime returns to higher education in Germany accounting for the redistribution through the tax-and-transfer system, cohort effects, and the effect of income pooling within households. For this purpose we build a dynamic microsimulation model that simulates individual life cycles of a young German cohort in terms of several key variables, such as employment, earnings, and household formation. To estimate the returns to higher education, we link our dynamic microsimulation model to a tax-benefit simulator that allows converting gross wages into disposable incomes. On average, we find private and fiscal returns that are substantially higher than current market interest rates. However, analyzing the distribution of returns we also find that there is a considerable share of young adults for whom we forecast vocational training, the alternative to higher education, to be financially more rewarding. We demonstrate how the tax-transfer system and income pooling within couple households affect private returns and decompose the fiscal returns into its major components.

Keywords: Higher education, Returns to education, Dynamic microsimulation

JEL classification: C53, I23, I26

*The authors thank Sebastian Camarero, Ronny Freier, Astrid Pape, Ronnie Schöb, Viktor Steiner, and seminar participants at Freie Universität Berlin, the 7th World Congress of the International Microsimulation Association, the 33rd Annual Conference of the European Society for Population Economics, the BIEN Annual Conference 2019, the VfS Annual Conference 2019, the 5th LEER Conference on Education Economics, the SOEP User Conference 2018, and the Spring Meeting of Young Economists 2018 for valuable comments.

†School of Business and Economics, Freie Universität Berlin, e-mail: benjamin.fischer@fu-berlin.de

‡School of Business and Economics, Freie Universität Berlin, e-mail: dominik.huegle@fu-berlin.de

1 Introduction

In the last decades an ever increasing share of young adults have entered higher education across the globe.[1] As a consequence, the OECD average of the adult population between 25 and 64 with a tertiary education degree has increased from 20% in 1998 to 37% in 2018 (see OECD, 2020). It seems likely that this trend will be increasing further in the future. Going to university, however, requires a sizable financial investment of the individual in two dimensions. Directly, in terms of the tuition fees to be paid. But also indirectly, by the income foregone under the alternative career, which could either be a direct labor market entry or, in countries like Germany, entering a shorter and often remunerated vocational training. At the same time, higher education also implies a substantial investment of the state as academic tuition is usually highly subsidized by taxpayer money. In addition, the state forgoes taxes and social security contributions for the time an individual is in academic training. The return which individuals and the state reap from this investment comes later as academic graduates usually earn more than non-graduates (hence, pay more taxes and social security contributions) and have a lower risk of being unemployed (i.e. receive less transfers).

Estimating the earnings premium of education is often based on studying how economies rewarded schooling during the life paths of older cohorts (see Psacharopoulos and Patrinos, 2018, for an overview). However, for current decision makers, i.e. young adults and the state who is subsidizing higher education, the returns that can be expected for a young cohort are crucial, accounting for generational trends and institutional shifts. Therefore, the central goal of this paper is to forecast the distribution of private and fiscal returns to higher education for young adults in Germany.[2] We take the life cycles of the 1980s cohort as the basis of the simulation. We choose this particular cohort for two reasons. First, the individuals of the 1980s cohort are young enough to serve as a benchmark for individuals currently choosing between different educational paths and for public decision makers choosing the amount of higher education subsidization. And second, when measuring cohorts in ten-year intervals, it is the youngest cohort for which the vast majority of individuals has already finished post-secondary education, so that we can observe their early labor market biographies. In addition, crucial transitions in terms of marriage, divorce, and fertility have already occurred for this cohort.

Estimating "returns to education" has a long history in the field of empirical economics. The

[1] Note that we use the terms "higher education", "tertiary education", and "academic training" interchangeably throughout this paper.

[2] Note that our approach of *forecasting* returns to education of a young cohort is not equivalent to estimating *ex ante* returns according to Cunha and Heckman (2007). Cunha and Heckman (2007) define ex ante returns as the returns individuals are expecting at the time they make their educational choice, given their (restricted) information set. Importantly, the latter also includes information which is unobserved by the econometrician, such as preferences and skills. Courtioux et al. (2014) show under which (rather restrictive) assumptions the estimated returns can be interpreted as true ex ante returns.

traditional approach is to estimate a Mincer (1974)-type (log) earnings equation with the schooling level in years and work experience as covariates, and to interpret the schooling coefficient as the internal rate of return to an additional year of schooling (see Psacharopoulos and Patrinos, 2018). However, as Heckman et al. (2006, 2008) have stressed, this coefficient only yields an (internal) rate of return to education under some strong assumptions, i.e. (i) there are no taxes (and transfers), (ii) no tuition costs and no earnings while in education, (iii) there is no loss in working life associated with education, (iv) earnings functions are multiplicatively separable in experience and education, i.e. log-earnings–experience profiles are parallel across education levels, and (v) marginal returns being equal to average returns. Moreover, when estimation of the Mincer equation is based on a synthetic cohort approach (where a single cross-section approximates a cohort's life cycle) (vi) changing skill price differentials across time are ruled out.

In the case of higher education, assumptions (i) and (iii) are clearly violated for most countries. Additionally, several studies present empirical evidence against parallelism of log earnings–experience profiles (iv) and against stationarity (vi) (e.g., for the U.S. see Katz and Autor, 1999 and Heckman et al., 2008; for Norway see Bhuller et al., 2017). Linearity of log wages in education (following from (iv) and (v)) has been rejected in several studies for the U.S., pointing towards a non-zero role of sheepskin effects (see Heckman et al., 2006, and the literature cited therein).

A second strand of the literature has explicitly relied on full individual life cycles to estimate private and fiscal returns to education, sometimes called the "full discounting" approach (see Psacharopoulos, 1995). Bhuller et al. (2017) and Nybom (2017) observe Norwegian and Swedish adults from young adulthood to retirement and are thereby able to compute ex-post life-cycle returns for these individuals. As such panel data that contain full employment biographies for the whole population do not exist for most countries, some studies rely on artificial life cycles instead (see OECD, 2019; Pfeiffer and Stichnoth, 2019; Levell and Shaw, 2015; and de La Fuente and Jimeno, 2009). These studies typically use recent cross-sectional data from which they construct full life-cycles, with the advantage that they can account for some of the problems of the traditional approach mentioned above, particularly by relaxing the assumptions (i)-(iv).

In our paper, we follow this second strand of the literature but use dynamic microsimulation models to simulate a number of individual life cycles of a young German cohort to estimate private and fiscal returns to higher education. Dynamic microsimulation implies simulating individual life cycles sequentially in terms of several key variables such as births, marriages, divorces, labor force participation, employment, and earnings. Importantly, this approach allows to incorporate taxes and transfers, account for the length of working life, introduce a flexible modeling of wages, and correct for observable changes across birth year cohorts. Furthermore, dynamic microsimulation has the advantage of capturing path dependencies and simulating heterogeneous life cycles. In particular, our simulation strategy follows Courtioux et al. (2014) and Courtioux and Lignon (2016) who have estimated private returns to higher education for France. Similarly, we build a dynamic microsimulation model for Germany.

In addition to modules for the main transitions in family composition and labor market participation, our model includes a tax-benefit component that allows to simulate taxes, transfers, and social security contributions which are key for computing both private net returns and fiscal returns. Moreover, explicitly modeling the partnering process further makes it possible to analyze how the household context shapes the returns. The literature points out that a substantial share of households fully or partially pool their income (Ponthieux and Meurs, 2015; Ponthieux, 2017; Beznoska, 2019). In this case, an individual's future consumption prospects also depend on the earnings of her future spouse(s) and hence might be taken into account in the return calculation. We therefore compare the returns under different degrees of income pooling. Finally, by simulating individual careers we are able to examine the distribution of returns to higher education.

We find that higher education yields, on average and assuming no income pooling, a positive gross return of 11.5% for men and 13.4% for women. The tax-and-transfer system shrinks these returns to 8.7% and 9.7%, respectively. Furthermore, these returns are slightly reduced if we assume that individuals fully share their income with their spouses. Being closely related to the private returns, fiscal returns are 8.4% for men and 9.9% for women. At the same time, analyzing the returns for an "average" biography masks considerable heterogeneities among individuals. For about one third of individuals, we forecast negative private net present values.[3] We show how these heterogeneities can be explained by differences in hourly wages as well as employment, marriage, and fertility histories.

The paper proceeds as follows. Section 2 explains the institutional background of post-secondary education in Germany. Section 3 describes how we define and compute returns to higher education. Section 4 introduces our dynamic microsimulation model and the data and Section 5 presents the validation results of the model. Section 6 shows the results, Section 7 discusses them and Section 8 concludes.

[3]For technical reasons, we use two different concepts to evaluate the financial gains from higher education, "returns" and "net present values". We refer the reader to the discussion in section 3.

2 Post-secondary education in Germany

2.1 Higher education and vocational training

In Germany, those who obtained a higher education entrance degree (*Hochschulreife*, henceforth HEED) from secondary school can opt for higher education. As an alternative, these individuals can also take up a vocational training (*Berufsausbildung*). Only a small share of individuals decide to enter the labor market without any of these two types of training.[4] Currently, about three out of four individuals with a HEED enter higher education, while virtually the rest takes up vocational training (Autorengruppe Bildungsberichtserstattung [ABB], 2018).

More precisely, we define "higher education" as attending either a university (*Universität*) or a university of applied sciences (*Hochschule für Angewandte Wissenschaften*). While the two types of institutions differ with respect to the content of tuition (universities of applied sciences have, for instance, a larger focus on practical applications than universities), they are similar in terms of study length. By law (*Hochschulrahmengesetz*), the official study durations are three to four years for bachelor and one to two years for master programs. In 2016, approximately 58% of new higher education entrees entered a university and 42% a university of applied sciences (ABB, 2018).[5]

We define "vocational training" as attending either school-based training or dual training. School-based training mainly takes place at a vocational school and usually does not involve any salary.[6] In contrast, dual training combines on-the-job training in a firm and classes at a vocational school, and trainees receive an apprentice's pay (*Ausbildungsvergütung*) which depends on training year, profession, and region (Beicht, 2018). Of the individuals who obtained a HEED and start a vocational training 66% are in the dual training and 30% in the school-based training system (ABB, 2018).[7] While the duration of vocational training programs ranges between two and 3.5 years, the vast majority of programs have a duration of three years (Frank and Walden, 2012).

While we focus on the life cycles of those who obtained a HEED in order to estimate the return to higher education, we also simulate life paths of individuals with other educational degrees that are common in Germany.[8] The first category is defined by not obtaining any post-secondary degree, i.e. individuals belonging to this class neither obtained a higher education nor a vocational training de-

[4]Following Biewen and Tapalaga (2017), this share is about 2% for the cohorts 1944-1986.

[5]Until the Bologna reforms in the early 2000s, the most common higher education degree at both types of institutions was the diploma (*Diplom*). Since then, the diploma has gradually been replaced by bachelor and master degrees. We assume that the diploma is equivalent to the combination of bachelor and master degree, since they are similar in terms of official study length and content.

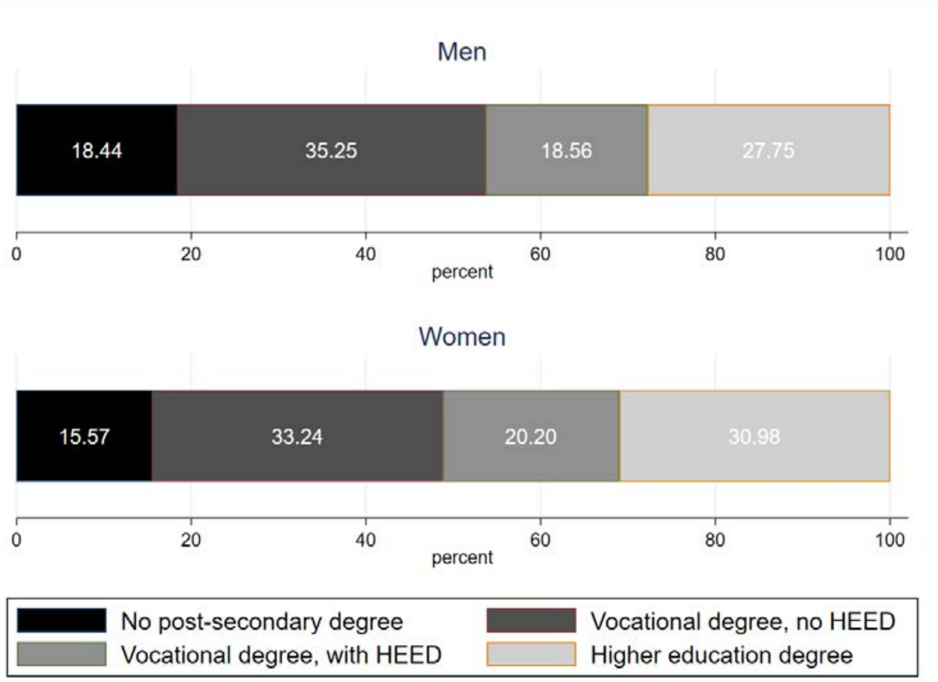
[6]Typical examples of a school-based training are health professions, such as medical or pharmaceutical technical assistants.

[7]About 4% enter some form of pre-vocational training.

[8]Individuals who belong to these educational categories are potential spouses of those with a HEED and therefore, for instance, contribute to the joint income of spouses.

gree. The second category comprises individuals without a higher education entrance degree but with a vocational training degree. Usually, these individuals graduated from a lower or middle secondary school track (ABB, 2018). Figure 1 displays the gender-specific shares of the educational categories as defined in this study, at ages 30–35 for the 1983–88 cohort, which are later used for the simulation.[9]

Figure 1: Distribution of post-secondary degrees, birth cohort 1983–88, in %



Notes: The figure displays the shares of combinations of secondary and post-secondary degrees for men and women in the age group 30–35 in 2018. *Vocational degree, no HEED* = Vocational degree without higher education entrance degree; *Vocational degree, HEED* = Vocational degree with higher education entrance degree. *Source*: Statistisches Bundesamt (2018), own calculations.

2.2 Funding of post-secondary education

In Germany, both academic and vocational training are heavily subsidized. While a place at a university or university of applied sciences currently costs, averaged across degrees and subjects, approximately 6500 Euros per year (Statistisches Bundesamt, 2017), currently no federal state collects tuition fees.[10] However, usually all students have to pay a small fee for administrative costs and the students' representatives (*Semesterbeitrag*), which also includes subsidized public transport and subsidized lunch at the university. A place at a vocational school is estimated to cost 4600 Euros per year on average. However, there is a considerable cost difference between a place in dual training (2900 Euros) and one in school-based training (7400 Euros). The difference is explained by the fact

[9]Note that this cohort classification does not fully correspond to the way we defined our cohort of interest (birth years 1980–89). However, the education distribution should not differ much.

[10]Some West German states had introduced tuition fees in 2006/2007, but abolished them afterwards.

that while dual training only partially takes place in a vocational school, school-based training means that students spent most of their training time in schools (Statistisches Bundesamt, 2017).

Beyond subsidized places in higher education and vocational training, students and vocational trainees also receive direct financial support through grants and loans, especially through the Federal Training Assistance Act, commonly referred to as *BAföG*. Whether an individual is eligible for support depends on own and parental wealth and income. As of 2019, the maximum monthly amount an individual could receive is 735 Euros for students and 590 Euros for individuals in vocational training. Usually about half of the amount received has to be paid back later during working life.

3 The lifetime returns to higher education

3.1 Private returns

The literature has predominantly used two measures to assess the returns to education in a life-cycle perspective. The first one is the (marginal) internal rate of return (IRR), which is the interest rate that equates the income streams under investment and non-investment.[11] The second one is the net present value (NPV), defined as the difference between the present values of all benefits and the (opportunity) costs that accrue to investing in education given a fixed interest rate r. As discussed above, in Germany, for individuals who already obtained a HEED, not pursuing any post-secondary path does not seem to be an attractive option, which is reflected in the fact that only very few individuals follow this path. Therefore, for these individuals we consider vocational training to be the relevant counterfactual for the higher education path. We use the above two measures to derive the educational returns under different income concepts. Specifically, we distinguish between gross and net income, and between different degrees of income pooling within couple households.

The reason to study the effect of income pooling is the following. Ponthieux and Meurs (2015), Ponthieux (2017), and Beznoska (2019) provide empirical evidence for income pooling within households. At the same time, Eika et al. (2019) show for Germany and other countries that there is positive educational assortative mating, i.e. individuals with similar education mate with one another more frequently than would be expected under random mating. This implies that there might be an indirect return to education: pursuing higher education is associated with a higher expected level of education of potential partners and therefore with higher expected (shared) household income (Courtioux and Lignon, 2016). Therefore, we take educational assortative mating into account and study the effect

[11] Originally, the term "internal rate of return" implied comparing an educational investment to a null project without costs and returns. In contrast, the "*marginal* internal rate of return" implied comparing a high-investment project with a low-investment project. In recent studies (see, e.g. Bhuller et al., 2017), however, the term "internal rate of return" is used when comparing two investment projects. See Carneiro and Heckman (2003) for a more detailed discussion.

income pooling has on the financial return to higher education.

We define $Y_{i,t}$, the income of individual i in period t, as follows:

$$\begin{aligned} Y_{i,t} &= (1-\rho)(labinc_{i,t} + capinc_{i,t} - taxes_{i,t} - contribs_{i,t} + benefits_{i,t}) \\ &\quad + 0.5\rho(labinc_{hh,t} + capinc_{hh,t} - taxes_{hh,t} - contribs_{hh,t} + benefits_{hh,t}) \\ &\quad + 0.5\tau_{i,t}(transfers_{hh,t}), \end{aligned} \tag{1}$$

with $0 \leq \rho \leq 1$ indicating the share of income that is pooled within individual i's household hh and $\tau = 1$ if married and $\tau = 2$ if unmarried. *labinc* is labor income from dependent employment (after employer's social security contributions) and self-employment. *capinc* is income from interest on savings. *taxes* include income tax and capital withholding tax payments.[12] *contribs* comprise employee's social security contributions as well as contributions to private health insurance contracts.[13] $benefits$ are individual-level benefits like unemployment benefits and parental leave benefits. All social transfers granted at the household level are included in $transfers$. For the gross income measure, all components except *labinc* and *capinc* are set to zero.

Importantly, we exclude pension insurance contributions and benefits from the analysis. This means that neither are contributions to the statutory pension system included in *contribs*, nor are received pensions included in $benefits$. One central feature of Germany's public pension system is that pension entitlements are, in general, equivalent to contributions paid, which implies a rather limited effect on the private return. More relevant with respect to the private return could be extraneous pension benefits such as the inclusion of time spent bringing up children in the pension insurance. However, adding a detailed simulation module for the pension system and several additional transition modules for pension-related outcomes is beyond the scope of this study.

We use the IRR as a measure to evaluate the return to higher education for a representative life cycle, which is defined by income streams averaged over gender and migration background. The reason is a technical one: Comparing incomes under higher education and vocational training, a unique interest rate that equates the two income streams (i.e. a unique root) is only guaranteed if, over the life cycle, there is exactly one change of sign. This is the case if, for instance, income from vocational training is larger than income from higher education up to an age threshold and smaller beyond that threshold. For (simulated) individual biographies, however, multiple sign changes can be expected to

[12] While we account for the value added tax when estimating fiscal returns (see below), we follow most of the literature on net private returns to education and exclude it from the calculation of private returns.

[13] This definition is in line with the incidence assumption that about 50% of the overall social security contributions are borne by the employees. While it deviates from the conventional assumption in the public finance literature that employees bear the full burden, it ensures that our results are comparable to the literature on educational returns, where most other studies work on the same incidence assumption.

be the rule rather than the exception. Hence, we cannot compute the IRR on an individual basis, but instead only use this measure when computing the return for an "average" individual, for whom there are no multiple sign changes.

The IRR of the average life cycle is obtained by solving the following equation:

$$\sum_{t=0}^{T} \frac{\overline{Y}_t^{HE} - \overline{Y}_t^{VOC}}{(1+IRR)^t} = 0 \tag{2}$$

where $\overline{Y}_t^{HE}$ and $\overline{Y}_t^{VOC}$ define the average incomes under the two educational paths, higher education (HE) and vocational training with higher education entrance degree (VOC), in period t and IRR is the internal rate of return. Hence, investing in higher education is financially beneficial if the IRR is larger than the market interest rate.

To assess the distribution of returns to higher education, we compute net present values (NPVs). The NPV of higher education is computed as the difference between the simulated lifetime income LTI_i^{HE} under higher education and its counterfactual LTI_i^{VOC}, i.e.

$$NPV_i^{private} = LTI_i^{HE} - LTI_i^{VOC}, \tag{3}$$

with lifetime incomes LTI_i being defined as

$$LTI_i^{HE} = \sum_{t=0}^{T} \frac{Y_{i,t}^{HE}}{(1+r)^t} \tag{4}$$

and

$$LTI_i^{VOC} = \sum_{t=0}^{T} \frac{\overline{Y}_{g,t}^{VOC}}{(1+r)^t}, \tag{5}$$

where $\overline{Y}_{g,t}^{VOC}$ is the average income of group g individual i belongs to, defined by migration background and gender, in period t (for a similar definition of counterfactual income see Courtioux et al. (2014)).

3.2 Fiscal returns

The concepts of IRR and NPV can be applied analogously to the measurement of fiscal returns. In each period t, the fiscal surplus $S_{i,t}$ that is generated by individual i is defined as the difference between public revenue and public expenditures with regard to that individual. We restrict the measure to budgetary components that can be expected to deviate strongly between individuals of different educational degrees and that can be simulated within our model framework.[14]

[14] For instance, fully capturing the fiscal effects of higher education on the social security system would imply modeling outcomes like health and life expectancy as well, which is beyond the scope of this paper. See Eide and Showalter (2011)

The fiscal surplus is defined as

$$S_{i,t} = taxes_{i,t} + ssc_{i,t} - benefits_{i,t} - 0.5\tau_{i,t}(transfers_{hh,t}) - pubexp_{i,t}, \tag{6}$$

with $\tau = 1$ if individual i is married and $\tau = 2$ if unmarried, $taxes$ comprising income, capital withholding and value added taxes, ssc overall (employer's and employee's) social security contributions, $benefits$ unemployment and parental leave benefits, $transfers$ social transfers like social assistance, housing allowance and *BAföG* payments, and *pubexp* public expenditures for education as well as other public expenditures. Importantly, employing the same argumentation as for private returns, we refrain from including the public pension system in the fiscal surplus measure.

When computing the fiscal surplus generated by higher educated individuals, we take into account the financing of health care expenditures. In Germany, dependent employees are typically insured under the statutory health insurance scheme. However, a non-negligible fraction of individuals are privately insured, in particular civil servants, the self-employed, and dependent employees with relatively high labor incomes. Therefore, within our simulation framework, the effect of higher education on the balance of public health insurance is twofold: it consists, first, of the effect on the probability to be insured under the public health insurance scheme and, second, conditional on being insured under the public health insurance scheme, of the effect on the level of contributions paid.

Estimating fiscal returns, one group that deserves special attention are civil servants and public employees. Civil servants are typically privately health insured in Germany, with premiums being subsidized by the state (*Beihilfe*). Moreover, in theory the fiscal return generated by higher educated civil servants and public employees would have to account for the effect on remuneration and the marginal revenue product of labor of these employees as well. To make things easier, throughout the simulation we assume the number, qualification, and remuneration of the publicly employed to be fixed. We exclude life cycles of civil servants from the computation of fiscal returns and treat life cycles of other public employees as if they were working in the private sector.

The fiscal IRR is then defined as the interest rate at which the present values of the average fiscal surplus are equal for the two paths of education:

$$\sum_{t=0}^{T} \frac{\overline{S}_t^{HE} - \overline{S}_t^{VOC}}{(1+IRR)^t} = 0. \tag{7}$$

for an overview of potential education-level related differences in health and life expectancy.

Finally, the fiscal NPV of higher education is defined as:

$$NPV_i^{fiscal} = LTS_i^{HE} - LTS_i^{VOC}, \tag{8}$$

with the actual lifetime surplus LTS_i being defined as

$$LTS_i^{HE} = \sum_{t=0}^{T} \frac{S_{i,t}^{HE}}{(1+r)^t} \tag{9}$$

and the counterfactual lifetime surplus as

$$LTS_i^{VOC} = \sum_{t=0}^{T} \frac{\overline{S}_{g,t}^{VOC}}{(1+r)^t}, \tag{10}$$

where $\overline{S}_{g,t}^{VOC}$ is the period t average surplus of group g that individual i belongs to, defined by migration background and gender. Hence, at a given interest rate r, the NPV is positive if the benefits of investing in higher education — higher tax revenues and social security contributions due to higher labor earnings and lower transfers due to less and shorter phases of unemployment — outweigh the costs of higher public expenditures and a delayed labor market entry.

Similar to other studies that estimate the fiscal returns to education, we do not take into account human capital externalities which might be an indirect channel through which educational investment impacts the public budget. Using a similar computational measure, de La Fuente and Jimeno (2009) therefore interpret their fiscal return estimate to be a lower bound of the true return. However, the literature has come to very ambiguous results as to whether these externalities are sizeable (see, for instance, Acemoglu and Angrist, 2000; Moretti, 2004; Ciccone and Peri, 2006). Also, we abstract from general-equilibrium effects of a potential increase in higher educational attainment on the (distribution of) marginal returns to higher education.

4 Dynamic Microsimulation Model

4.1 Projecting the lives of a young cohort

A number of studies have used projection techniques to forecast life cycles. Often, these studies aim at providing information on trends of socio-economic development under current versus alternative policies, or at evaluating the future performance of pension, health and long-term care systems, given economic and demographic trends (Li and O'Donoghue, 2013). For Germany, for instance, different life cycle models and projection techniques have been used to study future public pension entitlements of specific birth cohorts (Geyer and Steiner, 2014), to evaluate the lifetime monetary values of family policy measures (Bonin et al., 2016), to study inequality in lifetime income across cohorts (Bönke et al., 2020), and to assess how the tax-transfer system reduces the inequality in lifetime income (Haan et al., 2017).

In the literature on returns to education, different methodological approaches have been applied to generate artificial life cycles. The OECD, for instance, regularly provides estimates on the returns to education across countries using the so-called short-cut method (OECD, 2019). Here, employment biographies are constructed by averaging earnings of all individuals given a certain age, gender and education level. Another approach, often called "splicing", consists in drawing observations from a cross-section of individuals, conditional on age, education, and other characteristics, and stitching them together to form full life cycles. Applications of this method range from randomly drawing based on a small subset of variables only (Pfeiffer and Stichnoth, 2019, 2020) to versions based on elaborate matching approaches (Levell and Shaw, 2015). Compared to the simple OECD approach, the former variant has the advantage that the precision of these estimates can be assessed by looking at the distribution of generated life cycles and returns, while the latter more accurately reflects actual life cycles.

The approach we follow relies on dynamic microsimulation, that is, modeling the transition processes between different states and then simulating life cycles of individuals with different initial values. Dynamic microsimulation has the advantage of capturing path dependencies which are key to simulating heterogeneous life cycles and to analyze heterogeneous returns to higher education. Furthermore, microsimulation models are able to capture observable differences across cohorts.[15] Yet, there are, to our knowledge, only two studies for France that use dynamic microsimulation models to estimate the returns to education: Courtioux et al. (2014) model earnings, employment, and mortality to estimate the distribution of returns to higher education for different educational degrees in France. Courtioux and Lignon (2016) add a simulation module for household formation to disentangle (direct) labor market returns and (indirect) marriage market returns.

4.2 Model overview and data

In our model, we simulate biographies on a yearly basis from age 18 until the supposed retirement at age 67 (the status quo legal retirement age for the chosen cohort).[16] For the simulation, we take as given

[15] In contrast, dynamic microsimulation has its own drawbacks such as high data requirements, the complexity of modeling and a high computational burden.

[16] Since we simulate the trajectories of a distinct cohort over a relatively long time period, our model belongs to the class of "cohort models" as opposed to "population models", that model a population cross-section over a defined period of time (see Li and O'Donoghue, 2013).

the distribution of highest educational degrees of the 1983–88 birth cohort by gender and migration background, as presented in Figure 1.[17] Even though our focus is the comparison of life cycles under higher education and vocational training, given a higher education entrance degree (HEED), we also simulate the life courses of individuals with either of the two other educational outcomes presented in Figure 1: vocational degree without a HEED and no post-secondary degree at all. The reason is that these individuals serve as potential spouses for the individuals of the first two educational categories.

Our model proceeds in three stages: estimation, life cycle simulation, and tax-transfer simulation. We provide a brief overview of these stages first and present the details of each part thereafter.

Estimation We estimate transition probabilities regarding family formation and employment outcomes, aggregate cohort-specific targets for the respective states, and hourly wage regressions.

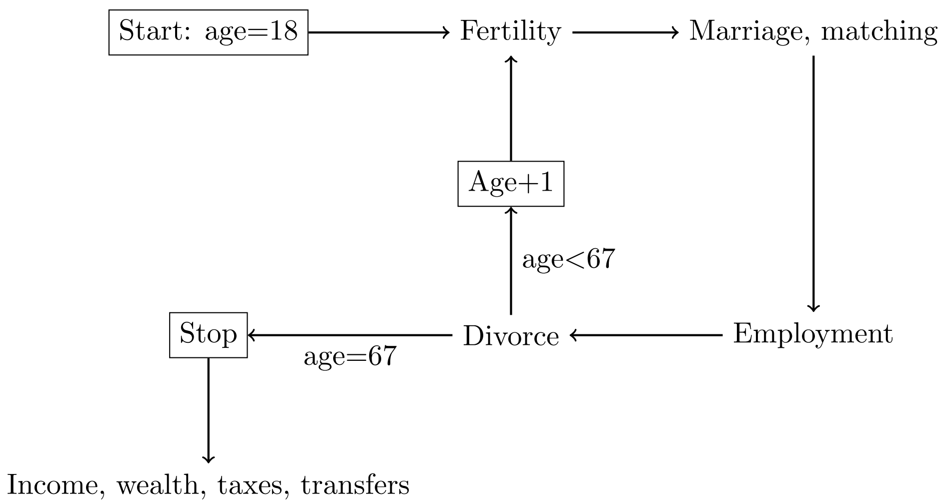
Life-cycle simulation In the simulation stage we sequentially simulate the cohort's life cycles based on the estimated parameters and using alignment methods. Figure 2 provides a graphical representation of the main processes. Model projections are based on *dynamic ageing*, implying a recursive process in which the individual age is updated and the trajectories in terms of several key variables are simulated year after year. This process is repeated until individuals reach the age of 67.[18] We also add an alignment (calibration) module to each transition process. This ensures that aggregate outcomes are kept in line with projected macro developments. Yet, we abstract from potential general equilibrium effects of changes in household formation and employment patterns and hold the estimated parameter values of our transition models constant throughout the simulation process, as is common in the dynamic microsimulation literature (Li and O'Donoghue, 2013). After having simulated the life cycles, we predict gross labor incomes conditional on education, simulated labor market experience and other covariates.

Tax-transfer simulation In the last stage, the resulting taxes, transfers, and social security contributions and finally disposable incomes are simulated using a trimmed version of the Tax-Benefit Microsimulation Model (STSM, see Steiner et al., 2012). Instead of applying the actual tax-transfer regulations experienced by the 1980s cohort, we use the rules of 2019 uniformly for all simulated age years. The reason is that we aim at forecasting the life cycle of a young cohort under the current tax-transfer regulations where the life cycles of the 1980s cohort serve as a benchmark. Further details on the simulation of taxes and transfers can be found in Section A.1 in the Appendix.

[17] In Figure 1 the distribution is not conditioned on migration background.

[18] Naturally, after the age of 45 fertility is virtually zero. Nevertheless, the process is technically simulated until the age of 67.

Figure 2: The simulation stage



Our main database is the German Socio-Economic Panel (SOEP). The SOEP is an annual, nationally representative longitudinal study of private households across Germany, with currently more than 30,000 individuals (Goebel et al., 2018). The SOEP data provides information on all household members, on topics such as household composition, occupational biographies, employment, and earnings. The laws of motion for household formation and employment and the wage regressions are estimated using the SOEP waves of the years 1984–2018 and individuals aged 18 to 66. We exclude individuals with missing or inconsistent information on their education levels and the refugee samples of the SOEP. For the alignment modules we additionally use official statistics of the Federal Statistical Office, which are based on administrative and census data.[19]

The next section describes the main components of our model in more detail.

4.3 Transition modules

We estimate discrete-choice models for household formation and employment transitions and linear regressions for wages. To capture differences in these processes across educational paths, discrete-choice models include education and training indicators as covariates and wage models are estimated separately by degree. In all transition models, we include year and birth cohort dummies, where a cohort spans ten years. Also, we control for living in the states of the former East Germany. All models are estimated separately by gender. For some categorical variables we include an orthogonalized transformation. This has the advantage that, while we can control for these variables in the

[19]While target estimation for employment states relies on SOEP data, marriage, divorce, and fertility targets use different data sources: Statistisches Bundesamt (2019b,c) for fertility, Statistisches Bundesamt (2004–2017a) for marriages, and Statistisches Bundesamt (2004–2017b) for divorces. The decision to base target estimation partly on aggregated data instead of SOEP micro data is motivated by the attempt to account for cohort differences in life cycle patterns, which is difficult with micro data for relatively rare events such as fertility, marriage and divorce.

estimations, they do not have to be simulated in the simulation stage of the model. Instead, the average effect over all categories is applied when simulating outcomes. In general, all estimations rely on a selection-on-observables assumption. Estimation results of the transition models are displayed in the Appendix.

4.3.1 Marriage and divorce

We model the probability of marrying as a function of age polynomials, educational training indicators, lagged births and child indicators, lagged employment indicators, accumulated years of unemployment, previous divorces, and migration background (for the detailed specification and estimated coefficients, see Table A1). Our microsimulation model also includes a matching process which allows to match women and men that were selected for a marriage. We do this by separating the pool of individuals that were simulated to marry into two groups: one group within which there is perfect educational assortative mating, i.e. highly educated men marry highly educated women, while low educated men marry low educated women. And another group, within which there is no assortative mating, i.e. men and women are randomly matched. The share of the respective groups is calibrated to match the empirically observed pattern as closely as possible.

As to divorce, we estimate a logit model for women only. We model the probability to divorce as a function of couple characteristics such as the age difference between the spouses, the time married, an indicator for having married before the age of thirty, the presence of children in the household, the employment states of the spouses, experience of previous divorces, and migration background (see Table A2 for the specification and estimated coefficients).

4.3.2 Fertility

In our fertility module, we estimate logit models for the probability of giving birth to a child, separately by marital status. For both married and unmarried women, this probability is modeled as a function of age, training indicators, the presence of children already living in the household, tenure, and migration background. The model for married women also includes years married and the lagged employment status (see Table A3 for the estimated coefficients).

4.3.3 Employment

We model employment as a three-step process, consisting of labor force participation, involuntary unemployment and working time estimations. In what follows we set out the general modeling approach. For the exact specification of the estimated models and coefficient estimates, see Tables A4–A7 in the

Appendix.

First, we describe labor force participation as a binary choice model that depends on lagged labor force participation, age, education levels, interactions thereof, and migration background. In addition, we allow women's labor force participation to depend on marriage status and on having given birth in previous periods while for men we include the presence of children in the household. We model the particularities of the entry into the labor market by modeling separate transition models for each of the first five years after graduation (estimation results not shown). Second, conditional on labor force participation, we model the probability to be involuntarily unemployed. We define involuntary unemployment as being without work but having actively searched for a new position within four weeks at the time of the survey and being able to start working within the following two weeks. The set of covariates is similar to above, additionally including cumulated years of employment and unemployment. Third, conditional on labor force participation and employment, each individual is modeled to be in one of several possible employment states, characterized by a discrete set of working hours. Hours classifications are based on the empirical distribution observed in the SOEP. For women, we model five employment states: Marginal employment (0–14 hours), reduced part-time work (15–24 hours), extended part-time (25–34 hours), full-time work (35–42 hours), and over-time (more than 42 hours). For men, we model three employment states: Part-time (0–34 hours), full-time (35–42 hours), and over-time work (more than 42 hours). We estimate multinomial logit models that, compared to the above employment-related model specifications, additionally control for tenure and self-employment.

4.3.4 Wages

Since, clearly, a strong link between hourly wage premiums to higher education and the full (private and fiscal) return to higher education exists, appropriate modeling of wages is key to our simulation analysis. In the baseline specification, we regress individual log gross hourly wages on fourth-order polynomials of age, experience, and tenure, and on indicators for industry sector, for being self-employed or a civil servant, migration background, and orthogonalized indicators for federal states.[20] Moreover, we include year-specific dummy variables to control for changes in macroeconomic factors (e.g., business cycles) in a flexible way. Estimation is conducted separately by gender and educational category. In an alternative specification we also include birth cohort dummies where birth cohorts span 20 years. Clearly, in a regression that includes age and year effects, cohort effects cannot be identified without relying on further assumptions.[21] Here, we follow the approach in Deaton (1997),

[20] We consider real wages with 2019 as the base year.

[21] When cohort measures are defined as dummy variables for several births years, identification of age, year and cohort effects is in principle possible through the variation in age within each year-cohort group.

which assumes that year effects are orthogonal to a linear trend and add up to zero over the observation period. In this setting, the year effects only capture cyclical movements around this trend.

By estimating wage regressions separately by education we do not rely on the assumption that education and experience are separable as in the original Mincer equation, which is, as argued above, a problematic assumption (see Bhuller et al., 2017). While empirical studies on the returns to education often refrain from questioning the validity of the assumptions implicit in estimating Mincer-type equations which are necessary to infer internal rates of return to education, a potential endogeneity bias in the estimated schooling coefficient is one of the most discussed topics in labor economics (Psacharopoulos and Patrinos, 2018). When reliable instruments are absent, unbiased estimation of the wage equation rests on a *selection-on-observables* assumption, i.e. the assumption that conditional on the other explanatory variables included in the wage regression, the level of schooling is not correlated with the error term, which might include variables like ability or motivation. We do not consider an Instrumental Variables (IV) strategy since there is no suitable instrument that predicts higher education attainment available in our data.[22] Moreover, it is unclear in which direction IV estimates would differ from OLS. For instance, Bhuller et al. (2017) find larger IRR estimates from IV than from OLS. Furthermore, they show that the biases arising from violations of the other key IRR assumptions (see Section 1) are empirically more important than the selection bias arising in OLS estimates. Finally, we alleviate the potential impact of an ability bias on our estimates by restricting our counterfactual group to those who have obtained a HEED. As the German education system is essentially tracking students according to their ability, we expect these individuals to be more similar to the group of academics than individuals without HEED.[23][24] The estimation results are displayed in Tables A8–A10 in the Appendix.

4.4 Alignment

We include an alignment module in our model as a tool for calibrating the model output, as is common in the dynamic microsimulation literature (see Li and O'Donoghue, 2014, for an overview of alignment methods). Alignment ensures that the results of our micro models are, on average, in line with aggregate generational trends. In addition, it allows us to incorporate accurate macro-level statistics of births, marriages, and divorces into our microsimulation model. The goal of the

[22] Kamhöfer et al. (2018), for instance, exploit college openings for their IV strategy. These openings, however, took place in the period 1958-1990 and hence only affected older cohorts than the one we are interested in.

[23] See Walker and Zhu (2011) and Glocker and Storck (2014) for a similar approach.

[24] Another potential empirical problem could be non-random selection into the labor force. However, previous studies that also estimate returns to education such as Glocker and Storck (2014) and Steiner and Lauer (2000) have found that applying Heckman corrections with different sets of exclusion restrictions yielded only minor changes in the estimated education coefficients.

alignment module is to predict age-specific targets for our cohort of interest such as the proportion of men/women marrying or being in the labor force at each age of the life cycle. While the targets of these variables are proportions between zero and one, using linear regression models would not restrict predicted target values to the unit interval. Therefore, we estimate fractional logit models to estimate our targets by quasi maximum likelihood. In fractional logit models, the conditional expectation of the dependent variable is modeled as a logistic function, just as in the binary logit model (see Papke and Wooldridge, 1996, 2008). Similarly, the proportions of each of the working hours categories, dependent on employment, are estimated using fractional multinomial logit models. This ensures that the predicted shares add up to unity.[25]

For each of the estimated target models, the dependent variable is defined on a cell level as the respective proportion by birth year, age and gender.[26] As explanatory variables we include age polynomials and either generational trend polynomials (generational trend defined as $birth\ year - 1930$) or birth year cohort dummies (grouped over ten-year intervals), as well as the overall unemployment rate. See Tables A12–A14 for estimation results including the full list of independent variables.

We predict target rates for each simulated age year of the 1980s cohort, based on the estimated parameters.[27] The projected targets are graphed in the Appendix in Figures A1–A3, together with observed patterns for the 1980s and older cohorts. Plotting predicted against observed life-cycle patterns is useful in two dimensions: first, it allows to investigate trends in household formation and employment patterns over past cohorts, and second, it enables us to check the reliability of our out-of-sample predictions. Again, data availability determines which segments of the life cycle are observed for a particular cohort. While for some transitions and states, trends in life-cycle patterns across cohorts appear to be rather negligible (e.g. the marriage rate of women), for others they are more noticeable. This especially holds for birth rates and labor force participation, but also for some working hours categories. Regarding women's working lives, the model captures the increases in the participation rate and in extended part-time employment over cohorts. Moreover, for men our model predicts a higher participation rate at older ages compared to previous cohorts, including a more prominent role of part-time work.[28]

[25]See Ramalho et al. (2011) for an overview of estimation strategies for fractional regression models and Mullahy (2015) for a discussion of the extension of these models to multivariate fractional data.

[26]The number of cells used for the estimation of the target models is restricted by data availability and differs over the targets. For instance, regarding births, for each birth year cohort between 1930 and 1967 there are 35 data points available for the age years between 15 and 49. For birth year cohorts from 1968 to 1999, only $n = 35 - (birthyear - 1967)$ data cells are available. In total, this results in 1,573 cells. While employment-specific target estimations are based on a similar number of cells, the target estimation for marriages and divorces relies on a lower cell number. Note that the proportion of individuals divorcing is only available in 5-year age intervals.

[27]For the predictions we assume an aggregate unemployment rate of 6%, which is the average unemployment rate over the last ten years and hence seems a plausible assumption for the life cycle of the young cohort.

[28]Note that for men we define weekly working hours of up to 34 hours as part-time work.

4.5 Simulation

In order to simulate family formation, family dissolution, and employment transitions, we follow an approach similar to Courtioux and Lignon (2016). In what follows, we provide an overview of how we simulate life cycles. A description of further modeling details and assumptions can be found in in Section A.1 of the Appendix.

Given the empirical distribution of gender, education, and migration background for the 1980s cohort, we simulate the life cycles of 5,000 artificial individuals from age 18 to 66. Individuals enter the labor market after their training phases, which we assume to be six years for academic training and three years for vocational training. The model implicitly captures the phenomenon of early retirement as some individuals might be simulated to leave the labor force before the age of 66. Potential pension payments, however, are not accounted for.

The procedure to select individuals for transitions works as follows:

1. Predict individual transition probabilities using the parameter estimates from the transition models.

2. Multiply each probability with a random draw from the unit interval.

3. Rank individuals according to these modified probabilities.

4. Based on this ranking, select individuals for transitions until the respective aggregate target rates are met.

This procedure is repeated for every transition process and each age, from 18 to 66.[29] It guarantees that the simulation reflects individual (education level)-specific differences in the transition probabilities (step 1) and additional variability that is not captured by our models (step 2). Aggregate shares are aligned to our estimated cohort-specific targets, which assures that we capture generational trends in household formation and employment behavior (steps 3 and 4).[30]

Having simulated the transitions in employment allows us simulate earnings over the life course. First, we predict log gross hourly wages for all ages given exogenous and simulated variables. To this prediction we add random draws from the distribution of the log wage residuals, conditional on gender, education, and self-employment status.[31] This procedure aligns the variance of simulated

[29] The algorithm is very similar to what has been called the *SBD* approach in Li and O'Donoghue (2014).

[30] For some employment states, measured persistence was very high for simulated life cycles compared to observed ones. Therefore, a calibrated share of individuals are selected randomly for transitions into those states. Specifically, this holds for 30 percent of the transitions into unemployment, and for 15 percent of the transitions into the three part-time employment states of women.

[31] We condition on self-employment status because we observe a higher variance in wages for the group of self-employed, compared to dependent employees.

wages to the variance of observed wages. Finally, to obtain hourly wages in levels, the resulting sum is exponentiated.[32] To compute labor earnings, hourly wages are multiplied by the observed average level of working hours given the simulated employment category.

5 Validation

Due to their complex model structure, dynamic microsimulation models are often regarded as a *black box* (see, e.g., Dekkers, 2016; Lütz and Stein, 2020). From the estimation results of the transition models one cannot infer whether the models perform well in terms of simulated life-cycle profiles. Therefore, we provide evidence on the predictive performance of our dynamic modeling approach by contrasting simulated and observed life-cycle patterns. Since life cycles of our simulated cohort are only observable up to their mid-thirties, older cohorts and results of previous simulation studies serve as a benchmark.

5.1 Autocorrelations

First, we check whether the level of persistence in household composition and employment outcomes in our simulated sample is plausible. We do this by showing autocorrelations of simulated and observed outcomes as well as average durations in different employment states. For observed outcomes, we pool all cohorts from 1950 to 1980.

Figures A4 and A5 depict simulated autocorrelations in employment and family formation outcomes for women and men, respectively. They are contrasted with the levels of autocorrelation observed in the SOEP data. The following findings indicate that the microsimulation model performs reasonably well in predicting life cycles of our target cohort. First, autocorrelations for both observed and simulated states are generally increasing in age, with exceptions at age years with major changes in the labor force participation taking place. For example, for women observed employment persistence decreases at age years with high birth rates. Moreover, autocorrelations shrink for our simulated cohort when university graduates enter the labor market.[33] Second, autocorrelations for the simulated trajectories are highest between the ages of fifty and sixty, which is in line with observed patterns. This holds particularly for higher-order autocorrelations.[34] Regarding employment, our model pre-

[32] Hence, we predict the gross hourly wage of individual i by using the formula $\hat{w}_i = exp(x_i'\hat{\beta})exp(\hat{u}_i) = exp(x_i'\hat{\beta} + \hat{u}_i)$, where x is a vector of covariates, $\hat{\beta}$ a vector of the estimated coefficients and $\hat{u}$ is the randomly assigned residual from the log wage regression.

[33] Due to the simplifying assumption regarding the duration of educational programs, which results in all persons with a particular schooling degree entering labor market at the same age, former (full-time) employees are partly crowded out. This should be of no major concern regarding the validity of the estimated financial gain of higher education, since primarily individuals with no or low vocational degrees are affected.

[34] A notable exception is the autocorrelation in the (un)employment state of men. While the pattern of short-term autocorrelation is similar for simulated and observed life cycles, simulated longer-term persistence is shrinking between

dicts a falling level of persistence thereafter, which is rooted in the raising share of persons switching to part-time employment or leaving the labor force.

Finally, we measure persistence by the average length of spells, which is included in Tables A15 and A16.[35] Our simulated data reflect the observation that persistence in unemployment is relatively low, while persistence in marriage and parent status is relatively high, compared to the other labor market states shown. For example, the average unemployment spell is about two years in our simulation, while the average marriage spell comprises about 25 years.[36] The fact that the model considerably over-predicts persistence in the marriage status of men (see Figure A5) is partly rooted in the distribution of marriages being based on the marriage pattern of women, which is due to the *closed* character of our cohort model.

5.2 Differences by education

As the focus of this study is on the financial gain of higher education degrees, it is essential that the microsimulation model succeeds in replicating differences in life cycles across educational groups.[37] Sequence analysis helps to visualize transition patterns and how they compare across distinct population groups (Abbott, 1995). For a graphical representation of how employment and family formation trajectories differ between education levels, we depict sequence index plots of 400 randomly drawn male and female life cycles.[38] Moreover, we provide descriptive statistics on the distribution and length of spells across educational groups. Finally, we describe how simulated gross hourly wages and labor earnings evolve over the life cycle.

ages forty and fifty whereas observed persistence increases.

[35] Note that, since our simulation stops at age 66, the last spell of each individual is right-censored. By including these spells, we generally underestimate the spell lengths.

[36] The average duration in unemployment is high compared to the average duration of completed spells reported by the German Federal Employment Agency, which is about 8.5 months (Bundesagentur für Arbeit, 2019). Measuring only registered unemployment, this number is not fully comparable to our estimate, however. Moreover, note that while the unit of time in our model is years, we set unemployment spells of one period to three months, as described in Section A.1, resulting in a median unemployment spell of three months in our simulation. Also, classification as being unemployed is based on the labor market state at the time of the interview. For instance, an individual might be unemployed in the month of the interview, then employed for eleven months, and then unemployed again in the month of the interview of the following year, leading to be classified as unemployed for two consecutive years. This leads to an over-prediction of persistence in unemployment.

[37] Recall that while aggregate employment and family formation patterns are set by the estimated targets, differences in transitions between education levels exclusively stem from the estimated transition models.

[38] Sequence index plots graph sequences as horizontal lines and sort these sequences according to a matching algorithm. To sort sequences, we apply the Needleman-Wunsch optimal matching algorithm and use the Levenshtein distance measure for evaluating the distance between two sequences. See Scherer (2001), who proposed sequence index plots as a tool to investigate early career patterns. The number of 400 life cycles considers the trade-off between having a representative sample and readable figures.

Employment

We start by validating simulated employment careers. Figure A6 depicts a sequence analysis for employment transitions. Additionally, Table A15 lists average years spent in labor market states over the life cycle across education groups and gender. In general, men's employment biographies are much more homogeneous than those of women. The majority of male paths are characterized by more or less permanent full-time episodes in the middle of their careers, with only short interruptions by part-time work or unemployment. On the contrary, frequent changes between employment states can be observed in the career paths of women. As expected, post-education labor market attachment increases with education. The share of men not working at young ages decreases with education. This divergence is less pronounced at the end of the employment biographies, but still existent. Women's careers, on the other hand, are characterized by more (often birth-related) employment breaks in the middle of their working lives. Long episodes characterized by near-permanent absence from the labor market are however strongly concentrated at women without any post-secondary degree. Life cycles with a high attachment to part-time employment are most common for women without a HEED but with a vocational degree. Very often these episodes start with the birth of the first child and continue until the exit from the labor market at older ages. Prevalence and length of such episodes decrease with further educational attainment. Interestingly, for higher educated women, episodes of part-time employment are less frequent compared to those with a vocational degree, while for men this pattern cannot be observed. In sum, these findings result in the number of working hours increasing with education.

The findings generally reflect what has been found in Geyer and Steiner (2014), who forecast employment biographies of a German cohort that is slightly older than the one studied here.[39] For men our number of simulated years of full-time employment by education is very similar compared to their study (for instance, for individuals with higher education, 32.5 years in our study and 32.2 (31.3) in Geyer and Steiner (2014) for West (East) Germany). At the same time, years in unemployment are lower in our study (1.7 vs. 2.0 (5.2)). For women, we generally simulate a higher attachment to the labor market and a higher share of part-time employed women, particularly for the middle to lower educated. For example, our model simulates 37.1 years of labor market participation for low educated women, compared to 27 (44) years for West (East) Germany in Geyer and Steiner (2014). This should be rooted in the generational trends of female labor market behavior predicted by our model.

[39] The authors show cumulated years for the 1967–71 cohort, separately for East and West Germany and do not estimate years in part-time employment for men. Importantly, their education classification differs from ours, which further complicates a comparison of the findings. Moreover, their simulation stops at the age of 65, while our simulation ends two years later in life, reflecting the rise in the statutory retirement age.

Family formation

Regarding transitions in family status, sequence analysis (see Figure A7) and descriptive statistics in Tables A16, A17, and A18 reveal the following patterns: In general, life cycles can be categorized into four major groups. The first group consists of individuals who never marry nor become parents. The second group comprises individuals who become parents and get married consecutively within one or two years and stay married thereafter, as well as parents that never marry. Heterogeneous life cycles with one or more divorces form the third group. Finally, a small group of individuals marry (often at higher ages) without becoming parents. The overall fertility rate is projected to be 1.69, which is in line with the "optimistic" G3 assumption in the current coordinated population projection of the German Federal Statistical Office (Statistisches Bundesamt, 2019a). In general, our model replicates the distribution of the number of births per woman correctly when compared to the distribution observed for women of the 1970s cohort in the SOEP. Women's age at first birth and age at first marriage are considerably higher compared to older cohorts, as expected.[40] The average duration of a marriage at the time of divorce is about 8.5 years in our simulation, which is considerably lower than what is reported in current cross-sectional administrative statistics, which is not surprising due to our simulations terminating at the age of 66.[41]

For women, age at first birth, age at first marriage, and age at first divorce increase with education. Moreover, the average number of marriages and births decreases with education and the shares of women who have zero marriages and zero births are highest among those with an academic degree. The pattern of childlessness is comparable to what Kreyenfeld and Konietzka (2017) and Bujard (2015) have observed for a slightly older German cohort of women, using German microcensus data. Interestingly, Kreyenfeld and Konietzka additionally show for this cohort that childlessness is largely related to whether a woman obtained an HEED or not, while the decision whether to pursue higher education is less relevant, which is also in line with our simulation results.

For men, we do not observe such linear relationships across education groups in the simulated life cycles. For instance, the share of men who never marry turns out to be inverse U-shaped in our simulations, which is not in line with what has been observed for older cohorts.[42]

[40] For instance, Frick et al. (2012) report 23.6–25.3 (22.6–24.6) as the range of mother's age at first birth (marriage) for the 1926–1965 birth cohorts, following a U-shaped pattern across cohorts.

[41] Statistisches Bundesamt (2004–2017b) reports an average duration of 15.0 marriage years for divorces in 2016.

[42] In particular, the average number of marriages is predicted to be highest for men without any post-secondary degree. This is partially driven by men with migration background who have a higher propensity to marry and are over-proportionally represented in the group of men without post-secondary degree.

Hourly wages and yearly earnings

Figures A11 and A12 in the Appendix show the evolution of the simulated hourly wages and yearly labor earnings in prices of 2019 over the life cycle, for specifications with and without estimation of birth cohort effects (cohort groups being defined in intervals of twenty years). On several dimensions, the simulated wage profiles are as expected. First, there is a clear ranking in terms of educational degrees. Academics have higher wages than individuals with a vocational degree over most parts of the life cycle. Second, wages are concave in age. Importantly, age–wage profiles not only differ between men and women but also between educational degrees. By construction, simulated age–earnings profiles show similar patterns as age–wage profiles. However, the decrease in earnings at the end of the employment career is more notable than for hourly wages. This is a consequence of the sharp reduction in working hours at these ages.

While the specifications with and without cohort effects deliver similar profiles for individuals with vocational degrees, there are larger differences for academics. For both men and women, wages and earnings are higher in the specification without cohort effects, yet this difference is stronger for women. We assess the plausibility of these simulated profiles by comparing observed hourly wages of different cohorts to simulated ones. Figure A13 in the Appendix shows this comparison for the specification without cohort effects. Simulated wages appear to approximate well wage profiles of recent cohorts. In contrast, for the specification with cohort effects (comparison with observed wages not shown), simulated wages appear rather low, particularly for academic women. Therefore, we use the wage specification without cohort effects as our preferred specification when analyzing the returns to higher education.

6 Simulation results

As discussed in Section 3.1, we use two measures to assess the returns to education: the IRR for an average employment biography and the NPV for the distribution of returns. Note that, importantly, the comparison group (i.e., the counterfactual) of an individual with a higher education degree is defined by the group of individuals who have a higher education entrance and a vocational training degree and are of the same gender and migration background. All results shown are obtained over 100 simulation runs and monetary values are in 2019 prices.

Table 1: Internal rates of return

	Men	Women
Gross return		
No pooling	11.5	13.4
	(1.0)	(0.8)
Full pooling	9.9	11.8
	(0.9)	(0.8)
Net return		
No pooling	8.7	9.7
	(0.7)	(0.5)
Full pooling	7.1	8.1
	(0.7)	(0.5)

Notes: Internal rates of return computed for average biographies. No pooling = no income pooling within households. Full pooling = complete income pooling within households. Standard errors are in parentheses. *Source*: Own calculations.

6.1 Private returns to higher education

Table 1 shows the estimated private IRR for an average life cycle. We provide results for different income concepts. First, to analyze the effect of the tax-transfer system on returns, we distinguish between returns based on gross versus net income. And second, we contrast the scenario of no income pooling of partners against the situation where all couples fully pool their income.[43]

Assuming no income pooling within households, we find a gross return of 11.5% for men and 13.4% for women. The higher returns for women compared to men are partially rooted in the hourly wage premium for academics which is larger for women. Under full income pooling, this return is reduced by about one and a half percentage points, to 9.9% for men and 11.8% for women. Intuitively, the returns to higher education under full income pooling means comparing the household incomes of academics with those of individuals with vocational degrees. The effect of income pooling on the estimated returns then crucially depends on the degree of assortative mating. If mating was completely random, the partner's income would be, on average, the same under all education levels. This would imply that an individual with a higher education degree would pool her income with the same partner income as an individual with a vocational degree, thereby mitigating the income advantage from higher education. Hence, we would expect that income pooling shrinks the returns to higher education under completely random mating. With some degree of assortative mating, however, higher educated individuals tend to have better educated spouses with higher earnings. Therefore, a larger degree of assortative mating compensates the negative effect of income pooling on the returns to a certain extent.

[43]We only present the returns under no and full income pooling as their returns are already close to each other. Furthermore, household surveys such as the EU-SILC suggest that a sizeable share of households falls into either of the two extremes (Ponthieux, 2017).

Table 2: Private NPVs, men

	Mean	Median	Share NPV< 0
	1,000 Euros		%
Gross income			
No pooling	311.4	241.0	33.9
	(35.7)	(37.2)	(2.7)
Full pooling	221.3	152.5	36.8
	(30.6)	(31.3)	(2.8)
Net income			
No pooling	172.5	130.9	36.2
	(22.2)	(23.9)	(2.6)
Full pooling	111.6	75.2	39.3
	(18.3)	(18.7)	(2.8)

Notes: No/full income pooling refers to income pooling between spouses. Standard errors are shown in parentheses. Discount rate of 2% applied. *Source*: Own calculations.

Table 3: Private NPVs, women

	Mean	Median	Share NPV< 0
	1,000 Euros		%
Gross income			
No pooling	345.8	249.3	27.4
	(25.1)	(25.5)	(2.3)
Full pooling	274.0	199.0	30.1
	(22.7)	(22.9)	(2.1)
Net income			
No pooling	183.0	127.8	30.8
	(14.8)	(15.4)	(2.4)
Full pooling	133.5	89.7	35.1
	(13.1)	(13.4)	(2.1)

Notes: No/full income pooling refers to income pooling between spouses. Standard errors shown in parentheses. Discount rate of 2% applied. *Source*: Own calculations.

Finally, the tax-and-transfer system reduces the returns to higher education further to 8.7 (7.1)% for men and 9.7 (8.1)% for women under no income pooling (full income pooling). This shows that the progressive elements of the tax-transfer system depress returns from education relatively strongly. This effect can be explained by looking at the average life cycles of academics and individuals with a vocational degree. While still in training, academics have, on average, an income disadvantage compared to individuals with vocational degrees. This income disadvantage is dampened by the progressive elements of the tax-transfer system. From graduation until retirement, however, the income advantage is inverse and academics earn substantially more. Importantly, this second effect dominates the first one and hence, the progressive elements reduce the returns to higher education.

In contrast to the IRR, the NPV can be computed for all simulated individuals. For this, we set

Table 4: Lifetime income components, in thousand Euros

	Women			*Men*		
	VOC	*HE*	Δ	*VOC*	*HE*	Δ
(1) Labor Income	665	1005	341	1174	1480	307
(2) Capital Income	18	23	5	25	30	5
(3) Income Tax	91	209	118	224	337	113
(4) Contributions	80	110	29	116	130	14
(5) Benefits	8	7	-1	7	5	-2
(6) Transfers	51	36	-15	24	14	-10
(1)+(2)-(3)-(4)+(5)+(6) *LTI, net*	570	753	183	889	1062	172

Notes: Numbers are discounted working-life averages of simulated income components in thousand Euros, differentiated by gender and education group, with *HE* higher education and *VOC* vocational training, assuming a 2% discount rate and no income pooling in married couple households. Contributions include both private health care and social security contributions (SSC). *LTI* denotes lifetime income. Regarding SSC, an incidence of 50% is assumed. *Source*: Own calculations.

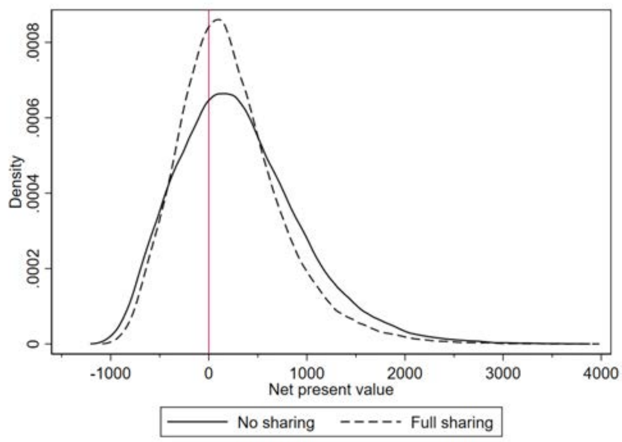
the discount rate to be 2%, similar to the OECD in recent analyses (OECD, 2019). Tables 2 and 3 present statistics on the financial gain of attaining higher education for our four income concepts. In general, the NPVs confirm the results from the IRR analysis. Assuming no income pooling, higher education increases men's gross lifetime earnings by 311,000 Euros, compared to 346,000 Euros for women. If one assumes that households fully pool their income, however, the returns shrink to 221,000 (274,000) Euros for men (women). Naturally, the tax-and-transfer system further reduces the returns. Under no income pooling, the NPV of a higher education degree is reduced to 173,000 (183,000) Euros for men (women). Assuming full income pooling we estimate the returns to be 112,000 and 134,000 Euros, respectively. In Figure A9, we plot the median NPV against different discount rates. Net NPVs are close to 250,000 Euros for a zero percent discount rate and then start converging to zero above discount rates of about seven percent.

Table 4 breaks down the returns just discussed into different components. It shows averages of income, taxes, social security contributions, benefits, and transfers by gender and education level. The largest part of returns comes from earned labor income. The difference in lifetime labor earnings, 307,000 Euros for men and 341,000 Euros for women, form the major part of the gross, no-pooling NPVs shown in Tables 2 and 3. In our model, returns due to higher capital income play a minor role for the overall private NPVs, and are simply a second-round effect from saved labor earnings.[44] The larger labor earnings under higher education result in more income taxes paid and, to a smaller extent, in higher social security contributions. At the same time, higher education decreases transfers and social security benefits received. Finally, the last line displays the increase in lifetime income under higher education, which corresponds to the private net NPV (no income pooling), as displayed

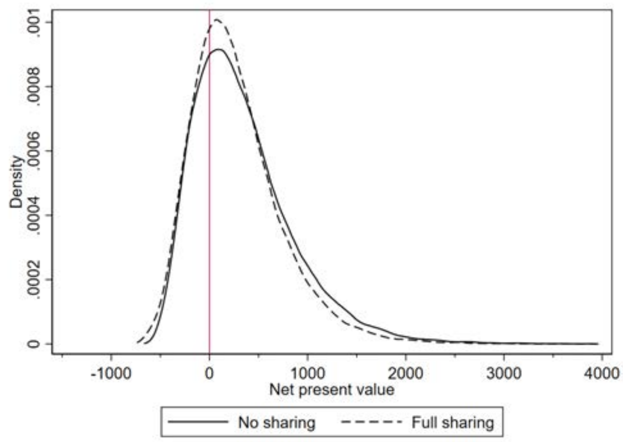
[44]We abstract from potential heterogeneities in capital investment returns.

Figure 3: Distribution of private net present values (in Thousands of Euros)

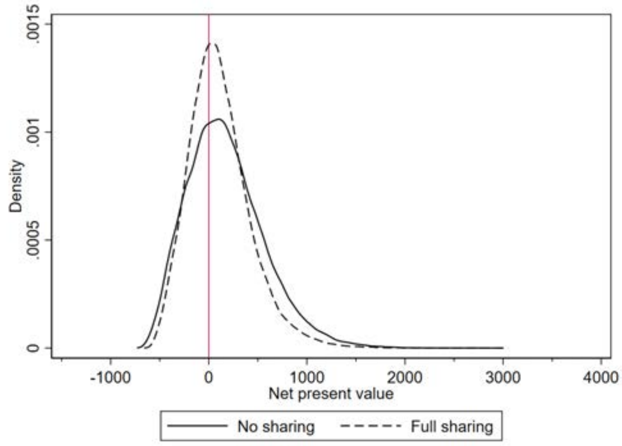
(a) Gross income, men



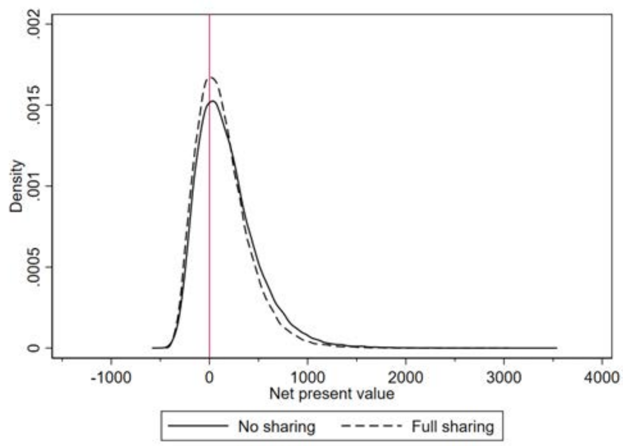
(b) Gross income, women



(c) Net income, men



(d) Net income, women



Notes: For the graphs, we pool all individuals from all 100 simulation runs. No/full sharing refers to income sharing/pooling between spouses. *Source*: Own simulations.

in Tables 2 and 3.

As argued above, the main advantage of using the NPV instead of the IRR measure is that heterogeneous returns can be analyzed. Figure 3 plots the distribution of returns for our four different income concepts.[45] They show that there is substantial heterogeneity in individual returns. On the one hand, a non-negligible fraction of individuals obtain a gross NPV larger than 1,000,000 Euros. At the same time, there is also a substantial share of life cycles with negative NPVs. For instance, assuming no income pooling and looking at gross incomes, we predict the share of negative NPVs to be 33.9% for men and 27.4% for women (see last column of Tables 2 and 3). Interestingly, income pooling compresses the distribution of returns, i.e. both strongly positive and strongly negative returns become less frequent.

What explains the heterogeneity in returns? To shed light on this question, it is helpful to consider

[45]Here, we pool all individuals from all runs. Figure A8 in the Appendix shows the distribution for each run separately.

Table 5: Life cycle characteristics by NPV deciles

	NPV, gross	LTI, gross	Hours	Wage	Yrs married	Births
Men						
1st	-628.6	629.7	31.4	13.8	15.9	0.7
2nd	-339.3	872.4	32.4	18.6	19.7	0.8
3rd	-146.7	1059.0	33.3	21.9	20.6	0.8
4th	15.5	1215.9	33.9	24.8	20.7	0.9
5th	166.5	1356.7	34.4	27.2	21.5	0.9
6th	318.3	1501.9	34.9	29.8	22.5	1.0
7th	483.7	1665.7	35.5	32.5	22.6	0.9
8th	684.6	1861.3	35.8	36.0	22.9	1.0
9th	956.8	2140.5	36.4	40.7	22.7	0.9
10th	1603.4	2796.0	37.3	51.9	22.4	1.0
Total	311.4	1509.9	34.5	29.7	21.1	0.9
Women						
1st	-323.8	373.3	19.6	12.9	21.2	2.0
2nd	-148.7	534.0	22.4	16.4	20.7	1.9
3rd	-27.3	652.9	23.9	18.8	19.9	1.8
4th	83.0	760.9	25.1	20.9	18.8	1.7
5th	192.7	870.2	26.2	23.0	17.9	1.6
6th	310.9	987.9	27.1	25.2	16.7	1.5
7th	444.4	1123.0	28.1	27.7	16.2	1.4
8th	609.7	1290.7	29.1	30.8	15.0	1.3
9th	854.0	1536.9	30.2	35.3	13.9	1.2
10th	1462.5	2153.6	31.7	47.0	12.2	1.1
Total	345.8	1028.3	26.3	25.8	17.2	1.6

Notes: NPV, LTI: Numbers are discounted working-life averages in thousand Euros, assuming a 2% discount rate and no income pooling in married couple households. *Hours* are average weekly working hours, *Wage* is the average gross hourly wage. *Yrs married* are average years married and *Births* is the average number of births. *Source*: Own calculations.

"typical" biographies with respect to the variables simulated in our model. In particular, there are patterns of household formation and employment that strongly correlate with the individual's return to higher education. Table 5 displays life cycle characteristics by NPV decile. For both men and women, working hours and hourly wages are almost monotonically increasing across deciles. For years married and births, however, the patterns differ by gender: For women, the higher the NPV decile, the lower the number of years married and the lower the number of births. For men, in contrast, those in higher deciles tend to be married longer and have fewer children, even though the absolute change across deciles is not as strong as for women.

It is important to understand that the mechanisms of household formation and employment should be seen as interdependent. In the simulation, marriage, divorce, fertility, and employment are sequentially determined and hence impact each other. As a result of the estimated transition models, for women being married and having children often goes along with working less, part-time and ultimately lower labor earnings. In contrast, for men being married and having children usually goes along with increasing employment, often working full-time and higher labor earnings.

6.2 Fiscal returns to education

Table 6 displays our estimates for the fiscal returns to higher education. As explained above, we exclude civil servants from the analysis of fiscal returns. Furthermore, we look at individual incomes, i.e. assuming no income pooling within couples.[46] We report the IRR for average life cycles as well as the central parameters of the distribution of NPVs for a discount rate of 2%. We find an IRR of 8.4% for men and 9.9% for women. In addition, we estimate the mean NPV to be about 143,000 Euros for men and 156,000 Euros for women. Similar to private returns, median NPVs are considerably lower, but still about 115,000 Euros. These finding suggest that, on average, sizable fiscal gains can be expected from an individual participating in higher education at current interest rate levels.

Table 6: Fiscal returns to higher education

	IRR	*NPV*		
		Mean	*Median*	*Share* < 0
	%	1,000 Euros		%
Men	8.4	142.6	115.1	35.8
	(0.9)	(19.1)	(21.0)	(2.7)
Women	9.9	155.7	111.6	30.7
	(0.7)	(14.8)	(16.3)	(2.5)

Notes: Standard errors in parentheses, obtained over 100 runs. For NPVs a discount rate of 2% is applied. *Source*: Own calculations.

In Table 7 we display the components of the fiscal surplus generated by average life cycles, separately by educational degrees. Also, for each component, we show the difference between higher education and vocational degrees. As for private returns, the difference in average surplus (shown in the 4th and 7th column of the bottom line) yields the average NPV for men and women, respectively. Revenue gains from higher income taxes (row 3) are the major driver of the fiscal returns to higher education, in the case of women (men) accounting for about 59% (71%) of the revenue-generating components.[47] The difference between men and women can be explained by the average labor income of men being considerably higher than that of women in combination with the progressive nature of income taxation in Germany. Since the average social security contributions rate is constant for

[46] Again, we account for the effect of joint income taxation on individual tax burdens in married couples, as described in detail in Section A.1 in the Appendix.

[47] These and all following shares can be obtained by dividing the respective monetary value of a component by the sum of all revenue-generating Δ.

Table 7: Lifetime fiscal surplus components

	Women			*Men*		
	VOC	*HE*	Δ	*VOC*	*HE*	Δ
(1) SSC employee	76	99	23	111	123	13
(2) SSC employer	63	75	12	101	102	1
(3) Income Tax	91	198	107	224	341	118
(4) VAT	65	84	19	106	127	21
(5) Benefits	8	7	-1	7	5	-2
(6) Transfers	51	37	-14	24	14	-10
(7) Relief public health care	2	4	2	2	4	2
(8) Educational expenses	14	37	24	14	37	24
(1)+(2)+(3)+(4)-(5)-(6)+(7)-(8)						
LTS	223	379	156	499	642	143

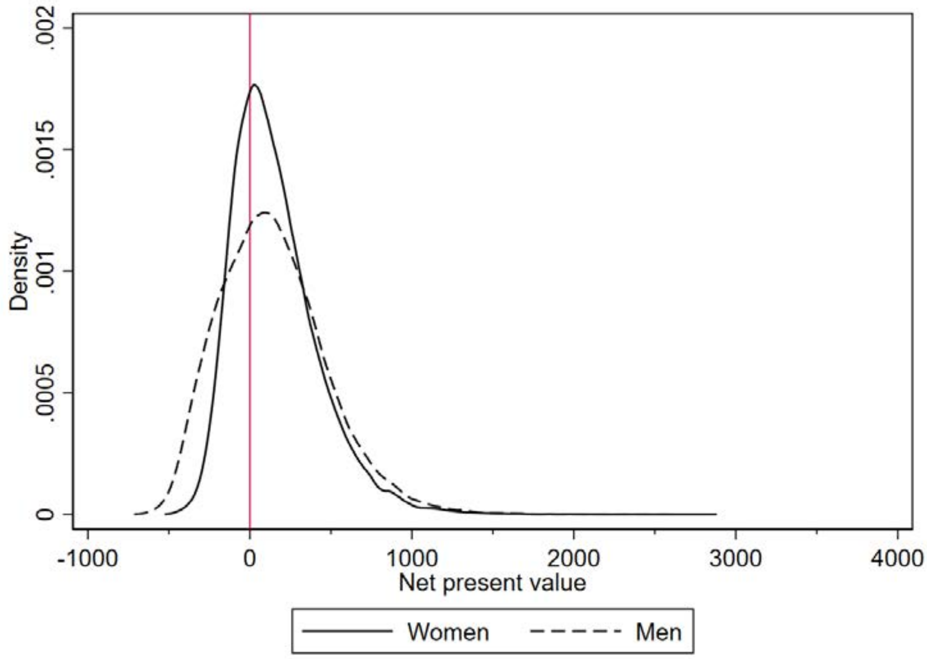
Notes: Numbers are discounted working-life averages of simulated income components in thousand Euros, differentiated by gender and education group, with *HE* higher education and *VOC* vocational training, assuming a 2% discount rate and no income pooling in married couple households. SSC denote social security contributions (pension system excluded). "Relief public health care" show public health care expenditures saved due to privately insured self-employed individuals. *Source*: Own calculations.

middle incomes and decreasing for high incomes, *SSC* (rows 1-2) contribute considerably less to the fiscal surplus components (19% and 8%) than income taxes. Also, as the share of privately health insured individuals is larger among the higher educated individuals, the rise in contributions with education is attenuated.[48] The social security contributions ceiling is the reason why the role *SSC* have for returns is less important for men than for women. As for value added taxes (row 4), we find their share to be about 11% for women and 13% for men. Due to the proportional character of the tax, the contributions to the overall fiscal surplus is limited. Lower transfers and benefits (rows 5-6) further increase the fiscal gains from higher education, while contributing no more than about 7–8% to the revenue generating components. In sum, the positive fiscal effects of taxes, *SSC*, and reduced transfers clearly outweigh the higher educational expenses and the opportunity costs due to academics' delayed labor market entry.

However, at the same time we predict negative fiscal NPVs for a sizable share of the higher educated individuals: slightly above (below) one third of the simulated life cycles of men (women) who pursue higher education imply fiscal losses when compared to the average counterfactual life cycles (see Table 6). This finding is also observable in Figure 4, which plots the distribution of fiscal returns for women and men. Again, both density functions are right-skewed, with women's returns somewhat more compressed.

[48] However, the self-employed who opt for the public health insurance scheme typically bear the full burden of contributions. In Table 7, these contributions are comprised in the position "SSC employee". Therefore, this position still rises faster with education than "SSC employer" does.

Figure 4: Distribution of the fiscal NPV of higher education



Notes: Density function of the fiscal NPV of higher education. A discount rate of 2 % is applied. *Source*: Own simulations.

7 Discussion

As explained above, there are other studies quantifying the lifetime returns to higher education. In general, it is difficult to compare our results to studies from other countries. One fundamental reason is the definition of the counterfactual to higher education. While Germany has a well-established system of vocational training which is a viable alternative for many young adults, this alternative is missing in other countries. Hence, we compare our results to two studies which also estimate the return to higher education for Germany, Pfeiffer and Stichnoth (2020) (henceforth PS) and OECD (2019) (henceforth OECD). Table 8 compares our return estimates to the ones presented in these studies. In sum, we find that our estimates for the private returns are similar to the ones in PS but considerably lower than the ones presented by the OECD. In contrast, for fiscal returns our estimates are larger than both OECD and PS, but closer to the ones reported in OECD.

Various reasons might be responsible for the differences found between our study and the others. For instance, PS exclude civil servants and self-employed from their estimation while we include both for private returns and only exclude civil servants for fiscal returns. In addition, PS assume that tax-transfer components of household income are divided equally between both partners, which might be a reason why they find lower fiscal returns (see the discussion above).

Table 8: Returns to higher education across studies

	Private				Fiscal
	Gross income		Net income		
	No pooling	Full pooling	No pooling	Full pooling	
This study	11.5, 13.4	9.9, 11.8	8.7, 9.7	7.1, 8.1	8.4, 9.9
Pfeiffer and Stichnoth (2020)	14.2	x	x	7.4	6.6
OECD (2019)	x	x	14, 16	x	6, 9

Notes: The table lists estimated IRRs to higher education for Germany (in %). When two numbers are shown, the first refers to women and the second refers to men. When one single number is shown, the return estimates for women and men are pooled. Income concepts for which no return estimates were computed are labeled with "x". *Source:* Own calculations, Pfeiffer and Stichnoth (2020), OECD (2019).

Net private returns to higher education estimated by the OECD appear surprisingly high. One reason for this should be that the OECD uses another comparison group for the group of academics. This comparison group is comprised of individuals with either higher education entrance degree or with vocational training degree. Hence, this approach precisely excludes our comparison group, i.e. those with both degrees. Another reason might be the modeling of the tax-and-transfer system. The OECD uses a very simplified model for estimating taxes and transfers for each individual. Transfers and benefits, for instance, are not simulated, which are more important for individuals with a vocational degree than for academics. Lastly, changes in methodology across OECD publications that are difficult to trace for the reader have produced substantial variation in results across recent publication years.

Finally, it should be noted that beyond the reasons mentioned above, there are fundamental differences in the modeling approaches. While PS and OECD use one single or a few recent cross-section(s), implicitly relying on the so-called synthetic cohort assumption, we account for time and cohort effects. An example is female labor force participation where we see considerable changes across birth cohorts.

One striking feature of our results is that the share of life cycles that yield negative private or fiscal NPVs to higher education is relatively high, for instance when compared to the estimates of Courtioux et al. (2014), who find an overall share of 3.5% negative private returns for France.[49] Again, the sharp difference to their results can be explained for the most part by their definition of the counterfactual income stream, which is based on *all* individuals without a higher education degree. Following this approach in our model would result in shifting the distribution of NPVs to the right, and decreasing the share of negative NPVs. However, as argued above, we think that our approach is more appropriate to estimate returns to higher education for Germany. Nevertheless, one might still

[49] Note that Courtioux et al. (2014) base their measure of negative returns on individual IRRs instead of NPVs. Even though the two measures are related, their results are not fully comparable.

wonder whether the proportion of negative NPVs is plausible. Indeed, we argue that these results have to be interpreted with some caution. First, our simulation strategy rests on the assumption that the wage residuals we impute for each individual are the result of a matching process between employee and employer. However, to some extent the empirical distribution of the wage residuals, or, more specifically their variance, also reflects measurement error in hourly wages. In our data, this can be expected in particular for hourly wages of self-employed individuals.[50] The measurement error in observed (log) hourly wages inflates the variance of residuals, and thereby the variance of simulated wages and of estimated NPVs. This is also reflected in the very low hourly wage levels we observe for the lowest NPV deciles (shown in Table 5). Second, we assume that all individuals, conditional on gender and migration background, have the same counterfactual. This implies that there is no correlation between the wage residuals drawn under higher education and those under vocational training. Hence, an academic who "draws" a high residual has the same counterfactual as an academic with a low residual. Assuming that there is some positive correlation between residuals would compress the distribution and hence imply a lower share of negative NPVs. A similar argument could be made about the correlation of economic sectors under both educational paths, for instance. In general, however, it is difficult to argue how the "correct" counterfactual would look like.

8 Conclusion

In this paper we use a dynamic microsimulation model to estimate the lifetime returns to higher education both for the individual and for the state. Going beyond most of the previous research, we account for generational trends, explore the effect of income pooling, and analyze the distribution of returns. At the individual level, we find private gross financial returns of 11.5% (13.4%) and fiscal returns of about 8.4% (9.9%) for men (women). We show how the tax-and-transfer system and the extent of income pooling within households shrink these returns. We find a large heterogeneity in returns, which is rooted in different hourly wages, employment biographies and patterns of household formation.

For a sizeable share of individuals we predict negative NPVs of an investment in higher education. It is important to emphasize, however, that these individuals are not necessarily worse off in terms of utility. The reason is that we do not take into account the non-monetary returns and costs of higher education in an individual's utility function. Several studies have investigated the effect of education on outcomes such as mental and physical health (Kamhöfer et al., 2018; Heckman et al., 2018b,a). In

[50]In the SOEP questionnaire, they are asked to estimate the income in the month before the interview, which can be expected to be fluctuating more sharply compared to dependent employees.

addition, (higher) education also has a consumption value in that studying comes along with "psychic" costs (or rewards) (Heckman et al., 2008) that differ among individuals. Moreover, many of the life cycles that yield a negative return in our simulations are characterized by more part-time work than the average life cycle under vocational training. Insofar as consumption and leisure time are normal goods, utility losses through lower income are compensated by leisure time induced utility gains.

Furthermore, as more data that follows individuals over a large parts of their life cycle becomes available, researchers will be able to compute cohorts' life-cycle returns. This might lead some researchers to assume that there is no need for simulating artificial life cycles. However, it should be clear that, by nature, fully observable life cycles imply an ex-post perspective: The returns are computed for older, already retired cohorts. If one is interested in the *expected* returns of more recent cohorts, then constructing artificial life cycles continues to be an indispensable exercise. Here lies a task for dynamic microsimulation even as long panel data sets become more readily available.

We see several potential extensions to our model. On the one hand, heterogeneity in educational careers and its impact on returns could be further analyzed. For instance, researchers will be able to model the careers of bachelor degree holders when a larger part of such life cycles will be observed. In addition, it might be interesting to explore the effects of dropout on the returns to higher education. On the other hand, future studies might account for the effect of the statutory pension, health and long-term care insurance over the whole life cycle, incorporating also differences in life expectancy and further health outcomes between education levels.

References

Abbott, A. (1995): "Sequence Analysis: New Methods for Old Ideas," *Annual Review of Sociology*, 21, 93–113.

Acemoglu, D. and J. Angrist (2000): "How large are human-capital externalities? Evidence from compulsory schooling laws," *NBER Macroeconomics Annual*, 15, 9–59.

Autorengruppe Bildungsberichtserstattung [ABB] (2018): *Bildung in Deutschland 2018: Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*, W. Bertelsmann Verlag.

Beicht, U. (2018): "Ausbildungsvergütungen 2017: Schwächerer Anstieg in West- und Ostdeutschland," Bundesinstitut für Berufsbildung, Bonn.

Beznoska, M. (2019): "Do couples pool their income? Evidence from demand system estimation for Germany," Discussion Papers 2019/3, Free University Berlin, School of Business & Economics.

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017): "Life-cycle earnings, education premiums, and internal rates of return," *Journal of Labor Economics*, 35, 993–1030.

Biewen, M. and M. Tapalaga (2017): "Life-cycle educational choices in a system with early tracking and 'second chance' options," *Economics of Education Review*, 56, 80–94.

Bönke, T., R. Glaubitz, K. Göbler, A. Harnack, A. Pape, and M. Wetter (2020): "Die Entwicklung und Prognose von Lebenserwerbseinkommen in Deutschland," Discussion Papers 2020/5, Free University Berlin, School of Business & Economics.

Böhringer, C. and W. Wiegard (2013): "Analyse der fiskalischen Auswirkungen des ermässigten Umsatzsteuersatzes in Deutschland unter Verwendung eines Simulationsmodells sowie der Wachstumseffekte von Straffungskonzepten," RWI Projektberichte, Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI), Essen.

Bonin, H., K. Reuss, and H. Stichnoth (2016): "The Monetary Value of Family Policy Measures in Germany over the Life Cycle: Evidence from a Dynamic Microsimulation Model," *CESifo Economic Studies*, 62, 650–671.

Brenke, K. and J. Pfannkuche (2018): "Konsum und Sparquote der privaten Haushalte hängen stark vom Erwerbsstatus, Einkommen und Alter ab," *DIW-Wochenbericht*, 85, 181–191.

Brzinsky-Fay, C., U. Kohler, and M. Luniak (2006): "Sequence analysis with Stata," *The Stata Journal*, 6, 435–460.

Bujard, M. (2015): "Kinderlosigkeit in Deutschland: Wie interagieren Bildung, Wohnort, Migrationshintergrund, Erwerbstätigkeit und Kohorte?" *Journal of Family Research*, 27, 270–296.

Bundesagentur für Arbeit (2019): "Der Arbeitsmarkt in Deutschland," Berichte: Blickpunkt Arbeitsmarkt– Monatsbericht zum Arbeits- und Ausbildungsmarkt, Nünberg.

Carneiro, P. M. and J. J. Heckman (2003): "Human Capital Policy," IZA Discussion Papers No. 821, Institute for the Study of Labor (IZA), Bonn.

Ciccone, A. and G. Peri (2006): "Identifying human-capital externalities: Theory with applications," *The Review of Economic Studies*, 73, 381–412.

Courtioux, P., S. Gregoir, and D. Houeto (2014): "Modelling the distribution of returns on higher education: A microsimulation approach," *Economic Modelling*, 38, 328–340.

Courtioux, P. and V. Lignon (2016): "A good career or a good marriage: The returns of higher education in France," *Economic Modelling*, 57, 221–237.

Cunha, F. and J. J. Heckman (2007): "Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling," *Labour Economics*, 14, 870–893.

de La Fuente, A. and J. F. Jimeno (2009): "The private and fiscal returns to schooling in the European Union," *Journal of the European Economic Association*, 7, 1319–1360.

Deaton, A. (1997): *The analysis of household surveys*, Washington, D.C.: The World Bank, chap. 2, 63–132.

Dekkers, G. (2016): "What are the Driving Forces behind Trends in Inequality among Pensioners? Validating MIDAS Belgium Using a Stylized Model," in *New Pathways in Microsimulation*, ed. by C. O. Gijs Dekkers and M. Keegan, London: Routledge, chap. 19, 287–304.

Eide, E. R. and M. H. Showalter (2011): "Estimating the relation between health and education: What do we know and what do we need to know?" *Economics of Education Review*, 30, 778–791.

Eika, L., M. Mogstad, and B. Zafar (2019): "Educational assortative mating and household income inequality," *Journal of Political Economy*, 127, 2795–2835.

Frank, I. and G. Walden (2012): "Ausbildungsdauer von Berufen – verbindliche Begrenzung auf drei Jahre?" *BWP-Berufsbildung in Wissenschaft und Praxis*, 4, 20–23.

Frick, J., M. Grabka, A. Rasner, M. Schmidt, M. Schuth, and C. Westermeier (2012): "Familienbiographische Verläufe im Kohortenvergleich," SOEPpapers on Multidisciplinary Panel Data Research 439, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Geyer, J. and V. Steiner (2014): "Future public pensions and changing employment patterns across birth cohorts," *Journal of Pension Economics and Finance*, 13, 172–209.

Glocker, D. and J. Storck (2014): "Risks and returns to educational fields–A financial asset approach to vocational and academic education," *Economics of Education Review*, 42, 109–129.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2018): "The German Socio-Economic Panel (SOEP)," *Jahrbücher für Nationalökonomie und Statistik*, 239, 345–360.

Haan, P., D. Kemptner, and V. L. Prowse (2017): "Insurance, redistribution, and the inequality of lifetime income," DIW Discussion Papers No. 1716, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018a): "The Nonmarket Benefits of Education and Ability," *Journal of Human Capital*, 12, 282–304.

——— (2018b): "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking," *Journal of Political Economy*, 126, 197–246.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, ed. by E. Hanushek and F. Welch, Amsterdam: Elsevier, vol. 1, chap. 7, 307–458.

——— (2008): "Earnings functions and rates of return," *Journal of Human Capital*, 2, 1–31.

Kamhöfer, D. A., H. Schmitz, and M. Westphal (2018): "Heterogeneity in Marginal Non-Monetary Returns to Higher Education," *Journal of the European Economic Association*, 17, 205–244.

Katz, L. and D. Autor (1999): "Changes in the wage structure and earnings inequality," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Amsterdam: Elsevier, vol. 3A, chap. 26, 1463–1555.

Kreyenfeld, M. and D. Konietzka (2017): "Childlessness in East and West Germany: Long-term trends and social disparities," in *Childlessness in Europe: Contexts, causes, and consequences*, ed. by M. Kreyenfeld and D. Konietzka, Cham: Springer International Publishing, Demographic Research Monographs, chap. 5, 97–114.

Levell, P. and J. Shaw (2015): “Constructing full adult life-cycles from short panels,” IFS Working Papers No. W15/01, Institute for Fiscal Studies (IFS), London.

Li, J. and C. O'Donoghue (2013): “A survey of dynamic microsimulation models: uses, model structure and methodology,” *International Journal of Microsimulation*, 6, 3–55.

——— (2014): “Evaluating Binary Alignment Methods in Microsimulation Models,” *Journal of Artificial Societies and Social Simulation*, 17.

Lütz, C. and P. Stein (2020): ““Validierung in dynamischen Mikrosimulationsmodellen”,” in *Mikrosimulationen*, ed. by M. Hannappel and J. Kopp, Wiesbaden: Springer Fachmedien, 141–176.

Mincer, J. (1974): *Schooling, Experience, and Earnings*, New York: Columbia University Press for NBER.

Moretti, E. (2004): “Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data,” *Journal of Econometrics*, 121, 175–212.

Mullahy, J. (2015): “Multivariate fractional regression estimation of econometric share models,” *Journal of Econometric Methods*, 4, 71–100.

Nybom, M. (2017): “The distribution of lifetime earnings returns to college,” *Journal of Labor Economics*, 35, 903–952.

OECD (2019): *Education at a glance 2019: OECD indicators*, Paris: OECD Publishing.

——— (2020): “Adult education level (indicator),” Online: https://data.oecd.org/eduatt/adult-education-level.htm#indicator-chart, accessed: Aug. 2020.

Papke, L. E. and J. M. Wooldridge (1996): “Econometric methods for fractional response variables with an application to 401 (k) plan participation rates,” *Journal of Applied Econometrics*, 11, 619–632.

——— (2008): “Panel data methods for fractional response variables with an application to test pass rates,” *Journal of Econometrics*, 145, 121–133.

Pfeiffer, F. and H. Stichnoth (2019): "Fiskalische und individuelle Nettoerträge und Renditen von Bildungsinvestitionen im jungen Erwachsenenalter," SOEPpapers on Multidisciplinary Panel Data Research No. 1022, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

——— (2020): "Fiscal and individual rates of return to university education with and without graduation," ZEW Discussion Papers No. 20-016, ZEW - Leibniz Centre for European Economic Research.

Ponthieux, S. (2017): "Intra-household pooling and sharing of resources: a tentative "modified" equivalised income," in *Monitoring social inclusion in Europe*, ed. by A. B. Atkinson, G. Anne-Catherine, and E. Marlier, Luxembourg: Eurostat, chap. 9, 175–190.

Ponthieux, S. and D. Meurs (2015): "Gender inequality," in *Handbook of income distribution*, ed. by A. B. Atkinson and F. Bourguignon, Elsevier, vol. 2, chap. 12, 981–1146.

Psacharopoulos, G. (1995): "The Profitability of Investment in Education: Concepts and Methods," Human Capital Development and Operations Policy Working Papers 63, World Bank, Washington, D.C.

Psacharopoulos, G. and H. A. Patrinos (2018): "Returns to investment in education: a decennial review of the global literature," *Education Economics*, 26, 445–458.

Ramalho, E. A., J. J. Ramalho, and J. M. Murteira (2011): "Alternative estimating and testing empirical strategies for fractional regression models," *Journal of Economic Surveys*, 25, 19–68.

Scherer, S. (2001): "Early career patterns: A comparison of Great Britain and West Germany," *European Sociological Review*, 17, 119–144.

Statistisches Bundesamt (2004–2017a): *Bevölkerung und Erwerbstätigkeit. Natürliche Bevölkerungsbewegung (Fachserie 1, Reihe 1.1)*, Wiesbaden.

——— (2004–2017b): *Bevölkerung und Erwerbstätigkeit. Statistik der rechtskräftigen Beschlüsse in Ehelösungssachen (Fachserie 1, Reihe 1.4)*, Wiesbaden.

Statistisches Bundesamt (2017): *Bildungsfinanzbericht 2017*, Wiesbaden.

Statistisches Bundesamt (2017): *Krankheitskosten (Fachserie 12 Reihe 7.2.1)*, Wiesbaden.

——— (2018): *Bildungsstand der Bevölkerung – Ergebnisse des Mikrozensus 2017*, Wiesbaden.

——— (2019a): "A changing population - Assumptions and results of the 14th coordinated population projection," Wiesbaden.

——— (2019b): “Cohort-specific birth rates (live births per 1000 women): Germany, woman’s year of birth, age.” Online: https://www-genesis.destatis.de/genesis/online/data, accessed: Jan. 2019.

——— (2019c): “Live births per 1000 women: Germany, years, age.” Online: https://www-genesis.destatis.de/genesis/online/data, accessed: Jan. 2019.

Steiner, V. and C. Lauer (2000): “Private Erträge von Bildungsinvestitionen in Deutschland,” ZEW Discussion Papers No. 00-18, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim.

Steiner, V., K. Wrohlich, P. Haan, and J. Geyer (2012): “Documentation of the tax-benefit microsimulation model STSM: Version 2012,” DIW Data Documentation, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

Walker, I. and Y. Zhu (2011): “Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales,” *Economics of Education Review*, 30, 1177–1186.

Appendix

A.1 Further simulation assumptions

Vocational and academic training

We assume that individuals spend six (three) years in higher education (vocational training). Furthermore, we ignore any potential educational breaks between graduation from school and the beginning of post-secondary education. We use the SOEP to compute mean levels of gross labor income during this period. Following this approach, we assign students in higher education a monthly income of 392 Euros, while vocational trainees receive 744 Euros. In addition, students in higher education (vocational training) receive 109 (252) Euros per month as a student grant (part of *BAföG*), which is the unconditional mean *BAföG* payment given as a grant. To quantify the fiscal returns, we assign each higher education (vocational training) place the mean cost of 6500 (4600) Euros per year.

Family formation

We assume that each individual is either single or married and if married, is living in a joint household with the spouse. Changes between these two states occur by marriage and divorce and new spouses are selected from within the base population.[51] Due to dynamic aging, spouses have to be of same age. Therefore, we align marriages uniformly to the marriage target of women.[52] When parents are simulated to divorce, their children are assumed to stay in the mother's household. Also, the children that are born to single women are not assigned to single men. Children are assumed to leave the household at age 18.

Employment

Labor market experience is updated every age year after the simulation of the current employment state. We assume that part-time employment states increase labor market experience by half a year while full-time states increase experience by one year. For tenure, we define an additional target that defines the age-specific share of working individuals that change their employer from one year to another, conditional on being employed in both periods. This target is defined separately by gender and educational degree. In our simulation, the target rates are used to randomly select working individuals to either stay with their current employer or change to a new employer.

We account for the empirical observation that episodes of involuntary unemployment tend to last

[51] In the microsimulation literature this is called a "closed model".

[52] We opt for the female instead of the male marriage pattern since timing of marriages is closely related to the timing of births, which are more relevant for female than for male labor supply.

less than a year for a large fraction of individuals entering unemployment. We do this by re-setting the length of one-year unemployment spells to three months. Accordingly, we assume that the other nine months of that year are equally split to the employment state of the preceding and the subsequent year, respectively.

Wages

Having estimated the hourly wage regressions we predict individual wages and store the residuals. At the beginning of each new employment episode, we assign each individual a random draw from the distribution of log wage residuals, depending on educational attainment, gender, and self-employment status. This can be interpreted as the result of a matching process between employee and employer. Moreover, individuals are randomly assigned to an economic sector as well as the position on the labor market (self-employed, civil servant or otherwise employed) based on the observed shares in the SOEP data and depending on educational attainment and gender. All monetary values are in 2019 prices.

Savings and wealth

For the sake of simplicity, we assume that in each period individuals accumulate savings according to an age-specific average savings rate as estimated by Brenke and Pfannkuche (2018), yielding capital income in each period, calculated under the hypothesis of a real interest of 2%. As in Haan et al. (2017), we assume that if, in a given period, the household's disposable income is below the minimum income level that social assistance guarantees, and if the social assistance wealth test is not passed, this household dis-saves the amount that is necessary to reach the minimum income level.

Taxes and social security contributions

The microsimulation takes into account the main features of the German income tax and social security contributions (SSC) regime. We simulate employee's and employer's SSC to health, long-term care, and unemployment insurance based on the individual status in the labor market. We exclude the pension system from our analysis. This implies that contributions as well as benefits related to social and private pension are set to zero throughout the simulated part of the life cycles.[53] Self-employed persons and civil servants are free to choose between private health insurance and voluntary public health insurance, therefore we randomly assign them to one of the two options according to income-quintile-specific shares observed in the SOEP. For those privately insured, we impute contributions

[53]Pension benefits are partly subject to (progressive) income taxation in Germany, which undermines our proportionality argument. At the same time, not accounting for the (regressive) effect of the deduction of old-age contributions should counteract this bias.

based on age-gender-specific average contributions reported in the SOEP. Since the share of privately insured individuals differs between levels of educational attainment, we also take the costs of public health insurance into account, i.e. for those privately insured, the public health system is assumed to save the age-gender-specific average annual cost of public health care.[54]

For the simulation of personal income taxes, we compute taxable income using the simulated income from self-employment, income from dependent employment, income from capital, and by deducting the lump-sum allowances and deductible expenses such as the simulated SSC. We further assume that married couples always opt for joint taxation of income due to the financial gains of income splitting in Germany. Concerning the intra-household division of the tax burden, we work on the assumption that all couple households split taxes according to tax class IV/IV and the so-called factor method (*Faktorverfahren*), which both reflects the financial gain of income splitting and income differences within a couple. We also apply the higher-yield test between child benefit and child tax allowance, as is implemented in the tax laws. Finally, for earnings from interest we apply the higher-yield test that compares tax levels to the flat rate withholding tax scheme.

For the simulation of value added taxes (VAT) we follow the simple approach proposed in Pfeiffer and Stichnoth (2019), who calculate fiscal revenue from VAT using simulated disposable income and a hypothetical uniform VAT level (16,712%) computed by Böhringer and Wiegard (2013) as the budget-neutral equivalent of the existing system incorporating two product-group specific VAT levels and a pseudo tax exemption for housing and insurance expenditures, inter alia. However, we define disposable income less savings as the relevant tax base.

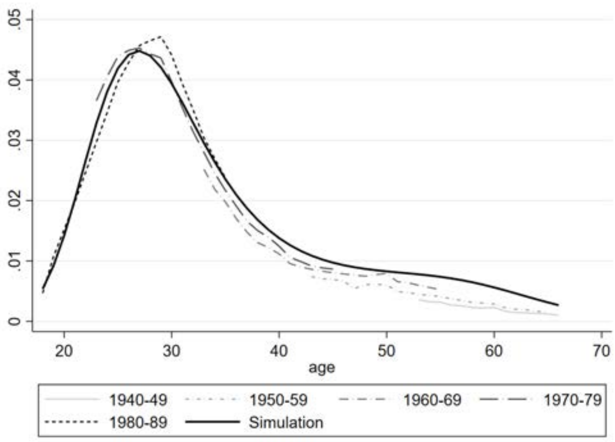
Welfare programs

We simulate unemployment benefits, parental leave allowances and public student grants on the individual level, and social assistance, housing benefits, child benefits and additional child benefits on the household level. Insofar as transfer rules refer to individual needs that are not simulated within the model (e.g. housing costs, heating costs), upper thresholds as specified by law are applied.

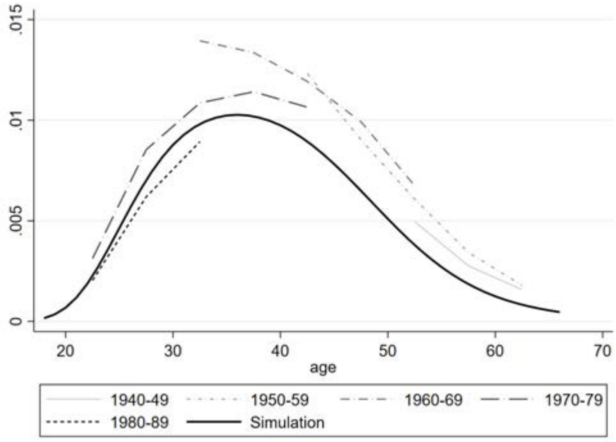
[54] Average health costs are taken from Statistisches Bundesamt (2017).

Figure A1: Observed vs. predicted patterns of family formation, women

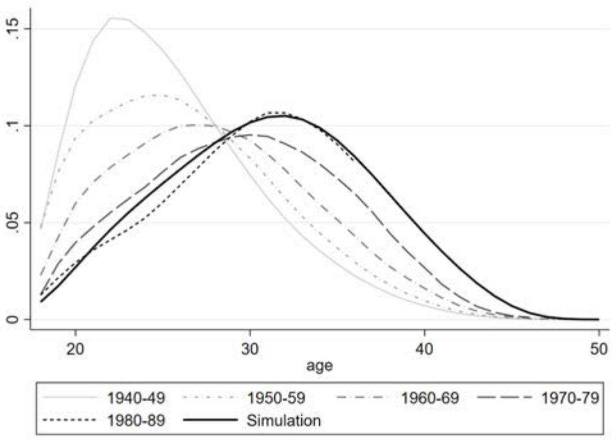
(a) Marriage rate



(b) Divorce rate



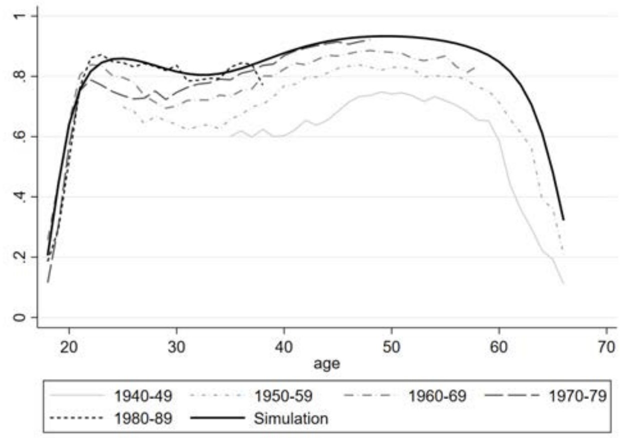
(c) Fertility rate



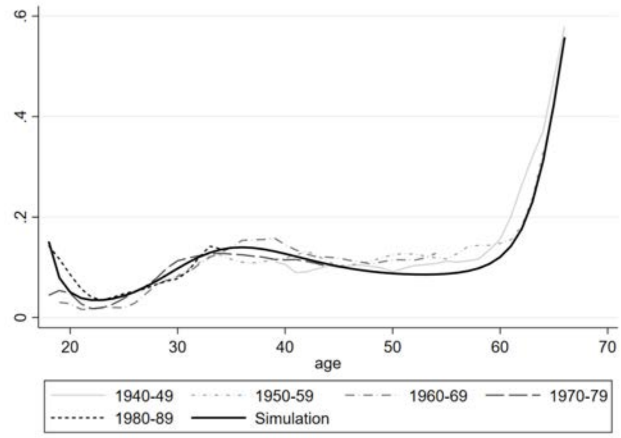
Notes: Observed aggregate marriage, divorce and fertility rates are displayed by birth cohort. For instance, the line "1970-1979" states the marriage, divorce, and fertility rates of the individuals born between 1970 and 1979 over the life cycle as observed in the data. Simulation: predicted life-cycle pattern for the 1980–89 cohort. *Source:* Calculations based on Statistisches Bundesamt (2019b), Statistisches Bundesamt (2019c), Statistisches Bundesamt (2004–2017a), Statistisches Bundesamt (2004–2017b), and own simulations.

Figure A2: Observed vs. predicted patterns of labor market participation, women

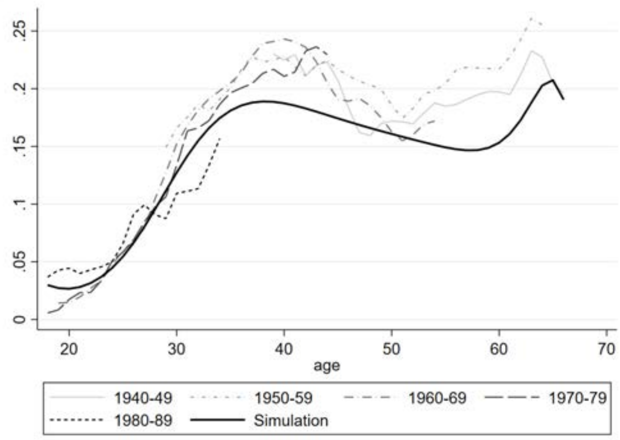
(a) Labor force participation rate



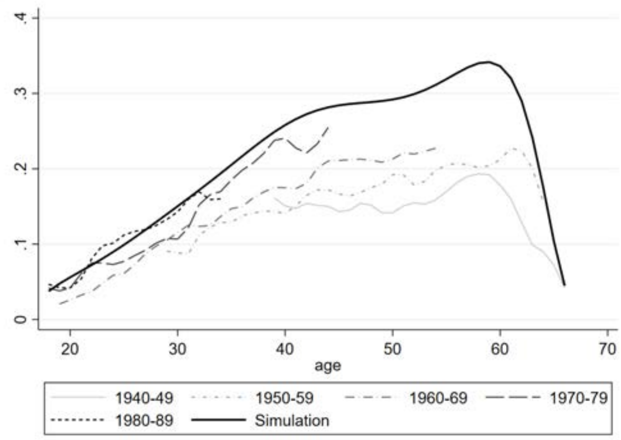
(b) Reduced part-time employment rate



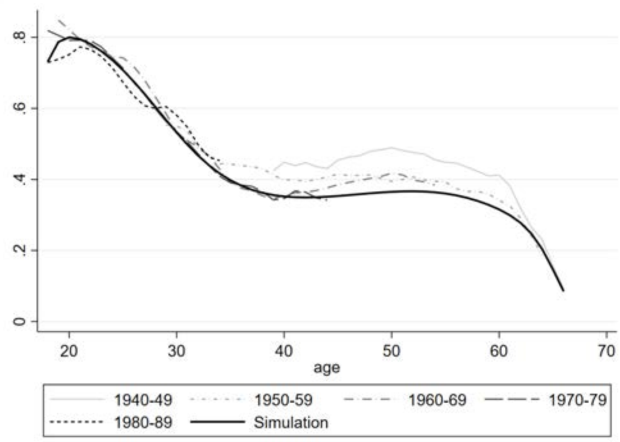
(c) Standard part-time employment rate



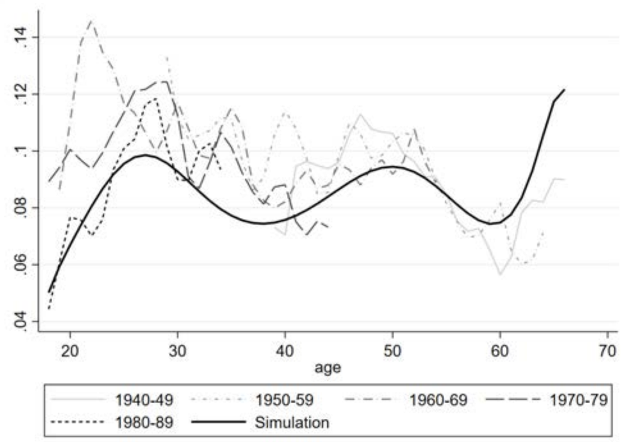
(d) Extended part-time employment rate



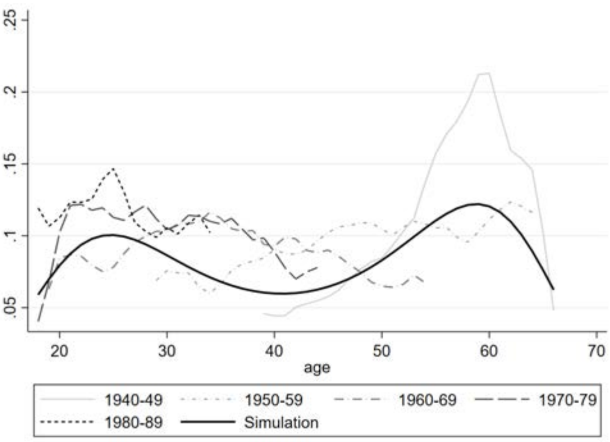
(e) Full-time employment rate



(f) Over-time employment rate



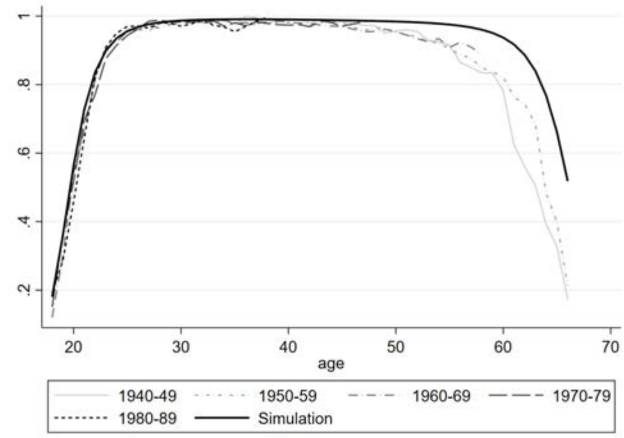
(g) Unemployment



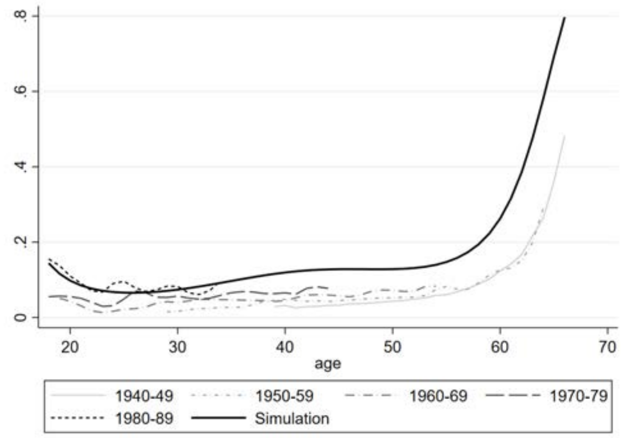
Notes: Observed aggregate rates are shown as three-year moving averages and by birth cohort. Employment shares in subfigures (b)–(g) conditional on labor force participation. Simulation: predicted life-cycle pattern for the 1980–89 cohort. *Source:* Own calculations based on SOEP v35, and own simulations.

Figure A3: Observed vs. predicted patterns of labor market participation, men

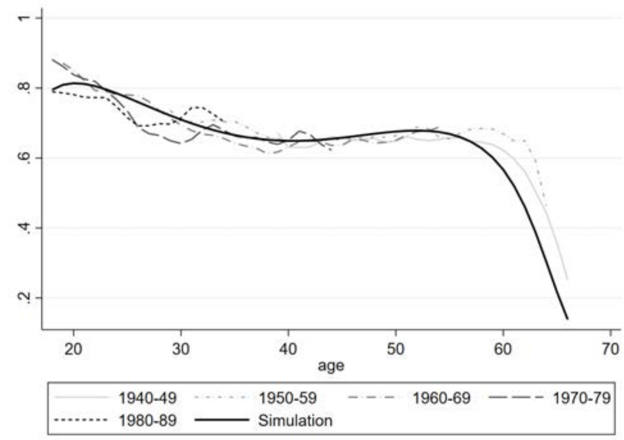
(a) Labor force participation rate



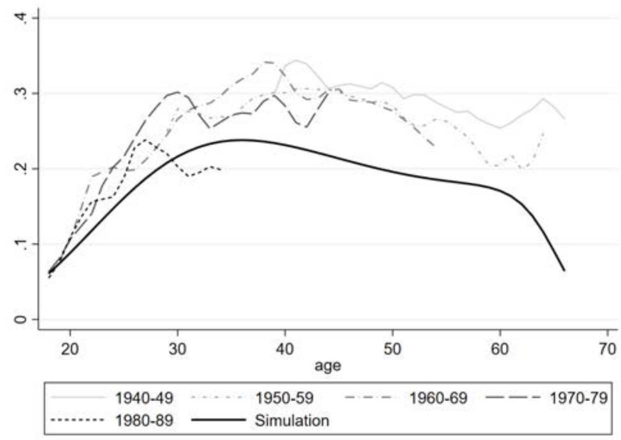
(b) Standard part-time employment rate



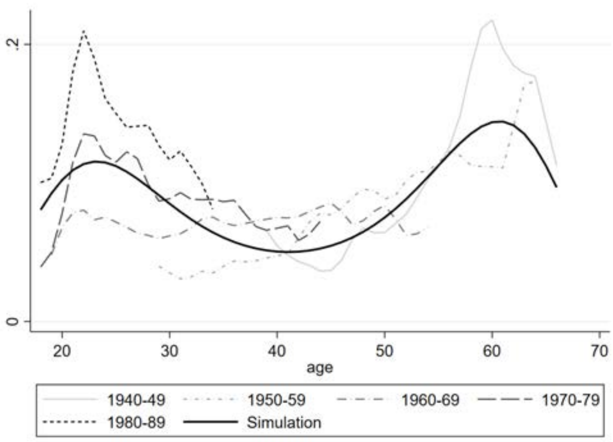
(c) Full-time employment rate



(d) Over-time employment rate



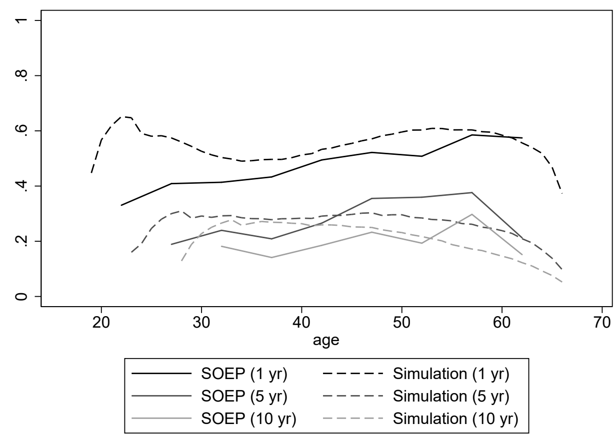
(e) Unemployment rate



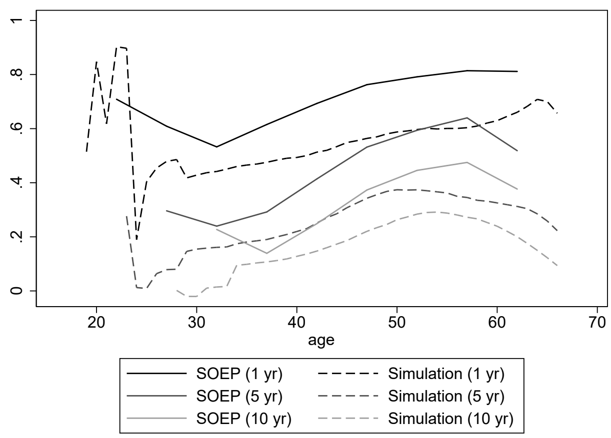
Notes: Observed aggregate rates are shown as three-year moving averages and by birth cohort. Employment shares in subfigures (b)–(e) conditional on labor force participation. Simulation: predicted life-cycle pattern for the 1980–89 cohort. *Source:* Own calculations based on SOEP v35, and own simulations.

Figure A4: Autocorrelations for employment and family status of women, simulated vs. observed.

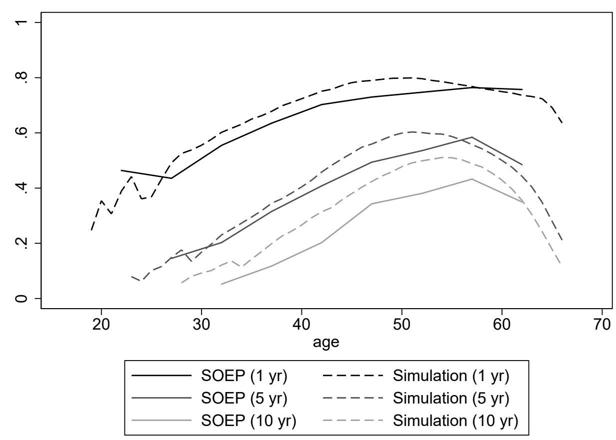
(a) Unemployment



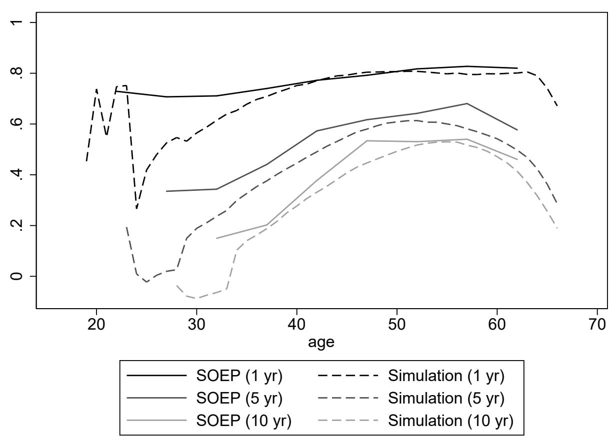
(b) Employment



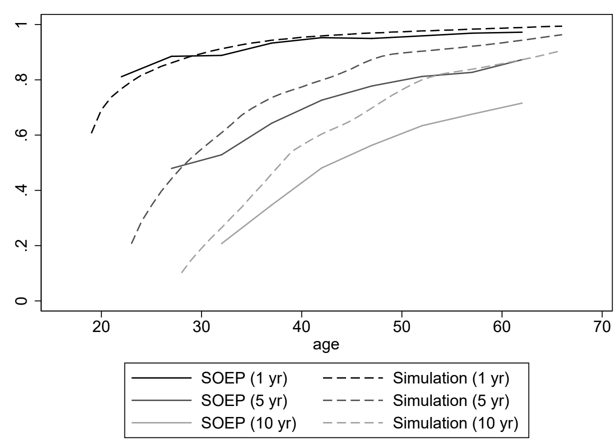
(c) Part-time employment



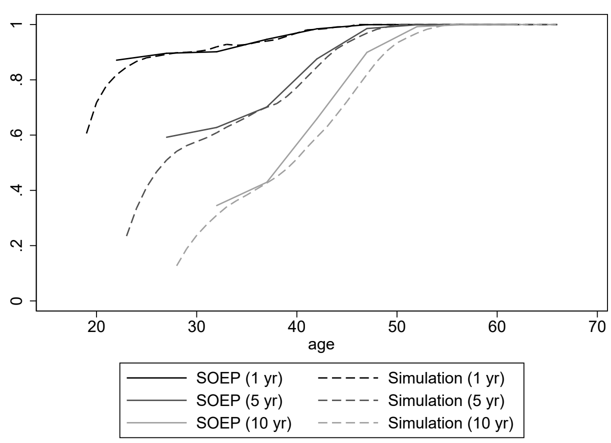
(d) Full-time employment



(e) Marriage status



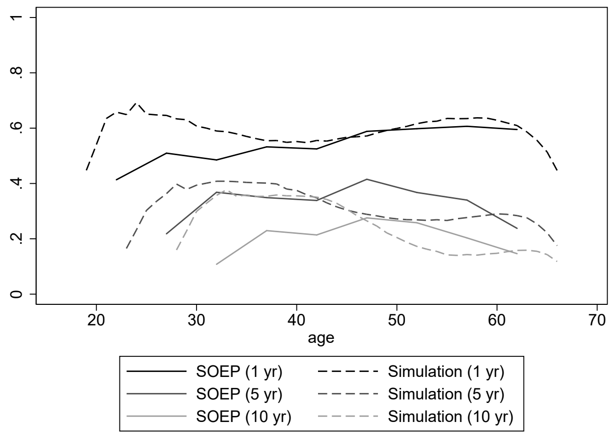
(f) Parent status



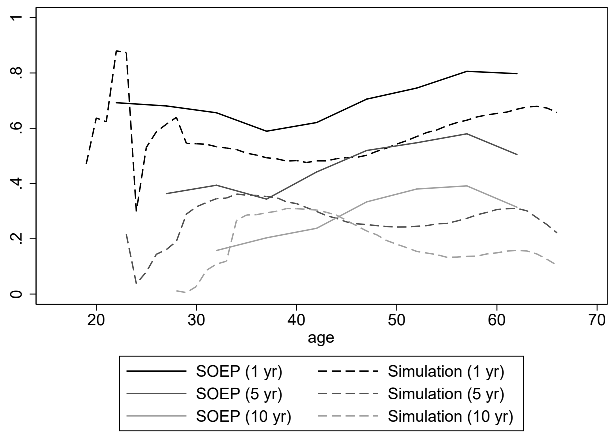
Notes: Autocorrelations with one-, five-, and ten-year lags. Observed values are moving averages of order three for cohorts starting from birth year 1950, weighted by the SOEP weighting factors. Simulated values are averaged over a set of 100 simulations. *Source:* Own calculations based on SOEP v35, and own simulations.

Figure A5: Autocorrelations for employment and family status of men, simulated vs. observed.

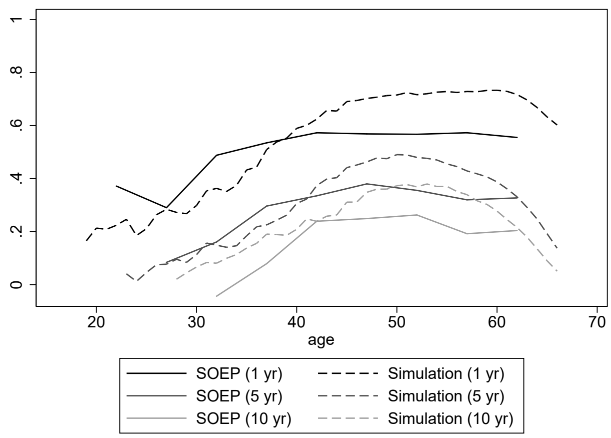
(a) Unemployment



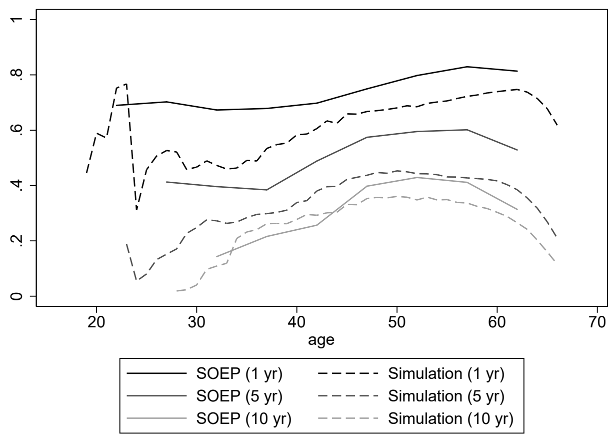
(b) Employment



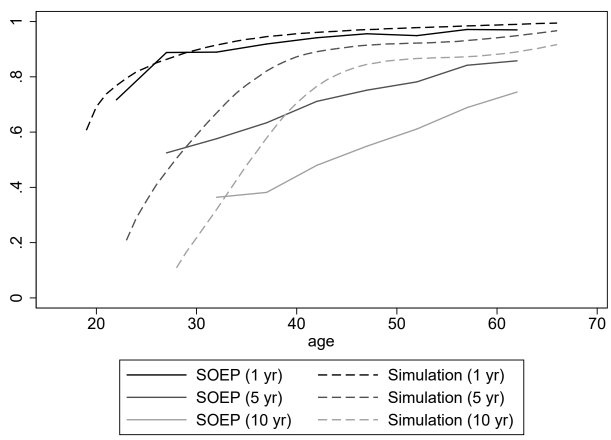
(c) Part-time employment



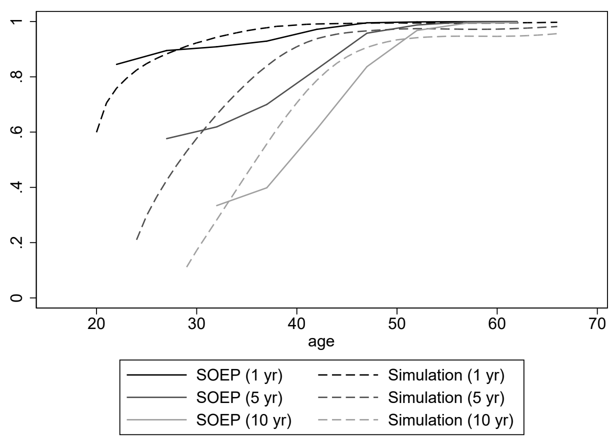
(d) Full-time employment



(e) Marriage status



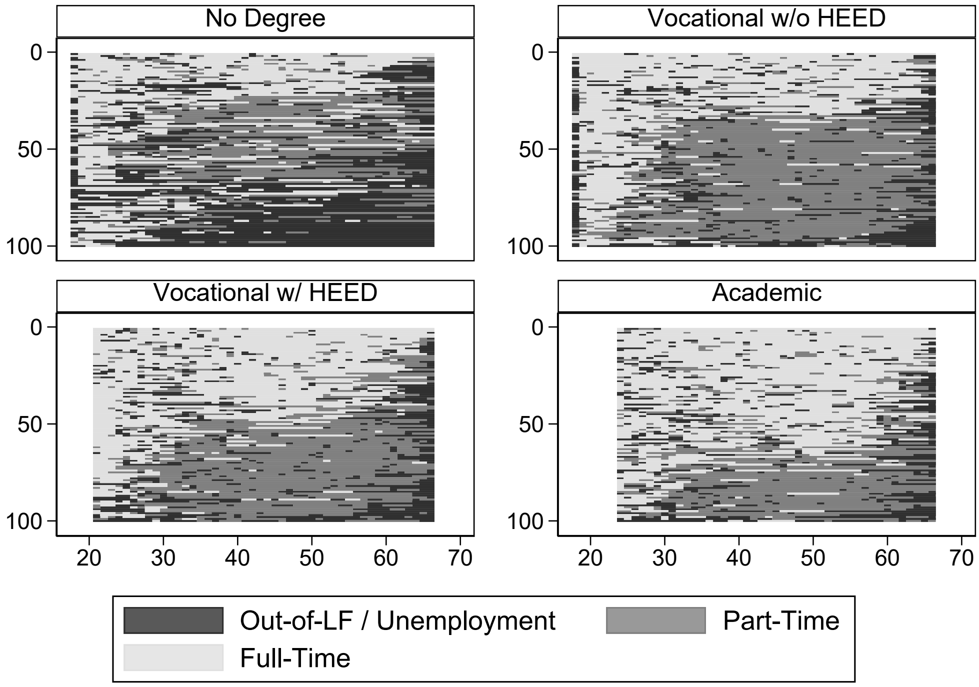
(f) Parent status



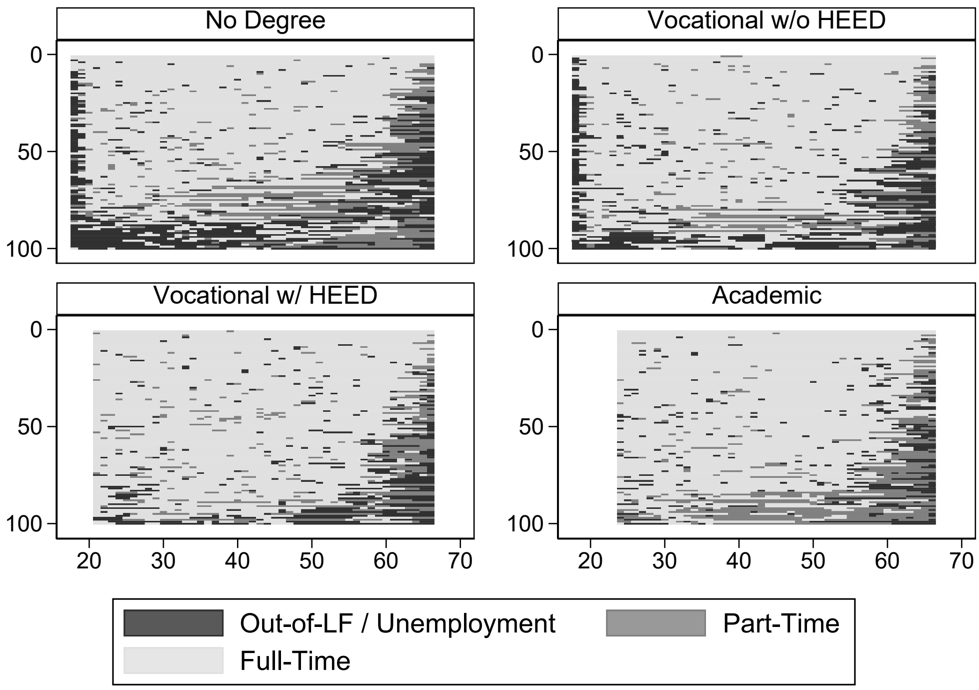
Notes: Autocorrelations with one-, five-, and ten-year lags. Observed values are moving averages of order three for cohorts starting from birth year 1950, weighted by the SOEP weighting factors. Simulated values are averaged over a set of 100 simulations. *Source:* Own calculations based on SOEP v35, and own simulations.

Figure A6: Sequence index plot for employment status by age

(a) Women



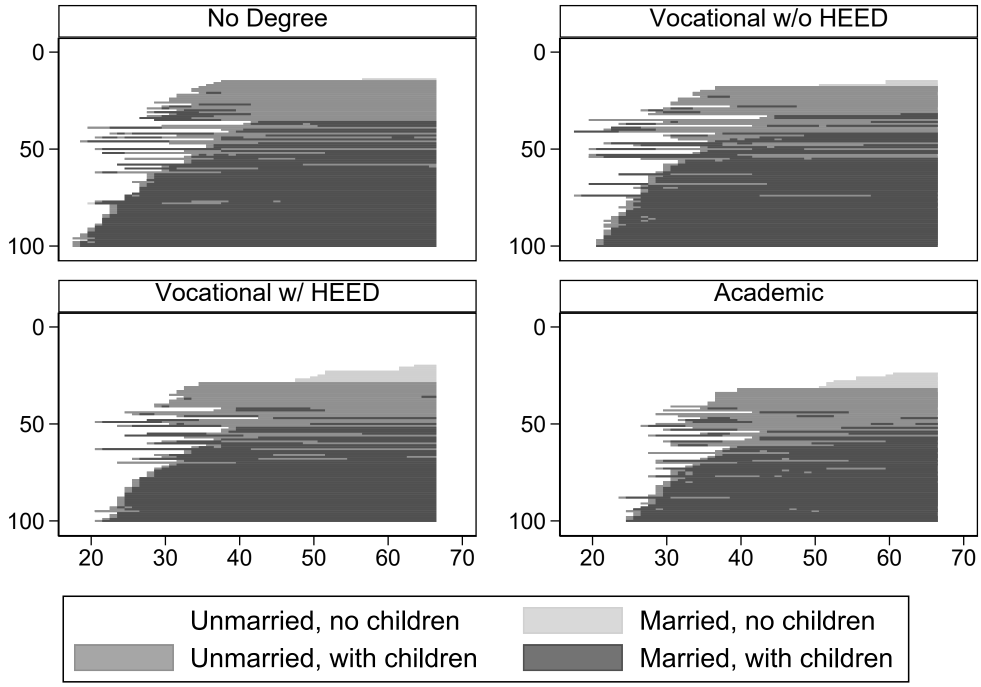
(b) Men



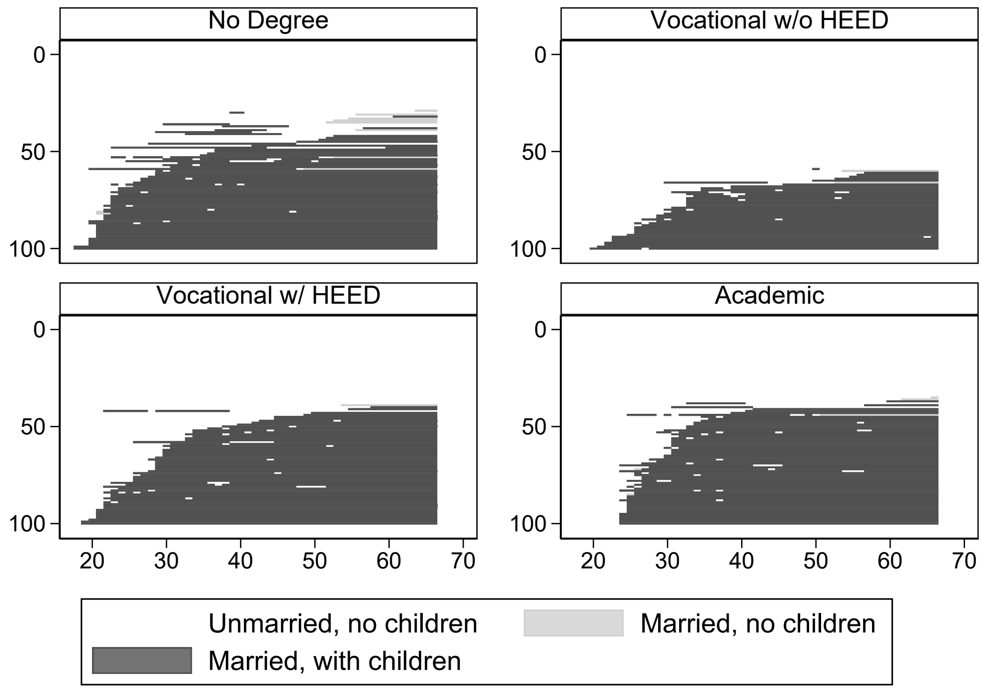
Notes: For each combination of gender and education level, a sequence index plot displays 100 life cycles randomly drawn from the universe of simulated individuals. "Out-of-LF / Unemployment" are episodes out of labor force (not including academic and vocational training) or in involuntary unemployment. "Part-Time" includes episodes of employment with 1-34 weekly working hours. Episodes in "Full-Time" comprise employment states with at least 35 weekly working hours. Sequences are sorted according to the Levenshtein distance to a reference individual with the highest number of years in full-time employment (Brzinsky-Fay et al., 2006). *Source*: Own simulations.

Figure A7: Sequence index plot for marriage and parental status by age

(a) Women



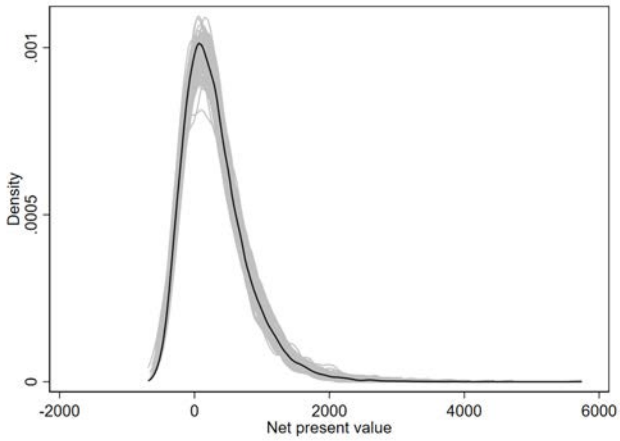
(b) Men



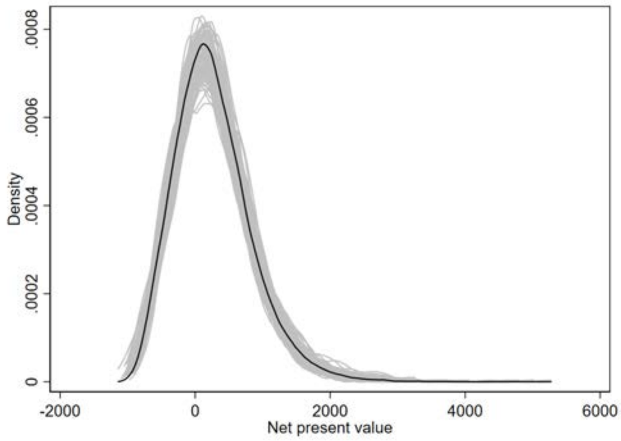
Notes: For each combination of gender and education level, a sequence index plot displays 100 life cycles randomly drawn from the universe of simulated individuals. Sequences are sorted according to the Levenshtein distance to a reference individual without any marriages or births over the life cycle (Brzinsky-Fay et al., 2006). As in our model children are assumed to stay in their mothers' households after parental divorce, "unmarried, with children" episodes are not possible for men by construction. *Source*: Own simulations.

Figure A8: Distribution of NPVs of higher education attainment, all runs separately

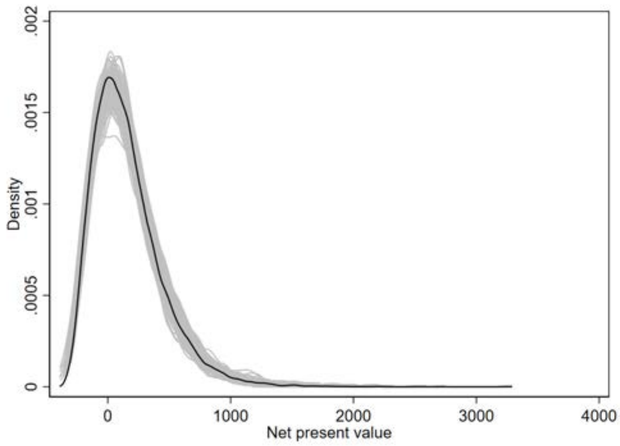
(a) Women, gross income



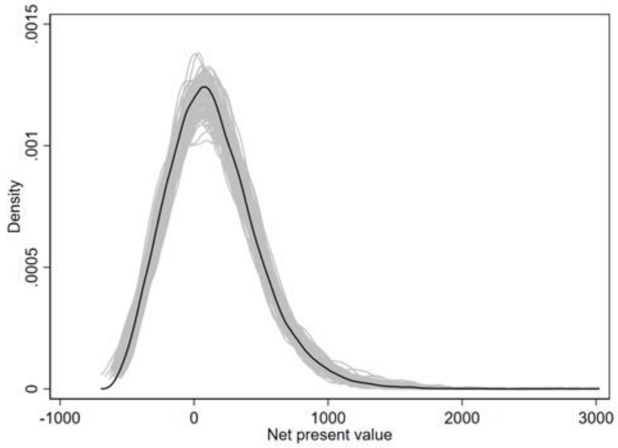
(b) Men, gross income



(c) Women, net income

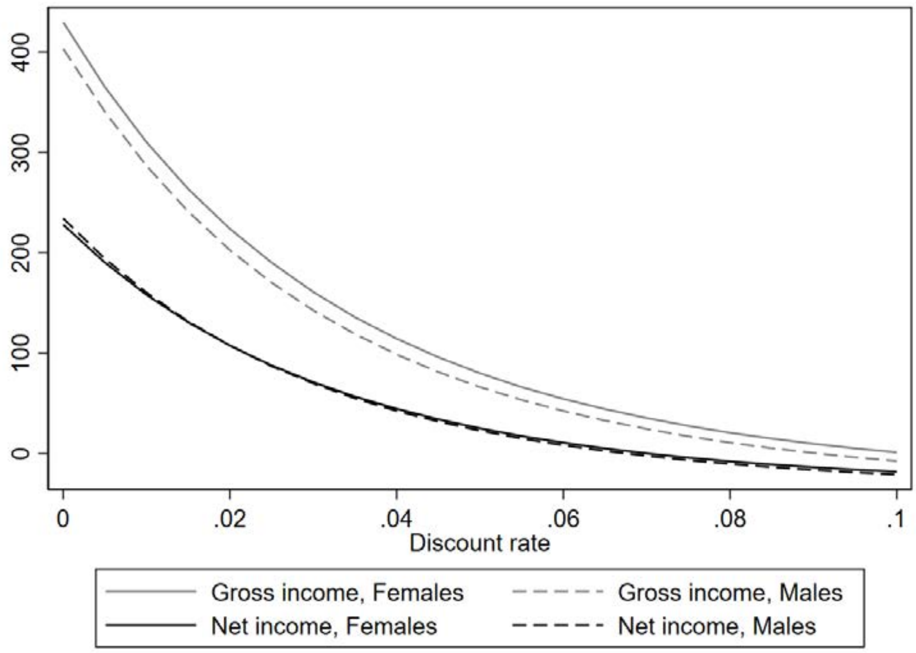


(d) Men, net income



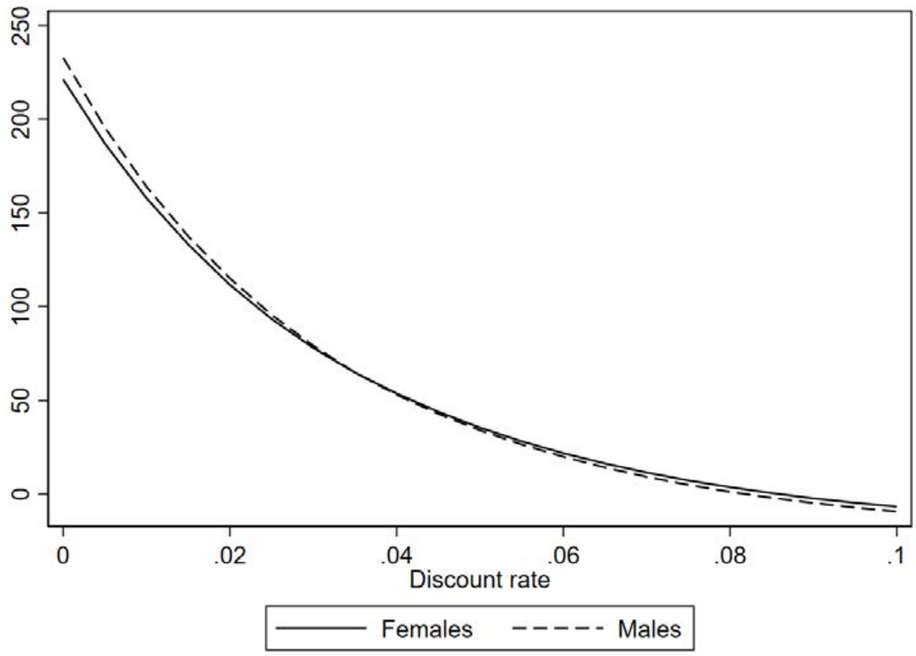
Notes: Density functions for individual simulation runs displayed in light grey, single density function for pooled runs in black. For married couples, 50% income pooling is assumed. *Source:* Own simulations.

Figure A9: Median private NPV by discount rate



Notes: Median private NPVs of higher education by discount rate and income concept, averaged over 100 runs. For married couples, 50% income pooling is assumed. *Source*: Own simulations.

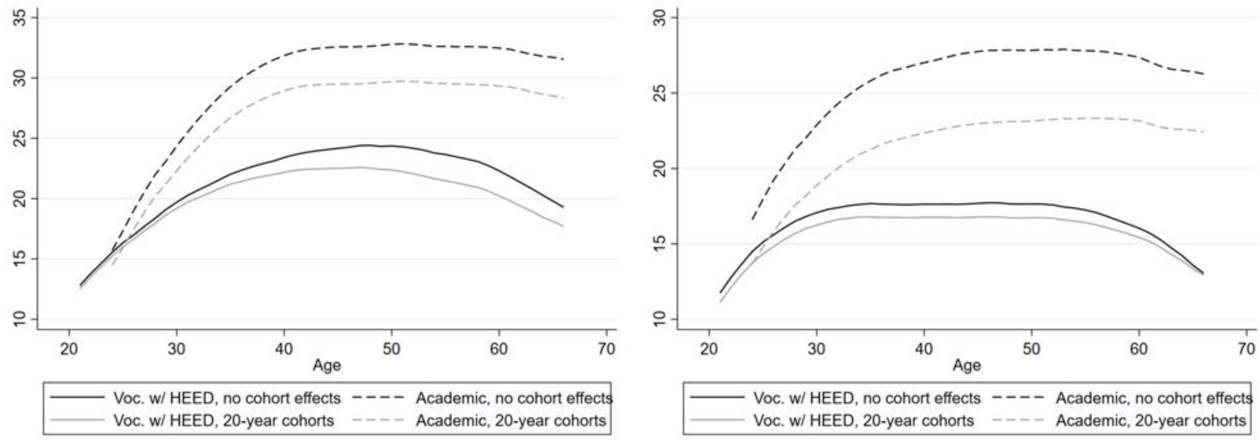
Figure A10: Median fiscal NPV by discount rate



Notes: Median fiscal NPVs of higher education by discount rate and income concept, averaged over 100 runs. *Source*: Own simulations.

Figure A11: Simulated hourly wage profiles by level of education and gender.

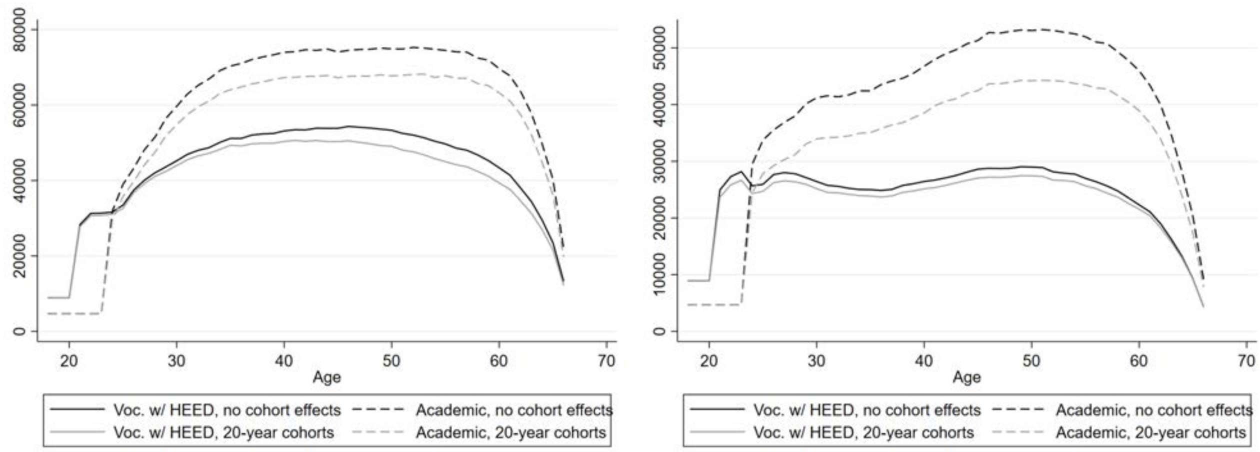
(a) Men (b) Women



Notes: The figure depicts simulated gross hourly wages in Euros in 2019 prices conditional on employment, which are based on different wage regressions, either including or excluding the estimation of birth cohort effects (cohort groups being defined in intervals of twenty years). "Voc. w/ HEED" = Vocational degree with higher education entrance degree. *Source*: Own simulations.

Figure A12: Simulated age-earnings profiles by level of education and gender.

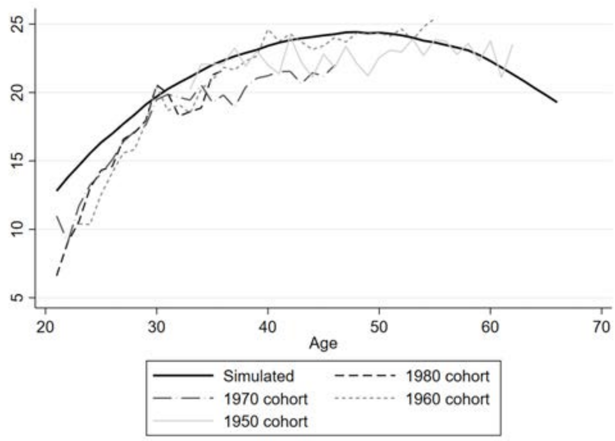
(a) Men (b) Women



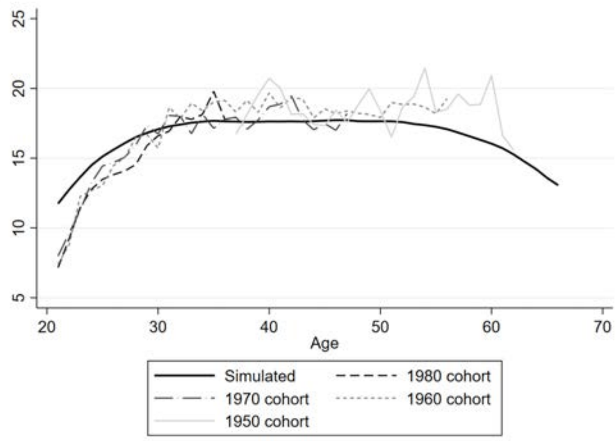
Notes: The figure depicts simulated gross labor earnings in Euros in 2019 prices, with hourly wages based on different wage regressions, either including or excluding the estimation of birth cohort effects (cohort groups being defined in intervals of twenty years). "Voc. w/ HEED" = Vocational degree with higher education entrance degree. *Source*: Own simulations.

Figure A13: Simulated and observed hourly wage profiles by level of education and gender.

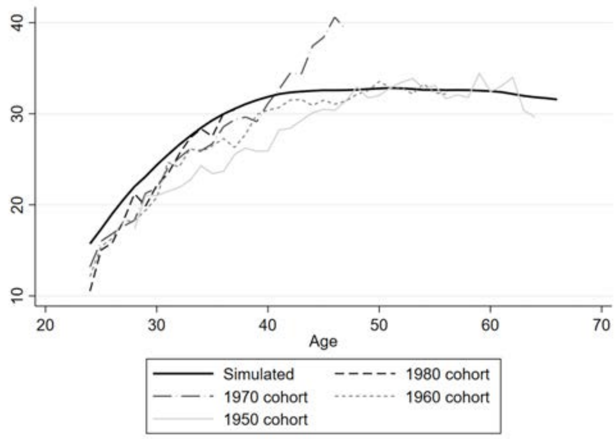
(a) Men, vocational degree with HEED



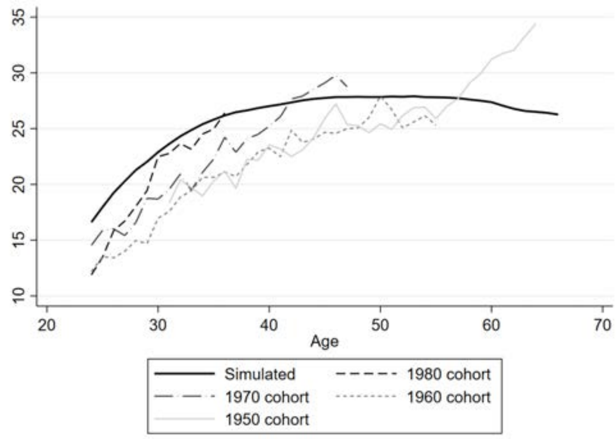
(b) Women, vocational degree with HEED



(c) Men, higher education degree



(d) Women, higher education degree



Notes: The figure depicts simulated and observed gross hourly wages in Euros. The shown simulated wage profiles are for the estimation without cohort effects. The observed wages are shown for different birth cohorts. *Source*: Own simulations, SOEP v35.

Table A1: Estimation results for the probability of marrying

	Women	Men
Age	4.521***	4.760***
	(0.542)	(0.609)
$Age^2/10$	-1.628***	-1.637***
	(0.209)	(0.229)
$Age^3/100$	0.251***	0.243***
	(0.0348)	(0.0370)
$Age^4/10{,}000$	-0.142***	-0.132***
	(0.0211)	(0.0218)
In academic training	-0.710***	-0.402**
	(0.176)	(0.158)
In vocational training	-0.585***	-0.0102
	(0.199)	(0.203)
Child under 7	0.385***	0.854***
	(0.0976)	(0.118)
Birth in t-1	1.357***	1.330***
	(0.117)	(0.137)
Birth in t-2	0.660***	0.696***
	(0.150)	(0.163)
Birth in t-3	0.344**	0.434***
	(0.147)	(0.161)
One divorce	0.492***	0.608***
	(0.0720)	(0.0838)
Two or more divorces	1.351***	1.249***
	(0.130)	(0.163)
Migration background	0.0903	0.336***
	(0.0729)	(0.0824)
Unemployed in t-1		-0.484***
		(0.142)
Years of unemployment	-0.0707***	-0.121***
	(0.0122)	(0.0258)
(Years of unemployment^2)/10		0.0464***
		(0.0147)
In labor force in t-1	0.126	
	(0.0965)	
In labor force in t-2	0.134	
	(0.0942)	
Constant	-48.95***	-54.53***
	(5.105)	(5.940)
Cohort dummies	yes	yes
Year dummies	yes	yes
Orthogonalized east dummy	yes	yes
N	38396	31841

Notes: Separate Logit estimations for women (column 1) and men (2). Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A2: Estimation results for the probability of divorcing, women

Age difference	0.0180
	(0.0146)
Last marriage before 30	0.820***
	(0.253)
Years married	0.0250**
	(0.0127)
One divorce	3.760***
	(0.187)
Child under 7	-0.347
	(0.298)
Husband not employed	1.033***
	(0.158)
Migration background	0.551***
	(0.162)
Constant	-9.996***
	(0.511)
Cohort dummies	yes
Year dummies	yes
Orthogonalized east dummy	yes
N	93842

Notes: Logit estimations using divorce of a married couple as the dependent variable. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A3: Estimation results for the probability of giving birth

	Married women	Single women
Age	0.580***	0.740***
	(0.0485)	(0.0503)
Age^2/10	-0.105***	-0.125***
	(0.00729)	(0.00797)
One child	-0.190***	0.441***
	(0.0652)	(0.0793)
Two children	-1.279***	0.333***
	(0.0794)	(0.109)
More than two children	-0.857***	0.527***
	(0.0932)	(0.139)
Years married	-0.0907***	
	(0.00663)	
In academic training	-0.946***	-1.219***
	(0.197)	(0.153)
In vocational training	-1.590***	-1.047***
	(0.326)	(0.162)
Migration background	0.125**	0.0242
	(0.0499)	(0.0825)
(Tenure in t-1)/10	0.426**	0.383*
	(0.175)	(0.206)
(Tenure in t-1)^2/100	-0.259**	-0.336**
	(0.103)	(0.145)
Empl. status in t-1 = 1	-0.144*	
	(0.0858)	
Empl. status in t-1 = 2	-0.331***	
	(0.0902)	
Empl. status in t-1 = 3	-0.198**	
	(0.0843)	
Empl. status in t-1 = 4	-0.297***	
	(0.0963)	
Empl. status in t-1 = 5	-0.421***	
	(0.0865)	
Empl. status in t-1 = 6	-0.621***	
	(0.133)	
Constant	-8.914***	-14.19***
	(0.823)	(0.878)
Cohort dummies	yes	yes
Year dummies	yes	yes
Orthogonalized east dummy	yes	yes
N	84670	53384

Notes: Separate Logit estimations for married (column 1) and single women (2). Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A4: Estimation results for the probability of labor force participation

	Women	Men
LFP=1 in t-1	2.913***	3.453***
	(0.0377)	(0.0606)
LFP=1 in t-2	1.188***	1.375***
	(0.0429)	(0.0819)
LFP=1 in t-3	0.545***	0.811***
	(0.0470)	(0.0961)
LFP=1 in t-4	0.419***	0.407***
	(0.0489)	(0.107)
LFP=1 in t-5	0.543***	0.492***
	(0.0448)	(0.107)
Age	0.0562*	0.201***
	(0.0308)	(0.0626)
$Age^2/10$	-0.0146***	-0.0318***
	(0.00326)	(0.00638)
Academic	-4.152***	1.319
	(1.205)	(2.770)
Vocational w/ HEED	-1.175	-4.139
	(1.205)	(2.560)
Vocational w/o HEED	-2.335***	-2.546*
	(0.742)	(1.514)
Academic x Age	0.246***	0.0993
	(0.0522)	(0.110)
Vocational w/ HEED x Age	0.0855	0.217**
	(0.0548)	(0.110)
Vocational w/o HEED x Age	0.137***	0.157**
	(0.0332)	(0.0646)
Academic x $Age^2/100$	-0.288***	-0.182*
	(0.0544)	(0.107)
Vocational w/ HEED x $Age^2/100$	-0.105*	-0.232**
	(0.0597)	(0.114)
Vocational w/o HEED x $Age^2/100$	-0.167***	-0.186***
	(0.0355)	(0.0668)
Birth in t	-2.722***	
	(0.0887)	
Birth in t-1	-4.667***	
	(0.0825)	
Birth in t-2	-0.845***	
	(0.0708)	
One child		0.202***
		(0.0607)
Two children		0.289***
		(0.0603)
More than two children		0.184**
		(0.0764)
Married	-0.202***	
	(0.0368)	
Migration background	-0.156***	-0.146**
	(0.0414)	(0.0700)
Constant	-2.410***	-5.585***
	(0.748)	(1.548)
Cohort dummies	yes	yes
Year dummies	yes	yes
Orthogonalized east dummy	yes	yes
N	75941	64296

Notes: Separate Logit estimations for women (column 1) and men (2). "HEED"= Higher education entrance degree. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A5: Estimation results for the probability of unemployment

	Women	Men
Years of unemployment	0.444***	0.524***
	(0.00843)	(0.0102)
(Years of unemployment^2)/10	-0.168***	-0.208***
	(0.00473)	(0.00578)
Years of employment	-0.115***	-0.0653***
	(0.00524)	(0.00764)
(Years of employment^2)/10	0.0234***	0.0108***
	(0.00129)	(0.00151)
Empl. status in t-1 = 1	0.698***	0.630***
	(0.0377)	(0.0525)
Empl. status in t-1 = 2	-2.150***	-1.784***
	(0.0563)	(0.0701)
Empl. status in t-1 = 3	-2.174***	-1.927***
	(0.0519)	(0.0514)
Empl. status in t-1 = 4	-2.128***	-2.227***
	(0.0528)	(0.0622)
Empl. status in t-1 = 5	-1.967***	
	(0.0431)	
Empl. status in t-1 = 6	-2.189***	
	(0.0767)	
Age	0.886***	0.685***
	(0.175)	(0.173)
Age^2/10	-0.324***	-0.272***
	(0.0675)	(0.0672)
Age^3/100	0.0528***	0.0462***
	(0.0111)	(0.0111)
Age^4/10,000	-0.0313***	-0.0276***
	(0.00665)	(0.00660)
Migration background	0.343***	0.433***
	(0.0322)	(0.0349)
Academic	-0.683***	-0.688***
	(0.0495)	(0.0599)
Vocational w/ HEED	-0.403***	-0.260***
	(0.0520)	(0.0651)
Vocational w/o HEED	-0.196***	-0.0427
	(0.0344)	(0.0390)
Married	-0.0745**	-0.213***
	(0.0292)	(0.0342)
Child under 7	0.378***	0.137***
	(0.0381)	(0.0425)
Constant	-8.914***	-6.506***
	(1.638)	(1.618)
Cohort dummies	yes	yes
Year dummies	yes	yes
Orthogonalized east dummy	yes	yes
N	114514	117409

Notes: Separate Logit estimations for women (column 1) and men (2). "HEED"= Higher education entrance degree. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A6: Employment states of women, Multinomial Logit estimates

	Employment state			
	Part-time	Ext. part-time	Full-time	Over-time
Reference category: Reduced part-time				
Empl. status in t-1 = 1	0.499***	0.823***	0.818***	0.947***
	(0.0625)	(0.0700)	(0.0641)	(0.103)
Empl. status in t-1 = 2	-0.849***	-1.368***	-1.784***	-1.077***
	(0.0458)	(0.0640)	(0.0670)	(0.108)
Empl. status in t-1 = 3	2.355***	1.483***	0.497***	1.098***
	(0.0483)	(0.0589)	(0.0625)	(0.0978)
Empl. status in t-1 = 4	1.885***	4.354***	2.666***	2.956***
	(0.0713)	(0.0736)	(0.0743)	(0.102)
Empl. status in t-1 = 5	1.247***	2.758***	4.725***	4.156***
	(0.0815)	(0.0810)	(0.0761)	(0.0984)
Empl. status in t-1 = 6	1.225***	2.434***	3.574***	5.299***
	(0.122)	(0.118)	(0.110)	(0.125)
Tenure	0.0262***	0.0285***	0.0303***	0.0207***
	(0.00214)	(0.00234)	(0.00242)	(0.00285)
Years of full-time experience	0.0370***	0.0552***	0.128***	0.112***
	(0.00536)	(0.00596)	(0.00627)	(0.00786)
(Years of full-time experience^2)/100	-0.0463***	-0.0391**	-0.160***	-0.143***
	(0.0171)	(0.0182)	(0.0183)	(0.0216)
Years of part-time experience	0.0315***	0.0241***	-0.113***	-0.0973***
	(0.00648)	(0.00723)	(0.00742)	(0.00909)
(Years of part-time experience^2)/100	-0.0305	0.0120	0.386***	0.351***
	(0.0205)	(0.0233)	(0.0248)	(0.0317)
Self-employed	-0.380***	-0.301***	-0.647***	1.262***
	(0.0564)	(0.0635)	(0.0654)	(0.0645)
Academic	1.807***	2.027***	2.337***	2.679***
	(0.247)	(0.265)	(0.253)	(0.298)
Vocational w/ HEED	0.856***	0.829***	0.691***	0.904***
	(0.250)	(0.272)	(0.261)	(0.316)
Vocational w/o HEED	0.584***	0.427**	-0.314*	0.182
	(0.164)	(0.182)	(0.170)	(0.212)
Age	0.876***	1.645***	1.380***	1.758***
	(0.236)	(0.254)	(0.233)	(0.290)
Age^2/10	-0.329***	-0.658***	-0.632***	-0.750***
	(0.0880)	(0.0959)	(0.0892)	(0.111)
Age^3/100	0.0537***	0.112***	0.115***	0.132***
	(0.0141)	(0.0155)	(0.0146)	(0.0182)
Age^4/10,000	-0.0327***	-0.0696***	-0.0748***	-0.0840***
	(0.00822)	(0.00912)	(0.00868)	(0.0108)
Academic x Age	-0.0262***	-0.0281***	-0.0356***	-0.0421***
	(0.00552)	(0.00597)	(0.00578)	(0.00682)
Vocational w/ HEED x Age	-0.0105*	-0.00941	-0.00920	-0.0160**
	(0.00577)	(0.00633)	(0.00622)	(0.00756)
Vocational w/o HEED x Age	-0.00924**	-0.00645	0.00587	-0.00493
	(0.00368)	(0.00416)	(0.00397)	(0.00496)
Migration background	-0.0456	0.0324	0.0752*	0.0500
	(0.0368)	(0.0416)	(0.0418)	(0.0524)
Married	-0.216***	-0.595***	-0.792***	-0.856***
	(0.0385)	(0.0402)	(0.0403)	(0.0466)
Child under 7	0.0120	-0.250***	-0.987***	-1.061***
	(0.0422)	(0.0474)	(0.0490)	(0.0647)
Constant	-9.130***	-15.39***	-9.029***	-15.62***
	(2.273)	(2.416)	(2.171)	(2.719)
Cohort dummies	Yes			
Year dummies	Yes			
Orthogonalized east dummies	Yes			
N	106656			

Notes: The table reports the coefficient estimates of a Multinomial Logit model with the employment status as the dependent variable. Employment status 2 is the base category. Each column reports the estimates of one category. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A7: Employment states of men, Multinomial Logit estimates

	Employment state Full-time	Over-time
Reference category: part-time		
Empl. status in t-1 = 1	0.537***	0.544***
	(0.0728)	(0.0923)
Empl. status in t-1 = 2	-1.136***	-0.899***
	(0.0656)	(0.0871)
Empl. status in t-1 = 3	2.880***	2.230***
	(0.0643)	(0.0807)
Empl. status in t-1 = 4	2.076***	3.705***
	(0.0739)	(0.0871)
Tenure	0.0176***	0.00261
	(0.00211)	(0.00221)
Years of full-time experience	0.0639***	0.0967***
	(0.00849)	(0.00931)
(Years of full-time experience^2)/100	0.0123	-0.0391**
	(0.0177)	(0.0193)
Years of part-time experience	-0.303***	-0.327***
	(0.0119)	(0.0134)
(Years of part-time experience^2)/100	1.322***	1.484***
	(0.0676)	(0.0757)
Self-employed	-1.496***	0.528***
	(0.0551)	(0.0521)
Academic	0.169***	0.350***
	(0.0606)	(0.0653)
Vocational w/ HEED	0.330***	0.295***
	(0.0768)	(0.0816)
Vocational w/o HEED	0.0712	0.189***
	(0.0536)	(0.0574)
Age	1.232***	2.106***
	(0.208)	(0.231)
Age^2/10	-0.496***	-0.802***
	(0.0806)	(0.0887)
Age^3/100	0.0812***	0.126***
	(0.0133)	(0.0145)
Age^4/10,000	-0.0493***	-0.0731***
	(0.00787)	(0.00860)
Married	0.0162	0.0947**
	(0.0414)	(0.0435)
Constant	-9.131***	-19.57***
	(1.933)	(2.166)
Cohort dummies	Yes	
Year dummies	Yes	
Orthogonalized east dummies	Yes	
N	106656	

Notes: The table reports the coefficient estimates of a Multinomial Logit model with the employment status as the dependent variable. Employment status 2 is the base category. Each column reports the estimates of one category. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1995–2018.

Table A8: OLS wage regressions, no cohort effects

	Women, academic	Men, academic	Women, vocational	Men, vocational
Experience/10	0.266***	0.489***	0.528***	0.230*
	(0.0955)	(0.0988)	(0.100)	(0.123)
$\text{Experience}^2/100$	-0.110	-0.299***	-0.255**	-0.0674
	(0.0930)	(0.0910)	(0.102)	(0.115)
$\text{Experience}^3/1{,}000$	0.0258	0.0709**	0.0635	0.0162
	(0.0347)	(0.0321)	(0.0394)	(0.0418)
$\text{Experience}^4/100{,}000$	-0.0233	-0.0637*	-0.0617	-0.0152
	(0.0429)	(0.0373)	(0.0495)	(0.0501)
Tenure/10	0.423***	0.387***	0.309***	0.413***
	(0.0585)	(0.0530)	(0.0573)	(0.0732)
$\text{Tenure}^2/100$	-0.257***	-0.245***	-0.109	-0.237***
	(0.0684)	(0.0595)	(0.0720)	(0.0873)
$\text{Tenure}^3/1{,}000$	0.0860***	0.0743***	0.0293	0.0681*
	(0.0285)	(0.0239)	(0.0314)	(0.0366)
$\text{Tenure}^4/100{,}000$	-0.110***	-0.0842***	-0.0302	-0.0665
	(0.0384)	(0.0309)	(0.0425)	(0.0490)
Age/10	3.403***	2.982**	2.585**	1.684
	(1.202)	(1.362)	(1.170)	(1.408)
$\text{Age}^2/100$	-1.009**	-0.849*	-0.950**	-0.472
	(0.436)	(0.476)	(0.442)	(0.529)
$\text{Age}^3/1{,}000$	0.129*	0.108	0.148**	0.0582
	(0.0685)	(0.0719)	(0.0724)	(0.0856)
$\text{Age}^4/100{,}000$	-0.0613	-0.0509	-0.0863**	-0.0294
	(0.0395)	(0.0397)	(0.0433)	(0.0508)
Migration background	-0.0929***	-0.133***	-0.145***	-0.248***
	(0.0254)	(0.0215)	(0.0174)	(0.0214)
Constant	-1.648	-1.212	-0.232	0.390
	(1.211)	(1.416)	(1.116)	(1.354)
Orthog. state dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
N	22273	26994	13571	11193

Notes: Separate wage OLS regressions based on gender and education group. Dependent variable is the gross hourly wage. Standard errors clustered on the individual level shown in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1984–2018.

Table A10: OLS wage regressions, 20-year cohorts

	Women, academic	Men, academic	Women, vocational	Men, vocational
Experience/10	0.271***	0.493***	0.517***	0.233*
	(0.0954)	(0.0988)	(0.101)	(0.122)
$\text{Experience}^2/100$	-0.120	-0.304***	-0.248**	-0.0737
	(0.0928)	(0.0911)	(0.103)	(0.114)
$\text{Experience}^3/1{,}000$	0.0300	0.0730**	0.0619	0.0194
	(0.0346)	(0.0321)	(0.0401)	(0.0414)
$\text{Experience}^4/100{,}000$	-0.0283	-0.0662*	-0.0604	-0.0190
	(0.0427)	(0.0373)	(0.0509)	(0.0496)
Tenure/10	0.424***	0.387***	0.306***	0.422***
	(0.0585)	(0.0530)	(0.0574)	(0.0731)
$\text{Tenure}^2/100$	-0.259***	-0.247***	-0.105	-0.245***
	(0.0683)	(0.0596)	(0.0722)	(0.0871)
$\text{Tenure}^3/1{,}000$	0.0878***	0.0750***	0.0275	0.0712*
	(0.0284)	(0.0240)	(0.0315)	(0.0366)
$\text{Tenure}^4/100{,}000$	-0.113***	-0.0853***	-0.0281	-0.0705
	(0.0383)	(0.0309)	(0.0428)	(0.0490)
Age/10	3.443***	3.028**	2.593**	1.639
	(1.201)	(1.360)	(1.176)	(1.391)
$\text{Age}^2/100$	-1.023**	-0.869*	-0.938**	-0.440
	(0.435)	(0.475)	(0.444)	(0.522)
$\text{Age}^3/1{,}000$	0.131*	0.111	0.144**	0.0489
	(0.0684)	(0.0718)	(0.0727)	(0.0845)
$\text{Age}^4/100{,}000$	-0.0622	-0.0527	-0.0820*	-0.0218
	(0.0395)	(0.0397)	(0.0436)	(0.0499)
Migration background	-0.0884***	-0.132***	-0.141***	-0.243***
	(0.0255)	(0.0216)	(0.0174)	(0.0215)
Cohort 1930	-0.0463	0.0227	-0.0709	0.0800
	(0.0490)	(0.0432)	(0.0671)	(0.0777)
Cohort 1950	0.00352	0.0213	0.0139	0.0688**
	(0.0268)	(0.0245)	(0.0276)	(0.0328)
Constant	-1.557	-1.080	-0.169	0.498
	(1.207)	(1.415)	(1.126)	(1.337)
Orthog. state dummies	yes	yes	yes	yes
Orthog. year dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
N	22273	26994	13571	11193

Notes: Separate wage OLS regressions based on gender and education group. Dependent variable is the gross hourly wage. Standard errors clustered on the individual level shown in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level. Cohort dummies are based on 20-year birth year cohorts, with base category the 1970–89 cohort.
Source: Own calculations based on SOEP v35, waves 1984–2018.

Table A12: Fractional regression models for age-specific household formation patterns of women

	Births	Marriages	Divorces
Age	54.711***	4.140***	10.094***
	(2.882)	(0.612)	(0.990)
$Age^2/10$	-43.939***	-1.482***	-5.430***
	(2.528)	(0.372)	(0.599)
$Age^3/10^2$	18.620***	0.176	1.550***
	(1.158)	(0.116)	(0.186)
$Age^4/10^4$	-43.924***	0.108	-2.456***
	(2.924)	(0.198)	(0.315)
$Age^5/10^6$	54.632***	-0.383**	2.033***
	(3.864)	(0.174)	(0.275)
$Age^6/10^8$	-28.031***	0.215***	-0.686***
	(2.090)	(0.062)	(0.097)
Birth year/10	-8.600***	-0.061	-0.143**
	(1.359)	(0.300)	(0.069)
Birth $year^2/10^2$	6.407***	0.125	
	(1.105)	(0.247)	
Birth $year^3/10^3$	-2.663***	0.067	
	(0.446)	(0.166)	
Birth $year^4/10^5$	5.856***	-0.432	
	(0.953)	(0.487)	
Birth $year^5/10^7$	-6.558***	0.795	
	(1.031)	(0.651)	
Birth $year^6/10^9$	2.941***	-0.503	
	(0.445)	(0.329)	
Age x birth year / 10^2	0.412***		
	(0.012)		
Unemployment rate/10^2	-10.168***	-0.053	-0.471
	(1.139)	(0.412)	(1.084)
Unemployment $rate^2/10^2$	0.497***		
	(0.057)		
Year $\geq$ 1991	-0.134***		
	(0.015)		
Constant	-278.566***	-43.611***	-81.888***
	(13.379)	(4.089)	(6.533)
N	1573	862	176

Notes: Fractional regression models with age and birth year specific shares as dependent variables. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.

Source: Own calculations based on Statistisches Bundesamt (2019b,c) for fertility, on Statistisches Bundesamt (2004–2017a) for marriages, and on Statistisches Bundesamt (2004–2017b) for divorces.

Table A13: Fractional regression models for age-specific LFP/UE patterns

	Labor force participation		Unemployment	
	Women	Men	Women	Men
Age	34.267***	4.733***	1.895***	1.938***
	(1.540)	(0.156)	(0.182)	(0.173)
$Age^2/10$	-21.369***	-1.618***	-0.776***	-0.813***
	(1.034)	(0.060)	(0.073)	(0.069)
$Age^3/10^2$	6.866***	0.242***	0.132***	0.140***
	(0.358)	(0.010)	(0.012)	(0.012)
$Age^4/10^4$	-12.030***	-0.137***	-0.080***	-0.084***
	(0.677)	(0.006)	(0.008)	(0.007)
$Age^5/10^6$	10.953***			
	(0.663)			
$Age^6/10^8$	-4.072***			
	(0.263)			
Birth year/10	0.947***	-0.189***		
	(0.027)	(0.053)		
Birth $year^2/10^2$	-0.089***			
	(0.004)			
Age x birth year / 10^2		0.092***		
		(0.009)		
Year $\geq$ 1991	-0.009	-0.186***	0.529***	0.672***
	(0.036)	(0.065)	(0.054)	(0.053)
Unemployment rate/10^2	-5.387***	-5.395***	8.473***	5.105***
	(0.574)	(0.974)	(0.808)	(0.814)
Constant	-221.453***	-46.340***	-19.649***	-19.216***
	(9.218)	(1.506)	(1.624)	(1.564)
N	1,699	1,699	1,698	1,699

Notes: Fractional regression models with age and birth year specific shares as dependent variables. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1984–2018.

Table A14: Fractional regression models for age-specific employment state

	Women				Men	
	Part-time	Ext. part-time	Full-time	Over-time	Full-time	Over-time
	Ref. category: Reduced part-time				*Ref. category: Part-time*	
Age	13.428***	33.586***	29.234***	24.628***	2.433***	2.912***
	(3.371)	(3.459)	(2.713)	(3.476)	(0.187)	(0.221)
$Age^2/10$	-8.157***	-21.809***	-18.514***	-14.806***	-1.006***	-1.097***
	(2.294)	(2.342)	(1.856)	(2.360)	(0.074)	(0.087)
$Age^3/10^2$	2.579***	7.324***	6.004***	4.525***	0.175***	0.178***
	(0.804)	(0.817)	(0.653)	(0.826)	(0.012)	(0.015)
$Age^4/10^4$	-4.489***	-13.471***	-10.631***	-7.478***	-0.110***	-0.106***
	(1.534)	(1.556)	(1.251)	(1.575)	(0.008)	(0.009)
$Age^5/10^6$	4.093***	12.922***	9.826***	6.386***		
	(1.515)	(1.534)	(1.240)	(1.556)		
$Age^6/10^8$	-1.535**	-5.070***	-3.726***	-2.218***		
	(0.607)	(0.614)	(0.498)	(0.623)		
Year $\geq$ 1991	-0.333***	-0.299***	-0.065	-0.415***	-0.281***	-0.413***
	(0.076)	(0.075)	(0.065)	(0.086)	(0.087)	(0.093)
Year $\geq$ 2001	-0.176**	-0.276***	-0.409***	-0.494***	-0.166**	-0.013
	(0.078)	(0.072)	(0.071)	(0.090)	(0.078)	(0.085)
Unemp. rate/10^2	-1.708	-3.390***	-3.076***	-1.460	-3.944***	-0.481
	(1.263)	(1.196)	(1.120)	(1.484)	(1.335)	(1.502)
Constant	-89.121***	-208.255***	-180.310***	-160.962***	-16.493***	-24.913***
	(19.849)	(20.531)	(15.898)	(20.590)	(1.693)	(2.001)
Cohort dummies	Yes				Yes	
N	1624				1624	

Notes: Fractional regression models with age and birth year specific shares as dependent variables. Standard errors in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level.
Source: Own calculations based on SOEP v35, waves 1984–2018.

Table A15: Years per labor market status over the life-cycle, by degree

	LFP		*UE*		*EMP*		*PT*		*FT*	
	Yrs	Lgth	Yrs	Lgth	Yrs	Lgth	Yrs	Lgth	Yrs	Lgth
Women										
No Degree	37.1	12.6	6.2	3.0	30.9	8.4	13.7	4.7	17.2	5.3
Voc w/o HEED	41.9	13.3	4.0	2.2	37.8	9.5	20.7	6.1	17.1	5.6
Voc w/ HEED	40.0	12.5	3.2	1.8	36.8	9.2	19.2	5.8	17.6	5.4
Academic	37.3	12.4	2.1	1.2	35.2	9.4	14.5	4.9	20.7	5.8
Overall	39.3	12.7	3.6	2.0	35.7	9.2	17.4	5.5	18.3	5.6
Men										
No Degree	45.2	30.7	6.5	3.2	38.7	13.9	6.7	3.0	32.0	9.5
Voc w/o HEED	45.8	31.5	4.9	2.6	40.8	15.3	5.9	2.7	35.0	10.4
Voc w/ HEED	44.2	29.7	3.5	2.1	40.7	15.9	5.4	2.6	35.3	10.7
Academic	41.1	28.6	1.7	1.0	39.4	16.6	6.9	3.3	32.5	10.7
Overall	44.1	30.2	4.1	2.3	40.0	15.5	6.2	2.9	33.8	10.3

Notes: "Yrs" are simulated average years spent in a specific labor market state. "Lgth" is the average length of a labor market spell. *LFP* stands for labor force participation, *UE* for unemployment, *EMP* for employment, *PT* for part-time employment, and *FT* for full-time employment. "HEED" means higher education entrance degree. *Source*: Own calculations.

Table A16: Number and length of marriages over the life-cycle, by degree

	Total	Share with ... marriages			Spell length		Age at 1st marr.
		0	1	≥2	overall	uncens.	
Women							
No Degree	0.85	0.26	0.63	0.11	27.89	8.80	29.36
Voc w/o HEED	0.84	0.30	0.58	0.12	25.47	8.70	29.86
Voc w/ HEED	0.83	0.30	0.59	0.11	24.48	8.40	32.59
Academic	0.78	0.33	0.57	0.10	22.06	8.03	36.55
Overall	0.82	0.30	0.59	0.11	24.65	8.47	32.32
Men							
No Degree	0.95	0.31	0.50	0.19	25.92	8.69	30.07
Voc w/o HEED	0.62	0.53	0.35	0.12	23.67	8.40	33.46
Voc w/ HEED	0.80	0.41	0.43	0.16	24.96	8.58	31.05
Academic	0.86	0.36	0.47	0.17	24.47	8.31	32.17
Overall	0.78	0.42	0.43	0.15	24.65	8.47	31.87

Notes: The table displays the average number of marriages, the share of individuals with a certain number of marriages, the average spell length of a marriage, and age at first marriage for the simulated cohort. Uncensored spells comprise spells that are observed to be divorced before the age of 67 only (no right-censored spells). *Source*: Own calculations.

Table A17: Women's births over the life-cycle, by degree

	Birth rate	Share with ... births				Age at 1st birth
		0	1	2	≥3	
No Degree	1.88	0.17	0.17	0.41	0.25	27.36
Voc w/o HEED	1.71	0.19	0.18	0.43	0.20	27.95
Voc w/ HEED	1.70	0.21	0.17	0.42	0.20	29.62
Academic	1.56	0.26	0.19	0.38	0.17	31.98
Overall	1.69	0.21	0.18	0.41	0.20	29.37

Notes: The table displays the average number of births, the share of women with a certain number of births, and average age at first birth for the simulated cohort. *Source*: Own calculations.

Table A18: Number of divorces over the life-cycle, by degree

	Divorce rate	Share with ... divorces			Age at 1st divorce
		0	1	≥2	
Women					
No Degree	0.24	0.79	0.19	0.02	36.24
Voc w/o HEED	0.28	0.75	0.22	0.03	36.84
Voc w/ HEED	0.25	0.78	0.20	0.02	38.77
Academic	0.21	0.81	0.17	0.02	41.31
Overall	0.25	0.78	0.20	0.02	38.36
Men					
No Degree	0.28	0.79	0.15	0.05	35.09
Voc w/o HEED	0.19	0.85	0.11	0.03	38.36
Voc w/ HEED	0.25	0.81	0.14	0.05	37.26
Academic	0.26	0.80	0.15	0.05	38.59
Overall	0.24	0.82	0.14	0.04	37.53

Notes: The table displays the average number of divorces, the share of individuals with a certain number of divorces, and average age at first divorce for the simulated cohort. *Source*: Own calculations.

Diskussionsbeiträge - Fachbereich Wirtschaftswissenschaft - Freie Universität Berlin
Discussion Paper - School of Business & Economics - Freie Universität Berlin

2020 erschienen:

2020/1 ASSEYER, Andreas: Wholesale price discrimination with regulatory asymmetry
Economics

2020/2 JESSEN, Robin und Johannes KÖNIG: Hours Risk and Wage Risk: Repercussions over the Life-Cycle
Economics

2020/3 ASSEYER, Andreas: Collusion and Delegation under Information Control
Economics

2020/4 ADAM, Marc Christopher: Liquidating Bankers' Acceptances: International Crisis, Doctrinal Conflict and American Exceptionalism in the Federal Reserve 1913-1932
Economics

2020/5 BÖNKE, Timm; Rick GLAUBITZ; Konstantin GÖBLER; Astrid HARNACK; Astrid PAPE und Miriam WETTER: Die Entwicklung und Prognose von Lebenserwerbseinkommen in Deutschland
Economics

2020/6 HUNDSDOERFER, Jochen und Eva MATTHAEI: Gender Discriminatory Taxes, Fairness Perception, and Labor Supply
FACTS

2020/7 ZHU, Junbing und Theocharis GRIGORIADIS: Chinese Dialects, Revolutionary War & Economic Performance
Economics

2020/8 POLUGODINA, Maria und Theocharis GRIGORIADIS: East Prussia 2.0: Persistent Regions, Rising Nations
Economics

2020/9 DOMBI, Akos; Theocharis GRIGORIADIS und Junbing ZHU: Antiquity and capitalism: the finance-growth perspective
Economics

2020/10 AHMED LAHSEN, Amina; Alan T. PIPER und Ida-Anna THIELE: Kim Jiyoung, Born 1982, and the Labour Market: Overeducation, Gender, Income and Life Satisfaction. Panel evidence from Korea
Economics

2020/11 COLEMAN, Winnie und Dieter NAUTZ: The Credibility of the ECB's Inflation Target in times of Corona: New Evidence from an Online Survey
Economics

2020/12 SYDOW, Jörg; Georg SCHREYÖGG und Jochen KOCH: Current interest in the theory of organizational path dependence : a short update on the occasion of the 2019 AMR Decade Award
Management

2020/13 DIECKELMANN, Daniel: Cross-Border Lending and the International Transmission of Banking Crises
Economics

2020/14 WOLF, Tobias: Welfare while working: How does the life satisfaction approach help to explain job search behavior?
Economics

2020/15 PIPER, Alan: Temps dip deeper: temporary employment and the midlife nadir in human well-being
Economics

2020/16 OLSEN, Nils: A literature overview on scheduling electric vehicles in public transport and location planning of the charging infrastructure
Information Systems

2020/17 WINKELMANN, Lars und Wenying YAO: Cojump Anchoring
Economics

2020/18 MARTÍNEZ-HERNÁNDEZ, Catalina: Disentangling the effects of multidimensional monetary policy on inflation and inflation expectations in the euro area
Economics

2020/19 DIEGEL, Max und Dieter NAUTZ: The role of long-term inflation expectations for the transmission of monetary policy shocks
Economics

2020/20 REISCH, Julian: State of the Art Overview on Automatic Railway Timetable Generation and Optimization
Information Systems