School-to-Work Transitions and Labor Market Outcomes

Inaugural-Dissertation zur Erlangung des
akademischen Grades eines Doktors der
Wirtschaftswissenschaft des Fachbereichs
Wirtschaftswissenschaft der Freien Universität Berlin

vorgelegt von

Dominik Hügle, M.Sc.

geboren in Emmendingen

Berlin, 2021

Dekan:

Prof. Dr. Dieter Nautz, Freie Universität Berlin

Erstgutachter:

Prof. Dr. Viktor Steiner, Freie Universität Berlin

Zweitgutachter:

Prof. Dr. Ronny Freier, Technische Hochschule Wildau

Tag der Disputation:

29. Juli 2021

Selbstständigkeitserklärung

Hiermit versiche ich, dass ich die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Literatur und folgenden Hilfsmitteln angefertigt habe: Stata, Microsoft Office, LaTeX und JabRef. Die Arbeit wurde in keinem anderen früheren Promotionsverfahren vorgelegt, angenommen oder beurteilt.

Zusammenarbeit mit Koautoren und Vorveröffentlichungen[1]

Kapitel 1: The Private and Fiscal Returns to Higher Education - A Simulation Approach for a Young German Cohort

- Koautor: Benjamin Fischer
- Eigenleistung: 50%
- Veröffentlicht als Diskussionspapier der Freien Universität Berlin (*School of Business and Economics Discussion Paper 2020/21*)
- DOI: `http://dx.doi.org/10.17169/refubium-28847`

Kapitel 2: The Decision to Enrol in Higher Education

- Keine Koautoren
- Eigenleistung: 100%
- Veröffentlicht als Diskussionspapier der Freien Universität Berlin (*School of Business and Economics Discussion Paper 2021/8*)
- DOI: `http://dx.doi.org/10.17169/refubium-29947`

Kapitel 3: Higher Education Funding in Germany - A Distributional Lifetime Perspective

- Keine Koautoren
- Eigenleistung: 100%
- Veröffentlicht als Diskussionspapier der Freien Universität Berlin (*School of Business and Economics Discussion Paper 2021/1*)
- DOI: `http://dx.doi.org/10.17169/refubium-28932`

[1] An allen Kapiteln wurden verglichen mit den Diskussionspapieren kleinere Veränderungen vorgenommen.

Danksagung

Zuallererst möchte ich meinem Doktorvater Viktor Steiner für seine Betreuung danken. Ich habe enorm von seiner Expertise und seinem regelmäßigen Feedback profitiert und methodisch sehr viel von ihm gelernt (was sich hoffentlich in dieser Arbeit zeigt). Darüber hinaus hat er sich immer für mich eingesetzt, als ich an seinem Lehrstuhl gearbeitet habe. Für all dies bin ich ihm sehr dankbar. Mein Dank gilt auch meinem Zweitbetreuer Ronny Freier, der mir mit jeder Menge wertvoller Ratschläge, wie ich meine Arbeit verbessern könnte, zur Seite stand und auf dessen Unterstützung ich immer zählen konnte. Neben meinen Betreuern möchte ich mich auch bei Bernd Fitzenberger bedanken, der mich in den ersten beiden BDPEMS-Jahren sehr unterstützt hat und mitverantwortlich für mein Interesse an der Bildungsökonomie und den Übergängen zwischen Schule und Arbeitsmarkt ist. Bedanken möchte ich mich auch generell beim BDPEMS, welches mir den Einstieg in die Promotion ermöglicht und mich mit einem Stipendium gefördert hat.

Die Promotionszeit wäre nicht annäherend das gewesen, was sie war, ohne die Menschen mit denen ich an den verschiedenen Forschungsinstitutionen Zeit verbringen durfte. Astrid und Benny waren die besten Kollegen, die man sich wünschen kann. Mit Benny hatte ich darüber hinaus das Vergnügen, an einem gemeinsamen Projekt zu arbeiten, woraus am Ende das erste Kapitel dieser Dissertation entstanden ist. Mein Dank gilt auch Robin und Sebastian, die für kurze Zeit meine Kollegen waren und mir sehr dabei geholfen haben, an der FU anzukommen. Konstantin war ein großartiger Hiwi, auf den man sich immer verlassen konnte. Danke an alle weiteren Doktorand*innen, mit denen ich einen Teil meiner Promotionszeit verbringen durfte, bei meiner BDPEMS-Kohorte, den Leuten vom DIW und den anderen Doktorand*innen der FU, vor allem bei Anja, Ksushi, Luisa, Merlin und allen weiteren Mitgliedern der Boltzmanngang. Die gemeinsamen “Kaffee"pausen werde ich nie vergessen.

Am Ende möchte ich bei meiner Familie, meinen Freunden und bei Diana für die Unterstützung bedanken, vor allem bei Heidi, die mich während meiner gesamten Bildungskarriere gefördert hat. Mein größter Dank gilt meinen Eltern. Ohne euch wäre nichts davon möglich gewesen. Deshalb widmet euer Bue euch diese Dissertation.

Dominik Hügle
Berlin, April 2021

Contents

Preface

This dissertation studies the economics of post-secondary education in Germany. Considering that young adults in Germany generally take up either higher education or vocational training (or sometimes both in a sequential manner) after they finish secondary education I focus on three fundamental questions: (i) “Does higher education pay off for the individual and the state?", (ii) “How strongly do earnings expectations influence the individual's choice between higher education and vocational training?", and (iii) “What are the distributional effects of higher education funding?"

Importantly, I analyze these questions from a *lifetime* perspective, i.e. considering the whole life cycle of an individual instead of focusing on one particular point of the life cycle (a certain age, for instance). In addition, the perspective taken in this dissertation is *forward looking*, in the sense that it takes the perspective of the individuals of a young cohort and their projected life cycles. While it seems self-evident to consider a forward-looking lifetime perspective to answer the questions of interest, such a perspective has rarely been taken in the literature. Most likely a main reason is that, by nature, observable life cycle data (until retirement, for instance) do not exist for younger cohorts. Hence, in order to take a lifetime perspective of a younger cohort one needs to generate “artificial" data reflecting a plausible life course of currently young adults. Here, a dynamic microsimulation model on the basis of the German Socio-Economic Panel (Goebel et al., 2018) is developed. The dynamic microsimulation model sequentially simulates an individual's life cycle in terms of several key variables such as employment and family formation (Li and O'Donoghue, 2013). This model is the foundation of the empirical work in this dissertation.

The first chapter, *The Private and Fiscal Returns to Higher Education – A Simulation Approach for a Young German Cohort*, explains in detail how the dynamic microsimulation model works. Essentially, it first estimates transitions models for the variables that are to be simulated and then uses the estimated parameters to simulate the individual life cycles from one year to the next. In addition, it contains a tax-transfer calculator that models the German tax-transfer system and allows to compute taxes, transfers, and social security contributions. Using the dynamic microsimulation model the first chapter then estimates the private and fiscal returns to higher education. We distinguish between gross and net income and different degrees of income pooling within households. For a typical biography, we find large positive internal rates of return (IRR) for both the individual and the state. At the same time, however, we also find that a substantial share of individuals would incur negative net present values (NPV).

Chapter two, *The Decision to Enrol in Higher Education*, studies the question how strongly the choice to enter higher education depends on the expectations of future income. Using the dynamic microsimulation model from chapter 1 I forecast an individual's expected life cycle given a specific educational choice. In addition to the dynamic microsimulation model and the SOEP data, I use the starting cohort 4 of the National Educational Panel Study (Blossfeld and Von Maurice, 2011), that follows 9th graders until after secondary school graduation. This allows me to estimate an educational choice model where individuals maximize lifetime utility by choosing between higher education and vocational training. Using the estimated parameters from the decision model I simulate the introduction of tuition fees and graduate taxes. I find that such reforms would only induce few people to change their educational decisions.

The third chapter, *Higher Education Funding in Germany – A Distributional Lifetime Perspective*, analyzes the distributional effects of higher education funding. For this I first compare the quantitative importance of different funding instruments, ranging from free tuition to subsidized health insurance for students. The analysis shows that free tuition is, by far, the most important instrument. However, there is a large heterogeneity by how much a student benefits from free tuition depending on her *field of study*. To connect the amount of benefits an individual receives from higher education funding, particularly free tuition, to the expected lifetime income of an individual, I use the dynamic microsimulation model and simulate the individual biographies. Finally, I use the decision model of chapter 2 and extend it to the case of multiple alternatives (with fields of study and vocational training being the alternatives). Using the estimated parameters I simulate how the choice between the fields would change under different tuition fee schemes. In line with the results of chapter 2, I find that the tuition fees would barely change the individuals' educational choices.

Chapter 1

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

1.1 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, as 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 Psacharopou-

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

los 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 chapter 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 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

[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.

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 this chapter, 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 chapter proceeds as follows. Section 1.2 explains the institutional background of post-secondary education in Germany. Section 1.3 describes how we define and compute returns to higher education. Section 1.4 introduces our dynamic microsimulation model and the data and Section 1.5 presents the validation results of the model. Section 1.6 shows the results, Section 1.7 discusses them and Section 1.8 concludes.

1.2 Post-secondary education in Germany

1.2.1 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, 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

[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 1.3.

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

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 (Autorengruppe Bildungsberichtserstattung, 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 (Autorengruppe Bildungsberichtserstattung, 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 degree. 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 (Autorengruppe Bildungsberichtserstattung, 2018). Figure 1.1 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]

[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.

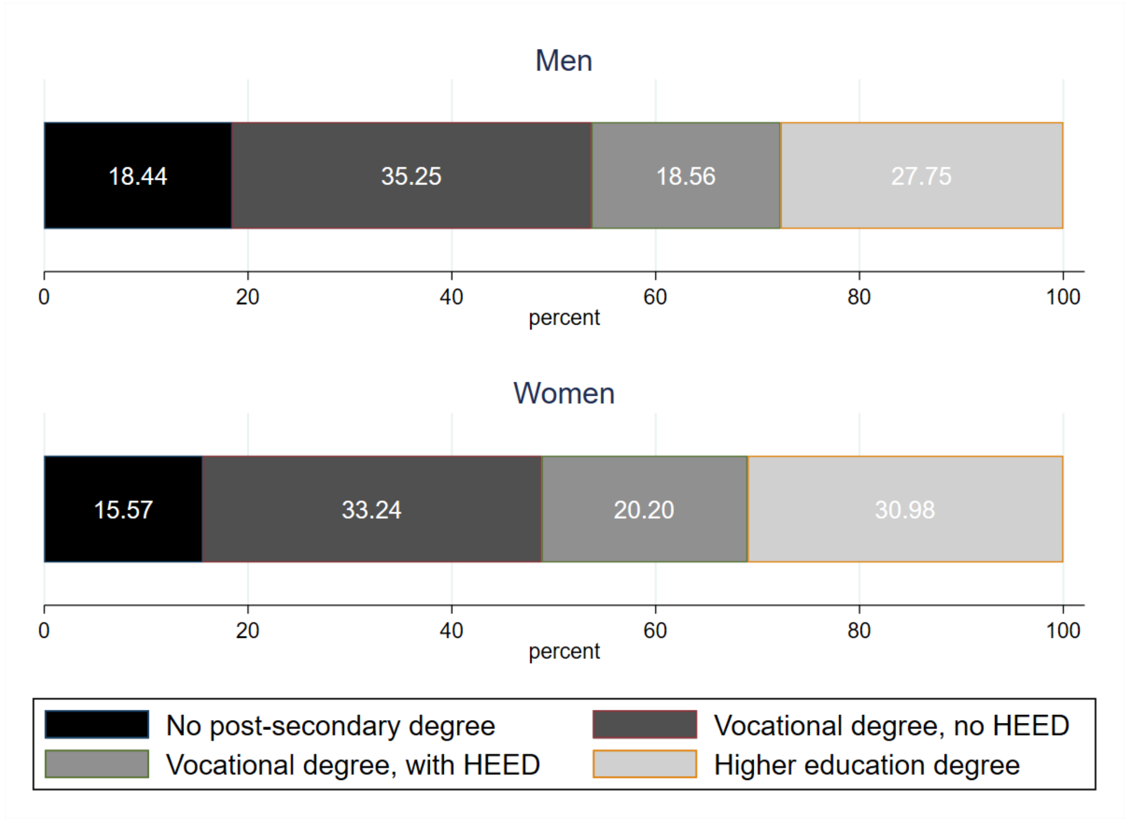
[6]Typical examples of a school-based training are health professions, such as medical or pharmaceutical technical assistants.

[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.

[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.

Figure 1.1: 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, with HEED* = Vocational degree with higher education entrance degree. *Source*: Statistisches Bundesamt (2018), own calculations.

1.2.2 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 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).

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

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.[11]

1.3 The lifetime returns to higher education

1.3.1 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.[12] 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

[11] There are other, quantitatively much less important, instruments of higher education funding which will be ignored here. For an overview, see Hügle (2021).

[12] Originally, the term "internal rate of return" implied comparing an educational investment to a null project without costs and benefits. 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 Heckman and Carneiro (2003) for a more detailed discussion.

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 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:

$$\begin{aligned} Y_{i,t} &= (1-\rho)(labinc_{i,t} + capinc_{i,t} - taxes_{i,t} - contribs_{i,t} + benefits_{i,t}) \\ &\quad + 0.5\rho(labinc_{hh,t} + capinc_{hh,t} - taxes_{hh,t} - contribs_{hh,t} + benefits_{hh,t}) \\ &\quad + 0.5\tau_{i,t}(transfers_{hh,t}), \end{aligned} \tag{1.1}$$

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.[13] $contribs$ comprise employee's social security contributions as well as contributions to private health insurance contracts.[14] $benefits$ 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

[13] 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.

[14] 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.

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 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:

$$\sum_{t=0}^{T} \frac{\overline{Y}_t^{HE} - \overline{Y}_t^{VOC}}{(1+IRR)^t} = 0 \tag{1.2}$$

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.

$$NPV_i^{private} = LTI_i^{HE} - LTI_i^{VOC}, \tag{1.3}$$

with lifetime incomes LTI_i being defined as

$$LTI_i^{HE} = \sum_{t=0}^{T} \frac{Y_{i,t}^{HE}}{(1+r)^t} \tag{1.4}$$

and

$$LTI_i^{VOC} = \sum_{t=0}^{T} \frac{\overline{Y}_{g,t}^{VOC}}{(1+r)^t}, \tag{1.5}$$

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)).

1.3.2 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.[15]

The fiscal surplus is defined as

$$S_{i,t} = taxes_{i,t} + ssc_{i,t} - benefits_{i,t} - 0.5\tau_{i,t}(transfers_{hh,t}) - pubexp_{i,t}, \tag{1.6}$$

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 $BAf\ddot{o}G$ 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

[15] 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 chapter. See Eide and Showalter (2011) for an overview of potential education-level related differences in health and life expectancy.

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:

$$\sum_{t=0}^{T} \frac{\overline{S}_t^{HE} - \overline{S}_t^{VOC}}{(1+IRR)^t} = 0. \tag{1.7}$$

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

$$NPV_i^{fiscal} = LTS_i^{HE} - LTS_i^{VOC}, \tag{1.8}$$

with the actual lifetime surplus being defined as

$$LTS_i^{HE} = \sum_{t=0}^{T} \frac{S_{i,t}^{HE}}{(1+r)^t} \tag{1.9}$$

and the counterfactual lifetime surplus as

$$LTS_i^{VOC} = \sum_{t=0}^{T} \frac{\overline{S}_{g,t}^{VOC}}{(1+r)^t}, \tag{1.10}$$

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.

1.4 Dynamic Microsimulation Model

1.4.1 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.[16] 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.

1.4.2 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).[17] For the simulation, we take as given the distribution of highest educational degrees of the 1983–88 birth cohort by gender and migration background, as presented in Figure 1.1.[18] 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 1.1: a 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 1.2 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

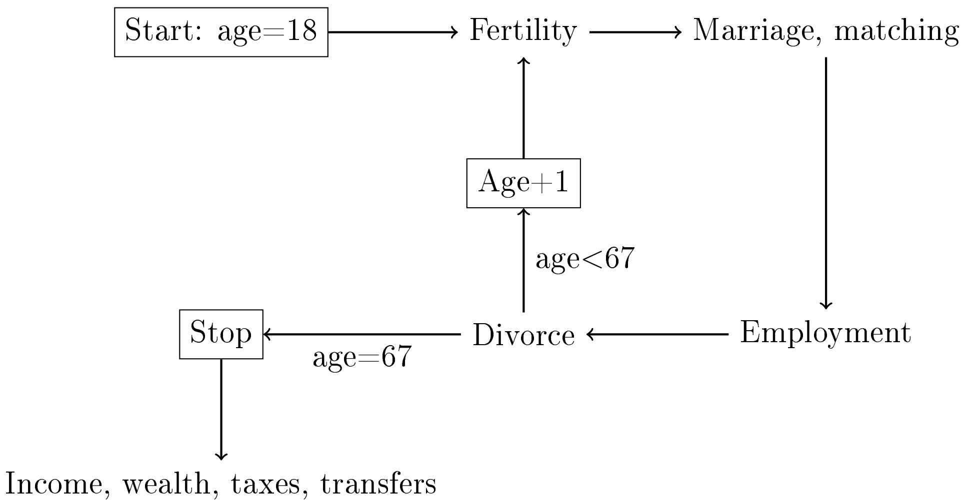
[16]In contrast, dynamic microsimulation has its own drawbacks such as high data requirements, the complexity of modeling and a high computational burden.

[17]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).

[18]In Figure 1.1 the distribution is not conditioned on migration background.

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.[19]

Figure 1.2: The simulation stage



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 the Appendix.

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

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

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

1.4.3 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 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.

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 1.A1 in the Appendix). 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

[20] 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.

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 1.A2 in the Appendix for the specification and estimated coefficients).

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 1.A3 in the Appendix for the estimated coefficients).

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 1.A4–1.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.

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.[21] 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.[22] Here, we follow the approach in Deaton (1997), 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

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

[22] 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.

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.[23] 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 1.1) 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.[24][25] The estimation results are displayed in Tables 1.A8–1.A9 in the Appendix.

1.4.4 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 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

[23] 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.

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

[25] 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.

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.[26]

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.[27] 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 1.A10–1.A12 (Appendix) 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.[28] The projected targets are graphed in the Appendix in Figures 1.A1–1.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.[29]

[26]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.

[27]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.

[28]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.

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

1.4.5 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 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.[30] 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).[31]

Having simulated the transitions in employment allows us to 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

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

[31] 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.

of the log wage residuals, conditional on gender, education, and self-employment status.[32] This procedure aligns the variance of simulated wages to the variance of observed wages. Finally, to obtain hourly wages in levels, the resulting sum is exponentiated.[33] To compute labor earnings, hourly wages are multiplied by the observed average level of working hours given the simulated employment category.

1.5 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.

1.5.1 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 1.A4 and 1.A5 (Appendix) 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

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

[33]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.

the labor market.[34] 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.[35] Regarding employment, our model predicts 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 1.A13 and 1.A14 (Appendix).[36] 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.[37] The fact that the model considerably over-predicts persistence in the marriage status of men (see Figure 1.A5 in the Appendix) 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.

1.5.2 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.[38] 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

[34] 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.

[35] 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 ages forty and fifty whereas observed persistence increases.

[36] 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.

[37] 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 the Appendix, 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.

[38] 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.

education levels, we depict sequence index plots of 400 randomly drawn male and female life cycles.[39] 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.

Employment

We start by validating simulated employment careers. Figure 1.A6 (Appendix) depicts a sequence analysis for employment transitions. Additionally, Table 1.A13(Appendix) 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.[40] For men our number of simulated years of full-time employment

[39]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.

[40]The authors show cumulated years for the 1967–71 cohort, separately for East and West Germany

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.

Family formation

Regarding transitions in family status, sequence analysis (see Figure 1.A7) and descriptive statistics in Tables 1.A14, 1.A15, and 1.A16 (Appendix) 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.[41] 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.[42]

For women, age at first birth, age at first marriage, and age at first divorce in-

---

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.

[41] 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.

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

crease 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.[43]

Hourly wages and yearly earnings

Figures 1.A8 and 1.A9 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 1.A10 in the Appendix shows this comparison for the specifica-

[43] 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.

tion 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.

1.6 Simulation results

As discussed in Section 1.3.1, 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 1.1: 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.

1.6.1 Private returns to higher education

Table 1.1 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.[44]

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 can be explained by two factors: First, by the hourly wage premium for academics which is larger for women. And second, by simulated employment where the gap between academics and individuals with vocational degrees is also larger for women than for men. Going from no to full income pooling, the 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.

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 the returns to 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,

[44]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 1.2: 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 1.3: 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.

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 the discount rate to be 2%, similar to the OECD in recent analyses (OECD, 2019). Tables 1.2 and 1.3 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 on average 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 1.A12 (Appendix), 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 1.4: 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.

Table 1.4 breaks down the NPVs 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 the NPVs 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 1.2 and 1.3. 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.[45] 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 in Tables 1.2 and 1.3.

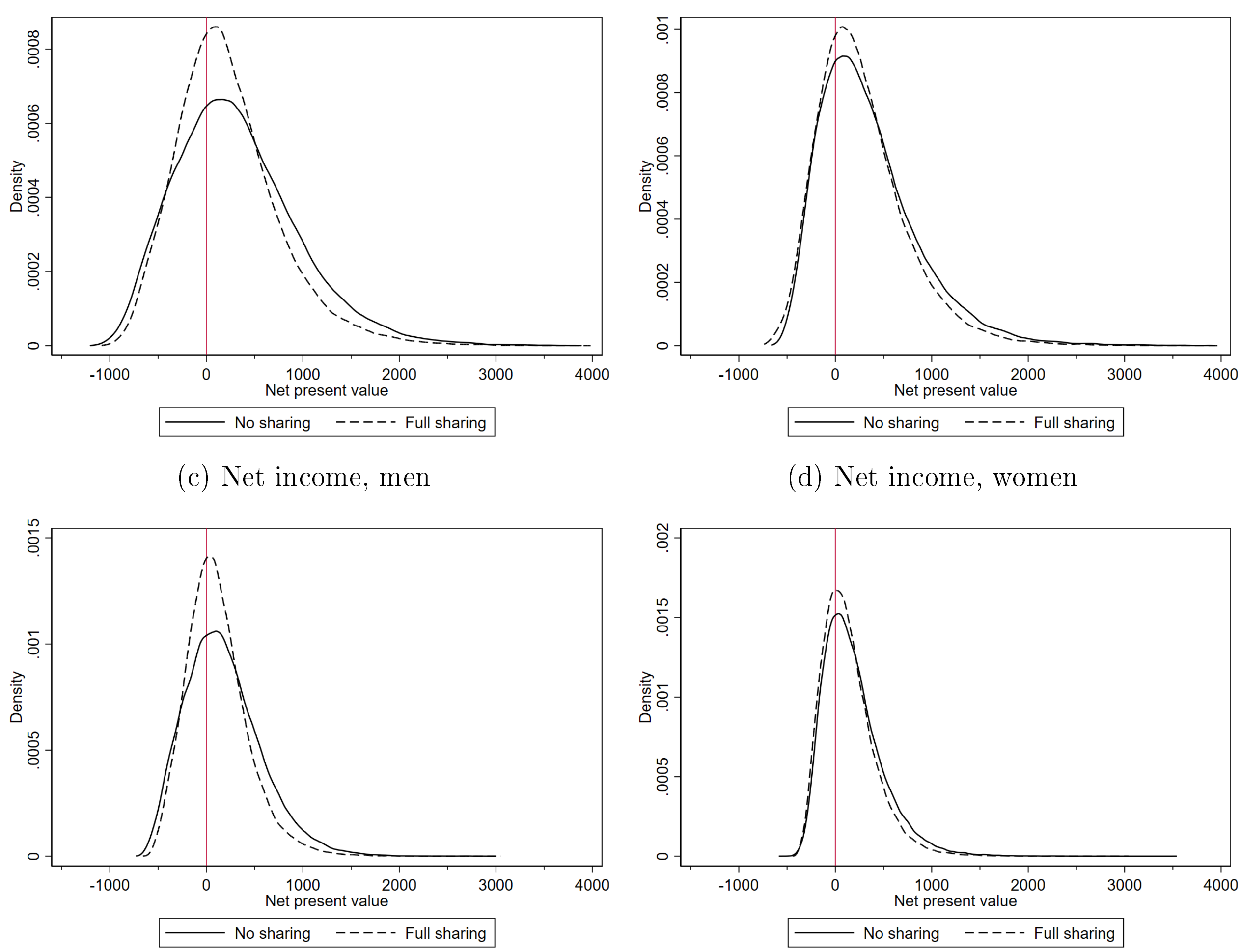
Figure 1.3: Distribution of private net present values (in Thousands of Euros)

(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.

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

As argued above, the main advantage of using the NPV instead of the IRR measure is that heterogeneous returns can be analyzed. Figure 1.3 plots the distribution of NPVs for our four different income concepts.[46] 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 1.2 and 1.3). 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 "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 1.5 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.

[46] Here, we pool all individuals from all runs. Figure 1.A11 in the Appendix shows the distribution for each run separately.

Table 1.5: 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.

1.6.2 Fiscal returns to education

Table 1.6 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.[47] 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 1.6: 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 1.7 we display the components of the fiscal surplus generated by average life cycles, separately by educational degrees and gender. 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

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

Table 1.7: 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.

revenue-generating components.[48] 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 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.[49] 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

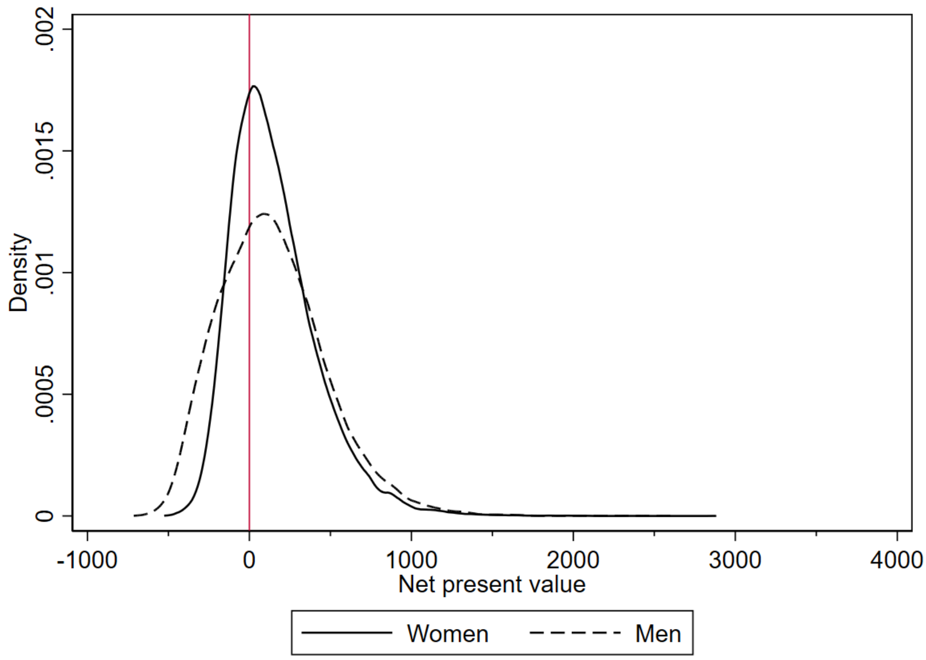
[48] These and all following shares can be obtained by dividing the respective monetary value of a component by the sum of all revenue-generating Δ.

[49] However, the self-employed who opt for the public health insurance scheme typically bear the full burden of contributions. In Table 1.7, these contributions are comprised in the position "SSC employee". Therefore, this position still rises faster with education than "SSC employer" does.

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 1.6). This finding is also observable in Figure 1.4, which plots the distribution of fiscal NPVs for women and men. Again, both density functions are right-skewed, with women's returns somewhat more compressed.

Figure 1.4: 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.

1.7 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 1.8 compares our IRR 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.

Table 1.8: 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/Stichnoth	14.2	x	x	7.4	6.6
OECD	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". *Pfeiffer/Stichnoth*=Pfeiffer and Stichnoth (2020); *OECD*= OECD (2019).
Source: Own calculations, Pfeiffer and Stichnoth (2020), OECD (2019).

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).

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 a higher education entrance degree or with a 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.[50] 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 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.[51] 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 1.5). 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.

[50] 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.

[51] 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.

1.8 Conclusion

In this chapter 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. Furthermore, we find a large heterogeneity in returns, which is rooted in different hourly wages, employment biographies and patterns of household formation.

For a sizable 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 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.

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.

Appendix

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.[52] Due to dynamic aging, spouses have to be of same age. Therefore, we align marriages uniformly to the marriage target of women.[53] 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

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

[53] 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.

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 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.[54] 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 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.[55]

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 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.

[54]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.

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

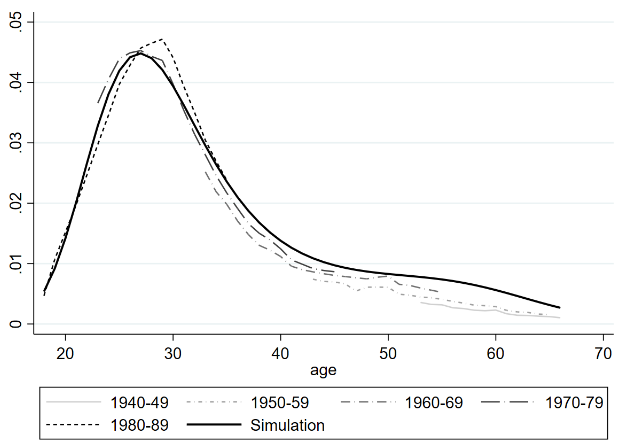
Welfare programs

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.

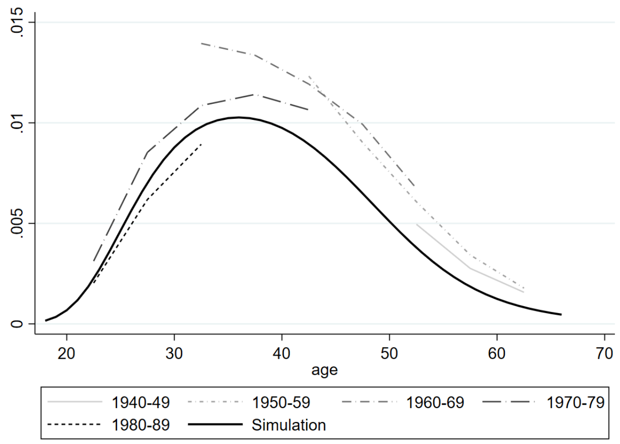
Figures and tables

Figure 1.A1: Observed vs. predicted patterns of family formation, women

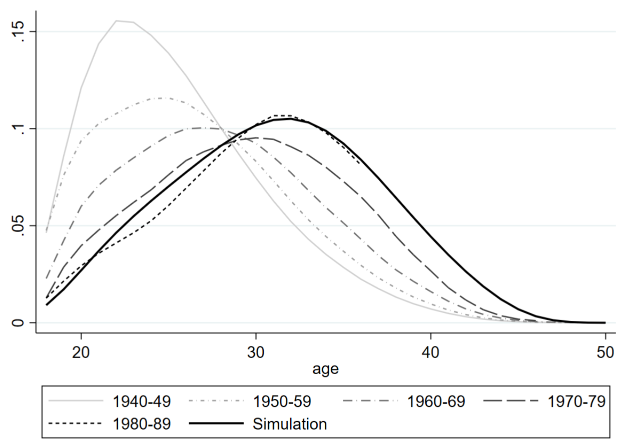
(a) Marriage rate



(b) Divorce rate



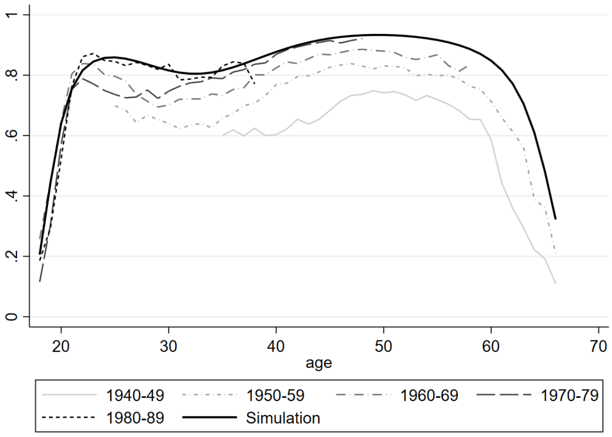
(c) Fertility rate



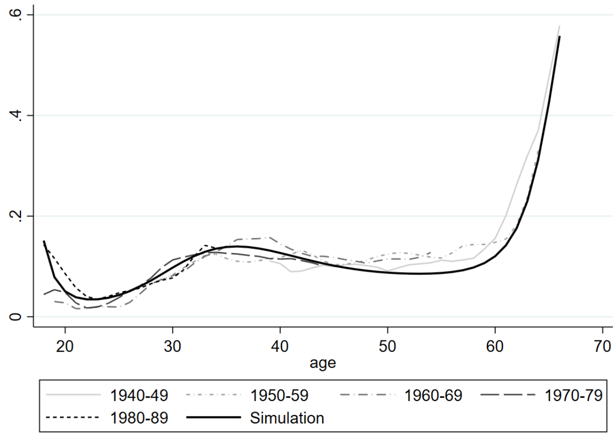
Notes: Observed aggregate marriage, divorce and fertility rates are displayed by birth cohort. For instance, the line "1970-1979" states the marriage, divorce, and fertility rates of the individuals born between 1970 and 1979 over the life cycle as observed in the data. Simulation: predicted life-cycle 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 1.A2: Observed vs. predicted patterns of labor market participation, women

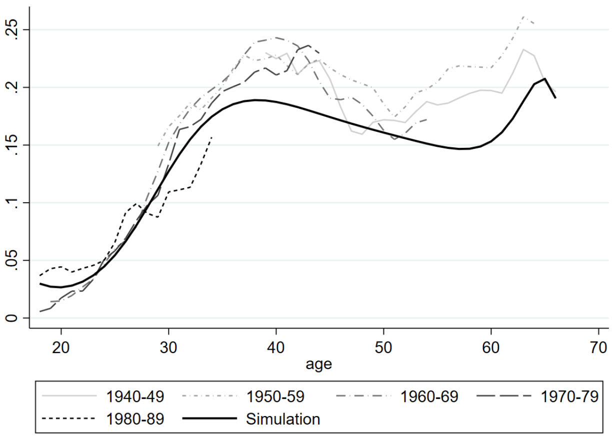
(a) Labor force participation rate



(b) Reduced part-time employment rate



(c) Standard part-time employment rate



(d) Extended part-time employment rate

(e) Full-time employment rate

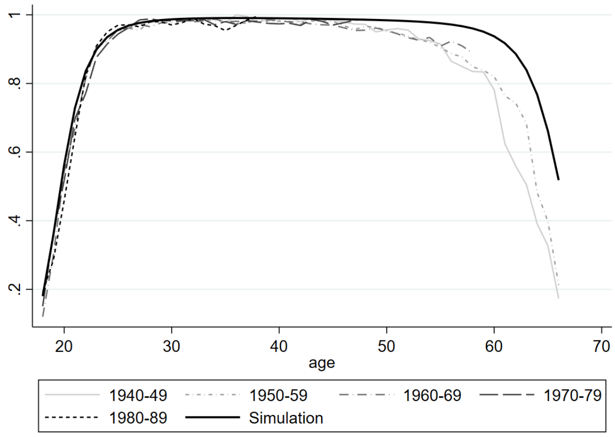
(f) Over-time employment rate

(g) Unemployment

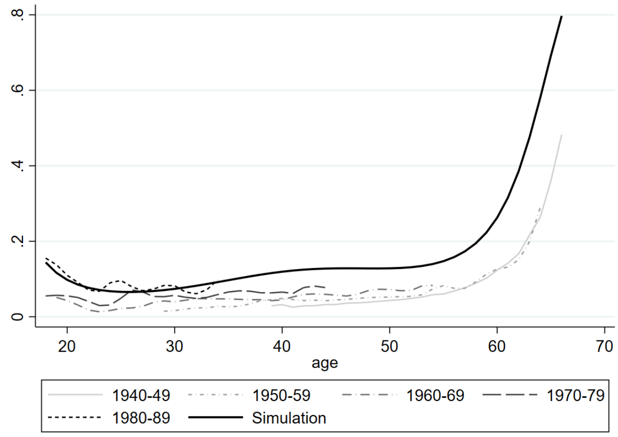
Notes: Observed aggregate rates are shown as three-year moving averages and by birth cohort. Employment shares in subfigures (b)–(g) conditional on labor force participation. Simulation: predicted life-cycle pattern for the 1980–89 cohort. *Source:* Own calculations based on SOEP v35, and own simulations.

Figure 1.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

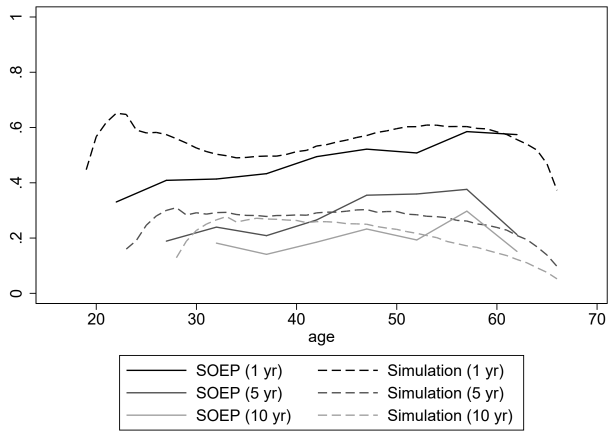
(d) Over-time employment rate

(e) Unemployment rate

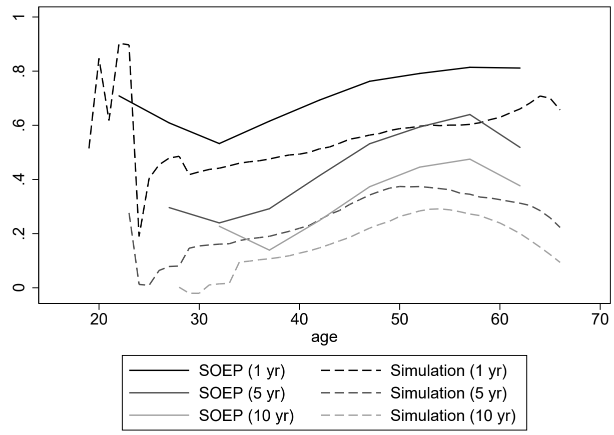
Notes: Observed aggregate rates are shown as three-year moving averages and by birth cohort. Employment shares in subfigures (b)–(e) conditional on labor force participation. Simulation: predicted life-cycle pattern for the 1980–89 cohort. *Source:* Own calculations based on SOEP v35, and own simulations.

Figure 1.A4: Autocorrelations for employment and family status of women, simulated vs. observed.

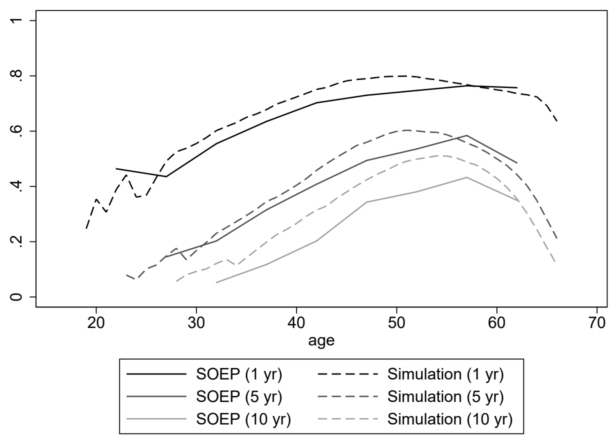
(a) Unemployment



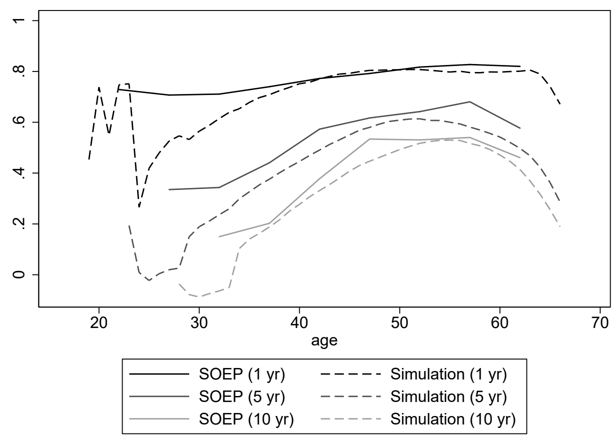
(b) Employment



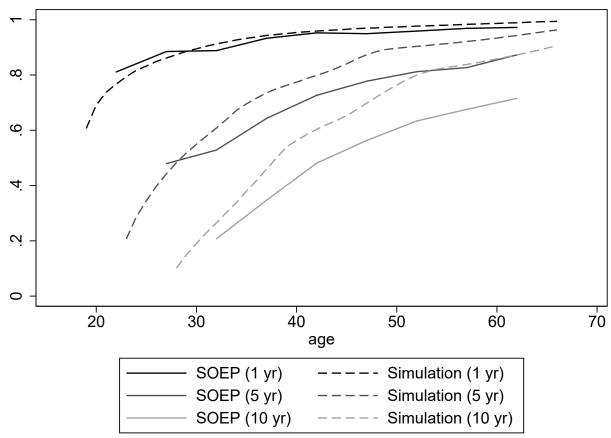
(c) Part-time employment



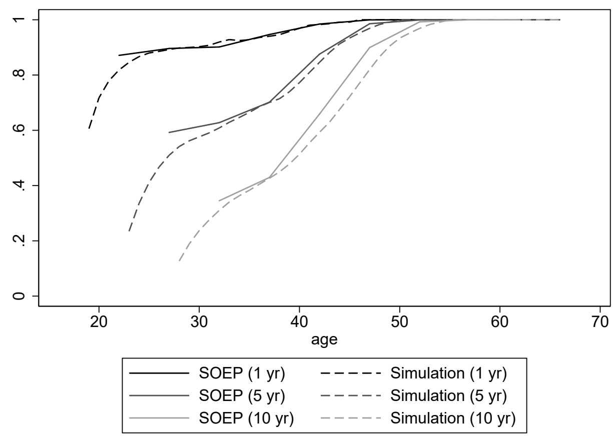
(d) Full-time employment



(e) Marriage status

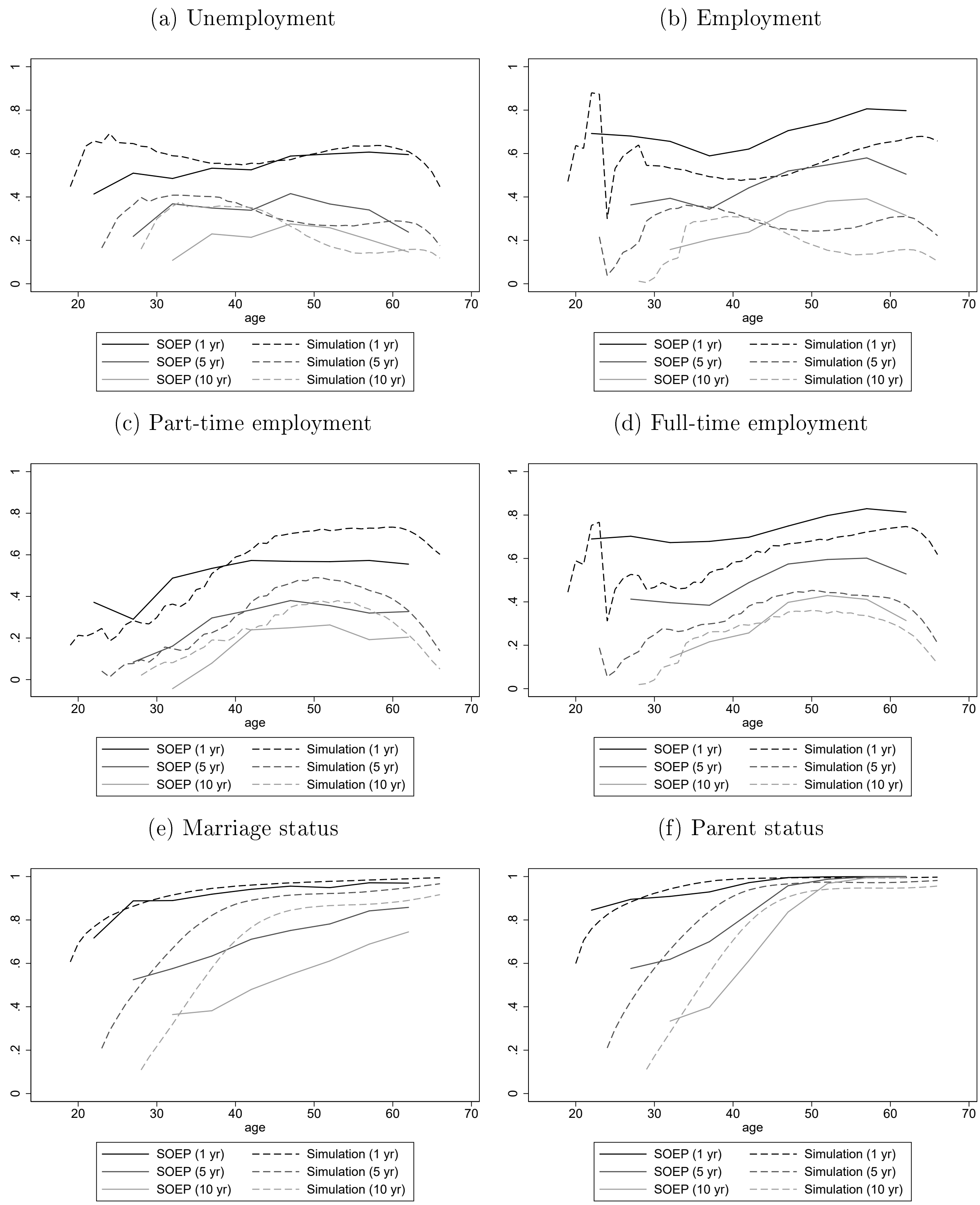


(f) Parent status



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.

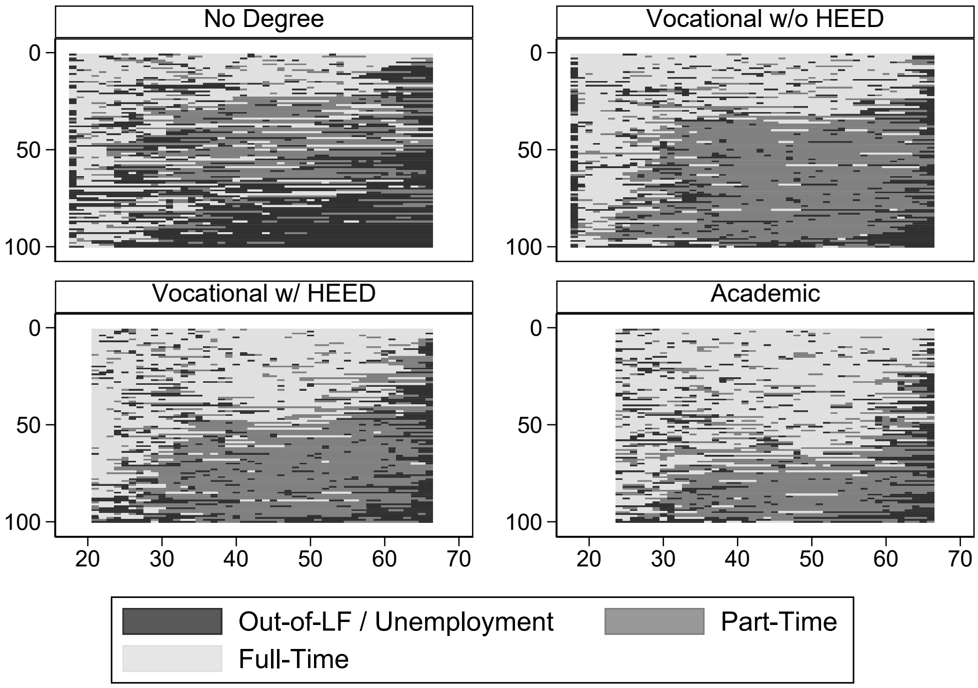
Figure 1.A5: Autocorrelations for employment and family status of men, simulated vs. observed.



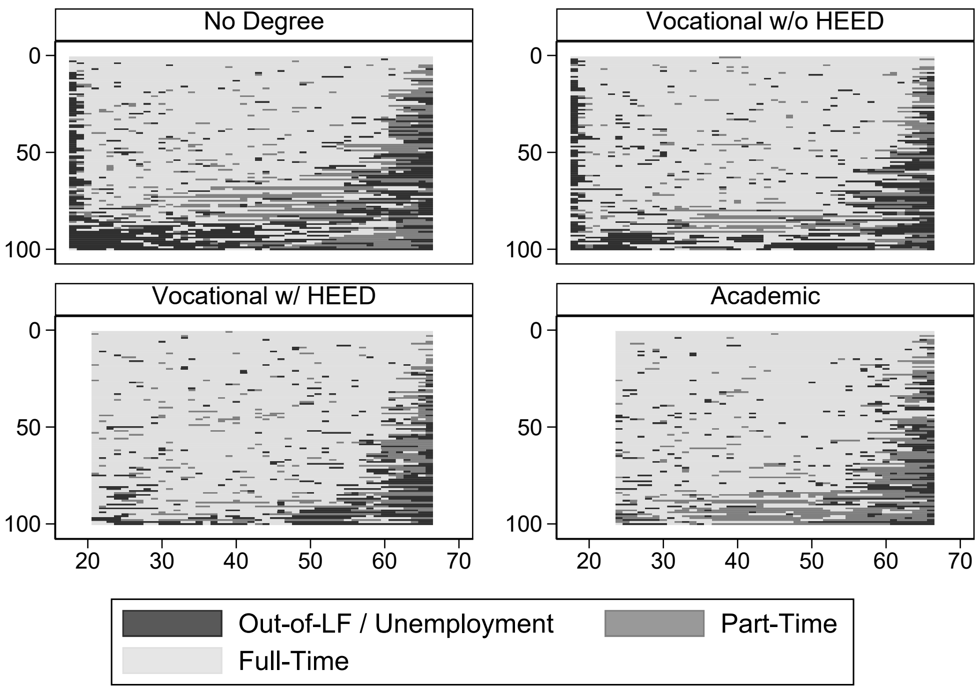
Notes: Autocorrelations with one-, five-, and ten-year lags. Observed values are moving averages of order three for cohorts starting from birth year 1950, weighted by the SOEP weighting factors. Simulated values are averaged over a set of 100 simulations. *Source:* Own calculations based on SOEP v35,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

Figure 1.A6: Sequence index plot for employment status by age

(a) Women



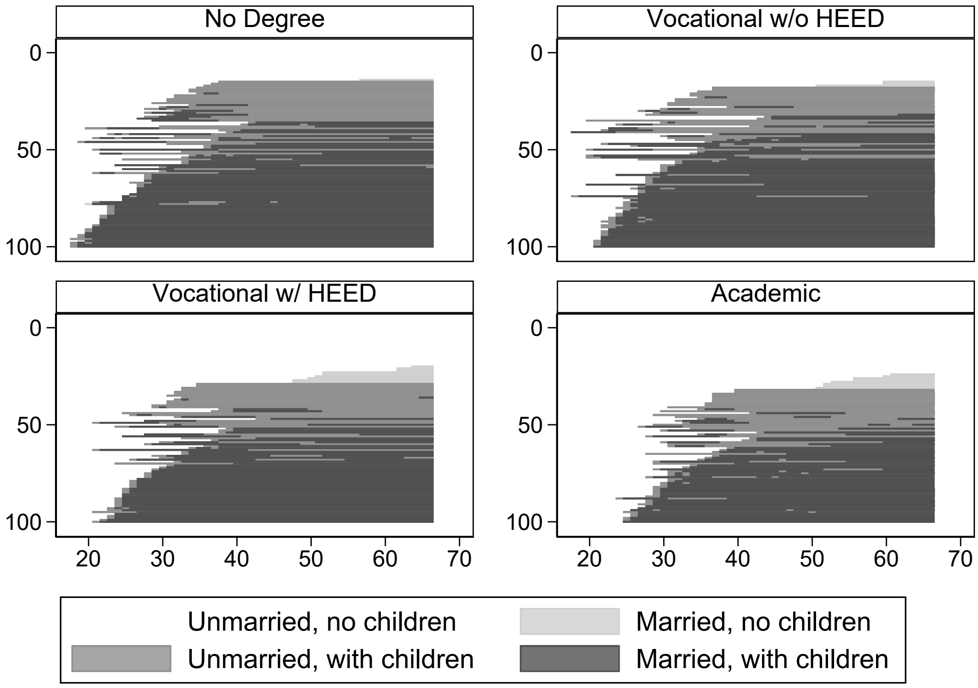
(b) Men



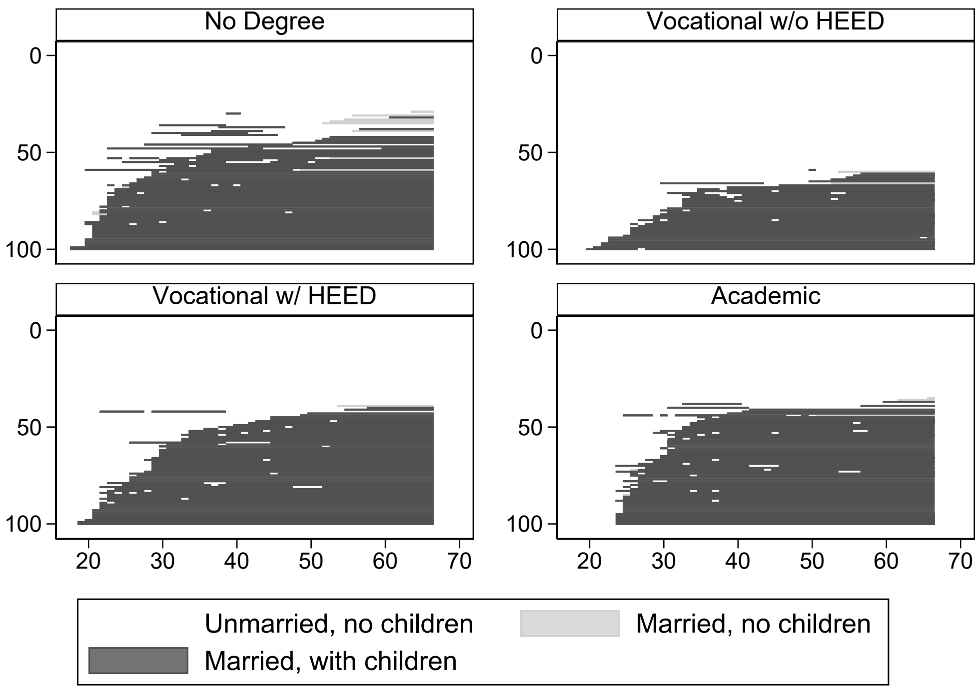
Notes: For each combination of gender and education level, a sequence index plot displays 100 life cycles randomly drawn from the universe of simulated individuals. "Out-of-LF / Unemployment" are episodes out of labor force (not including academic and vocational training) or in involuntary unemployment. "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 1.A7: Sequence index plot for marriage and parental status by age

(a) Women



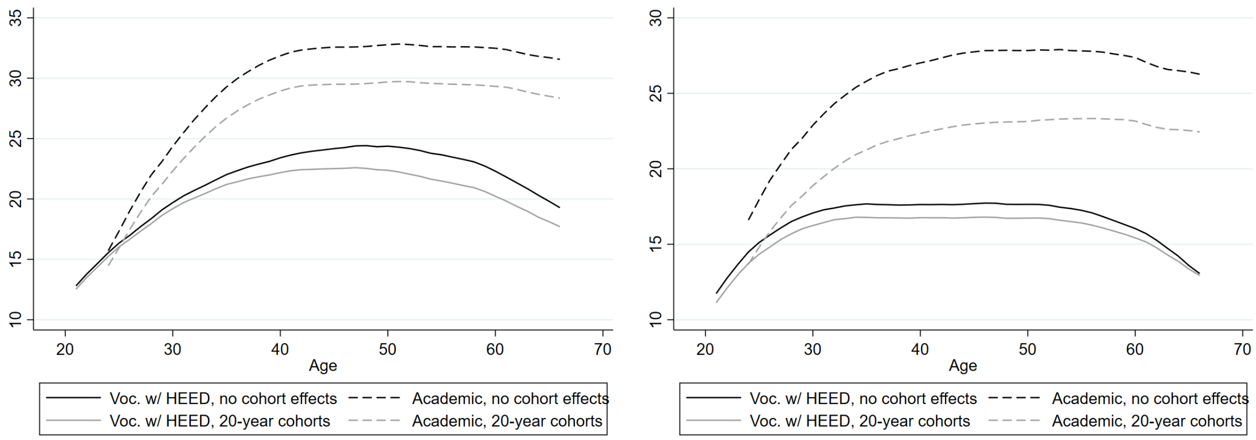
(b) Men



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 1.A8: Simulated hourly wage profiles by level of education and gender.

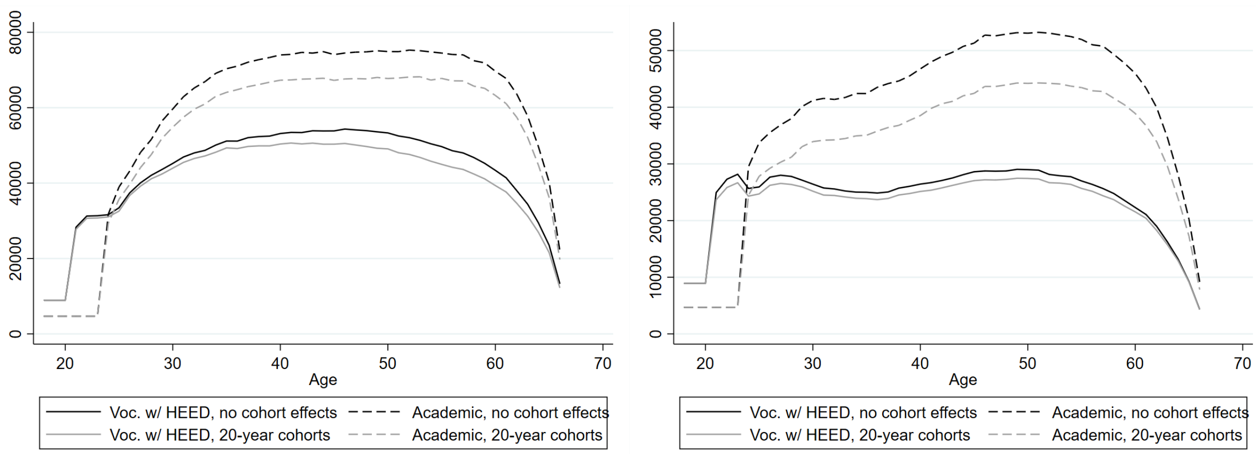
(a) Men (b) Women



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 1.A9: Simulated age-earnings profiles by level of education and gender.

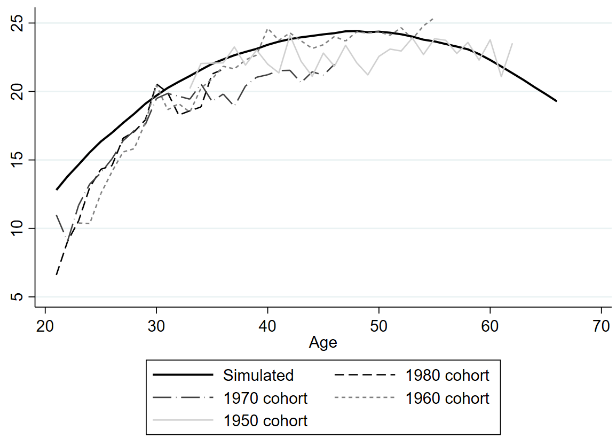
(a) Men (b) Women



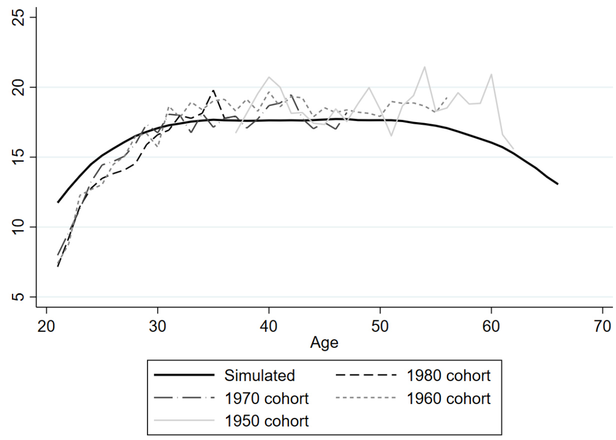
Notes: The figure depicts simulated gross labor earnings in Euros in 2019 prices, with hourly wages based on different wage regressions, either including or excluding the estimation of birth cohort effects (cohort groups being defined in intervals of twenty years). “Voc. w/ HEED" = Vocational degree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

Figure 1.A10: Simulated and observed hourly wage profiles by level of education and gender.

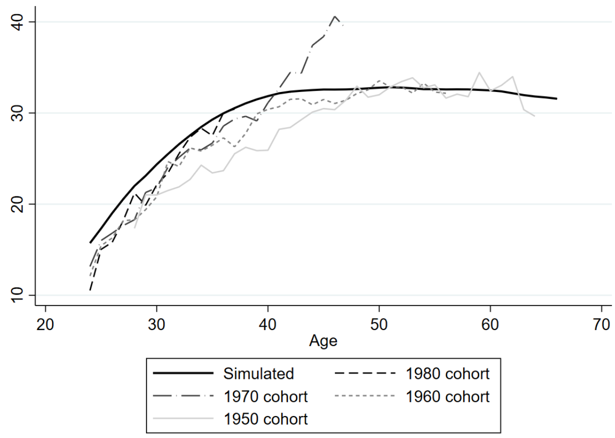
(a) Men, vocational degree with HEED



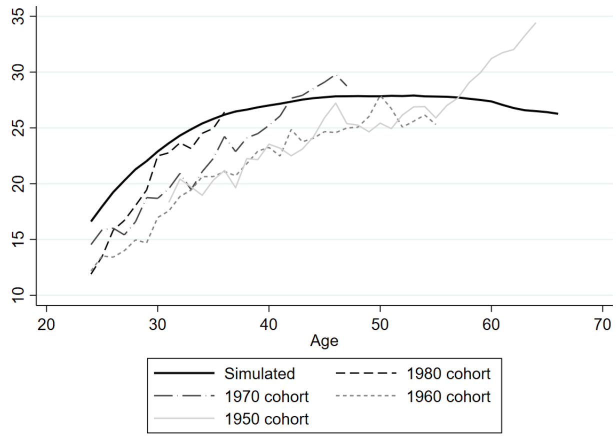
(b) Women, vocational degree with HEED



(c) Men, higher education degree

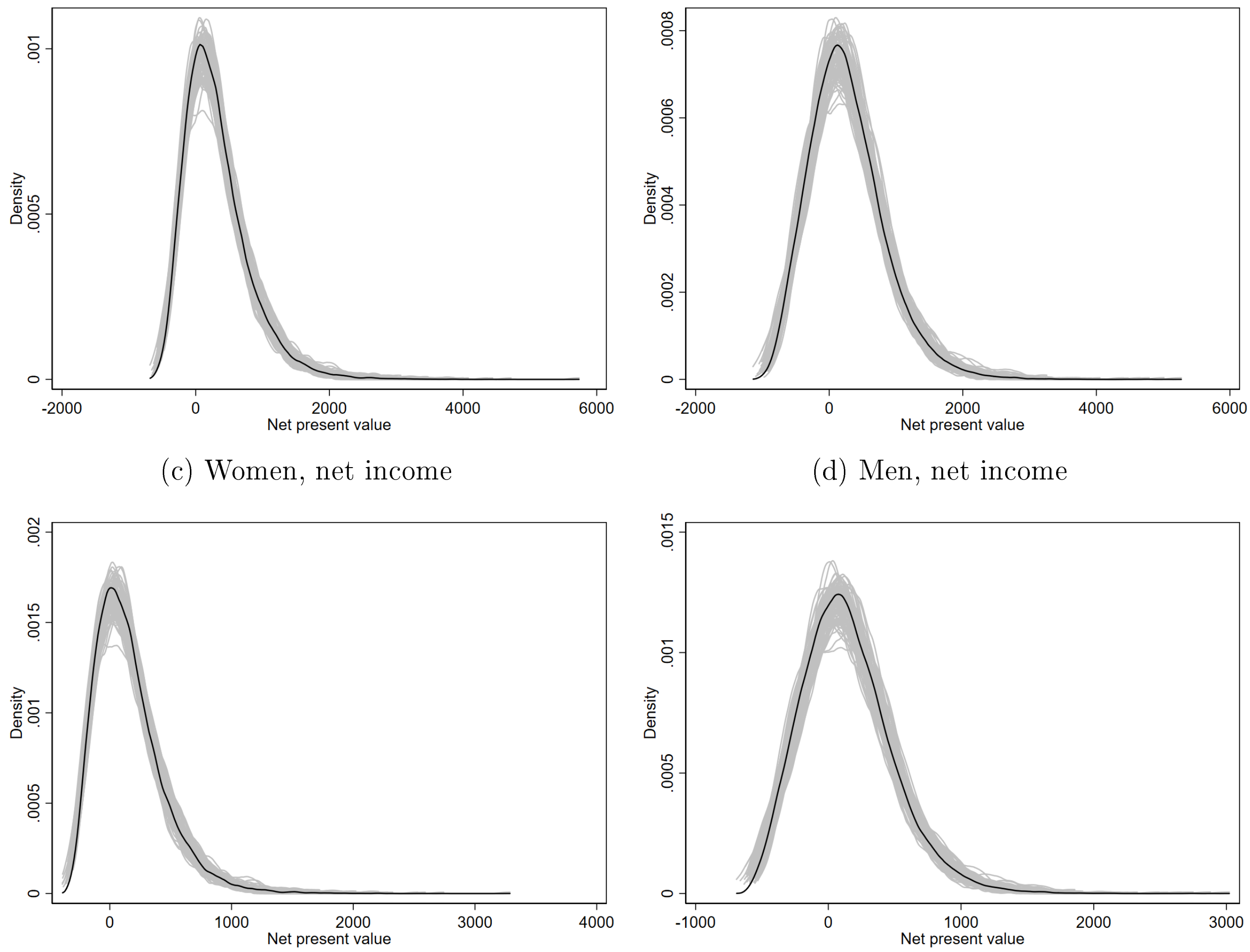


(d) Women, higher education degree



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.

Figure 1.A11: Distribution of NPVs of higher education attainment, all runs separately



(a) Women, gross income

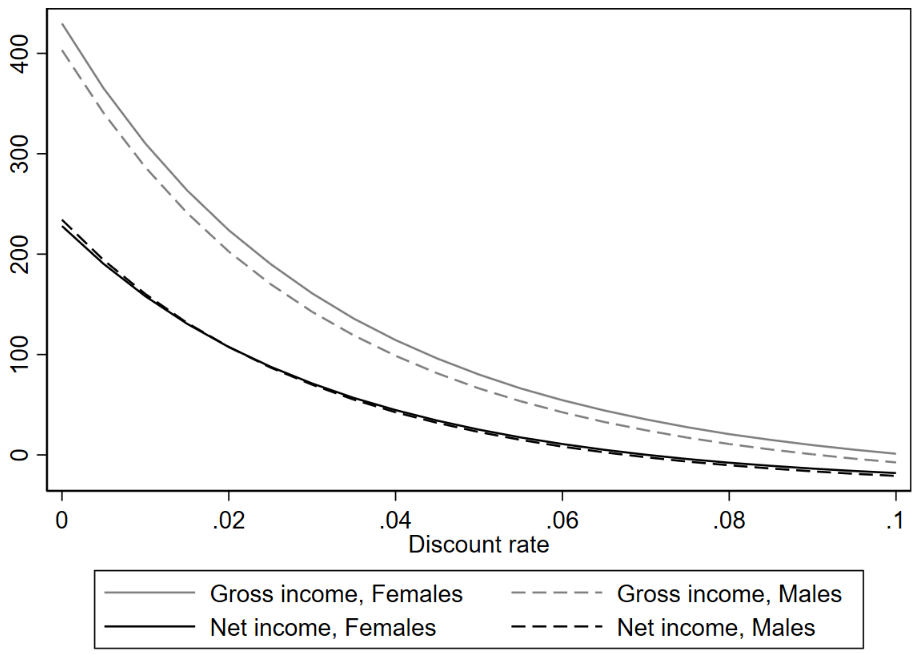
(b) Men, gross income

(c) Women, net income

(d) Men, net income

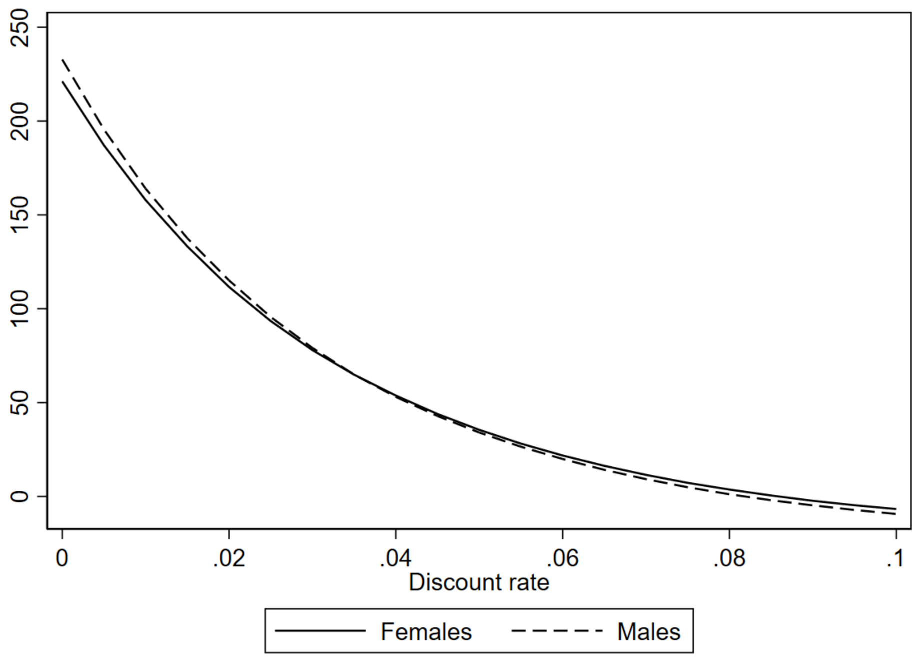
Notes: Density functions for individual simulation runs displayed in light grey, single density function for pooled runs in black. For married couples, 50% income pooling is assumed. *Source:* Own simulations.

Figure 1.A12: 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 1.A13: 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.

Table 1.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 1.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 1.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)
$(\text{Tenure in t-1})^2/100$	-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 1.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 1.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 1.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 1.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 1.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)
Experience^2/100	-0.110	-0.299***	-0.255**	-0.0674
	(0.0930)	(0.0910)	(0.102)	(0.115)
Experience^3/1,000	0.0258	0.0709**	0.0635	0.0162
	(0.0347)	(0.0321)	(0.0394)	(0.0418)
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)
Tenure^2/100	-0.257***	-0.245***	-0.109	-0.237***
	(0.0684)	(0.0595)	(0.0720)	(0.0873)
Tenure^3/1,000	0.0860***	0.0743***	0.0293	0.0681*
	(0.0285)	(0.0239)	(0.0314)	(0.0366)
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)
Age^2/100	-1.009**	-0.849*	-0.950**	-0.472
	(0.436)	(0.476)	(0.442)	(0.529)
Age^3/1,000	0.129*	0.108	0.148**	0.0582
	(0.0685)	(0.0719)	(0.0724)	(0.0856)
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 1.A9: 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)
Experience^2/100	-0.120	-0.304***	-0.248**	-0.0737
	(0.0928)	(0.0911)	(0.103)	(0.114)
Experience^3/1,000	0.0300	0.0730**	0.0619	0.0194
	(0.0346)	(0.0321)	(0.0401)	(0.0414)
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)
Tenure^2/100	-0.259***	-0.247***	-0.105	-0.245***
	(0.0683)	(0.0596)	(0.0722)	(0.0871)
Tenure^3/1,000	0.0878***	0.0750***	0.0275	0.0712*
	(0.0284)	(0.0240)	(0.0315)	(0.0366)
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)
Age^2/100	-1.023**	-0.869*	-0.938**	-0.440
	(0.435)	(0.475)	(0.444)	(0.522)
Age^3/1,000	0.131*	0.111	0.144**	0.0489
	(0.0684)	(0.0718)	(0.0727)	(0.0845)
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 1.A10: 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 1.A11: 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 1.A12: 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 1.A13: 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 1.A14: 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 1.A15: 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 1.A16: 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.

Chapter 2

The Decision to Enrol in Higher Education

2.1 Introduction

Policy makers around the world believe that human capital is a key factor in determining a country's economic success. However, it is ultimately individuals who decide on how much to invest in their education. Hence, policy makers have an interest in understanding how individuals make their educational decisions and how they can provide incentives to influence these decisions. This chapter deals with the question how individual decisions about entering higher education depend on their earnings expectations. Expected earnings can be modified by public policy, either directly through education policy (e.g. the abolition or introduction of tuition fees), or more indirectly, for instance through tax policy.

Analyzing the relationship between educational decisions and (expectations of) life-time earnings goes back to Mincer (1958), Becker (1962), and Ben-Porath (1967). Since then, numerous studies have analyzed how different earnings expectations lead to different decisions concerning (higher) education. One key challenge for these studies is the question how earnings expectations are formed. Some studies, for instance, assume that individuals make educational decisions based on the ex-post realizations of their income profiles, i.e. that individuals are able to perfectly forecast their future earnings. Others have argued that is it more realistic that individuals face some uncertainty with respect to their future earnings profiles. Hence, individuals act upon a limited information set available at the time of the educational decision rather than perfectly forecasting the future (Cunha et al., 2005; Cunha and Heckman, 2007). In the latter case, further assumptions have to be made as to how these earnings are forecasted. Some studies in this spirit have, for instance, assumed that individuals forecast their earnings based on older individuals' trajectories who are otherwise similar to them (see Wilson et al., 2005; Giannelli and Monfardini, 2003; and Flannery and O'Donoghue, 2013). The individuals' earnings expectations are then predicted using regression techniques.

To analyze the role expected earnings play for the higher education choice of young adults in Germany, I follow this literature and estimate a microeconometric model in which individuals maximize their expected life-time utility by deciding whether or not to take up academic training.[1] For the educational decision model, I use a German micro data set that follows a recent cohort of secondary school graduates and observes their educational trajectories after having completed upper secondary school. To forecast an individual's expected life cycle given an educational choice I use a dynamic microsimulation model (Fischer and Hügle, 2020). The dynamic microsimulation model uses survey and administrative data to estimate transition models in employment and family formation and then simulates individual transitions over the life cycle based on the estimated parameters. Estimating a tax function[2] I then translate the forecasted gross into individual net incomes. Importantly, I account for the fact that the current post-secondary education system in Germany contains multiple paths individuals might take. For instance, after entering higher education an individual might either leave the education system with a master degree, a bachelor degree, or no academic degree at all. Similarly, an individual might take into account taking up higher education after having finished a vocational training.

Finally, when estimating the educational decision model, I also take into account non-pecuniary factors such as cognitive skills and parental education that have been shown to be strong predictors of educational decisions (Black and Devereux, 2010). Having estimated the educational decision model, I use the estimates to simulate the introduction of different tuition fee and graduate tax schemes. Estimating the microeconometric model, I find an earnings elasticity of about 0.75, i.e. if expected net lifetime earnings of higher education graduates were to increase by 10%, on average the likelihood of entering higher education would increase by 7.5%. Yet, this elasticity would also imply that only few individuals would change their educational choice due to the introduction of tuition fees or graduate taxes.

The remainder of the chapter is as follows. Section 2.2 describes the institu-

[1] Note that I use the terms "higher education" and "academic training" interchangeably.

[2] The tax function also accounts for social security contributions. For the sake of simplicity, I use the term "tax function" throughout this chapter.

tional background of the higher education decision and introduces the microeconometric model. Section 2.3 presents the data and section 2.4 explains the regressions and the dynamic microsimulation model. Section 2.5 then presents the estimation of the educational choice model and the simulation results and section 2.6 concludes.

2.2 The higher education decision

2.2.1 Higher education and vocational training in Germany

Currently, 52% of recent German secondary school graduates have a higher education entrance degree (*Hochschulreife*). In general, these individuals face the decision between going to higher education or starting vocational training (*Berufsausbildung*)[3]. Higher education includes university (*Universität*) and university of applied sciences (*Hochschule für angewandte Wissenschaften*)[4] and vocational training comprises dual training (a combination of firm-based training and vocational school) and purely school-based training.[5]

Even though individuals with a higher education entrance degree can be assumed to choose among two options, higher education and vocational training, each option comes along with multiple paths that might potentially be realized from an ex-ante perspective. I model the most frequent of these pathways assuming that these are the potential pathways individuals take into account when making their educational choices. Figure 2.1 sums up these potential paths.

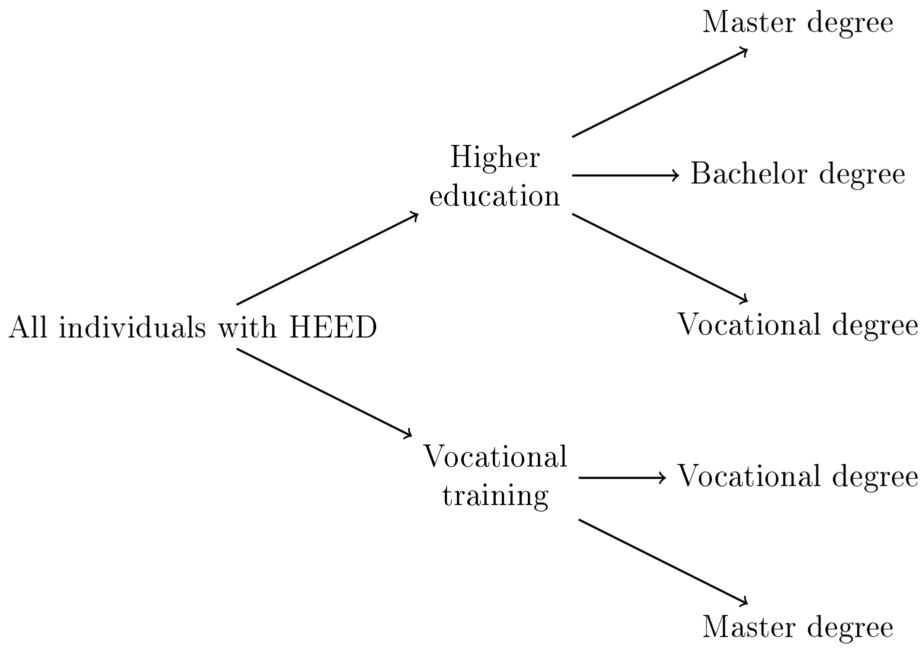
I assume that there exist three potential paths after having entered higher education: Obtaining a master degree, obtaining a bachelor degree (and not a master degree) and obtaining a vocational degree after having dropped out of higher education. While the master degree is the equivalent to the former *Diploma*, that used to be the most common degree in Germany before the Bologna reforms, it is estimated that a sizable fraction of 35% (Autorengruppe Bildungsberichtserstattung, 2016) do not enter a master program after graduating with a bachelor degree. Finally, there is a considerable risk of dropping out of higher education as the average dropout rate in bachelor degrees

[3]In principle, individuals could also enter the labor market directly without any post-secondary training. However, this does not seem to be an attractive option and almost no individuals choose this path.

[4]Currently, approximately 58% of new higher education entrees attend a university and 42% a university of applied sciences (Autorengruppe Bildungsberichtserstattung, 2018).

[5]Of the individuals who obtained a higher education entrance degree and start a vocational training 66% are in the dual training and 30% in the school-based training system. About 4% enter some form of pre-vocational training (Autorengruppe Bildungsberichtserstattung, 2018).

Figure 2.1: Potentially realizable paths of education



Note: *HEED*=Higher education entrance degree

across all subjects is 28% (Autorengruppe Bildungsberichtserstattung, 2018). I assume that individuals who drop out of higher education enter vocational training and obtain a vocational degree.[6]

For individuals who enter vocational training after their higher education entrance degree, I assume that there are only two realizable paths. Either the individual obtains a vocational training degree and leaves the education system entirely or she moves on to higher education and finishes with a master degree. Clearly, also other paths, such as dropping out of academic or vocational training and not obtaining any post-secondary degree, would theoretically be possible. However, they are rather rare so I deem it plausible that individuals do not take them into account when making their educational choice.

[6]Clearly, a part of the 28% who do not finish their bachelor studies enter another study program. Due to a lack of data, however, it is difficult to assess the share of these students. I therefore assume that individuals deciding about whether to enter higher education and assessing the dropout risk of higher education make the simplifying assumption that with a probability of 28% they drop out of higher education and enter vocational training.

2.2.2 A model of the higher education decision

I assume that individual i associates with each alternative $e = \{he, voc\}$ (i.e. higher education or vocational training) a life-time utility

$$V_i^e = \alpha^e LTI_i^e + x_i'\beta^e + \varepsilon_i^e \tag{2.1}$$

where LTI_i^e is the net lifetime income individual i expects to earn when choosing alternative e and x is a vector of other variables which are potentially important in explaining the higher education decision such as parental education and a measure for cognitive skills. Finally, ε captures all the determinants of life-time utility that cannot be observed by the researcher. It is assumed that ε is uncorrelated with the other terms on the right-hand side.[7]

Following the above discussion that each choice (i.e. higher education or vocational training) is associated with multiple realizable paths, expected lifetime income of entering higher education (he) and vocational training (voc) can be expressed as:

$$\begin{aligned}
LTI_i^{he} &= prob^{he_master} LTI_i^{he_master} \\
&\quad + prob^{he_bachelor} LTI_i^{he_bachelor} \\
&\quad + (1 - prob^{he_master} - prob^{he_bachelor}) LTI_i^{he_voc} && (2.2) \\
LTI_i^{voc} &= prob^{voc_voc} LTI_i^{voc_voc} \\
&\quad + (1 - prob^{voc_voc}) LTI_i^{voc_master} && (2.3) \\
& && (2.4)
\end{aligned}$$

where $prob^{c\_d}$ refers to the probability that the individual will choose c and leave the education system with degree d. For instance, $prob^{he\_bachelor}$ is the probability that the individual enters higher education (he) and leaves the system with a bachelor degree ($bachelor$). Hence, the expected lifetime income is just a probablity-weighted sum of lifetime incomes under different realizations.

Finally, the probability that individual i enters higher education can be written as

$$Pr(he_i = 1) = Pr(V_i^{he} > V_i^{voc}) = F(\alpha(LTI_i^{he} - LTI_i^{voc}) + x_i'\beta) \tag{2.5}$$

[7]Essentially, equation 2.1 assumes that individuals are risk-neutral. In general, the model could be extended to allow for risk aversion. Fossen and Glocker (2017, 2011) for instance, freely estimate such a parameter. However, estimating such a model with the data used in this chapter did not prove successful, as the estimated risk aversion parameters had implausibly large confidence intervals and were very sensitive to the slightest modification of the model. Therefore, I only estimate the model assuming risk-neutrality.

where $\alpha = \alpha^{he} - \alpha^{voc}$ and $\beta = \beta^{he} - \beta^{voc}$. $W_i = LTI_i^{he} - LTI_i^{voc}$ is the difference in the expected net lifetime incomes between the two alternatives for individual i. Assuming that the error term ε_i^e is EV(1) distributed and the difference $\varepsilon_i^{he} - \varepsilon_i^{voc}$ therefore is logistically distributed, the likelihood function is given by:

$$L = \prod_{i=1}^{N} F(\alpha W_i + x_i'\beta)^{he_i}(1 - F(\alpha W_i + x_i'\beta))^{(1-he_i)} \tag{2.6}$$

and the log-likelihood function is

$$LL = \sum_{i=1}^{N}(he_i\ log(F(\alpha W_i + x_i'\beta)) + (1 - he_i)\ log(1 - F(\alpha W_i + x_i'\beta))) \tag{2.7}$$

In the Logit estimation, equation (2.7) is maximized with respect to the income weight α and the parameters contained in β.

2.3 Data

I use two main data sets for the estimation, the National Educational Panel Study (NEPS) (Blossfeld and Von Maurice, 2011) and the Socio-Economic Panel (SOEP) (Goebel et al., 2018). The NEPS follows educational trajectories of different starting cohorts (SC), from newborns (SC1) to adults (SC6). I use the SC4, which, starting in 2010, has been following the educational careers of about 13,000 pupils starting in 9th grade. The educational decisions after secondary school of the SC4 cohort are the ones analyzed in this chapter. Table 2.A1 in the Appendix shows the descriptive statistics for the final sample with which the decision model will be estimated.

The SOEP is a household panel that started in 1984 and currently surveys about 30,000 individuals. I use the SOEP for two purposes. First, the SOEP is the main data base for the dynamic microsimulation model outlined in Fischer and Hügle (2020) (see Section 2.4.3 for more details). Second, I use the SOEP waves from 2000 to 2012 to estimate wage parameters and a tax function by which life-time income profiles are constructed. Restricting the data set to waves until 2012 is due to the timing of the education decisions analyzed in this study: The NEPS cohort of interest was in 11th grade in 2012. I assume that this is around the time when these individuals made their educational decisions and hence the point in time from which they draw their information set. Finally, in order to make assumptions about training length, income during training, and dropout probability I draw on additional sources, particularly on the Autorengruppe Bildungsberichtserstattung (2014, 2016, 2018).

2.4 Parameter estimation and life-cycle simulation

In this section, I describe the estimation of wage regressions, the tax function, and how the life cycles are constructed using dynamic microsimulation. Together with the forecasted life cycles, the parameters of the wage regressions are used to predict hourly wages and annual labor earnings over the life cycle. The tax function is then used to translate gross into net earnings.

2.4.1 Gross hourly wages

To predict earnings over the life cycle, I estimate Mincer-type wage regressions separately for education (i.e. for individuals with higher education and individuals with vocational degree and higher education entrance degree)[8] and gender. The estimating equations are defined as

$$log(wage_{it}^{he}) = x_{it}'\beta^{he} + \varepsilon_{it}^{he} \tag{2.8}$$

$$log(wage_{it}^{voc}) = x_{it}'\beta^{voc} + \varepsilon_{it}^{voc} \tag{2.9}$$

where equation 2.8 is estimated using the sample of higher education graduates and equation 2.9 using individuals with vocational degree and higher education entrance degree.[9] Importantly, I only use observations with a master or an equivalent degree for estimating equation 2.8.[10] $log(wage)_{it}$ is the log gross hourly wage of individual i in year t. x is a vector of covariates including a fourth-order polynomial of labor market experience, an indicator for migration background, nine industry dummies and dummies for civil service and self-employment, and a vector of year dummies. In addition, x also includes a vector of dummies for the German states. They are fundamental for generating the variation in expected lifetime income gaps between academic and vocational training across states and hence across individuals. The idea is that when making the educational choice each individual faces different expected lifetime income gaps between academic and vocational training because of the state she lives in.[11] Finally, equation 2.8 also controls for having a university of applied science degree.

[8] Note that, in this study, the term *vocational degree* implies a higher education entrance degree, even though it is not always explicitly stated.

[9] Using the parameter estimates of the log wage equation, the hourly wage of individual i can be computed using the formula $\hat{w}_i = exp(x_i'\hat{\beta} + 0.5\hat{\sigma}^2)$, where x is the vector of covariates, $\hat{\beta}$ is the vector of coefficient estimates of the log wage equation, and $\hat{\sigma}^2$ is an (unbiased) estimator of the model error in the log wage regression (Cameron and Trivedi, 2009).

[10] For the estimation of the bachelor wage penalty, see Section 2.4.1.

[11] Section 2.6 will discuss this issue further.

The equations are estimated by OLS. There are two selection issues that need to be addressed. The first is non-random selection into education, i.e. into higher education and vocational training, as individuals are “choosing" their education levels. Another potential selection bias might arise due to non-random selection into the labor force, i.e. the fact that the estimation samples only contain working individuals for whom an hourly wage can be computed. A natural solution for these two problems is the estimation of selection-corrected wage equations. This means that one first estimates selection equations for the education and work choices using Probit models and then adds selection correction terms to the set of x variables in the wage equations.

For the estimation of selection-corrected wage equations exclusion restrictions are required, i.e. variables that affect the education and work choices but do not directly enter the wage equations.[12] Here, I follow the literature and use marital status and dummies for the presence of children in the household between the ages 0 and 5 and between 6 and 17 as exclusion restrictions for the selection into work (for a similar approach see Steiner and Lauer (2000) and Fossen and Glocker (2017, 2011)). For the selection into education, I follow Fossen and Glocker (2017, 2011) and use parental variables before the individual graduates from secondary school such as indicator variables for parental education, for whether they work, and for whether they were born in Germany. However, one should bear in mind, particularly with respect to the selection-into-education corrections, that the advantage of using the selection corrections crucially depends on the validity of the exclusion restrictions. It is plausible to assume that variables that are related to parental attitudes, behavior, and characteristics (such as parental education and whether parents work) might be correlated with the unobservables in the wage equation, such as an individual's ability and motivation. For these reasons, I use the wage specification without selection corrections as my main specification, but also report the results using two additional wage specifications: one where I only use a selection correction for work and one where I use a selection correction for both education and work. The latter is, due to the argument above, the least preferred specification. The main results of the chapter, however, such as the elasticity of the educational choice with respect to lifetime income, barely depend on which specification is used, as will become clear below.

Tables 2.A2-2.A5 in the Appendix display the regression results for the selection and wage regressions. For the main specification without selection corrections, there are wage penalties for having a migration background between 12% (men and women with higher education) and 21% (men with vocational degrees) and a penalty for having

[12] Technically, the model could also be identified without exclusion restrictions due to the non-linearity of the selection correction terms in the observable variables.

a university of applied sciences degree (compared to university) between 14% (men) and 19% (women).

Bachelor wage penalty

As individuals potentially finish their academic career with a bachelor degree, we need to make assumptions concerning the wage profile of such graduates. In order to estimate a potential hourly-wage penalty of bachelor relative to master degrees, I use SOEP waves from 2010 to 2012[13] and estimate a similar wage equation to (2.8). I find a bachelor wage penalty of 10.1% for men and 13.5% for women. This is comparable to the estimate of Christoph et al. (2017) who use administrative data and find a wage penalty of about 10% at age 30.

2.4.2 The tax function

As, by assumption, lifetime utility is a function of individual *net* income, it is necessary to translate expected gross into net incomes. To do this, I approximate the tax-and-contributions system[14] of the year 2012 (by assumption the year of the educational choice) by estimating the function

$$\begin{aligned} taxrate_{it} = \quad & \beta_0 + \beta_1 grossinc_{it} + \beta_2 grossinc_{it}^2 + \beta_3 grossinc_{it}^3 + \beta_4 grossinc_{it}^4 \\ & \beta_5 grossinc_{it}^5 + \beta_6 nr\_children_{it} + \beta_7 married_{it} + \varepsilon_{it} \end{aligned} \tag{2.10}$$

with data for the years 2010-2012 where $taxrate_{it}$ is the tax rate of individual i in period t[15], *grossinc* is the individual annual gross labor income, *married* a dummy for being married, and *nr_ children* is the number of children.

Table 2.A6 (Appendix) displays the estimated coefficients and Figure 2.A1 (Appendix) plots the predicted average tax rates for different annual labor incomes for an unmarried individual without children. Somewhat surprisingly, the curve of the average tax rate is downward sloping starting at an individual annual labor income of about 75,000 Euros. However, this part of the slope concerns no individual as the maximum predicted annual earnings of any individual is about 70,000 Euros.

[13]For the estimation I use the SOEP's ISCED11 classification that distinguishes between master and bachelor degrees and is only available from 2010 on.

[14]For simplicity, I ignore transfers in this analysis. Yet, as Fischer and Hügle (2020) show, their quantitative importance for individuals with higher education entrance degree is very limited compared to taxes and social security contributions.

[15]$taxrate = \frac{\text{individual annual gross labor income} - \text{individual annual net labor income}}{\text{individual annual gross labor income}}$

2.4.3 Life-cycle simulation

Having estimated the corresponding hourly wage and tax parameters, the next step is to forecast the individual life cycles. Here, we need to make assumptions about the individuals' perceptions of their potential training trajectories, such as the probabilities of different realized paths, training length, and the earnings while in training. I make these assumptions based on different aggregate statistics as of 2012 (or before), as this was the time period when individuals had to decide about the enrolment in higher education.

In general, I assume that all individuals make their decision whether or not to enter higher education at the age of 20[16] and then make a transition into one of the two alternatives. In academic training, an individual drops out with 28% (Autorengruppe Bildungsberichtserstattung, 2014). Given she finishes the bachelor degree, she will leave higher education with a probability of 36%, and move on to graduate with a master degree with 64% (Autorengruppe Bildungsberichtserstattung, 2018). In 2012, the average duration until graduation with a master degree or diploma was 11.2 semesters (Autorengruppe Bildungsberichtserstattung, 2018). I therefore assume that if the individual continues after the bachelor, she will leave the education system with a master degree after six years. I further assume that an individual has net earnings of 474 Euros while in academic training which is the average of the sum of labor earnings and student grants in Germany (Middendorff et al. (2017) and own calculations).

If, in contrast, an individual enters vocational training, she is assumed to finish after three years of training, the official duration of most such training programs. With a probability of 35%, she will afterwards take up academic training and I assume that she finishes with a master degree in six years. Furthermore, I assume that while in vocational training an individual has net earnings of 632 Euros which is the weighted average of those who earn salaries (*Ausbildungsvergütung*) in dual training and those who receive pupil grants in school-based training.

After graduation, individuals are assumed to enter the labor market and retire at the age of 67, the official retirement age for this cohort in Germany. In order to simulate the individual life cycles in terms of employment and family formation (i.e. marriage, fertity, and divorce) I use a modified version of the dynamic microsimulation model outlined in Fischer and Hügle (2020). The modified dynamic microsimulation model has two stages: In the first stage, transition models for the processes of employment and family formation are estimated via different discrete-choice models. The

[16]The median age of entry into higher education was 19,7 in 2012 (Autorengruppe Bildungsberichtserstattung, 2018).

key explanatory variables of these models are indicators that capture the academic and vocational degrees and indicators for being in academic or vocational training. In addition, the models control for migration background, dummies for the federal states, year dummies, and lagged variables of employment states and family formation. All transition models are estimated using SOEP data. In order to guarantee that the simulated individual transitions will follow predicted aggregate trends, so-called *fractional regression models* (Papke and Wooldridge, 1996, 2008) are estimated. These fractional regression models use the shares of different employment states, and birth, marriage, and divorce rates as dependent variables and regress them on polynomials of age and cohort dummies. While the fractional regression models for employment are estimated using SOEP data, administrative data are used for the models of family formation.

In the second stage, individual transitions are sequentially simulated, starting at age 18. For this, using the parameter estimates of the transitions models one first predicts individual fertility, marriage, and employment probabilities. Then these predictions are multiplied with random draws from the unit interval. Finally, individuals are selected for transitions based on these modified probabilities until the aggregate targets (which are predicted using the fractional regression estimates) are met. In the end, each individual in the decision sample is assigned the age-specific means of the simulated variables conditional on gender, migration background, and federal state.

Having forecasted the individual life cycles in terms of employment and family formation, I can then simulate the annual gross earnings over the life cycle by using the estimates of the hourly wage regressions. Importantly, each individual has then one estimate for the expected lifetime income under each realizable path. Annual net income can then be computed using the yearly gross income, the estimated tax parameters, and the (simulated) presence of children and marital status. Tables 2.A7 and 2.A8 in the Appendix show the average predicted net and gross life-time earnings for higher education and vocational training separately for men and women and the different wage specifications.[17]

As expected, life-time earnings are substantially higher for men than for women across all paths and finishing with a master degree (HE) is associated with higher life-time earnings than vocational training (VOC). Comparing the different wage specifications, with and without selection corrections, the tables show that the selection corrections lead to reduced simulated lifetime earnings, particularly for the specification with selection corrections for both education and work. However, one should bear

[17] For the sake of clarity, I only present the two "standard" paths here, i.e. choosing higher education and finishing with a master degree and choosing vocational training and finishing with a vocational degree.

in mind that the latter specification is the one that should be seen with caution with respect to its validity. Finally, the implied average tax rate is much larger than the one for similar gross incomes simulated in Fischer and Hügle (2020). For instance, while the implied average tax rate for gross incomes of 1.427 Mio. Euros (the simulated lifetime income of men under higher education in the base specification) is about 38.5%, Fischer and Hügle (2020) report an estimate of about 31% for a similar income level. The difference, however, can mainly be explained by the fact that Fischer and Hügle (2020) exclude pension insurance contributions from their analysis.

2.5 Results

2.5.1 Decision model: Parameter estimation

Table 2.1 shows the estimates of the Logit model outlined in equation 2.7 for the three wage specifications discussed above. The main variable of interest is the difference in expected net lifetime income between entering higher education and vocational training, $\Delta LTI = LTI^{he} - LTI^{voc}$. In addition, I control for other determinants of the higher education decision: Gender, parental education, parental occupation, migration background, and cognitive skills. For parental education, I define three categories with respect to the parent with the highest education level: No higher education entrance degree, a higher education entrance degree (but no higher education degree), and a higher education degree. Hence, if one parent has a higher education degree and the other has no higher education entrance degree, the parents are classified as having a higher education degree. Similarly, parental occupation, which serves as a proxy for parental income, has three categories which are defined with respect to the parent with the highest EGP class: High (e.g. managers, high-ranked civil servants, highly qualified white collar workers), medium (e.g. qualified white collar workers, master craftsmen) and low. Hence, if one parent has a high EGP class and the other has a medium class, the parents are classified as having a high EGP class.[18] Cognitive skills are measured by the tested competencies in different fields such as perceptual speed, reasoning, and numeracy skills.

The statistically significant positive coefficient estimate of ΔLTI indicates that an increase in the expected net lifetime-income gap between higher education and vocational training increases the probability of entering higher education. In order to interpret its magnitude, one can use the parameter estimates and predict how the

[18] See Biewen and Tapalaga (2017) for a similar categorization.

Table 2.1: Enrolment decision: Logit estimates

	No selection correction	Selection into work	Selection into work and education
$\Delta LTI/10{,}000$	0.0863***	0.0925***	0.0956***
	(0.0340)	(0.0355)	(0.0366)
Female	0.1710***	0.1625***	0.0300
	(0.0737)	(0.0731)	(0.0832)
Parents: HE entrance deg.	0.4783***	0.4794***	0.3876***
	(0.0899)	(0.0899)	(0.0950)
Parents: HE deg.	0.9701***	0.9703***	0.8834***
	(0.0932)	(0.0932)	(0.0975)
Parents: Medium occ.	0.0379	0.0379	0.0441
	(0.1164)	(0.1165)	(0.1165)
Parents: Max. occ.	0.2878***	0.2877***	0.2960***
	(0.1251)	(0.1251)	(0.1251)
Migration background	0.5360***	0.5318***	0.5049***
	(0.1503)	(0.1505)	(0.1536)
Cognitive skills	0.7878***	0.7877***	0.7891***
	(0.0483)	(0.0483)	(0.0483)
Constant	−0.6545***	−0.6364***	−0.4935***
	(0.1573)	(0.1511)	(0.1224)
N	4,106	4,106	4,106

Notes: This table displays the Logit coefficients and standard errors (in parentheses) of estimating the log-likelihood function (equation 2.7). The different columns represent different underlying wage specifications used to simulate lifetime incomes. *No selection*= lifetime incomes are based on the wage specifications without selection correction. *Selection into work*= lifetime incomes are based on wage specification with selection correction for selection into work. *Selection into work and education*= lifetime incomes are based on wage specification with selection correction for selection into work and education. $\Delta LTI = LTI^{he} - LTI^{voc}$ is the difference in the expected individual net lifetime income between higher education and vocational training. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: NEPS, SOEP, own calculations.

enrolment probabilities would react if net lifetime incomes for those with a higher education degree would increase by 10%. Using the model without selection corrections in the wage equations, I find that such a 10% increase would rise the higher education enrolment probability, on average, by about 7.5%, which implies an “elasticity" of 0.75.[19][20][21]

[19]The elasticities for the different underlying wage specifications are: 0,747 (no selection correction), 0.779 (correction for selection into work), 0.749 (correction for selection into work and education).

[20]This quantity can be computed by first computing the relative change in the individual choice probabilities after increasing the net lifetime income of academics by 10% and then averaging the relative change over all individuals.

[21]This elasticity is very close to the elasticities found in Fossen and Glocker (2017), which are around 0.8.

For the control variables, one can use the estimated coefficients and compute the average marginal effects (not shown in the tables). For the base specification this computation shows that parental education has a strong positive impact: Having at least one parent with a higher education entrance degree increases the probability of entering higher education by about 9 percentage points, while having at least one parent with an academic degree increases this probability by about 18 percentage points (both effects compared to the base category of not having at least one parent with a higher education entrance degree). Parental occupation, in contrast, only has a small effect on higher education enrolment: Having at least one parent with high occupational status increases the likelihood to enter academic training by about 5.5 percentage points.[22] An increase in cognitive skills by one standard deviation increases the likelihood of entering academic training by about 15 percentage points. In addition, there is a strong positive effect of having a migration background on the probability of higher education enrolment of more than 9.5 percentage points.

2.5.2 Simulation of tution fees and graduate taxes

In order to assess how strongly a change in educational policies would impact the higher education decision, I simulate different tuition fee and graduate tax scenarios. As to tuition fees, I simulate three different scenarios: annual fees of (i) 2,000 Euros, (ii) 4,000 Euros, and (iii) 6,000 Euros. While the latter is close to the actual average cost of tuition per student and year, the first two scenarios would be more likely to receive public support.[23] I assume a system with deferred repayment and income-contingent loans, i.e. individuals gradually pay back their debt if their individual net income exceeds a certain threshold. Such a system has been in place in some Western countries such as England, Australia, and New Zealand and has been described in the theoretical literature as being superior in terms of efficiency and equity compared to a system where fees are to be paid up-front (see, for instance, Barr, 2004, and Chapman, 2006). A main reason is that up-front tuition fees might cause liquidity constraints and particularly deter individuals from low socio-economic background from enrolling. In addition, Lergetporer and Woessmann (2019) find that designing tuition fees as deferred income-contingent payments would substantially increase public support for fees.

I set the net income threshold above which an individual has to pay back tuition

[22] The effects of parental education and occupation are similar in magnitude to Biewen and Tapalaga (2017).

[23] Some West German states temporarily collected tuition fees starting in 2006/2007. These fees were usually about 500 Euros per semester, and hence 1,000 Euros per year.

debt to 20,000 Euros and the repayment rate, i.e. the share of net income above the threshold that has to be paid back, to 0.2. Furthermore, I assume that there are no interest rates. Another feature of tuition fees and the main difference to the nature of graduate taxes is that the maximum amount an individual would have to pay back over her lifetime is limited, for instance to 36,000 Euros if tuition fees are 6,000 Euros annually and if the individual studies for 6 years.

In contrast to tuition fees, graduate taxes imply that each higher education graduate pays a share on her individual net income, independently of the total amount already paid, i.e. the total amount of graduate taxes depend on income earned over the lifetime.[24] This implies that a graduate tax might imply a much higher total debt over the lifetime. Graduate taxes have been discussed as an alternative to tuition fees, especially in Great Britain.[25] Here, I simulate three different scenarios: A graduate tax of (i) 1%, (ii) 2%, and (iii) 3% of individual net income.

Tables 2.2 and 2.3 show the simulation results. It becomes clear that none of the tuition fee or graduate tax scenarios would dramatically change the average enrolment probabilities. As for tuition fees, the largest reform, the introduction of annual tuition fees of 6,000 Euros, would only reduce the average enrolment probability from 67.34% to 65.86%. For graduate taxes, the results are similar in magnitude. Comparing the two different policies one finds that the effects tuition fees of about 6,000 Euros per year have on enrolment correspond to the effects a graduate tax of 2% has on enrolment. The reason is that these two schemes imply a similar financial burden. A larger graduate tax of 3% would cause larger responses and a reduction in the enrolment probabilities of more than 3%.

That the effects of the simulated policies on higher education enrolment are limited in magnitude seems plausible given that the total amount of debt an individual would accumulate in systems with tuition fees or graduate taxes is quite small relative to the expected lifetime income. However, these results partially contrast recent research of the effects of tuition fees in Germany (see, for instance, Bietenbeck et al., 2020; Bruckmeier and Wigger, 2014; and Hübner, 2012). Exploiting the variation in the introduction of fees across the German states starting in the mid-2000s, these studies estimate the effect on enrolment via difference-in-differences estimations. While Bietenbeck et al. (2020) and Hübner (2012) find that the introduction of tuition fees reduced enrolment by about 3.9 and 2.7 percentage points, respectively, Bruckmeier and Wigger (2014) finds a smaller negative effect of 0.9 percentage points which is not

[24]Some proposals of graduate taxes also imply that the total amount paid is limited to some level. In this simulation exercise, I assume that there is no such limit.

[25]Supporters of graduate taxes include, for instance, the former UK prime minister Gordon Brown.

Table 2.2: Effect of different tuition fee schemes

Scenario	Probability (in %)	abs. change (in percentage points)	rel. change (in percent)
Base	67.34		
2,000 Euros/year	66.73	-0.61	-0.91
4,000 Euros/year	66.24	-1.10	-1.63
6,000 Euros/year	65.86	-1.48	-2.20

Notes: This table displays the effects of different tuition fee schemes on the average enrolment probabilities. *Base* = Base scenario. The lines below describe the different tuition fee scenarios. For instance, *2,000 Euros/year* describes the effects of an introduction of annual tuition fees of 2,000 Euros. *Probability* = Probability of enroling in higher education. *abs. change* = absolute change in enrolment probability. *rel. change* = relative change in enrolment probability.
Source: NEPS, SOEP, own calculations.

Table 2.3: Effects of different graduate tax schemes

Scenario	Probability (in %)	abs. change (in percentage points)	rel. change (in percent)
Base	67.34		
1% graduate tax	66.65	-0.69	-1.03
2% graduate tax	65.94	-1.40	-2.07
3% graduate tax	65.24	-2.10	-3.13

Notes: This table displays the effects of different graduate tax schemes on the average enrolment probabilities. *Base* = Base scenario. The lines below describe the different graduate tax scenarios. For instance, *1% graduate tax* describes the effects of an introduction of a graduate tax of 1%. *Probability* = Probability of enroling in higher education. *abs. change* = absolute change in enrolment probability. *rel. change* = relative change in enrolment probability.
Source: NEPS, SOEP, own calculations.

statistically significant. At a first glance it might be surprising that these studies find a much larger negative response of tuition fees on the enrolment decision, given that the tuition fees analyzed in those studies where mostly 1,000 Euros per year. However, there is a key difference between the hypothetical tuition fee reforms analyzed here and the actual ones implemented in the mid-2000s: While the mid-2000s reforms made students pay their fees up-front, the fee scheme simulated here would include a deferred payment, and only if income exceeds a certain threshold. As precisely the up-front nature of the tuition fees of the mid-2000s might have prevented individuals from lower socio-economic backgrounds from higher education, it is plausible to assume that the tuition fee reforms of the mid-2000s should imply larger negative effects on enrolment than the schemes considered here.

As argued above, the hypothetical fee system with deferred payment analyzed here is closer to the English tuition fee system. In 1998, England introduced tuition fees and increased them in 2006 and further in 2012. At the same time, however, England also increased the financial support for students leading to a similar system to the one analyzed here. Indeed, Murphy et al. (2019) and Azmat and Simion (2020) find only small effects of the introduction (in 1998) and the increase of tuition fees (in 2006 and 2012) on enrolment. In fact, these studies find that the introduction of sizable tuition fees decreased the enrolment probability of individuals of higher socio-economic background much more than those of lower socio-economic background. The reason, the authors state, is that while individuals of lower socio-economic background received generous financial support those of high socio-economic background often did not and had to bear the tuition fees by themselves. These studies suggest that the effects of enrolment one can expect from the introduction of tuition fees crucially depend on whether tuition fee (re)payment is up-front or deferred and whether repayment is income-contingent.

2.6 Conclusion

This chapter analyzes the role of expected earnings for the decision to enrol in higher education. To do so, I forecast life cycles using dynamic microsimulation and regression techniques. Then, I estimate a microeconometric model where individuals maximize expected life-time utility by choosing whether to participate in academic training. I assume that, while making their decision, individuals take into account that there is uncertainty with respect to the educational path they will follow in the future. I find an elasticity of about 0.75, i.e. a 10% increase in expected individual net lifetime income for higher education degrees would increase the average likelihood of entering higher education by about 7.5%. Finally, I simulate different tuition fee and graduate tax scenarios. I find that tuition fees of a “plausible" size would cause only small changes in enrolment behavior.

One argument in favor of tuition fees has been that tuition fees could help academic institutions increase the quality of their education. The analysis in this chapter suggests that governments could raise some additional revenue for higher education by introducing or increasing tuition fees or graduate taxes without deterring many students from entering higher education. However, the recent literature on the German experience with tuition fees in the mid-2000, also suggests that the effect of tuition fees crucially depends on whether they are to be paid up-front or whether payment is deferred and combined with income-contingent loans. While up-front fees may indeed have a strong negative effect on enrolment, deferred fees might be much more favorable preventing liquidity constraints of students from lower socio-economic backgrounds.

Key to the internal validity of the approach used in this chapter are assumptions on how individuals form expectations, particularly about future earnings, but also concerning academic dropout risks etc. However, validating these assumptions is a difficult task, especially as data sets which contain individuals' subjective expectations about future outcomes are only scarcely available. Yet, there is a growing body of research that tries to capture such expectations (see e.g. Wiswall and Zafar (2015) and Arcidiacono et al. (2020) for recent examples). Comparing "objective" and subjective expectations is a promising area for future research. It would be particularly interesting to conduct an analysis of the heterogeneity of such expectations. Here, I assumed for instance, that individuals have the same probability of obtaining a master degree, given that they already have obtained a bachelor degree. It might be, however, that there are structural differences between individuals.

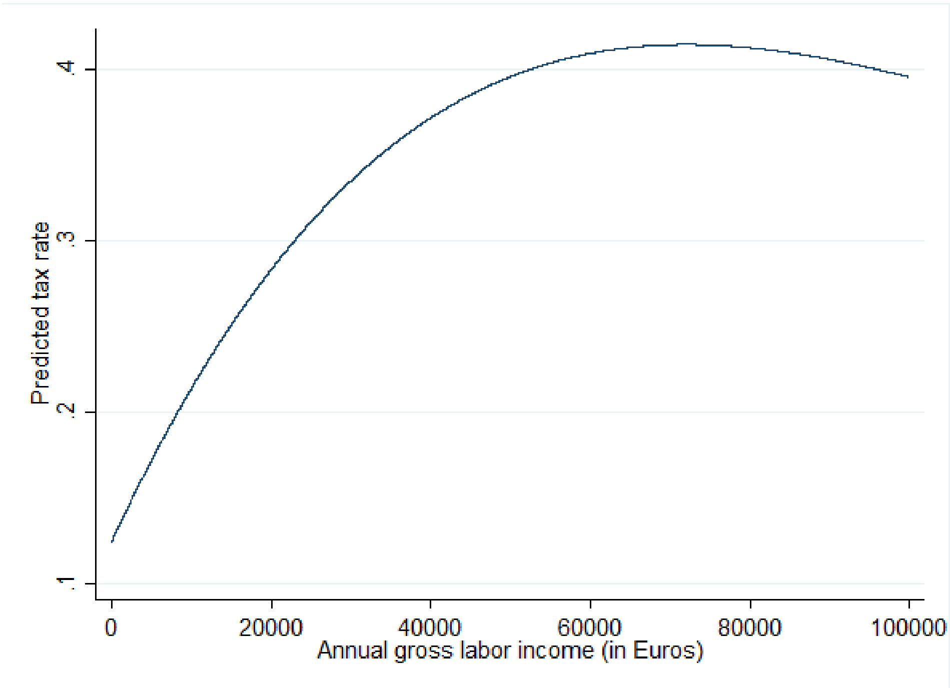
A similar aspect concerns the assumption in how far an individual's wage expectations are determined by the state or region she lives in. Here, as in the studies with a similar approach in the literature, I assume that individuals basically form their wage expectations for higher education and vocational training based on the wages they observe in their state. This assumption might be too strong for individuals who expect to move to other regions in Germany and therefore have different expectations than individuals from the same state who plan to stay in their state. Future research could aim at investigating these heterogeneities.

Another avenue for future research is the consideration of earnings risk in the educational decisions. A few studies in the literature have modeled earnings and earnings risk jointly (see Fossen and Glocker, 2017; Fossen and Glocker, 2011; and Buchinsky and Leslie, 2010). The idea is that individuals might take into account that their choices are "risky" in different dimensions: For instance, individuals might not associate an educational choice with a single estimate of lifetime income but rather a whole distribution of expected earnings and be uncertain as to which "draw" of this distribution will be realized. Future research might use the simulated life cycles and combine them with a more elaborate decision making model that allows for different degrees of risk aversion.

Finally, it should be noted that individuals who enrol in higher education also decide on a subject they want to specialize and major in (see Altonji et al. (2016) for an overview over this literature). In the case of Germany, Hügle (2021) analyzes this choice by modeling *fields of study*. It would be a promising avenue of research to analyze how these two neighboring fields of the literature could be combined, i.e. to model how individuals decide jointly about studying in general and choosing specific fields of study or training programs in particular.

Appendix

Figure 2.A1: Tax function



Notes: This graph shows the predicted tax function for an unmarried individual without children. “Taxes" also include social security contributions.
Source: SOEP, own calculations.

Table 2.A1: Descriptive statistics, NEPS sample

	Mean	Std. dev.
Female	0.54	0.50
Migration background	0.07	0.26
Transition into higher education	0.67	0.47
Cognitive skills	0.64	0.83
Parental education: No HEED	0.39	0.49
Parental education: HEED	0.24	0.43
Parental education: Academic degree	0.36	0.48
Parental occupation: Low	0.11	0.31
Parental occupation: Middle	0.44	0.50
Parental occupation: High	0.45	0.50
Parents germanborn: n.a.	0.29	0.46
Parents germanborn: no	0.06	0.23
Parents germanborn: yes	0.65	0.48
Father working at age 15: n.a.	0.16	0.37
Father working at age 15: no	0.05	0.22
Father working at age 15: yes	0.79	0.41
Mother working at age 15: n.a.	0.14	0.35
Mother working at age 15: no	0.11	0.31
Mother working at age 15: yes	0.75	0.43
N	4,106	

Notes: This table displays mean and standard deviation of the variables used in the analysis for the NEPS sample. Abbreviations: *HEED*= Higher education entrance degree.
Source: NEPS, own calculations.

Table 2.A2: Selection into education and work, Probit estimates

	Men, education	Women, education	Men, work	Women, work
main				
Parental education: HEED	0.859***	0.926***		
	(0.0121)	(0.0115)		
Parental education: n.a.	-0.398***	-0.205***		
	(0.0161)	(0.0152)		
Father working	0.0626***	0.197***		
	(0.0197)	(0.0204)		
Father working: n.a.	0.250***	0.167***		
	(0.0232)	(0.0231)		
Mother working	-0.157***	0.0796***		
	(0.0113)	(0.0116)		
Mother working: n.a.	-0.186***	0.0857***		
	(0.0125)	(0.0123)		
Parents germanborn	0.211***	0.0391*		
	(0.0222)	(0.0219)		
Parents germanborn: n.a.	0.577***	0.260***		
	(0.0218)	(0.0215)		
Experience^2/100			0.0139	0.460***
			(0.0164)	(0.0222)
Experience^3/1,000			-0.00833	-0.165***
			(0.00911)	(0.0150)
Experience^4/100,000			-0.00720	0.167***
			(0.0131)	(0.0252)
Migration background			-0.392***	-0.202***
			(0.0133)	(0.0122)
Married			0.511***	-0.809***
			(0.0122)	(0.0125)
Children aged 0-5 in hh.			-0.0939***	-0.848***
			(0.0130)	(0.0113)
Children aged 6-17 in hh.			-0.200***	0.188***
			(0.0108)	(0.0105)
Constant	-1.179***	-1.370***	-1.169***	1.610***
	(0.0269)	(0.0275)	(0.0293)	(0.0367)
State dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	99247	110592	96666	109368

Notes: This table displays the results of Probit regressions. Each column represents a separate regression. Columns 1 and 2 represent the selection-into-education equation. The dependent variable is a binary indicator whether the individual has an academic degree. Columns 3 and 4 represent the selection-into-work equation. The dependent variable is a binary indicator whether the individual is working.
Source: SOEP, own calculations.

Table 2.A3: Wage regressions, no selection correction

	Men, HE	Men, VOC	Women, HE	Women, VOC
Experience/10	1.189***	0.690***	0.718***	0.546***
	(0.0461)	(0.0664)	(0.0482)	(0.0553)
Experience^2/100	-0.676***	-0.282***	-0.416***	-0.264***
	(0.0449)	(0.0665)	(0.0569)	(0.0678)
Experience^3/1,000	0.172***	0.0504**	0.109***	0.0640**
	(0.0163)	(0.0245)	(0.0237)	(0.0291)
Experience^4/100,000	-0.163***	-0.0307	-0.108***	-0.0647
	(0.0195)	(0.0298)	(0.0316)	(0.0397)
Migration background	-0.123***	-0.214***	-0.117***	-0.135***
	(0.0109)	(0.0124)	(0.0134)	(0.0117)
UAS	-0.137***		-0.185***	
	(0.00662)		(0.00851)	
Constant	2.401***	2.078***	2.455***	2.193***
	(0.0318)	(0.0482)	(0.0308)	(0.0325)
State dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	15408	6391	11381	7332

Notes: This table displays the results of linear regressions without selection corrections using the log hourly gross wage as the dependent variable. Each column represents a separate regression. *HE* (*VOC*) implies that the estimation sample is based on individuals with a higher education (vocational training) degree. *UAS*=University of applied sciences degree.
Source: SOEP, own calculations.

Table 2.A4: Wage regressions, selection-into-work correction

	Men, HE	Men, VOC	Women, HE	Women, VOC
Experience/10	1.144***	0.648***	0.700***	0.530***
	(0.0464)	(0.0671)	(0.0482)	(0.0559)
Experience^2/100	-0.649***	-0.255***	-0.368***	-0.237***
	(0.0450)	(0.0668)	(0.0571)	(0.0690)
Experience^3/1,000	0.164***	0.0420*	0.0913***	0.0521*
	(0.0163)	(0.0246)	(0.0238)	(0.0297)
Experience^4/100,000	-0.152***	-0.0195	-0.0885***	-0.0483
	(0.0196)	(0.0299)	(0.0316)	(0.0406)
Migration background	-0.0708***	-0.174***	-0.133***	-0.138***
	(0.0126)	(0.0153)	(0.0135)	(0.0121)
Correction term work	-0.180***	-0.151***	0.179***	0.0390*
	(0.0219)	(0.0341)	(0.0196)	(0.0214)
UAS	-0.134***		-0.184***	
	(0.00664)		(0.00849)	
Constant	2.667***	2.309***	2.409***	2.188***
	(0.0453)	(0.0711)	(0.0311)	(0.0328)
State dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	15356	6378	11348	7297

Notes: This table displays the results of linear regressions with a selection-into-work correction using the log hourly gross wage as the dependent variable. Each column represents a separate regression. *HE* (*VOC*) implies that the estimation sample is based on individuals with a higher education (vocational training) degree. *UAS*=University of applied sciences degree.
Source: SOEP, own calculations.

Table 2.A5: Wage regressions, selection-into-work and selection-into-education correction

	Men, HE	Men, VOC	Women, HE	Women, VOC
Experience/10	1.146***	0.644***	0.708***	0.535***
	(0.0463)	(0.0670)	(0.0481)	(0.0559)
Experience^2/100	-0.650***	-0.250***	-0.370***	-0.241***
	(0.0449)	(0.0667)	(0.0570)	(0.0690)
Experience^3/1,000	0.164***	0.0400	0.0903***	0.0532*
	(0.0163)	(0.0246)	(0.0237)	(0.0297)
Experience^4/100,000	-0.152***	-0.0171	-0.0859***	-0.0493
	(0.0195)	(0.0299)	(0.0316)	(0.0406)
Migration background	-0.0644***	-0.169***	-0.128***	-0.135***
	(0.0126)	(0.0154)	(0.0135)	(0.0121)
Correction term work	-0.180***	-0.154***	0.179***	0.0392*
	(0.0218)	(0.0341)	(0.0195)	(0.0214)
Correction term education	-0.299***	-0.225***	-0.0780***	-0.0302*
	(0.0376)	(0.0676)	(0.0120)	(0.0159)
UAS	-0.126***		-0.177***	
	(0.00670)		(0.00854)	
Constant	2.842***	2.466***	2.501***	2.228***
	(0.0503)	(0.0852)	(0.0341)	(0.0391)
State dummies	yes	yes	yes	yes
Industry dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	15356	6378	11348	7297

Notes: This table displays the results of linear regressions with a selection-into-work and a selection-into-education correction using the log hourly gross wage as the dependent variable. Each column represents a separate regression. *HE* (*VOC*) implies that the estimation sample is based on individuals with a higher education (vocational training) degree. *UAS*=University of applied sciences degree.
Source: SOEP, own calculations.

Table 2.A6: The tax function

Gross income/10^5	1.007***
	(0.00921)
Gross income^2/10^{10}	-1.160***
	(0.0161)
Gross income^3/10^{15}	0.506***
	(0.00900)
Gross income^4/10^{21}	-0.873***
	(0.0185)
Gross income^5/10^{27}	0.500***
	(0.0120)
Married	-0.00877***
	(0.00123)
Number of children	-0.0112***
	(0.000514)
Constant	0.124***
	(0.00159)
N	47570

Notes: This table displays the results of linear regressions using the average tax rate as the dependent variable.
Source: SOEP, own calculations.

Table 2.A7: Simulated life-time earnings (in thousand Euros), men

	No selection	Only work	Educ + work
HE, gross	1,427	1,366	1,181
HE, net	876	842	743
VOC, gross	1,249	1,208	1,086
VOC, net	785	764	700

Notes: This table displays the simulated mean life-time earnings of men for higher education and vocational training in prices of 2016. The different columns represent different underlying wage specifications. *No selection*= No selection correction. *Only work*=Selection-into-work correction. *Educ+work*=Corrections for both selection into work and selection into education. *HE*=Master degree, *VOC*=Vocational training degree, *Gross (net)* refers to gross (net) labor earnings.
Source: NEPS, SOEP, own calculations.

Table 2.A8: Simulated life-time earnings (in thousand Euros), women

	No selection	Only work	Educ + work
HE, gross	923	901	907
HE, net	610	598	602
VOC, gross	725	722	726
VOC, net	510	508	511

Notes: This table displays the simulated mean life-time earnings of women for higher education and vocational training in prices of 2016. The different columns represent different underlying wage specifications. *No selection*= No selection correction. *Only work*=Selection-into-work correction. *Educ+work*=Corrections for both selection into work and selection into education. *HE*=Master degree, *VOC*=Vocational training degree, *Gross (net)* refers to gross (net) labor earnings.
Source: NEPS, SOEP, own calculations.

Chapter 3

Higher Education Funding in Germany - A Distributional Lifetime Perspective

3.1 Introduction

Who should pay for higher education has been a controversial topic for decades. Some industrialized countries, such as as England, Australia, and the US commit students to pay a substantial share of the cost of tuition themselves. Countries such as France or Germany, in contrast, offer free or highly subsidized tuition and finance higher education mostly through taxes. In addition, countries also differ with respect to the presence of additional instruments of higher education funding, such as student loans and grants, and subsidized health insurance.

One main reason why the financing of higher education is a fundamentally controversial topic are the distributional concerns that go along with it. The literature which analyzes the distributional effects of higher education funding can be classified into “cross-sectional" and “longitudinal" studies. Cross-sectional studies analyze how higher education funding redistributes between households along the current income distribution. Naturally, households with students receive benefits from subsidized tuition while those without students do not. A substantial part of this literature analyzes the distributional impact of subsidized tuition by computing the benefits each household receives (see, for instance, Callan et al., 2008; and Koutsampelas and Tsakloglou, 2015). In contrast, some studies go further and compute net transfers by contrasting the received benefits with the taxes each households (implicitly) pays for higher education (Barbaro, 2002).[1]

While the cross-sectional perspective might deliver interesting insights into how higher education funding redistributes between households of different socio-economic status, it does not take into account the heterogeneous life courses of students and non-students. This is exemplified by the fact that in a cross-sectional perspective a substantial fraction of students who live outside their parental households are at the lower end of the income distribution. While potentially being classified as "poor" given their current living standard, students usually have a higher expected lifetime income than non-students. Analyses with a longitudinal perspective, in contrast, address the lifetime perspective of higher education funding. These studies focus on the current young cohort itself and analyze the distributional effect of higher education funding within this cohort (and between academics and non-academics, for instance).

Analogously to Barbaro (2002) in the cross-sectional perspective, Borgloh et al. (2008) compare the benefits young adults receive from higher education funding to the taxes they (implicitly) pay for higher education over the life cycle. Dearden et al. (2008) analyze the distributional impact of a reform of the British higher education funding system. They focus on the subpopulation of students and analyze how the reforms redistributed transfers between students of high and low lifetime income and of different parental income. Finally, studies such as Harding (1995) and Chapman and Sinning (2014) indirectly focus on the distributional effects of higher education funding in that they simulate individual tuition fee repayments over the life cycle.

However, these studies are the only recent studies with such a longitudinal focus I am aware of. Most likely, the reason for this scarcity are the data requirements: If researchers want to analyze the distributional effects of higher education funding within a young cohort, then the latter's life cycle needs to be forecasted. Borgloh et al. (2008), for instance, generate such life cycles by averaging individual earnings based on gender, age, and field of study. Harding (1995), in contrast, uses a dynamic microsimulation model developed for Australia (Harding, 1991).

In this chapter, I apply a longitudinal perspective to analyze the distributional impact of higher education funding for a young cohort in Germany. To generate artificial life cycles, I use the dynamic microsimulation model developed in Fischer and Hügle (2020) and modify it for the purpose of this study. Dynamic Microsimulation has the advantage that it accounts for path dependencies in employment and family

[1] The assumption is that a fraction of an individual's tax payment is used to finance higher education. This fraction is assumed to be the same as the fraction higher education funding has on the total public expenditures.

formation (such as marriage and fertility) and for cohort and time effects when forecasting life cycles. In addition, it provides an entire distribution of life cycles for each set of initial values.

Importantly, my analysis distinguishes between different fields of study. The reason is that both tuition cost and expected lifetime income vary considerably across these fields. Borgloh et al. (2008) is the only other study that distinguishes fields of study I am aware of. Due to data limitation, however, they had to impute the fields of study from the current occupation of an individual. Instead of this indirect measure, I use more recent data in which the field of study was directly surveyed.

The distributional analysis has three parts: In the first part, I analyze the current system of higher education funding and compare the quantitative importance of different funding instruments such as free tuition, student loans and grants, and subsidized health insurance. Then, I compare these benefits across fields of study. In the second part, I add the lifetime perspective to the analysis and show how different deciles of the lifetime income distribution financially benefit from higher education funding. Here, I focus on the individuals with a higher education entrance degree in Germany, i.e. those individuals who are usually confronted with the choice between higher education and vocational training. In the last part, I show the potential distributional consequences of an alternative system with tuition fees and income-contingent loans. A system with tuition fees and income-contingent loans implies that students receive loans to cover tuition fees and living cost while they are studying, but have to pay back (part of) their tuition cost over the life course if their income exceeds a certain threshold. Such systems have been in place in various Western countries, such as England, Australia, and New Zealand (Britton et al., 2019) and have also been discussed in Germany (see Lergetporer and Woessmann, 2019; and Chapman and Sinning, 2014).

I find that, while students benefit from multiple higher education funding instruments, free tuition is, by far, the quantitatively most important one. Due to the heterogeneity in the cost of tuition across fields of study, however, there is a large range in how much students benefit from free tuition. For instance, while students of medicine receive free tuition worth of about 160,000 Euros over the lifetime, students of humanities and the social sciences only receive about 20,000 Euros.

Focusing on the young adults with a higher education entrance degree, I find that the share of academics continuously increases across lifetime income deciles and hence higher deciles benefit more from higher education funding than lower deciles. At the same time, there is a large heterogeneity within the groups of academics and between fields of study. For instance, those with the highest expected lifetime incomes, students from the fields of medicine and math/natural sciences, are the ones who benefit most from higher education funding as their fields are particularly costly.

Finally, I show that the introduction of different tuition fee regimes with income-contingent loans would have little distributional consequences relative to net lifetime income. Consequentially, I estimate that the hypothetical tuition fee reforms analyzed in this study would barely distort the educational decisions of young adults.

The chapter proceeds as follows. Section 3.2 explains the data sets used in the analysis. Section 3.3 describes the fields of study and the distributions of students over these fields. Section 3.4 explains the different instruments of higher education funding. Section 3.5 describes the approach of dynamic microsimulation. Results are presented in section 3.6 and section 3.7 concludes.

3.2 Data

In this study, I mainly rely on two data sets, the German Socio-Economic Panel (SOEP) and the National Educational Panel Study (NEPS). The SOEP is an annual, nationally representative longitudinal study of private households across Germany and currently surveys about 30,000 people in 15,000 households (Goebel et al., 2018). I use the SOEP primarily to estimate the dynamic microsimulation model (see Section 3.5), but partially also to compute the value of different instruments of higher education funding (see Section 3.4).

The NEPS (Blossfeld and Von Maurice, 2011) is a longitudinal study that surveys individual educational trajectories, decisions, and competences. It has different starting cohorts, defined by the point in time the NEPS starts observing the cohort, from newborns to adults, and follows each of these cohorts along the life span. I draw on three of these starting cohorts (SC): SC4 (9th graders), SC5 (students), and SC6 (adults). Using these data sets, I estimate the shares of students in different educational paths (see Section 3.3), the values of different instruments of higher education funding (see Section 3.4), and the educational choice model to assess likely behavioral responses of tuition fee reforms (see Section 3.6).

3.3 Higher education in Germany

I follow Borgloh et al. (2008) and the Federal Statistical Office (Statistisches Bundesamt, 2020) and cluster all individual study programs (*Studiengänge*) into "fields of study" (*Studienbereiche*).[2] Table 3.1 lists these fields of study and gives examples of specific individual study programs that fall into each field.[3]

Table 3.1: Fields of study

Field of study	Examples of individual programs
Engineering	Mechanical eng., Electrical eng., architecture, computer sc.
Humanities	Philosophy, history, linguistics
Math and natural sciences	Mathematics, biology, chemistry, physics, geography
Medicine	Human medicine, dentistry, veterinary medicine
Social sciences	Economics, law, political science, sociology, psychology
Other	Agricultural sc., science of forestry, nutritional sc., sports, arts

Notes: This table displays the list of fields of study and examples of corresponding individual study programs. Note that for convenciency, I decided to use a shorter name for "social sciences" (*Rechts-, Wirtschafts-, und Sozialwissenschaften*) rather than translating the longer name the Federal Statistical Office uses.

Figure 3.1 describes the distribution of first-year students in 2018 over these fields. While almost one half of all women (45%) are enrolled in the social sciences, only one third of men choose this field. In contrast, 42% of all men are enrolled in engineering, while for women this share is only about 13%.

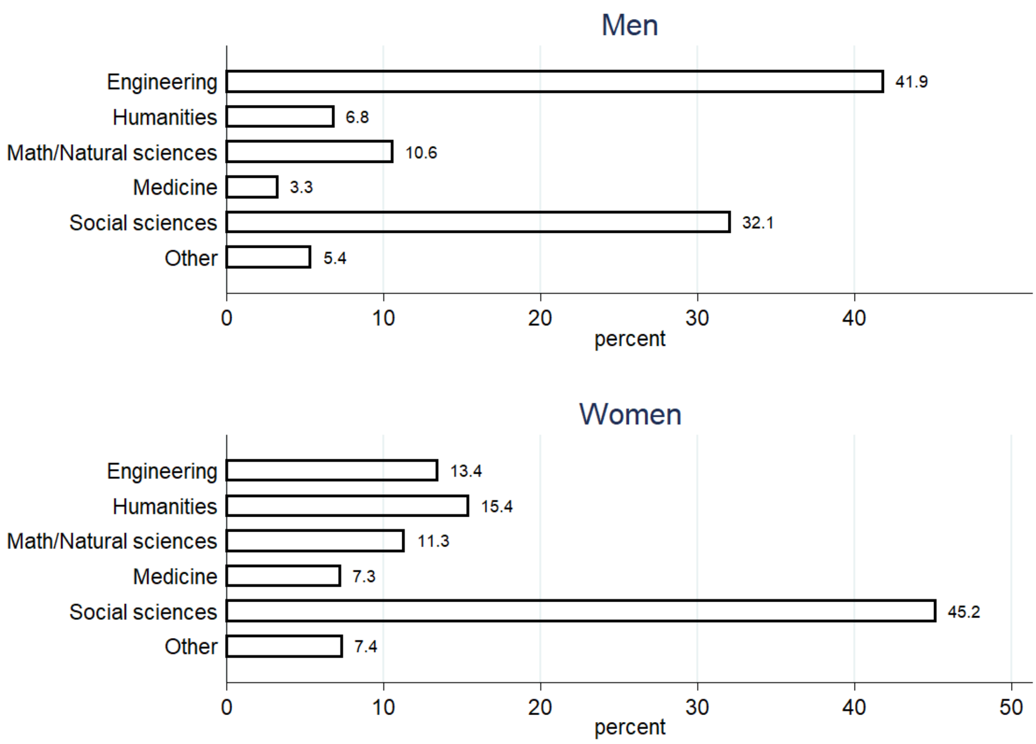
Depending on the field of study, students can earn different degrees. As a consequence of the mid-2000 Bologna reforms, German higher education institutions now largely award bachelor and master degrees. In addition, some study programs still finish with state exams (*Staatsexamen*), such as medicine, law, and teacher training. As among the latter only medicine represents a field of study by itself, I assume for the analysis that studying the field of medicine leads to a state exam while in all other fields either a bachelor or a master degree is obtained. For these other fields, I estimate on the basis of the NEPS student cohort that the field-specific share of students who leave the education system with a bachelor degree ranges from 11% (social sciences) to 21% (math/natural sciences).

Apart from the earned degrees, there are two other dimensions on which heterogeneity between fields might be considered. The first is length of study. I therefore estimate the average study length until graduation for each field of study using the NEPS student cohort. I find that for all fields of study presented above, the average study length to obtain a master degree or a state exam is approximately 11 semesters. I therefore do not consider potential heterogeneity in study length between fields of

[2]Note that I add sports and arts to the residual category of "other", while Borgloh et al. (2008) exclude these individual study programs.

[3]This level of aggregation was chosen as the SOEP would contain only a small number of observations for individual study programs and the Federal Statistical Office (Statistisches Bundesamt, 2020) only lists the tuition cost on this level of aggregation.

Figure 3.1: Distribution of academics over fields of study



Notes: The graphs display the distribution of first-year students in 2018 over fields of study by gender.
Source: Statistisches Bundesamt (2019d), own calculations.

study.[4]

A second dimension of potential heterogeneity is the risk of "dropout". It is a well-established fact that roughly about one third of bachelor students do not finish their studies. However, it is plausible to assume that a large fraction of these "dropouts" are just changing the study program, often within the same field of study, as defined above. In the latter case, "dropping out" just implies prolonging the length of study, while the student obtains a degree in the same field of study.[5] Yet, the frequency of these paths is difficult to assess with current data. Therefore, I ignore the risk of dropout.[6]

[4]For simplicity, I also ignore potential heterogeneity in study lengths within fields of study.

[5]An example would be a student who first enrols in mechanical engineering, then drops out and enrols afterwards in electrical engineering (i.e. another individual program within the same field of study.)

[6]However, using the relatively few field-specific dropout observations in the NEPS, I estimate that in all fields of study the share of dropouts who do not re-enter another study program is below 10%.

3.4 The instruments of higher education funding

In this section, I explain the different instruments of higher education funding in more detail. Using SOEP and NEPS data, I then estimate the monetary values of each instrument conditional on the field of study.

3.4.1 Free tuition

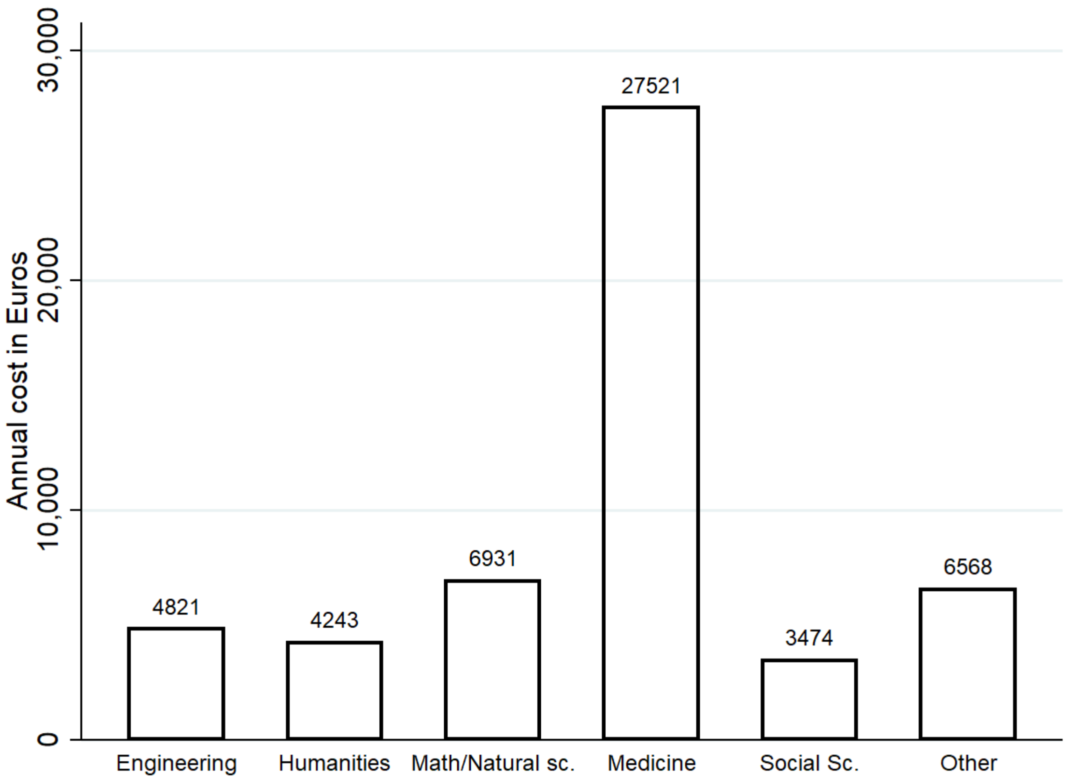
Currently, no German state collects tuition fees for the first degree, i.e. for the first bachelor and a consecutive master degree or a state exam.[7] Hence, the state incurs a sizable cost for each student. Figure 3.2 shows the cost of tuition per student and year. These costs can be calculated by using the costs for tuition and research reported by the Federal Statistical Office (Statistisches Bundesamt, 2018) and the so-called "coefficients for research and tuition" (*Forschungs- und Entwicklungskoeffizienten*) to subtract the cost associated with research activities. In terms of tuition cost, there is substantial heterogeneity across fields of study. With an annual cost of about 27,500 Euros, higher education places in medicine are, by far, the most expensive ones. In contrast, places in the social sciences cost only about 3,500 Euros per student and year.

3.4.2 Student loans and grants (BAföG)

Apart from free tuition, a key feature of higher education funding in Germany is the provision of grants and loans through the *Bundesausbildungsförderungsgesetz* (henceforth *BAföG*). The amount of BAföG support a student has access to during the standard period of study depends on own and parental income and wealth and on whether the students lives in the parental household. Currently, the maximum monthly BAföG payment is 744 Euros (excluding supplements for insurance). Half of the BAföG payment is usually given as a grant, while the other half is provided as an interest-free loan. In 2016, 13.9% of all students received BAföG. Using the NEPS, I estimate the annual average BAföG payments given as a grant by field of study. I find that that there is only little heterogeneity between fields, with all fields of study having an annual average between 630 (engineering) and 775 Euros (other).

[7] A few states collect fees in other cases, for instance for second degrees or for long-term students. I will ignore this here.

Figure 3.2: Tuition cost by field of study



Notes: Costs are in prices of 2019. The costs can be calculated by using the costs for tuition and research reported by the Federal Statistical Office (Statistisches Bundesamt, 2018) and using the "coefficients for research and tuition" (*Forschungs- und Entwicklungskoeffizienten*) to subtract the cost associated with research activities.
Source: Statistisches Bundesamt (2018), own calculations.

3.4.3 Child benefits

While their children are still in educational training and not older than 25, parents are entitled to child benefits. Currently, parents receive monthly 204 Euros for each of the first two children and slightly more for each additional child. As vocational trainees usually finish their education at a younger age than students, this implies that families of students receive child benefits for a longer time period. Assuming that parents transfer the received child benefits to their children, I assign each student an annual child benefit of 2,448 Euros (204 Euros/month $\cdot$ 12 months).

3.4.4 Health insurance

Students have different types of health insurances. Until the age of 25, they are usually covered by the public health insurance of their parents. If this is not the case, one alternative is the student health insurance with a reduced insurance contribution of currently 109 Euros per month. In addition, students might be compulsorily insured in the public health insurance, for instance, if they work more than 20 hours per week. Finally, some students are also privately insured.

I assume that only students who are either insured through their parents or who have a student health insurance are subsidized. To compute this benefit, I further assume that if students were not subsidized, they would be treated as individuals with low income who are part of the public health insurance. For the latter, the health care contribution rate of 14,6% is multiplied with the minimum assessment threshold, which is currently 1,062 Euros, equaling a contribution of 155 Euros. Hence, the benefit for individuals in the family insurance is 155 Euros, while the benefit of individuals in the student insurance is 46 Euros (155-109). Using the SOEP, I estimate the share of students in each type of health insurance and compute the weighted average of the health insurance subsidy. I find that, on average, students are subsidized by 1,228 Euros annually.

3.4.5 Education tax allowance

Another subsidy is the education tax allowance (*Ausbildungsfreibetrag*) which is currently 924 Euros per year. To receive this allowance, the student must live outside the parental household and be at most 25 years old. For simplicity, and as the subsidy implied by the allowance is relatively small compared to other instruments, I assume that the parents of all students receive the allowance.

To estimate the average benefit of receiving the allowance, I use the SOEP and the current income tax formula and estimate the tax burden of adults in the age range 40-65 who live in the same household as students under two scenarios: with and without the allowance. I assume income tax splitting of couples and no further allowances. The difference in the tax burden is then the financial benefit due to the education tax allowance. I find that, on average, the allowance implies an annual benefit of 201 Euros per student.

3.4.6 Further instruments

Further instruments of higher education funding include grants to organizations for the promotion of young talent (*Begabtenförderwerke*), the Germany scholarship (*Deutschlandstipendium*), funding of educational competitions, and student exchange programs.

I sum up the amounts spent for these instruments and assign each student the average benefit of 235 Euros per year. I do not consider student housing promotion as there is no recent publicly available data on it.

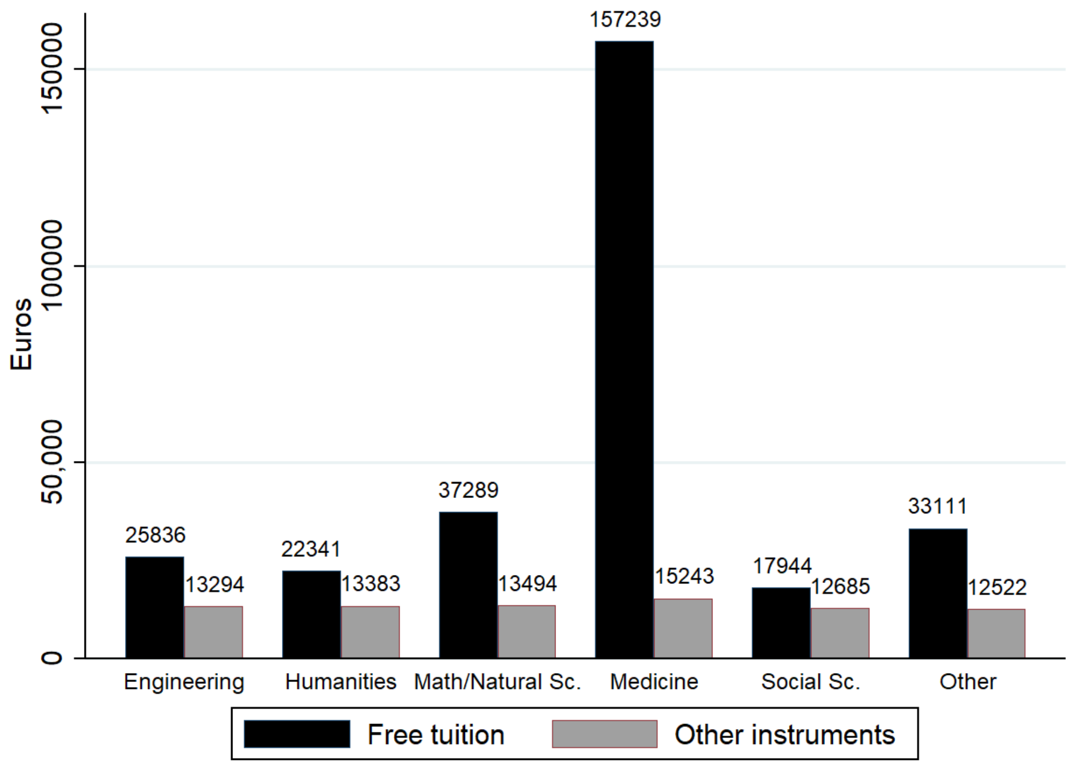
Another potential instrument that might be taken into account are credit points for higher education in the calculation of civil service pensions (Borgloh et al., 2008). The idea is that for those academics who enter civil service, the years of studying can be counted as credit points. Importantly, however, a civil servant reaches the maximum rate of pension with 40 years of service. In the dynamic microsimulation model (explained in the next section), however, I assume that all academics enter the labor market at age 24 and retire at age 67. Hence, the academics who are simulated to enter civil service cannot count their study time as credit points.

3.4.7 The present value of higher education funding instruments

Finally, Figure 3.3 shows the present value of lifetime benefits by field of study, assuming a discount rate of 2% and a duration of academic training of six years. As vocational trainees also receive all funding instruments except free tuition, I follow Borgloh et al. (2008) and only consider these instruments for the last three years of academic training (assuming that vocational trainees have finished their training after three years). Free tuition is the quantitatively most important funding instrument. For medicine, its value accumulates to more than 150,000 Euros over the lifetime. For other fields, it is substantially smaller, ranging from less than 20,000 Euros (social sciences) to about 37,000 Euros (math/natural sciences). In contrast, the value of all other instruments considered jointly only amounts to about 13,000 Euros over the lifetime.

The goal of the remainder of this chapter is to analyze who receives these benefits in terms of expected lifetime income quantiles. Thereby, I will focus on the individuals with a higher education entrance degree who face the choice between taking up academic or vocational training. The next section explains how their life cycles are simulated.

Figure 3.3: Lifetime benefits by field



Notes: The graph shows the present values of the higher education funding instruments by field of study in prices of 2019. “Other instruments" comprises the instruments student loans and grants (BAföG), child benefits, health insurance, education tax allowance and further instruments, as explained above. The instruments are explained in more detail in Section 3.4. The discount rate is 2%.
Source: SOEP, NEPS, own calculations.

3.5 Dynamic microsimulation

3.5.1 Simulating the lives of a young cohort

A key challenge to assess the distributional effects of higher education funding in longitudinal studies is the generation of life cycles. For this purpose, I use the dynamic microsimulation model outlined in Fischer and Hügle (2020). In contrast to short-cut approaches which rely on cross-sectional data, a dynamic microsimulation model has the advantage of disentangling time and cohort effects when projecting life cycles from past observations. In addition, such models are able capture path dependencies and can thereby model heterogeneous life cycles.

The model simulates life cycles for the 1980s cohort in terms of wages, employment, and household formation (fertility, marriage, and divorce) from age 18 to 67 (the official retirement age in Germany). While there is an obvious mechanic link between wages, employment, and lifetime earnings, the processes of fertility, marriage, and divorce impact employment decisions of individuals and hence, ultimately, also lifetime earnings. Simulating the life courses of the 1980s cohort serves two purposes: First, important life cycle events such as labor market entry, births, and marriages can already be observed for this cohort. And second, the 1980s cohort is young enough to argue that its life courses are a plausible benchmark for a younger cohort which is about to enter higher education or vocational training.

In total, I simulate the life course of 500,000 individuals.[8] As in Fischer and Hügle (2020), each individual is assigned one of four general post-secondary educational categories: No post-secondary degree, a vocational degree without a higher education entrance degree, a vocational degree with a higher education entrance degree, and a higher education degree. While the analysis focuses on the individuals with a higher education entrance degree, individuals of the other two educational categories serve as potential spouses in the simulation.

The distribution of post-secondary education levels in the simulation matches the one reported in the Mikrozensus for the highest educational degrees of the 1983–88 birth cohort.[9] Figure 3.4 describes this educational distribution. The individuals with higher education are then further assigned a field of study and either a bachelor degree, a master degree (or a state exam in case of being assigned the field of medicine).[10] The distribution of fields of study within the group of academics follows the one described in Figure 3.1. It is assumed that students study for six years if they leave the education system with a master degree or a state exam and three years if they only obtain a bachelor degree. Vocational training also has a duration of three years.

Hence, there are two levels of aggregation: A more aggregated level of post-secondary education levels which, in particular, distinguishes vocational degrees and higher education, and a more disaggregated level within higher education, which distinguishes field of studies and bachelor/master degrees. For the modeling of transitions in employment and household formation (fertility, marriage, and divorce), I use only the more aggregated level. This implies that I assume that academics follow the same expected patterns, independently of their field of study or whether they only obtained a bachelor degree.[11] For the modeling of wages, however, I differentiate further into

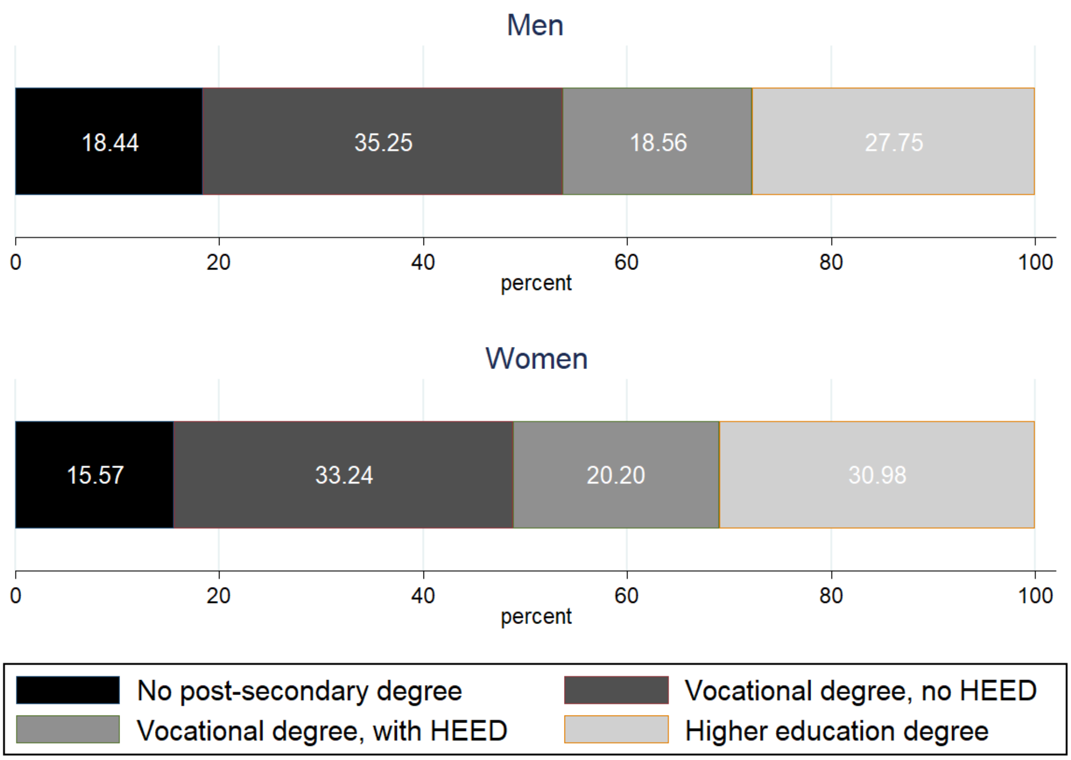
[8] Due to computational reasons, I simulate 100 runs with 5,000 individuals and pool them for the distributional analysis.

[9] The empirical distribution on which we base our simulation comes from German *Mikrozensus* data (see Statistisches Bundesamt, 2018, for details).

[10] From now on, I will also refer to state exams by the term *master degree*.

[11] With the only difference being that bachelor graduates who do not continue to a master program

Figure 3.4: Distribution of post-secondary education levels



Notes: The bars show the distribution of post-secondary education levels in the cohort 1983–1988 by gender. “HEED"= Higher education entrance degree.
Source: Statistisches Bundesamt (2018), own calculations.

bachelor and master degrees, and different fields of study. This means that I allow for different wage profiles between academics depending on their field of study and bachelor/master degrees. The dynamic microsimulation model proceeds in three stages: Parameter estimation, life-cycle simulation, and tax-transfer simulation. In the following, each stage is described in more detail (for additional information see Fischer and Hügle, 2020).

Parameter estimation

The objective of this part of the dynamic microsimulation model is to estimate parameters that can then be used to simulate life cycles. More precisely, I estimate transition probabilities for household formation (marriage, divorce, and fertility) and employment categories, hourly wage regressions, and aggregate cohort-specific targets for household formation and employment.

enter the labor market earlier.

The transition probabilities for household formation and employment are estimated by discrete-choice models and include dummies for the post-secondary education levels or indicators for being in academic and vocational training, indicators for migration background, age polynomials and different sets of variables that capture past employment biographies and past life-cycle events such as births, marriages, and divorces. All models are estimated separately by gender.

For marriage, I use two processes: First, the marriage probability is modeled estimating a binary logit model for all unmarried individuals. Second, in order to account for educational assortative mating, the individuals which were simulated to marry in the first process are matched based on an empirically observed matrix of marriage frequencies across the four post-secondary education levels. Divorce is estimated using a binary choice model on the household level. Finally, the probability of giving birth is estimated separately for married and unmarried women.

Employment is modeled as a three-step process: First, I use a binary logit model to estimate the probability of individual labor force participation. Second, I estimate a binary choice model for involuntary unemployment conditional on labor force participation. Third, multinomial logit models are used to estimate the probability of choosing specific working hours categories. For women, I distinguish five hours categories: Marginal employment (0-14 weekly working hours), reduced part-time work (15-24 hours), extended part-time (25-34 hours), full-time (35-42 hours), and overtime work (more than 42 hours). For men, I model three employment hours categories: Part-time (0-34 hours), full-time (35-42 hours), and over-time work (more than 42 hours).[12]

In the wage estimations, I regress individual log gross hourly wages on fourth-order polynomials of age, experience, and tenure, an indicator for migration background, orthogonalized indicators for federal states and year dummies separately by gender and post-secondary education level. Compared to the specification used in Fischer and Hügle (2020), I modify the regression for higher education graduates and include field of study and bachelor degree dummies. This ensures that wage differences across fields of study and a potential wage penalty of bachelor compared to master degrees are captured. Table 3.A1 in the Appendix presents the coefficient estimates. For the group of academics, I find a bachelor penalty of about 16% for women and 13.5% for men. As to the fields of study, I find large wage premiums of approximately 36-42% for studying medicine relative to study programs in the residual category “other"

[12] In addition, I account for the particularities of labor market entry by modeling separate transitions for each of the first five years after graduation.

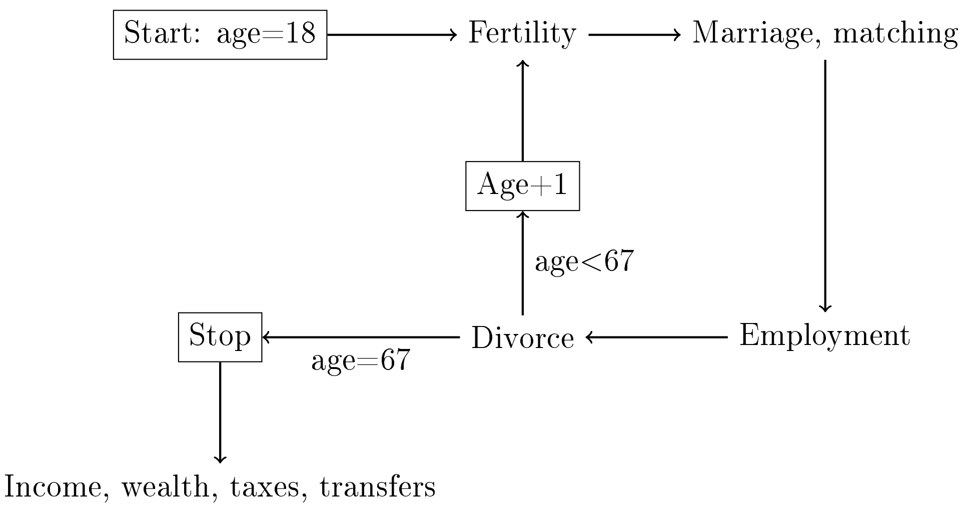
(the base category). Studying humanities, in contrast, only results in a 2-7% premium relative to this base category.

Finally, I estimate a set of aggregate cohort-specific targets for employment and household formation. This so-called *alignment* ensures that in the aggregate, the simulated transitions meet some projected trend. To estimate these targets I use binary and multinomial fractional regression models. All models include polynomials of age and generational trends (generational trend is defined as *birth year* - 1930) in order to capture cohort effects, and control for the overall unemployment rate. Using the coefficient estimates, I predict age-specific target rates in employment and household formation for the 1980s cohort over the life cycle.[13]

Life-cycle simulation

In the simulation stage, I use the estimated parameters and the projected aggregate targets from the estimation stage and simulate the cohort's life cycle. Starting at age 18, the individual age is updated year-by-year and the transitions in employment and household formation are simulated as described in Figure 3.5. More precisely, the

Figure 3.5: The simulation stage



procedure to select individuals for transitions (fertility, marriage, employment, and divorce) works as follows:

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

[13] The results of the transition models and the target estimation for employment and household formation are the same as presented in Fischer and Hügle (2020).

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.

After having simulated the transitions in employment and household formation, I can simulate the hourly wages. In order to align the variance of simulated wages to the variance of observed wages, individuals are assigned log wage residuals conditional on gender and education, which are added to the predicted log wages. Exponentiating this sum then results in the the predicted hourly wages.[14] Finally, to obtain labor earnings hourly wages are multiplied with the level of working hours given the simulated employment category.

Tax-and-transfer simulation

In order to translate the simulated gross labor earnings into disposable incomes, I use the modified version of the STSM as described in Fischer and Hügle (2020). The STSM is a module that describes the main features of the German tax-transfer regime. Using the rules of the tax-transfer regime as of 2019 and given the simulated life cycles, I compute taxes, transfers, and social security contributions. To compute taxable income, I take into account the simulated incomes from dependent employment and self-employment. In addition, I assume that individuals accumulate savings according to age-specific savings rates as estimated by Brenke and Pfannkuche (2018). Furthermore, I assume that married individuals file for joint taxation and that couples split taxes according to tax class IV/IV and the so-called factor method (*Faktorverfahren*). As to social security contributions, I consider contributions to health, long-term care, and unemployment insurance but not to the pension system.[15] Finally, I simulate unemployment benefits, parental leave allowances, social assistance, housing benefits, child benefits, and additional child benefits at the household level.

[14]Hence, the hourly gross wage of individual i is calculated as $\hat{w}_i = exp(x_i'\hat{\beta})exp(\hat{u}_i) = exp(x_i'\hat{\beta}+\hat{u}_i)$ where x is a vector of covariates, $\hat{\beta}$ is a vector of coefficient estimates, and $\hat{u}_i$ is a randomly drawn residual from the log wage regression conditional on gender and education.

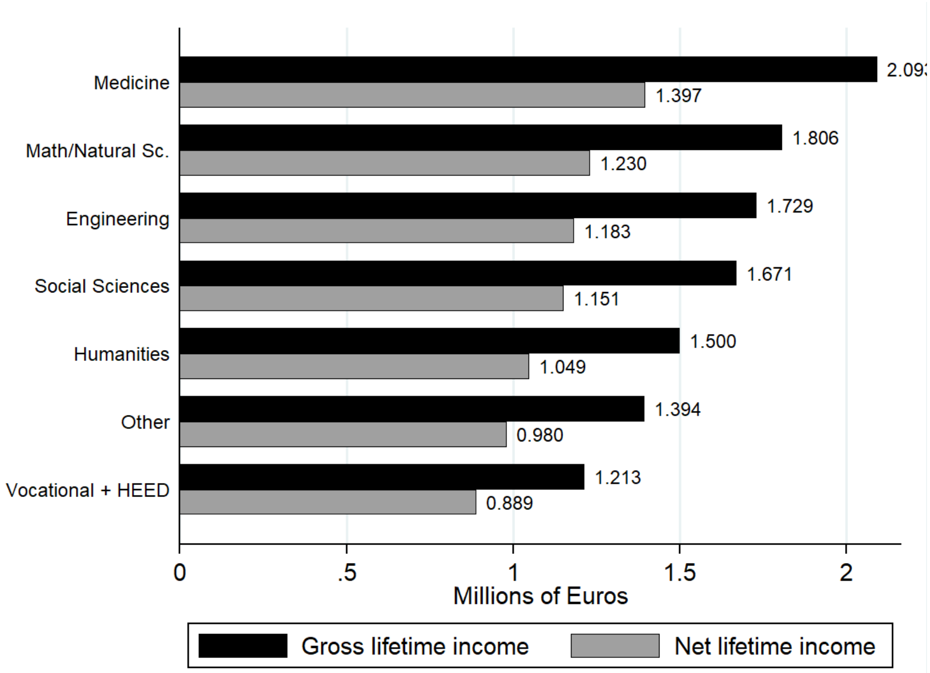
[15]Modeling the details of the pension system and pension-related transitions is beyond the scope of this study. However, pension entitlements are generally equivalent to the contributions paid.

3.6 Results

3.6.1 Simulated lifetime incomes

Figures 3.6 and 3.7 plot the average simulated individual gross and net lifetime incomes by field of study (including vocational degrees with higher education entrance degree).[16] The differences in lifetime incomes between fields of study essentially result from the differences in simulated hourly wages.[17] The rankings of fields in terms of lifetime earnings are similar for men and women. For both men and women, the average lifetime incomes of medicine graduates exceed those of other disciplines by far.

Figure 3.6: Average simulated gross and net lifetime incomes by field of study, men



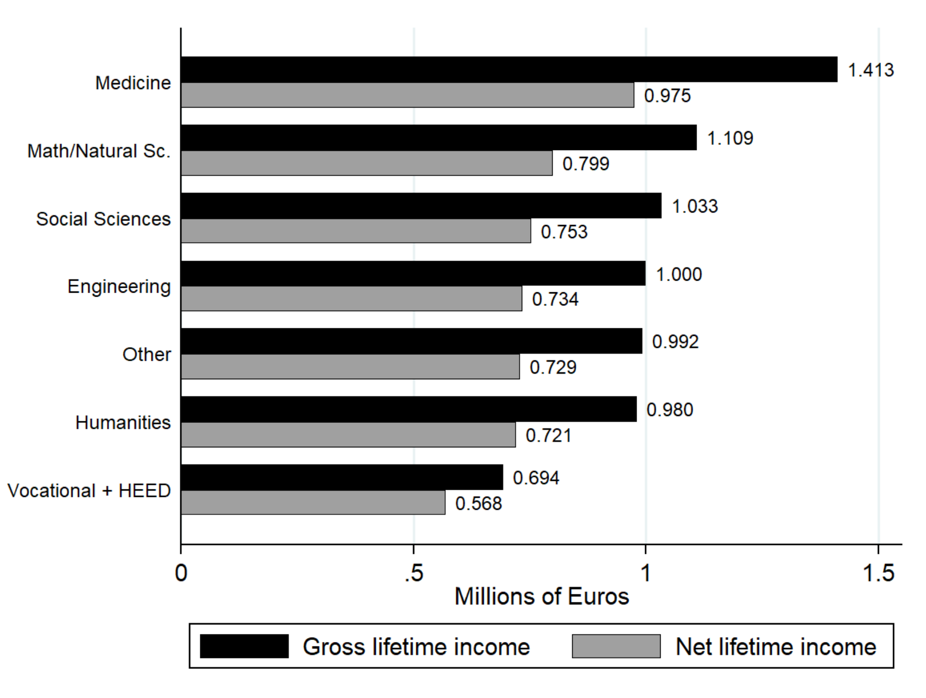
Notes: The figure depicts simulated individual gross and net lifetime incomes by field of study and education level in prices of 2019 for men. "Gross lifetime income" refers to the present value of annual gross income, discounted at 2%. Similarly, "net lifetime income" refers to the present value of annual net income, accounting for taxes, transfers, and social security contributions, discounted at 2%. I assume that individuals do not pool their incomes with their partners but that married couples file for joint taxation. "HEED"= higher education entrance degree.
Source: Own simulations.

Male medicine graduates have, on average, discounted gross earnings of more

[16]Importantly, I assume that married individuals file for joint taxation (as explained in Section 3.5.1) but do not pool their individual net incomes.

[17]Figures 3.A1-3.A4 in the Appendix depict the simulated average hourly wages and annual earnings over the life cycle by field of study and gender.

Figure 3.7: Average simulated gross and net lifetime incomes by field of study, women



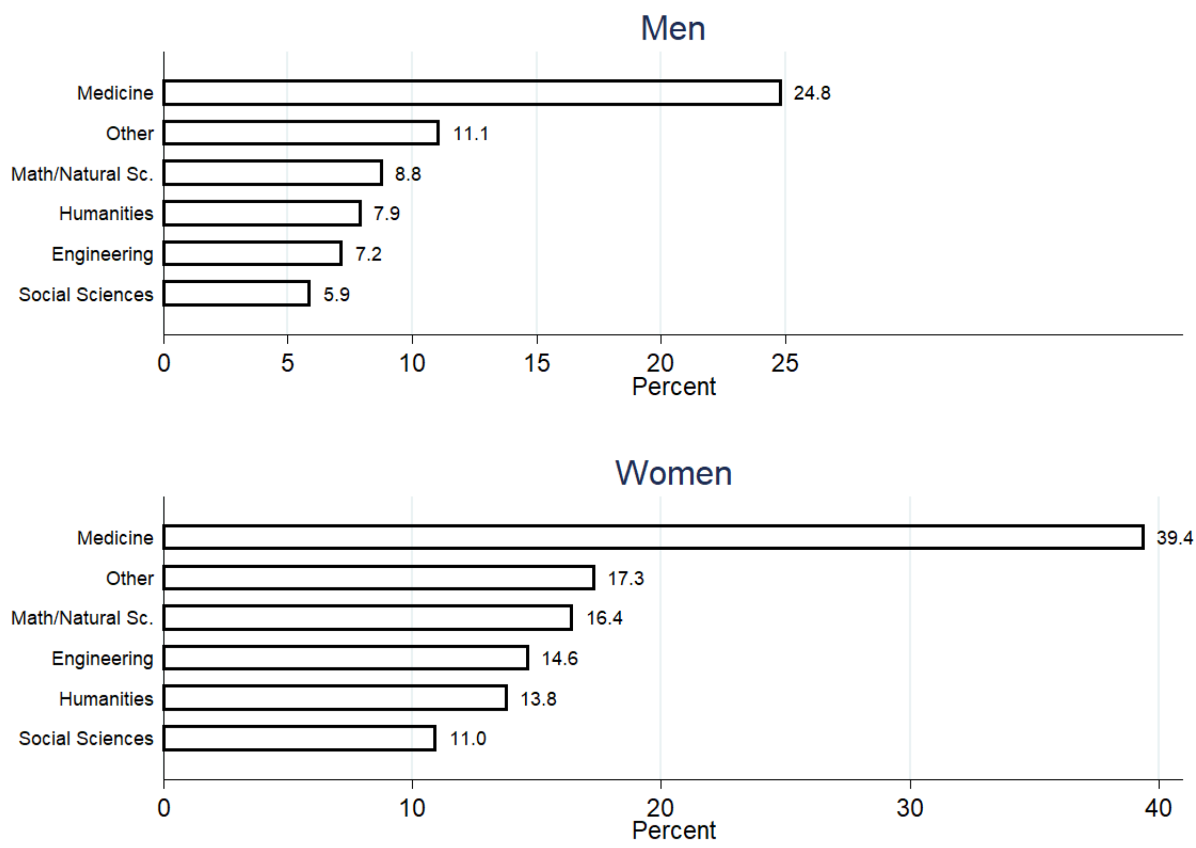
Notes: The figure depicts simulated individual gross and net lifetime incomes by field of study and education level in prices of 2019 for women. "Gross lifetime income" refers to the present value of annual gross income, discounted at 2%. Similarly, "net lifetime income" refers to the present value of annual net income, accounting for taxes, transfers, and social security contributions, discounted at 2%. I assume that individuals do not pool their incomes with their partners but that married couples file for joint taxation. "HEED"= higher education entrance degree.
Source: Own simulations.

than 2 million Euros over the lifetime, while female medicine graduates earn about 1.4 million. The tax-and-transfer system reduces these earnings to about 1.4 million Euros for men and 1 million for women. Individuals with degrees in the field of math and natural sciences are second in this ranking with male graduates earning about 1.8 million and female graduates accumulating 1.1 million. While for men, engineers earn substantially more than men in the remaining fields, for women there is no clear ranking. Finally, on average individuals with vocational degrees earn less than graduates of all fields of study with men earning about 1.2 million and women approximately 0.7 million.[18][19]

[18]The gross lifetime income gap between men and women could be decomposed into two factors: The differences in hourly wages (as shown in Figures 3.A1 and 3.A2 in the Appendix) and differences in the hours worked over the life cycle. While men with a vocational or academic degree are simulated to work around 80,000 hours over the life cycle, women with either a vocational or an academic degree are simulated to work around 60,000 hours.

[19]For the sake of clarity, I do not show here the profiles of the two other educational groups: No post-secondary degree and vocational degree without higher education entrance degree. The profiles of these groups are below the one for vocational training with higher education entrance degree.

Figure 3.8: Higher education funding relative to taxes and contributions paid over the life cycle



Notes: The graphs display the amount of higher education funding received relative to the taxes and social security contributions paid (minus transfers) by fields of study and gender.
Source: Own calculations.

Finally, one can set the net financial burden of taxes and contributions paid minus transfers received over the life cycle into perspective by comparing them with the lifetime subsidies received from higher education funding (as analyzed in Section 3.4). Figure 3.8 shows the amount of higher education funding relative to the taxes and contributions paid (minus transfers received) over the life cycle. Both male and female medicine graduates benefit the most from higher education funding, even relative to the taxes and contributions they pay over the life cycle. Graduates of programs from the residual “other" category are second in this ranking which is due to their relatively costly tuition (compared to the remaining fields) and to their relatively low amount of taxes paid. Beyond the heterogeneity across fields, the graph also shows that women benefit much more from higher education funding when compared to the taxes and contributions they pay, even though the absolute amount of funding (conditional on field of study) is assumed to be the same for men and women.[20]

3.6.2 The current system: Benefits by lifetime income

Having estimated gross and net lifetime incomes for all individuals in the cohort, we can now analyze how individuals of different lifetime income quantiles benefit from the higher education funding instruments. Importantly, I condition on a higher education entrance degree in this analysis. Figure 3.9 plots the value of the funding instruments across net lifetime income deciles.

As expected, the value of all instruments considered jointly increases with decile. While for men (women), the lowest decile benefits from higher education funding by about 15,000 (slightly less than 20,000) Euros, the highest decile receives about 40,000 (50,000) Euros. One main reason is that the higher the decile, the larger (approximately) the share of academics who benefit from these funding instruments. Figure 3.10 shows how the share of academics evolves across deciles. For both men and women, the share of academics increases from about 40 percent in the bottom decile to more than 80 percent in the top decile.

A second reason for the increase in benefits across deciles is the composition of fields of study within the group of academics. The higher the decile, the larger the share of graduates from both medicine and math/natural sciences, the two most costly fields of study. Finally, the fact that the increase in benefits across deciles is steeper for women than for men can be explained by the fact that the share of medicine students is more than double for women (7.3%) compared to men (3.3%) (as was shown in Figure 3.1).

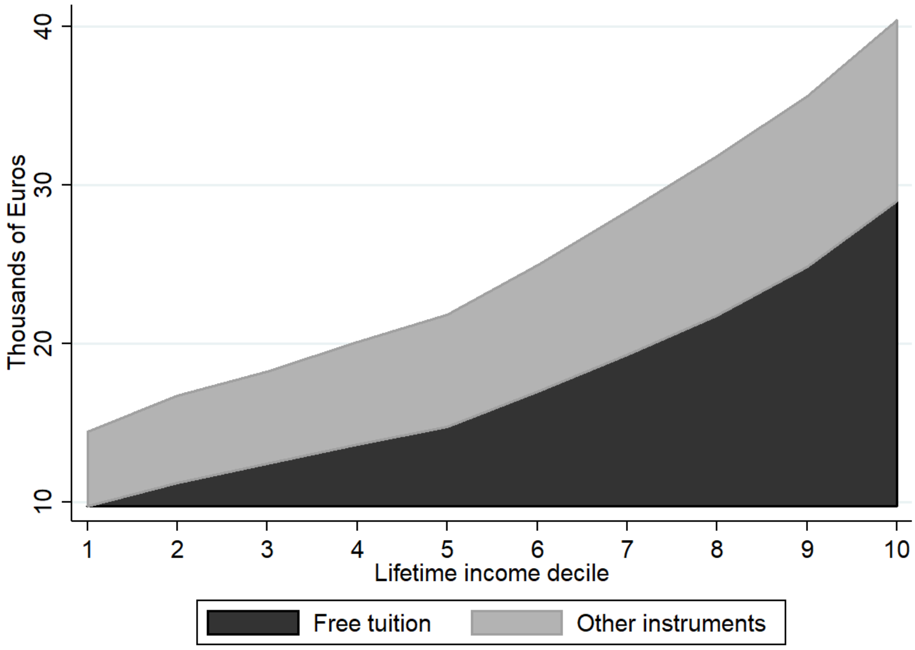
3.6.3 Simulation of an alternative tuition fee/ICL system

In this section, I will analyze a simulation of an alternative tuition fee system with income-contingent loans (ICLs). Tuition fees with ICLs mean that students gradually pay back (part of) the cost of tuition after graduation, given that their income exceeds a defined threshold. ICLs exist in various industrialized countries, such as England, Australia, and New Zealand and there is a vast heterogeneity in the systems' characteristics (see Britton et al., 2019, for a survey of these differences).

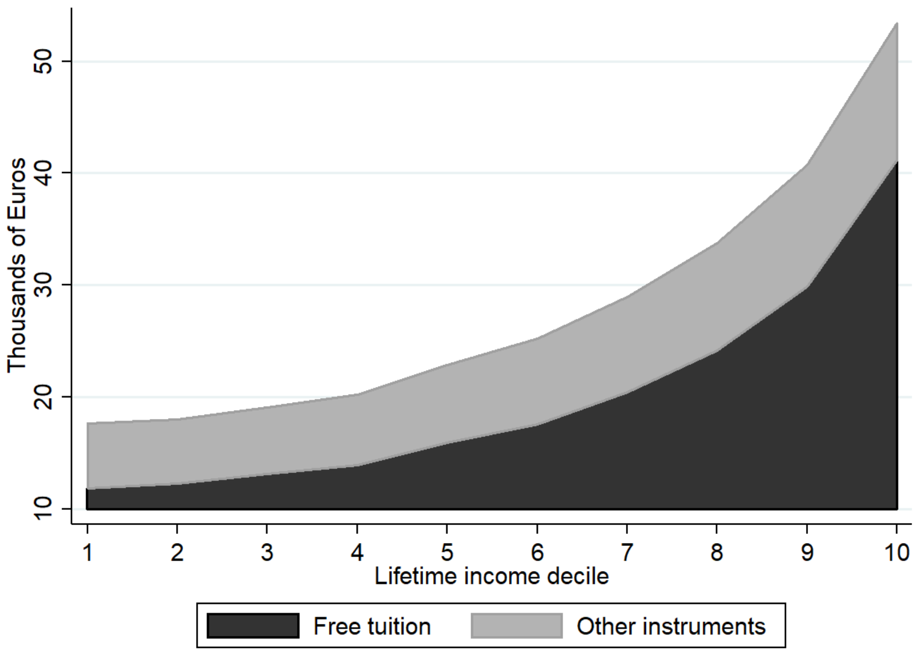
[20] As argued above, to compare the subsidies received from higher education funding with the taxes paid used for the financing of higher education, one would need to make an assumption concerning the share of taxes that is used to finance higher education.

Figure 3.9: Benefits by net lifetime income decile

(a) Men

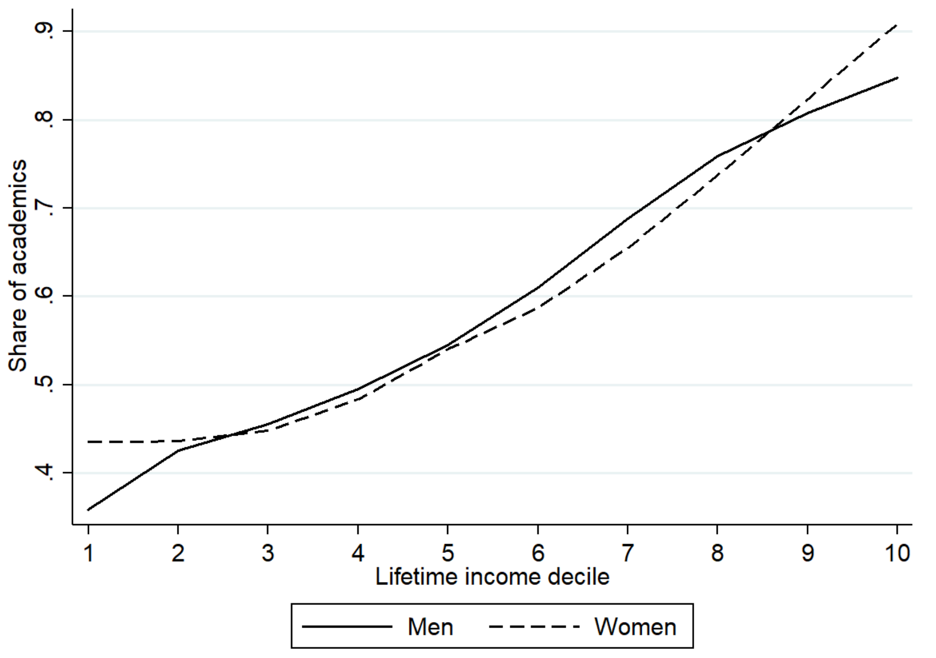


(b) Women



Notes: The figure depicts the present values of the individual higher education funding instruments by net lifetime income decile. “Other instruments" comprises the instruments student loans and grants (BAföG), child benefits, health insurance, education tax allowance and further instruments, as explained above. For further information about the individual instruments, see section 3.4. Panel (a) shows the results for men, panel (b) shows the results for women.
Source: Own simulations.

Figure 3.10: Share of academics by net lifetime income decile



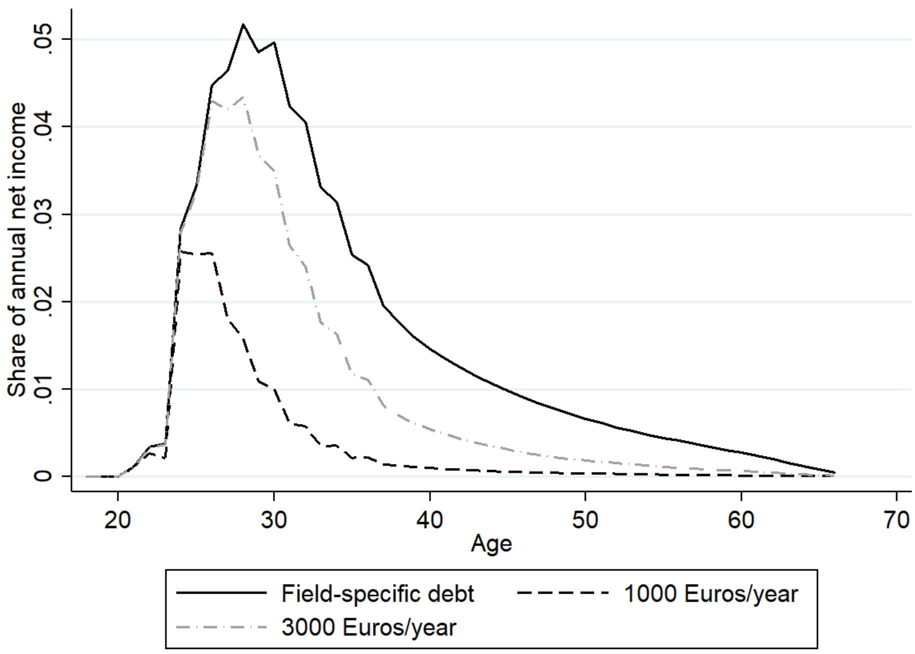
Notes: The figure depicts the share of academics by net lifetime income decile for men and women.
Source: Own simulations.

I consider such a system rather than a system with up-front payments as implemented in the mid-2000s in Germany since several authors have emphasized the advantage of deferred income-contingent payments in terms of efficiency and equity (see, for instance, Barr, 2004, and Chapman, 2006): ICLs may, for instance, reduce the risk of liquidity constraints for prospective students. As individuals from low-income parental households are more likely to encounter such liquidity constraints, ICLs may also be favorable in terms of intergenerational mobility. Furthermore, societal support seems to be larger for tuition fees with ICLs than for tuition fees with up-front payments: Using survey experiments for Germany, Lergetporer and Woessmann (2019) find that designing tuition fees as deferred income-contingent payments as opposed to up-front payments would considerably increase the support for fees (and indeed create a strong majority favoring the existence of tuition fees in general).[21]

For simplicity, I consider a straightforward repayment scheme where the individual net income threshold is set to be 20,000 Euros and the repayment rate is 20% of marginal income, i.e. the individual net income above the threshold. I further assume that there is no interest rate. In terms of the size of tuition fees to be paid I consider

[21] The literature also discusses other potential systems of tuition fees and repayment such as simple loan systems where repayment is not contingent on income or graduate taxes. I will not consider these instruments here.

Figure 3.11: Simulated repayment schedules over the life cycle



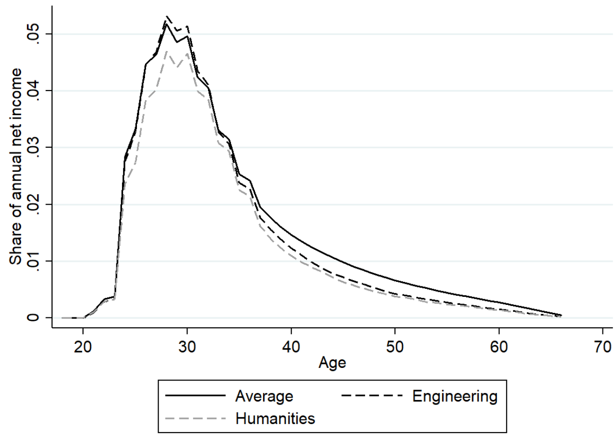
Notes: The figure shows the average simulated repayment schedules for the different tuition schemes over the life cycle. To compute the field-specific average, men and women are pooled with a weight of 50% each.
Source: Own simulations.

three different levels: (i) 1000 Euros per year, (ii) 3000 Euros per year, and (iii) tuition fees that equal the full cost of each field of study. The level of the first scheme is comparable to the tuition fees that were introduced in the mid-2000s in Germany. The key difference is that in the ICL system tuition fees would not have to be paid instantaneously, but only after graduation and only if earned income is sufficiently high. While the third scheme might seem extreme, it gives an idea about the (maximum) range of possibilities for tuition fees. At the same time, the debt levels accumulated under the third scheme correspond to the free tuition benefits (as shown in Figure 3.3).

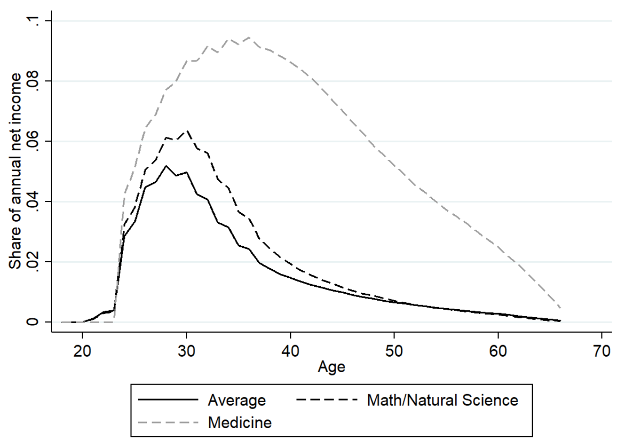
Figure 3.11 shows the simulated average repayment schedules, as a share of annual net income, over the life cycle (pooled for men and women). It distinguishes between the three tuition schemes considered before (1,000 Euros, 3,000 Euros, and field-specific fees). Naturally, the larger the initial debt level, the larger the repayments by a given age and the longer the repayment duration. Hence, while on average tuition fees of 1,000 Euros per year would imply a repayment maximum of about 2.5% of annual net income, field-specific tuition fees would cause average repayments of up to 5% of net income in a given year. In addition, while under the 1000-Euros fee individuals have paid back almost the whole debt before age 40, field-specific fees would imply that repayment has to continue until around retirement.

Figure 3.12: Simulated repayment schedules over the life cycle

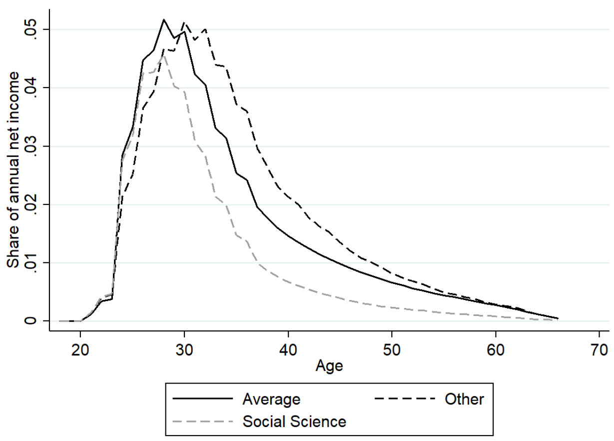
(a) Engineering, Humanities



(b) Math and Natural sciences, Medicine



(c) Social sciences, Other



Notes: The figure shows the average simulated repayment schedules for field-specific tuition fees over the life cycle. Each panel shows the repayment schedule of two fields of study and the average over all fields for comparison. To compute the field-specific average, men and women are pooled with a weight of 50% each.
Source: Own simulations.

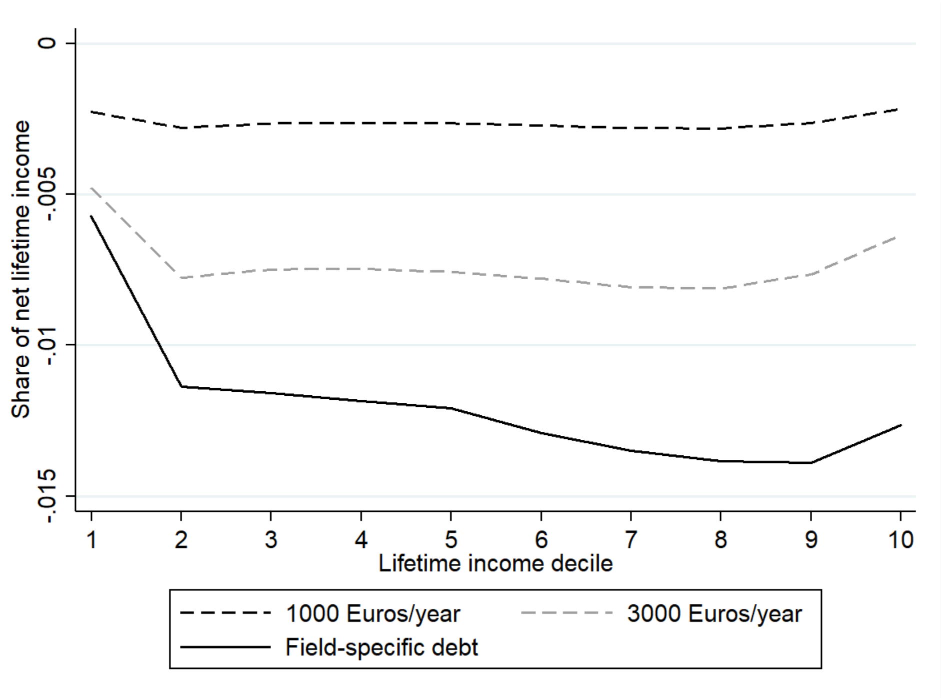
Figure 3.12 shows the average repayment schedule of each field under field-specific fees. It becomes clear that there would be substantial heterogeneity across fields of study. Most notably would graduates of medicine face repayments that are much larger than the profiles of graduates from other disciplines.[22] The graduates of mathematics and natural sciences have the second highest repayment burden in young ages being a result of their relatively large cost of tuition and their high expected earnings.

[22]It should be noted here, that there is also a substantial difference in repayment schedules for medicine between men and women. As male medicine graduates have much higher net earnings, they pay off their debt much earlier. Consequentially, the fact that repayment continues until around retirement is mainly driven by women.

Interestingly, despite having a similar cost of tuition, graduates of the residual category “other" have a below-average relative repayment when young but do have to continue repaying more towards older ages. The reason is that they have much lower projected earnings over the life cycle.

Finally, Figures 3.13 and 3.14 show the distributional effects of changing from the current system without tuition fees to one of the tuition fee and ICL systems described above. Naturally, the system with 1,000 Euros per year would have the lowest distributional impact. Here, the share of the net lifetime income spent for tuition would be below 0.5% for all deciles. The curve for field-specific tuition fees is essentially mirroring the area for free tuition in Figure 3.9 where the value of the higher education funding instruments was plotted against lifetime income deciles (in absolute terms).[23]

Figure 3.13: Simulated distributional effects of a tuition fee/ICL system, men

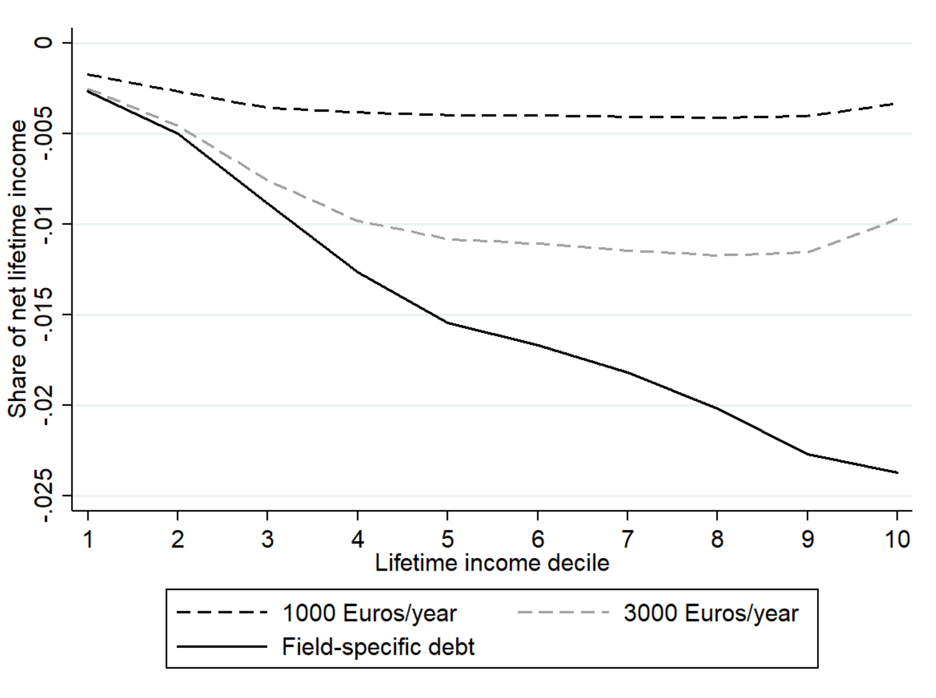


Notes: The figure depicts the distributional effects of the three different tuition fee schemes described above, without lump-sum redistribution of the revenue from paid tuition. *Source*: Own simulations.

The negative distributional effect increases across deciles, again being a result of both the increasing share of academics and the changing composition of the different fields of study within each decile. At the same time, the share of net lifetime income spent on tuition would be small even for the highest deciles, reaching a maximum of

[23]The only difference is that not all individuals pay back their full debt over the life cycle.

Figure 3.14: Simulated distributional effects of a tuition fee/ICL system, women



Notes: The figure depicts the distributional effects of the three different tuition fee schemes described above, without lump-sum redistribution of the revenue from paid tuition. *Source*: Own simulations.

less than 1.5% for men and less than 2.5% for women. The gender difference is mainly due to the fact that female graduates have a much lower projected net lifetime income and hence debt is larger relative to lifetime income. In addition, the gender difference in the distributional effects for the top deciles can be explained by the share of medicine graduates (which is more than double for women compared to men).

3.6.4 Behavioral responses

In order to assess the likely behavioral effects of the hypothetical tuition fee reforms described above, I estimate a Conditional Logit model. This model is a modified version of the one used in Hügle (2021) where the decision to enrol in higher education is estimated as a binary choice between higher education and vocational training. Here, I assume that the individuals with a higher education entrance degree can choose between different fields of study and vocational training.[24]

I assume that individuals associate a level of utility with each educational path

[24] For those individuals with a higher education entrance degree, vocational training usually is the only attractive alternative to higher education.

(comprising all fields of study plus vocational training), such that:

$$U_{ij} = x'_{ij}\beta + \varepsilon_{ij},\ j = 1, ..., J. \tag{3.1}$$

where U_{ij} is the utility level individual i associates with educational path j, x_{ij} describes the characteristics and attributes of individual i and alternative j, and ε is the error term. Most importantly, x_{ij} contains individual i's simulated individual net lifetime income under alternative j. To simulate lifetime incomes I use again the dynamic microsimulation model described above and estimate slightly modified wage regressions interacting the fields of study with dummies for the German states creating regional variation in net lifetime incomes. In addition, x contains variables for parental education, parental occupation, migration, a measure for cognitive skills, and gender as control variables.[25] Assuming that ε_{ij} is i.i.d. EV(1) distributed, the probability that i chooses j is:

$$P(y_i = j) = \frac{exp(x'_{ij}\beta)}{exp(x'_{i1}\beta) + ... + exp(x'_{iJ}\beta)},\ j = 1, ..., J. \tag{3.2}$$

I use the NEPS SC4 (9th graders) data set to estimate the model. Table 3.A2 (Appendix) presents the Conditional Logit coefficient estimates. Table 3.A3 (Appendix) shows the results from using the estimation sample to simulate changes in net lifetime income separately by field of study (including vocational training). I find elasticities in the range of 0.4–0.8, i.e. a 10% increase in the expected net lifetime income in field j, ceteris paribus, increases the average choice probability of this field by 4–8%. The literature on the elasticity of field choice with respect to expected income has produced a large range of estimates (see, for instance, the survey paper by Altonji et al., 2016). An elasticity of 0.6, for instance, appears to be rather in the upper part of those estimates.[26]

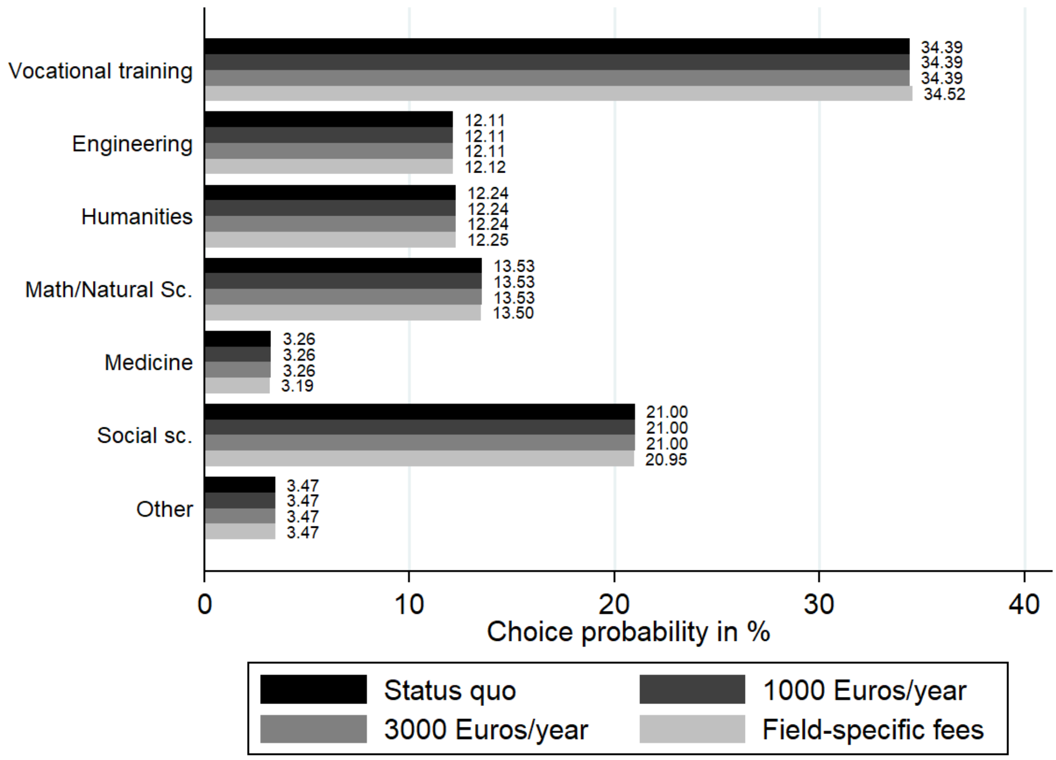
Figure 3.15 shows the average predicted probabilities under the different tuition schemes for the NEPS sample. In general, only marginal changes can be found. For the ICL schemes with tuition fees of 1,000 or 3,000 Euros annually, there are almost no changes. In addition, even for field-specific fees, responses are quite small. The average predicted probability to choose medicine, whose students would be financially affected most strongly by field-specific fees, would only drop from 3.26% to 3.19%.

The main reason these predicted changes are so small is that while the redistri-

[25] These variables were interacted with alternative-specific dummies to make them varying over alternatives.

[26] Recent estimates range from 0.01 (Beffy et al., 2012) to 0.67 (Long et al., 2015).

Figure 3.15: Behavioral responses to the introduction of different types of tuition fees



Notes: The graph shows the average predicted choice probabilities for the estimation sample using the Conditional Logit estimates as presented in 3.A2 in the Appendix. “Status quo" is the current system without tuition fees. “1,000 (3,000) Euros/year" refer to the tuition fee scheme with a tuition level of 1,000 (3,000) euros annually. Finally, “Field-specific fees" is the tuition scheme were tuition fees equal the full cost of tuition.
Source: SOEP, NEPS, own calculations.

bution implied by the proposed tuition fee reforms may appear large in its absolute size it is minor in terms of net lifetime income (as shown above). Clearly, the ICL scheme with fees that fully cover the cost of tuition are an extreme case. For alternative schemes that lie closer to the 1,000 and 3,000 Euro schemes, the model would predict almost no behavioral responses.

3.7 Conclusion and discussion

This chapter analyzes the instruments of higher education funding from a distributional perspective. In the first part, I assess the quantitative importance of different funding instruments for different fields of study. I find that free tuition is the most important instrument with a present value between 20,000 and close to 160,000 Euros depending on the field of study. In the second part, I simulate the life cycles of a young cohort in order to analyze who benefits from higher education funding in terms of expected lifetime income. I show that the benefit from higher education funding is increasing by decile, being a result of the increasing share of academics and the composition of different fields of study in higher deciles. In the third part, I analyze the distributional consequences of different hypothetical tuition fee schemes. I find that they would imply almost no behavioral responses in terms of the field choice. The reason is that the redistribution implied by the hypothetical reforms is only marginal relative to net lifetime income.

Differentiating between fields of study proved fruitful due to their differences in tuition cost and expected lifetime incomes of their graduates. Yet, a more detailed analysis would be important, as there is still sizable heterogeneity in earnings within the fields of study as defined here. However, there is a lack of panel data with a larger amount of observations for such individual study programs (*Studiengänge*). As such data becomes available, it might be worthwhile to differentiate further within fields of study. Furthermore, with a larger amount of panel data available, one could account for field-specific employment patterns and household transitions.

Fields of study also differ substantially in the share of their graduates who become self-employed. On the one hand, I account for this fact when computing social security contributions through the tax-and-transfer simulation. On the other hand, I do not consider the fact that self-employed academics, for instance physicians with a medical practice, often have a sizable investment early in their careers and then rely on this investment as a retirement provision in later years.[27] While the dynamic microsimlation model used in this chapter models some general form of age-specific savings, future research could include savings, investment, and (in a more detailed way) capital income that depend on the field of study.

On a more general level, future research could also connect the two main approaches in the analysis of the distributional effects of higher education funding, the cross-sectional and the longitudinal approach. More specifically, the longitudinal approach could be extended by connecting the individuals of the young cohort, whose life cycles are being forecasted, with the parental generation. This would essentially be a contribution to the field of intergenerational mobility.

Finally, it is necessary to put the analysis into the perspective of education fund-

[27]Using German microcensus data for the years 2005 to 2009, Glocker and Storck (2014), for instance, calculate that about 93% of all individuals with a PhD in dentistry are self-employed.

ing in general: While the analysis of this chapter shows the importance of different higher education funding instruments and who benefits from these instruments, it should be taken into account that individuals in post-secondary training other than higher education are also subsidized to some degree, for instance through free tuition in vocational schools. Hence, extending the current analysis to post-secondary education in general would be an interesting avenue for future research.

Appendix

Table 3.A1: OLS wage regressions

	Women, academic	Men, academic
Experience/10	0.481***	0.455***
	(0.106)	(0.110)
$\text{Experience}^2/100$	-0.198*	-0.0862
	(0.106)	(0.0981)
$\text{Experience}^3/1{,}000$	0.0381	0.00472
	(0.0398)	(0.0330)
$\text{Experience}^4/100{,}000$	-0.0252	-0.000919
	(0.0491)	(0.0369)
Tenure/10	0.494***	0.302***
	(0.0681)	(0.0601)
$\text{Tenure}^2/100$	-0.305***	-0.172***
	(0.0783)	(0.0663)
$\text{Tenure}^3/1{,}000$	0.104***	0.0463*
	(0.0323)	(0.0261)
$\text{Tenure}^4/100{,}000$	-0.126***	-0.0445
	(0.0424)	(0.0330)
Age/10	4.743***	5.154***
	(1.421)	(1.655)
$\text{Age}^2/100$	-1.522***	-1.474**
	(0.515)	(0.577)
$\text{Age}^3/1{,}000$	0.211***	0.177**
	(0.0808)	(0.0872)
$\text{Age}^4/100{,}000$	-0.110**	-0.0769
	(0.0464)	(0.0482)
Humanities	0.0230	0.0765*
	(0.0363)	(0.0428)
Social Sciences	0.0693**	0.179***
	(0.0305)	(0.0330)
Math/Natural Sciences	0.152***	0.258***
	(0.0377)	(0.0355)
Medicine	0.358***	0.416***
	(0.0402)	(0.0415)
Engineering	0.0380	0.215***
	(0.0414)	(0.0327)
Only bachelor degree	-0.158***	-0.135***
	(0.0174)	(0.0162)
Migration background	0.0104	-0.0438**
	(0.0235)	(0.0201)
Constant	-3.067**	-4.193**
	(1.457)	(1.742)
Orthog. state dummies	yes	yes
Year dummies	yes	yes
N	15701	19639

Notes: This table presents the coefficients of OLS wage regressions separately for men and women with higher education degrees. Dependent variable is the log gross hourly wage. The base category for the fields of study is is the residual category “other". Standard errors clustered on the individual level shown in parentheses. * / ** / ***: statistically significantly different from zero at the 10%- / 5%- / 1%-level. All estimations include dummies for survey year, and orthogonalized dummies for federal states.
Source: Own calculations based on SOEP v35, waves 1985–2018.

Table 3.A2: Conditional Logit estimates

Lifetime income/1000	0.00105**
	(2.79)
alternative_a1	-0.655***
	(-5.02)
Parents: Low education_a1	-0.811***
	(-5.60)
Parents: Medium education_a1	-0.665***
	(4.25)
Migration _Background_a1	0.902***
	(4.15)
Parents: Low occupation_a1	-0.0921
	(-0.46)
Parents: Medium occupation_a1	-0.381***
	(-2.95)
Cognitive_skills_1_a1	1.005 ***
	(13.33)
Female_a1	-1.076***
	(-8.46)
alternative_a2	-1.394***
	(-9.38)
Parents: Low education_a2	-1.087***
	(-7.53)
Parents: Medium education_a2	-0.405**
	(-2.77)
Migration _Background_a2	0.190
	(0.78)
Parents: Low occupation_a2	-0.177
	(-0.87)
Parents: Medium occupation_a2	0.105
	(0.85)
Cognitive_skills_1_a2	-0.561***
	(7.63)
Female_a2	1.027***
	(7.48)
alternative_a3	-0.866***
	(-6.60)
Parents: Low education_a3	-0.989***
	(-6.99)
Parents: Medium education_a3	-0.557***
	(-3.81)
Migration _Background_a3	1.020***
	(4.89)
Parents: Low occupation_a3	-0.164
	(-0.81)
Parents: Medium occupation_a3	-0.257*
	(-2.08)
Cognitive_skills_1_a3	1.109***
	(15.16)
Female_a3	-0.587***
	(-4.91)
alternative_a4	-3.138***
	(-11.76)
Parents: Low education_a4	-1.469***
	(-5.59)
Parents: Medium education_a4	-0.776***
	(-3.14)
Migration _Background_a4	1.414***
	(4.03)
Parents: Low occupation_a4	-0.737
	(-1.66)
Parents: Medium occupation_a4	-0.276
	(-1.30)
Cognitive_skills_1_a4	1.235***
	(9.76)
Female_a4	0.856***
	(3.95)
alternative_a5	-2.194***

	(-9.63)
Parents: Low education_a5	-0.916***
	(-3.95)
Parents: Medium education_a5	-0.568*
	(-2.33)
Migration _Background_a5	0.428
	(1.08)
Parents: Low occupation_a5	-0.635
	(-1.68)
Parents: Medium occupation_a5	-0.311
	(-1.53)
Cognitive_skills_1_a5	0.592***
	(4.91)
Female_a5	0.495*
	(2.37)
alternative_a6	-0.256*
	(-2.32)
Parents: Low education_a6	-0.925***
	(-7.76)
Parents: Medium education_a6	-0.406**
	(-3.26)
Migration _Background_a6	0.482*
	(2.55)
Parents: Low occupation_a6	-0.429*
	(-2.51)
Parents: Medium occupation_a6	0.286**
	(-2.77)
Cognitive_skills_1_a6	0.602***
	(9.86)
Female_a6	0.110
	(1.04)
N	26741

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[a] The table displays the coefficients estimates of the Conditional Logit model. “alternative_a1"-“alternative_a6" indicate alternative-specific dummy variables and suffixes like “_a1" to “_a6" indicate interactions with characteristics such as parental education or cognitive skills with the alternative-specific dummies. The alternatives are numbered as their alphabetic ordering: (1) Engineering, (2) Humanities (3) Math/Natural Sciences (4) Medicine (5) Other (6) Social Sciences, and vocational training is the base category.

[b] Source: SOEP, NEPS, own calculations.

Table 3.A3: Predicted elasticities for each field

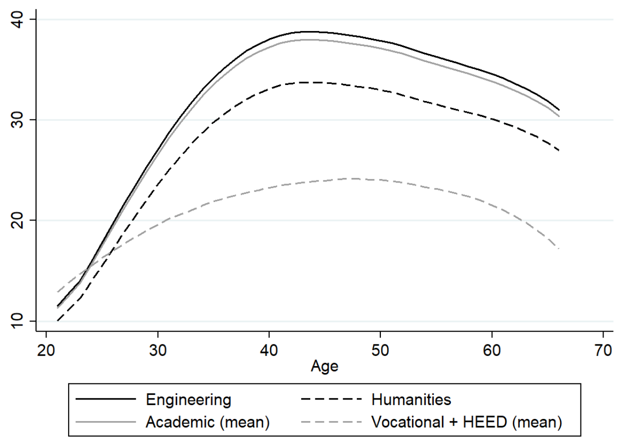
Field	Change in probability in % if net lifetime income increases by 10%
Engineering	6.51
Humanities	6.19
Math and Natural sciences	6.82
Medicine	8.68
Social sciences	5.34
Other	6.73
Vocational	4.48

Notes: The table lists the average relative changes in probabilities for a 10% increase in net lifetime income for each field (including vocational training). The quantities presented are equal to 10 times the elasticity.

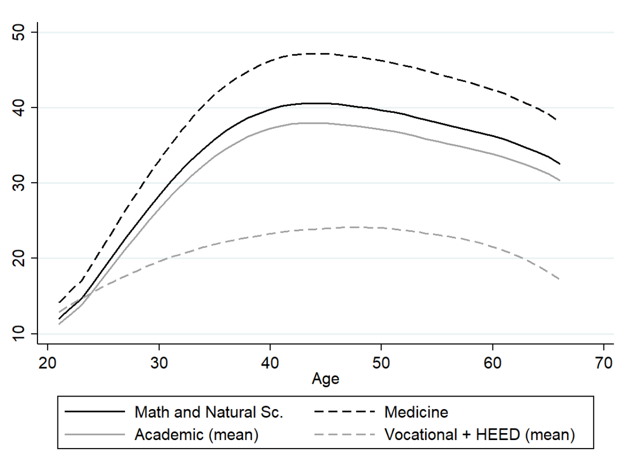
Sources: NEPS, own calculations.

Figure 3.A1: Simulated hourly wage profiles by field of study, men

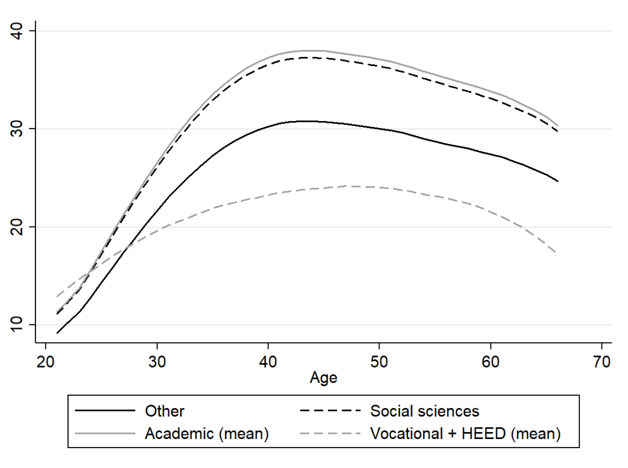
(a) Engineering, Humanities



(b) Math and Natural Sciences, Medicine



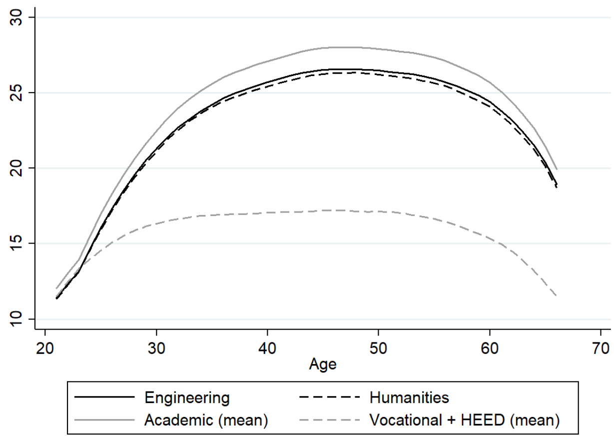
(c) Social sciences, Other fields



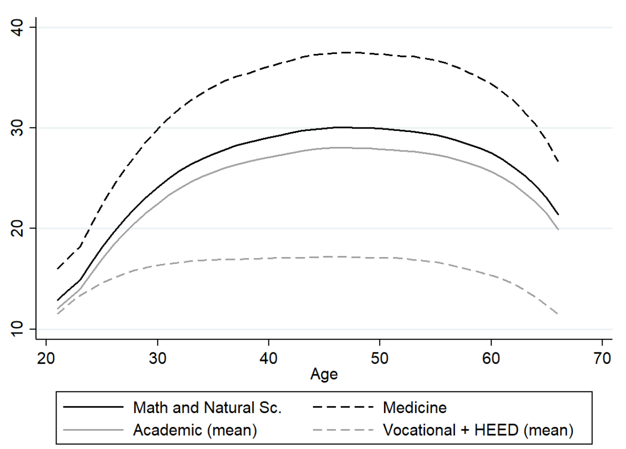
Notes: The figure depicts simulated gross hourly wages by field of study for men in Euros. Panel (a) is for the fields engineering and humanities, panel (b) is for math and natural sciences and for medicine, and panel (c) is for social sciences and other fields.
Source: Own simulations.

Figure 3.A2: Simulated hourly wage profiles by field of study, women

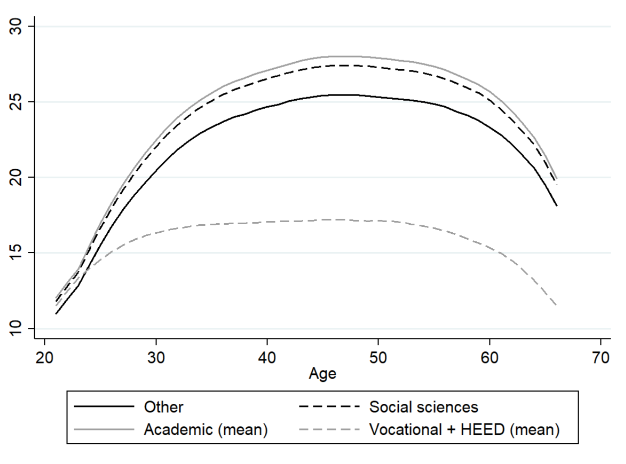
(a) Engineering, Humanities



(b) Math and Natural Sciences, Medicine



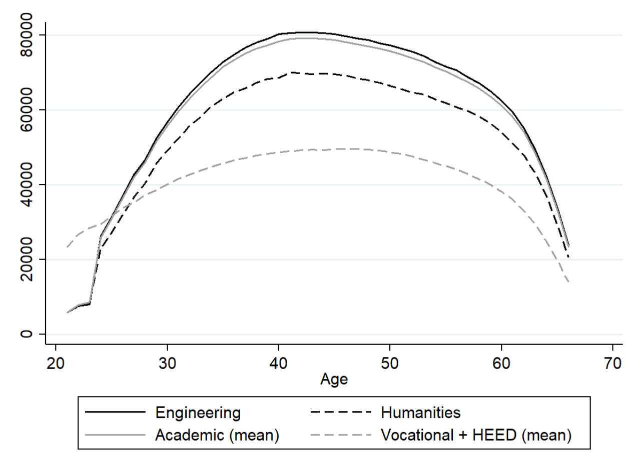
(c) Social sciences, Other fields



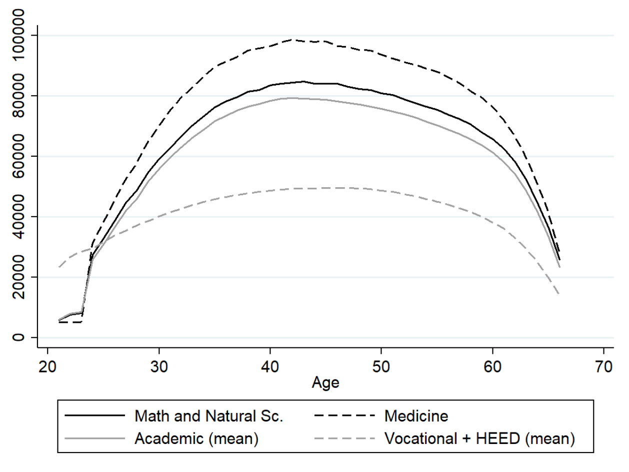
Notes: The figure depicts simulated gross hourly wages by field of study for women in Euros. Panel (a) is for the fields engineering and humanities, panel (b) is for math and natural sciences and for medicine, and panel (c) is for social sciences and other fields.
Source: Own simulations.

Figure 3.A3: Simulated annual labor earnings profiles by field of study, men

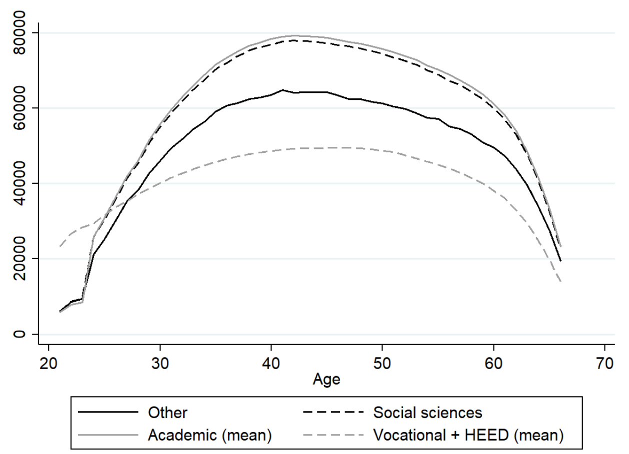
(a) Engineering, Humanities



(b) Math and Natural Sciences, Medicine



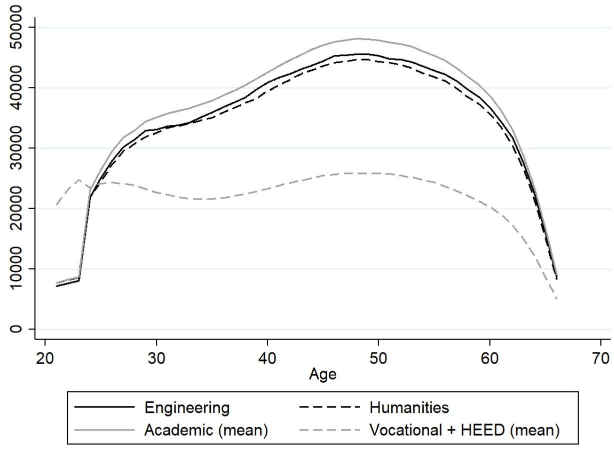
(c) Social sciences, Other fields



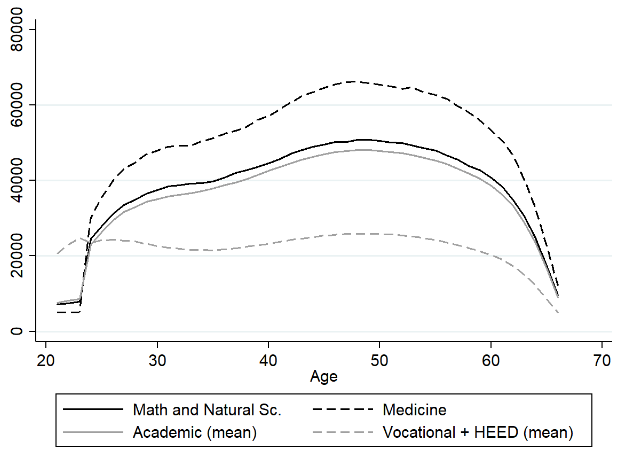
Notes: The figure depicts simulated annual labor earnings by field of study for men in Euros. Panel (a) is for the fields engineering and humanities, panel (b) is for math and natural sciences and for medicine, and panel (c) is for social sciences and other fields.
Source: Own simulations.

Figure 3.A4: Simulated annual labor earnings profiles by field of study, women

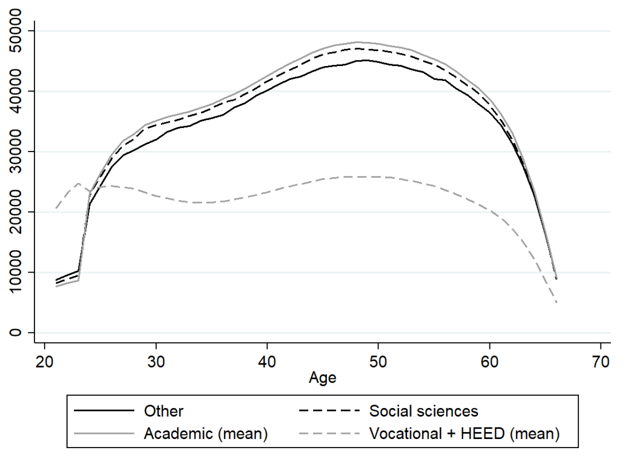
(a) Engineering, Humanities



(b) Math and Natural Sciences, Medicine



(c) Social sciences, Other fields



Notes: The figure depicts simulated annual labor earnings by field of study for women in Euros. Panel (a) is for the fields engineering and humanities, panel (b) is for math and natural sciences and for medicine, and panel (c) is for social sciences and other fields.
Source: Own simulations.

Bibliography

Abbott, A. (1995): "Sequence Analysis: New Methods for Old Ideas," *Annual Review of Sociology*, 21(1), 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.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016): "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, Elsevier, vol. 5, chap. 7, 305–396.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020): "Ex Ante Returns and Occupational Choice," *Journal of Political Economy*, 128(12), 4475–4522.

Autorengruppe Bildungsberichtserstattung (2014): *Bildung in Deutschland 2014: Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*, W. Bertelsmann Verlag.

——— (2016): *Bildung in Deutschland 2016: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*, W. Bertelsmann Verlag.

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

Azmat, G. and Ş. Simion (2020): "Charging for Higher Education: Estimating the Impact on Inequality and Student Outcomes," *The B.E. Journal of Economic Analysis & Policy*, 21(1), 175–239.

Barbaro, S. (2002): "The Distributional Impact of Subsidies to Higher Education - Empirical Evidence from Germany," *FinanzArchiv/Public Finance Analysis*, 59(4), 458–478.

Barr, N. (2004): "Higher Education Funding," *Oxford Review of Economic Policy*, 20(2), 264–283.

Becker, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70(5), 9–49.

Beffy, M., D. Fougere, and A. Maurel (2012): "Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter?" *Review of Economics and Statistics*, 94(1), 334–347.

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

Ben-Porath, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), 352–365.

Beznoska, M. (2019): "Do Couples Pool Their Income? Evidence from Demand System Estimation for Germany," Discussion Papers 2019/3, Freie Universität Berlin, School of Business & Economics.

Bhuller, M., M. Mogstad, and K. G. Salvanes (2017): "Life-Cycle Earnings, Education Pemiums, and Internal Rates of Return," *Journal of Labor Economics*, 35(4), 993–1030.

Bietenbeck, J., J. Marcus, and F. Weinhardt (2020): "Tuition Fees and Educational Attainment," Discussion Papers 1900, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

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.

Black, S. E. and P. J. Devereux (2010): "Recent Developments in Intergenerational Mobility," Working Paper 15889, National Bureau of Economic Research (NBER), Cambridge.

Blossfeld, H.-P. and J. Von Maurice (2011): "Education as a Lifelong Process – The German National Educational Panel Study (NEPS)," *Zeitschrift für Erziehungswissenschaft*, 14(2), 19–34.

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, Freie Universität 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*, no. 111426 in RWI Projektberichte, RWI – Leibniz-Institut für Wirtschaftsforschung, 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(4), 650–671.

Borgloh, S., F. Kupferschmidt, and B. U. Wigger (2008): "Verteilungseffekte der öffentlichen Finanzierung der Hochschulbildung in Deutschland: Eine Längsschnittbetrachtung auf der Basis des Sozioökonomischen Panels," *Jahrbücher für Nationalökonomie und Statistik*, 228(1), 25–48.

Brenke, K. and J. Pfannkuche (2018): "Konsum und Sparquote der privaten Haushalte hängen stark vom Erwerbsstatus, Einkommen und Alter ab," *DIW-Wochenbericht*, 85(10), 181–191.

Britton, J., L. van der Erve, and T. Higgins (2019): "Income Contingent Student Loan Design: Lessons from Around the World," *Economics of Education Review*, 71, 65–82.

Bruckmeier, K. and B. U. Wigger (2014): "The Effects of Tuition Fees on Transition from High School to University in Germany," *Economics of Education Review*, 41, 14–23.

Brzinsky-Fay, C., U. Kohler, and M. Luniak (2006): "Sequence Analysis with Stata," *The Stata Journal*, 6(4), 435–460.

Buchinsky, M. and P. Leslie (2010): "Educational Attainment and the Changing US Wage Structure: Dynamic Implications on Young Individuals' Choices," *Journal of Labor Economics*, 28(3), 541–594.

Bujard, M. (2015): "Kinderlosigkeit in Deutschland: Wie interagieren Bildung, Wohnort, Migrationshintergrund, Erwerbstätigkeit und Kohorte?" *JFR - Journal of Family Research*, 27(3), 270–296.

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

Callan, T., T. Smeeding, and P. Tsakloglou (2008): "Short-run Distributional Effects of Public Education Transfers to Tertiary Education Students in Seven European Countries," *Education Economics*, 16(3), 275–288.

Cameron, A. C. and P. K. Trivedi (2009): *Microeconometrics Using Stata*, vol. 1, Stata Press, College Station, Texas.

Chapman, B. (2006): "Income Contingent Loans for Higher Education: International Reforms," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek and F. Welch, Elsevier, vol. 2, chap. 25, 1435–1503.

Chapman, B. and M. Sinning (2014): "Student Loan Reforms for German Higher Education: Financing Tuition Fees," *Education Economics*, 22(6), 569–588.

Christoph, B., U. Leber, and H. Stüber (2017): "Einkommen von Bachelor- und anderen Hochschulabsolventen: Höhere Abschlüsse zahlen sich mit dem Alter zunehmend aus," IAB-Kurzbericht No. 13/2017, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Ciccone, A. and G. Peri (2006): "Identifying Human-Capital Externalities: Theory With Applications," *The Review of Economic Studies*, 73(2), 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., J. Heckman, and S. Navarro (2005): "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," *Oxford Economic Papers*, 57(2), 191–261.

Cunha, F. and J. J. Heckman (2007): "Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling," *Labour Economics*, 14(6), 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(6), 1319–1360.

Dearden, L., E. Fitzsimons, A. Goodman, and G. Kaplan (2008): "Higher Education Funding Reforms in England: The Distributional Effects and the Shifting Balance of Costs," *The Economic Journal*, 118(526), F100–F125.

Deaton, A. (1997): *The Analysis of Household Surveys*, The World Bank, chap. Econometric Issues for Survey Data, 63–132.

Dekkers, G. (2016): *New Pathways in Microsimulation*, Routledge, chap. What are the Driving Forces behind Trends in Inequality among Pensioners? Validating MIDAS Belgium Using a Stylized Model, 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(5), 778–791.

Eika, L., M. Mogstad, and B. Zafar (2019): "Educational Assortative Mating and Household Income Inequality," *Journal of Political Economy*, 127(6), 2795–2835.

Fischer, B. and D. Hügle (2020): "The Private and Fiscal Returns to Higher Education: A Simulation Approach for a Young German Cohort," Discussion Papers 2020/21, Freie Universität Berlin, School of Business and Economics.

Flannery, D. and C. O'Donoghue (2013): "The Demand for Higher Education: A Static Structural Approach Accounting for Individual Heterogeneity and Nesting Patterns," *Economics of Education Review*, 34, 243–257.

Fossen, F. M. and D. Glocker (2011): "Expected Future Earnings, Taxation, and University Enrollment," *International Tax and Public Finance*, 18(6), 688–723.

——— (2017): "Stated and Revealed Heterogeneous Risk Preferences in Educational Choice," *European Economic Review*, 97, 1–25.

Frank, I. and G. Walden (2012): "Ausbildungsdauer von Berufen–verbindliche Begrenzung auf drei Jahre?" *BWP-Berufsbildung in Wissenschaft und Praxis*, 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(2), 172–209.

GIANNELLI, G. C. AND C. MONFARDINI (2003): “Joint Decisions on Household Membership and Human Capital Accumulation of Youths. The Role of Expected Earnings and Local Markets,” *Journal of Population Economics*, 16(2), 265–285.

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(2), 345–360.

HAAN, P., D. KEMPTNER, AND V. L. PROWSE (2017): “Insurance, Redistribution, and the Inequality of Lifetime Income,” Discussion Papers 1716, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

HARDING, A. (1991): “Lifetime Income Distribution and Redistribution in Australia: Applications of a Dynamic Cohort Microsimulation Model.” Ph.D. thesis, London School of Economics and Political Science, United Kingdom.

——— (1995): “Financing Higher Education: An Assessment of Income-Contingent Loan Options and Repayment Patterns over the Life Cycle,” *Education Economics*, 3(2), 173–203.

HECKMAN, J. AND P. CARNEIRO (2003): “Human Capital Policy,” Working Paper 9495, National Bureau of Economic Research (NBER), Cambridge.

HECKMAN, J. J., J. E. HUMPHRIES, AND G. VERAMENDI (2018a): “The Nonmarket Benefits of Education and Ability,” *Journal of Human Capital*, 12(2), 282–304.

——— (2018b): “Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking,” *Journal of Political Economy*, 126(S1), S197–S246.

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, Elsevier, vol. 1, chap. 7, 307–458.

——— (2008): “Earnings Functions and Rates of Return,” *Journal of Human Capital*, 2(1), 1–31.

HÜBNER, M. (2012): “Do Tuition Fees Affect Enrollment Behavior? Evidence from a ‘Natural Experiment’ in Germany,” *Economics of Education Review*, 31(6), 949–960.

Hügle, D. (2021): "The Decision to Enrol in Higher Education," Discussion Papers 2021/8, Freie Universität Berlin, School of Business and Economics.

Hügle, D. (2021): "Higher Education Funding in Germany: A Distributional Lifetime Perspective," Discussion Papers 2021/1, Freie Universität Berlin, School of Business and Economics.

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(1), 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, Elsevier, vol. 3, Part A, chap. 26, 1463–1555, 1st ed.

Koutsampelas, C. and P. Tsakloglou (2015): "The Progressivity of Public Education in Greece: Empirical Findings and Policy Implications," *Education Economics*, 23(5), 596–611.

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, Springer, 97–114.

Lergetporer, P. and L. Woessmann (2019): "The Political Economy of Higher Education Finance: How Information and Design Afect Public Preferences for Tuition," CESifo Working Paper No. 7536, CESifo, Munich.

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(2), 3–55.

——— (2014): "Evaluating Binary Alignment Methods in Microsimulation Models," *Journal of Artificial Societies and Social Simulation*, 17(1).

Long, M. C., D. Goldhaber, and N. Huntington-Klein (2015): "Do Completed College Majors Respond to Changes in Wages?" *Economics of Education Review*, 49, 1–14.

Lütz, C. and P. Stein (2020): *Mikrosimulationen*, Springer, chap. Validierung in dynamischen Mikrosimulationsmodellen, 141–176.

Middendorff, E., B. Apolinarski, K. Becker, P. Bornkessel, T. Brandt, S. Heissenberg, and J. Poskowsky (2017): *Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016 - 21. Sozialerhebung des Deutschen Studentenwerks*, Bundesministerium für Bildung und Forschung (BMBF), Berlin.

Mincer, J. (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66(4), 281–302.

——— (1974): *Schooling, Experience, and Earnings*, National Bureau of Economic Research (NBER), Cambridge.

Moretti, E. (2004): "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data," *Journal of Econometrics*, 121(1-2), 175–212.

Mullahy, J. (2015): "Multivariate Fractional Regression Estimation of Econometric Share Models," *Journal of Econometric Methods*, 4(1), 71–100.

Murphy, R., J. Scott-Clayton, and G. Wyness (2019): "The End of Free College in England: Implications for Enrolments, Equity, and Quality," *Economics of Education Review*, 71, 7–22.

Nybom, M. (2017): "The Distribution of Lifetime Earnings Returns to College," *Journal of Labor Economics*, 35(4), 903–952.

OECD (2019): *Education at a Glance 2019: OECD Indicators*, OECD.

——— (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(6), 619–632.

——— (2008): "Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates," *Journal of Econometrics*, 145(1-2), 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): *Monitoring Social Inclusion in Europe*, Eurostat, chap. Intra-household pooling and sharing of resources: a tentative "modified" equivalised income, 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. 2A, 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, The 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(5), 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(1), 19–68.

Scherer, S. (2001): "Early Career Patterns: A Comparison of Great Britain and West Germany," *European Sociological Review*, 17(2), 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 (Fachserie1, Reihe 1.4)*, Wiesbaden.

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

Statistisches Bundesamt (2017): "Krankheitskostenrechnung, Krankheitskosten je Einwohner in Euro für Deutschland," Wiesbaden.

STATISTISCHES BUNDESAMT (2018): *Bildungsfinanzbericht 2018*, Wiesbaden.

STATISTISCHES BUNDESAMT (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.

——— (2019d): *Studierende an Hochschulen - Fachserie 11, Reihe 4.1*, Wiesbaden.

——— (2020): *Studierende an Hochschulen - Fächersystematik*, Wiesbaden.

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," DIW Data Documentation 63, 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(6), 1177–1186.

WILSON, K., B. WOLFE, AND R. HAVEMAN (2005): "The Role of Expectations in Adolescent Schooling Choices: Do Youths Respond to Economic Incentives?" *Economic Inquiry*, 43(3), 467–492.

WISWALL, M. AND B. ZAFAR (2015): "Determinants of College Major Choice: Identification Using an Information Experiment," *The Review of Economic Studies*, 82(2), 791–824.

List of Tables

List of Figures

English Summary (Abstracts)

Chapter 1: The Private and Fiscal Returns to Higher Education - A Simulation Approach for a Young German Cohort

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.

JEL: C53, I23, I26
Keywords: Higher education, Returns to education, Dynamic microsimulation

Chapter 2: The Decision to Enrol in Higher Education

In this paper, I analyze how the higher education decision of young adults in Germany depends on their expected future earnings. For this, I estimate a microeconometric model in which individuals maximize life-time utility by choosing whether or not to enter higher education. To forecast individual life cycles in terms of employment, earnings, and family formation under higher education and its alternative, vocational training, I use a dynamic microsimulation model and regression techniques. I take into

account that while individuals generally choose between two options, higher education and vocational training, they are aware of multiple potential realizations under both options, such as leaving higher education with a bachelor degree or taking up higher education after first having earned a vocational degree. Using the estimates from the decision model, I simulate the introduction of different tuition fee and graduate tax scenarios. I find that the impact of these education policies on the higher education decision is limited and only few individuals would change their educational decisions as a reaction to these policies.

JEL: C53, I23

Keywords: Educational choice, Higher education, Dynamic microsimulation

Chapter 3: Higher Education Funding in Germany - A Distributional Lifetime Perspective

This paper analyzes higher education funding in Germany from a distributional perspective. For this, I first compare the quantitative importance of different funding instruments, from free tuition to subsidized health insurance for students. I show that free tuition is, by far, the most important instrument. Then, I take a lifetime perspective and assess how individuals of different expected lifetime incomes benefit from higher education funding. I distinguish between different fields of study as there are large differences in both the expected lifetime earnings of graduating from a specific field and the social cost of tuition associated with each field. Finally, I focus exclusively on the instrument of subsidized tuition and simulate the introduction of different tuition fee schemes with income-contingent loans. While the distributional effects would be sizable in absolute terms, I estimate that they would cause few individuals to change their educational decisions.

JEL: C53, I22, I23

Keywords: Higher education, Education finance, Dynamic microsimulation

Deutsche Zusammenfassung

Diese Dissertation befasst sich mit der Ökonomie der postsekundären Bildung in Deutschland. In Anbetracht der Tatsache, dass junge Erwachsene in Deutschland nach dem Abschluss der Sekundarstufe in der Regel entweder ein Studium oder eine Berufsausbildung (oder manchmal auch beides) aufnehmen, werden drei grundlegende Fragen untersucht: (i) "Lohnt sich ein Hochschulstudium für das Individuum und den Staat?", (ii) "Wie stark beeinflussen die Einkommenserwartungen die Entscheidung des Individuums zwischen Studium und Berufsausbildung?", und (iii) "Welche Verteilungswirkungen hat die Hochschulfinanzierung?"

Ein wesentlicher Aspekt dieser Dissertation ist, dass diese Fragen aus einer "Lebenszeitperspektive" analysiert werden, d.h. der gesamte Lebenszyklus eines Individuums wird betrachtet, und nicht nur ein bestimmter Punkt im Lebenszyklus des Individuums (z.B. zu einem bestimmten Alter). Darüber hinaus liegt dieser Dissertation eine *vorausschauende* Perspektive zugrunde, in dem Sinne, dass sie die Perspektive der Individuen einer jungen Kohorte und ihrer projizierten Lebenszyklen einnimmt. Während eine vorausschauende Lebenszeitperspektive für die Analyse der genannten Fragen naheliegend erscheint, wurde eine solche Perspektive in der Literatur selten eingenommen. Einer der Hauptgründe dürfte sein, dass für jüngere Kohorten naturgemäß keine tatsächlich beobachtbaren Lebenszyklusdaten (bis zum Renteneintritt beispielsweise) existieren. Um eine Lebenszeitperspektive einer jüngeren Kohorte einnehmen zu können, müssen somit "künstliche" Daten generiert werden, die einen plausiblen Lebensverlauf von heute jungen Erwachsenen widerspiegeln. In dieser Dissertation wird dafür ein dynamisches Mikrosimulationsmodell auf Basis des Sozio-oekonomischen Panels (Goebel et al., 2018) entwickelt. Das dynamische Mikrosimulationsmodell simuliert sequentiell den Lebenszyklus eines Individuums in Bezug auf mehrere Schlüsselvariablen wie die Beschäftigung und Haushaltsentscheidungen (Li and O'Donoghue, 2013). Dieses Modell ist die Grundlage für die empirische Arbeit in dieser Dissertation.

Das erste Kapitel, *The Private and Fiscal Returns to Higher Education - A Simulation Approach for a Young German Cohort*, erklärt im Detail, wie das dynamische Mikrosimulationsmodell funktioniert. Im Wesentlichen schätzt es zunächst Übergangsmodelle für die zu simulierenden Variablen und verwendet dann die geschätzten Parameter, um die individuellen Lebenszyklen von einem Jahr zum nächsten zu simulieren. Darüber hinaus enthält es einen Steuer-Transfer-Rechner, der das deutsche Steuer-Transfer-System abbildet und die Berechnung von Steuern, Transfers und Sozialversicherungsbeiträgen ermöglicht. Mit Hilfe des dynamischen Mikrosimulationsmodells werden dann im ersten Kapitel die privaten und fiskalischen Renditen des Hochschulstudiums geschätzt. Bei der Analyse der Renditen unterscheiden wir zwischen Brutto- und Nettoeinkommen und verschiedenen Graden von "Einkommenspooling" innerhalb von Haushalten. Für eine typische Biographie finden wir stark positive Renditen (interne Zinsfüße) sowohl für das Individuum als auch für den Staat. Gleichzeitig finden wir aber auch einen beträchtlichen Anteil von Individuen, deren Investition ins Studium einen negativen Kapitalwert aufweist, für die also eine Berufsausbildung finanziell lohnender wäre als ein Studium.

Kapitel zwei, *The Decision to Enrol in Higher Education*, untersucht die Frage, wie stark die Entscheidung, ein Hochschulstudium aufzunehmen, von den Erwartungen an das zukünftige Einkommen abhängt. Unter Verwendung des dynamischen Mikrosimulationsmodells aus Kapitel 1 wird der erwartetete Lebenszyklus eines Individuums bei einer bestimmten Bildungswahl simuliert. Zusätzlich zum dynamischen Mikrosimulationsmodell und den SOEP-Daten nutze ich die Startkohorte 4 des Nationalen Bildungspanels (Blossfeld and Von Maurice, 2011), die Neuntklässler bis nach dem sekundären Schulabschluss verfolgt. Diese Daten ermöglichen mir, ein Modell der Bildungsentscheidung zu schätzen, bei dem Individuen ihren Lebenszeitnutzen maximieren, indem sie zwischen Hochschulstudium und Berufsausbildung wählen. Mit den geschätzten Parametern aus dem Modell simuliere ich anschließend die Einführung von Studiengebühren und Absolventensteuern. Es zeigt sich, dass solche Reformen nur wenige junge Menschen dazu veranlassen würden, ihre Bildungsentscheidungen zu ändern.

Das dritte Kapitel, *Higher Education Funding in Germany - A Distributional Lifetime Perspective*, analysiert die Verteilungswirkungen der Hochschulfinanzierung. Dazu vergleiche ich zunächst die quantitative Bedeutung verschiedener Finanzierungsinstrumente, von kostenloser Hochschulbildung bis zur subventionierten Krankenversicherung für Studierende. Die Analyse zeigt, dass kostenlose Hochschulbildung das mit Abstand wichtigste Instrument ist. Gleichzeitig hängt das Ausmaß, wie sehr Studierende von der kostenlosen Hochschulbildung profitieren, sehr von ihrer Studienrichtung ab. Um einen Zusammenhang zwischen der Höhe des finanziellen Vorteils aus der Hochschulfinanzierung, insbesondere der Studiengebührenfreiheit, und dem erwarteten Lebenseinkommen herzustellen, verwende ich das dynamische Mikrosimulationsmodell und simuliere die individuellen Biografien. Schließlich wird auf das Entscheidungsmodell aus Kapitel 2 zurückgegriffen und für den Fall multipler Alternativen (wobei die verschiedenen Studienfelder und die Berufsausbildung die Alternativen sind) erweitert. Mit den geschätzten Parametern simuliere ich, wie sich die Wahl zwischen den Feldern bei unterschiedlichen Szenarien von Studiengebühren verändern würde. In Übereinstimmung mit den Ergebnissen aus Kapitel 2 zeigt sich, dass die Studiengebühren die Bildungsentscheidungen der Individuen kaum verändern würden.