UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

What does successful university graduation signal to employers?

A factorial survey experiment on sheepskin effects

Neugebauer, Martin; Heisig, Jan Paul; Bol, Thijs

DOI
10.1093/esr/jcaf028

Publication date
2026

Document Version
Final published version

Published in
European Sociological Review

License
CC BY-NC

Link to publication

Citation for published version (APA):
Neugebauer, M., Heisig, J. P., & Bol, T. (2026). What does successful university graduation signal to employers? A factorial survey experiment on sheepskin effects. *European Sociological Review*, *42*(2), 165–182. https://doi.org/10.1093/esr/jcaf028

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:14 Nov 2026

European Sociological Review, 2026, **42**, 165–182
https://doi.org/10.1093/esr/jcaf028
Advance access publication 5 June 2025
Original Article

OXFORD

What does successful university graduation signal to employers? A factorial survey experiment on sheepskin effects

Martin Neugebauer[1,*], Jan Paul Heisig[2,3] and Thijs Bol[4]

[1]Karlsruhe University of Education, 76133 Karlsruhe, Germany
[2]WZB Berlin Social Science Center, 10785 Berlin, Germany
[3]Department of Political and Social Sciences, Freie Universität Berlin, 14195 Berlin, Germany
[4]Department of Sociology, University of Amsterdam, 1018 WV Amsterdam, The Netherlands
[*]Corresponding author. Email: martin.neugebauer@ph-karlsruhe.de

Higher education graduates enjoy substantial labour market advantages over similar individuals who attended higher education but did not complete a degree. We use a hiring survey experiment with 335 German employers to explore possible explanations for these 'sheepskin effects', while addressing concerns about unobserved confounding in observational studies. Across 2680 hypothetical job applicants, employers were nearly 1.8 times more likely to invite graduates for an interview than otherwise identical non-completers and were also willing to pay graduates substantially higher starting salaries. Using a unique survey module on employers' perceptions, we show that the average employer perceives degree-holders to outperform non-completers in terms of occupation-specific and non-cognitive skills but not in terms of general cognitive skills. These employer perceptions predict hypothetical hiring behaviour in that those who view graduates more favourably showed a stronger preference for this group in the survey experiment. We discuss these results in relation to signalling, human capital, and credentialism explanations of sheepskin effects.

Introduction

The literature on labour market returns to education consistently finds that completion of educational programs (i.e. graduation) is associated with disproportionate advantages in the labour market (e.g. Hungerford and Solon, 1987; Jaeger and Page, 1996; Flores-Lagunes and Light, 2010; Bol and Van de Werfhorst, 2011). Carrying through with the final year of education—the 'diploma year'—has a substantial payoff that cannot be readily explained by the additional amount of schooling acquired through program completion. This phenomenon, often referred to as the 'sheepskin effect'[1], is mirrored by the marked labour market penalties associated with dropout from (higher) education (Matkovic and Kogan, 2012; Scholten and Tieben, 2017; Berlingieri and Bolz, 2020; Giani *et al*., 2020; Klein *et al*., 2021). Better understanding the magnitude and sources of sheepskin effects could therefore also inform efforts to improve the labour market prospects of this relatively disadvantaged group.

As an associational regularity, sheepskin effects are both well-established and substantial in size. A recent study of 18 European countries reported that, after controlling for key demographics, higher education (HE) graduates enjoy hourly wage advantages between 10.7 (Denmark) and 34.6 (Germany) log points over non-completers (Berlingieri and Bolz, 2020: Table 2). For the United States, Flores-Lagunes and Light (2010) found that the hourly starting wages of college graduates were 16.1 to 22.5 log points higher than those of college non-completers after controlling for years of education and other covariates.[2]

These findings are consistent and robust, but they mostly come from observational studies based on supply-side (employee) data, which entails two key limitations. First, graduates might differ from non-completers in terms of characteristics that are missing in most observational studies yet observable to employers in the hiring process (e.g. grade point average [GPA] or job-relevant internship experience).

Received: November 2023; **revised:** April 2025; **accepted:** May 2025
© The Author(s) 2025. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

Second, employer perceptions and decision-making largely remain a black box, despite being central to the theoretical interpretation of sheepskin effects (Di Stasio and Van De Werfhorst, 2016; Bills *et al*., 2017; Neugebauer and Daniel, 2022).

To address these issues, we conducted a multifactorial survey experiment with 335 German employers who had recently sought to recruit employees for high-skill information technology (IT) and business jobs. In the experiment, employers assessed hypothetical job applicants who had either completed a degree or left higher education before program completion. In addition, our background questionnaire asked employers to indicate the extent to which graduates outperform non-completers in domains such as analytic competencies, perseverance, or occupation-specific skills. We use these unique data to contribute novel evidence on three interrelated research questions: *RQ1*: How large are sheepskin effects in a survey experimental setting? *RQ2*: What do 'sheepskins' signal, that is, in what domains do graduates outperform non-completers according to employers? *RQ3*: Do employer perceptions affect hypothetical hiring behaviour in the survey experiment?

Our first contribution is to provide experimental demand-side results on the magnitude of sheepskin effects. Our results complement—and in many ways improve upon—the observational supply-side evidence that dominates the extant literature: The experimental setup allows us to control the amount of information that is available to employers and to fix or experimentally manipulate key worker characteristics such as GPA, study-abroad experience, or job-specific skills. It thus minimizes concerns about unobserved differences between graduates and non-completers and more credibly identifies the independent effect of the 'sheepskin'.

Our second contribution is to provide novel evidence on how employers and recruiters actually interpret degree completion. While there is ample evidence for the existence of sheepskin effects, we do not know *why* employers prefer graduates over non-completers. What is it that the sheepskin signals? Our survey module on employers' perceptions of graduates vis-à-vis non-completers sheds much-needed light on this question.

Third and last, we show that employer perceptions matter in that they predict preferences for different types of candidates in the hiring experiment. This indicates that employers' self-reported perceptions are more than cheap talk and likely consequential for actual hiring behaviour. This realization makes them a promising target for further research efforts to understand the sources of education-related labour market inequalities and for interventions that seek to improve the employment prospects of HE non-completers.

Supply-side evidence and theoretical perspectives on sheepskin effects

Most prior empirical studies of sheepskin effects are based on supply-side (i.e. employee) data and find ample support for substantial graduation premia, particularly for the United States (Hungerford and Solon, 1987; Jaeger and Page, 1996; Flores-Lagunes and Light, 2010; Payne, 2023) but also for other countries (Pons and Blanco, 2005; Mora and Muro, 2008; Bol and Van de Werfhorst, 2011). This evidence typically takes one of two forms: in the absence of information on actual degree completion, scholars like Hungerford and Solon (1987) have pointed to marked discontinuities in the relationship between wages or earnings and years of education at thresholds corresponding to typical years of graduation, such as 12 for high school and 16 for college graduation in the US (see also Belman and Heywood, 1997). A second approach uses information on actual graduation and typically finds that it is strongly positively related to labour market attainment after accounting for time spent in education (Jaeger and Page, 1996; Flores-Lagunes and Light, 2010; Bol and Van de Werfhorst, 2011).

While these empirical regularities are largely uncontested, there is a lot of debate about their interpretation. One set of concerns revolves around *unobserved confounding*. Do the premia observed in supply-side analyses really represent a (causal) effect of the sheepskin? Or do they reflect differences between graduates and non-completers that are not recorded in most data sets but relatively easy to observe for potential employers? For example, graduates and non-completers might differ in terms of high school and university GPA, course choice, or extracurricular activities and work experiences, all of which are commonly reported in application materials but very imperfectly observed in most social science data sets.

Our first research question thus focuses on the importance of degrees in a survey experimental setting that allows us to control the information available to employers and render the average graduate comparable to the average non-completer. Do we continue to find large degree premia in this controlled setting?

Another recurrent theme in the literature on sheepskin effects is their theoretical interpretation. As noted by Bills (2003), Huntington-Klein (2020), and many others, sheepskin effects are often seen as evidence for signalling explanations. The idea here is that graduation might be a proxy for worker characteristics that are not only missing from most datasets but difficult to ascertain for employers as well. For example, graduates and non-completers might differ with respect to cognitive ability or non-cognitive skills. Such group differences would provide employers with a reason to

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

screen potential employees on the basis of their formal qualifications and engage in a kind of positive statistical discrimination of graduates. In the strong version of this argument, graduation signals pre-existing differences in skills that are valued by employers but are not actually taught or cultivated in the course of educational programs. That is, there is 'sorting' (Weiss, 1995) into graduation on the basis of characteristics such as general trainability, intelligence, or non-cognitive skills like motivation or 'grit' (i.e. an individual's passion and perseverance for a longer-term goal, see Duckworth *et al*., 2007).

Closely connected to this interpretation is a criticism of human capital models of labour market returns to education (Becker, 1964; Mincer, 1970). According to human capital theory, time spent in education pays off because it enhances the job-relevant skills of students and thereby, eventually, their productivity. For many, a human capital interpretation of sheepskin effects, therefore, comes down to the claim that the concluding phase of an educational program has a disproportionate impact on the skills of students relative to earlier years. But why should this be the case? Why should the final months of, for example, a college program have a much larger impact on the skills and subsequent productivity of students than the three preceding years?

While these arguments (and rhetorical questions) against human capital explanations have some immediate plausibility, there are also considerations that may help to reconcile the existence of sheepskin effects with human capital theory. First, in Germany as in many other countries, the concluding months of higher education programs often involve special tasks that might have a disproportionate impact on human capital (e.g. thesis writing, high-stakes final examinations). Second, Chiswick (1973) and others have argued that graduates might differ from non-completers in being more efficient learners. If that is the case, students who will eventually graduate will accumulate human capital at higher rates than eventual non-completers throughout their course of study, including the initial months and years of an educational program. Human capital differences between graduates and non-completers would then reflect the combined effect of graduates' faster learning during schooling episodes shared by both groups as well as the additional skill increments experienced in the course of program completion.[3]

A third influential perspective on sheepskin effects reads it as evidence of credentialism (Collins, 1979; Di Stasio and Van De Werfhorst, 2016). According to Bills (2003: 452), the 'credentialist thesis holds that formal schooling leads to socio-economic success not because of the superior skills and knowledge of the more highly educated, but rather because of the ability of the highly educated to control access to elite positions'. This interpretation is closely linked to the concept of occupational closure (Weeden, 2002; Bol and Van de Werfhorst, 2011; Bol and Weeden, 2014). As stated in Sørensen's influential rent-extraction theory of social class, 'educational credentials, used as rationing devices […], create monopoly rents to those holding the credentials' (2000: 1544). This account appears particularly plausible when access to (well-paying) occupations is restricted to those holding specific degrees and/or when the opportunity to obtain degrees with substantial labour market values is heavily rationed (Ketel *et al*., 2016).

Demand-side evidence and contributions of the present study

Following Bills *et al*. (2017), we argue that our understanding of sheepskin effects can be greatly advanced by paying closer attention to the demand side, that is, to employer perceptions and behaviour. Employers are key gatekeepers, eventually making the decision on who ends up with which job and on what conditions. With the observational supply-side data on employee characteristics and outcomes used in most prior research on sheepskin effects, employer behaviour remains unobserved and evidence for alternative theoretical interpretations tends to be rather indirect.[4]

Our paper adds to a growing literature that uses field and survey experimental designs to provide direct evidence on employer perceptions and behaviour toward HE attendees. Previous work has investigated the effects of for-profit degrees (Deterding and Pedulla, 2016), abroad experience (Petzold, 2017), and various other markers of occupation-specific and general skills (e.g. Humburg and van der Velden, 2015; Piopiunik *et al*., 2020). However, these studies have been restricted to variation among graduates and are therefore not informative about the extent and sources of sheepskin effects.

To our knowledge, only two prior studies have included both graduate and non-completer profiles in their experimental conditions. Di Stasio and van de Werfhorst (2016) administered applicant vignettes to 72 employers in the IT sector in the United Kingdom and the Netherlands. One of their treatments described applicants as having completed their program on time, with a delay of two years, or not at all. Having a degree (whether completed on time or with delay) had a large positive effect on reported hiring propensities in their analysis. Di Stasio and van de Werfhorst also asked their respondents if education signals 'job-relevant skills' or 'trainability' and whether the reputation of the attended school signals 'technical ability'. Yet, while these questions do provide initial evidence on relevant employer perceptions, the authors did not

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

attempt to link perceptions to hiring intentions in the survey experiment and were generally limited by their relatively small employer sample.

Neugebauer and Daniel (2022) drew on a factorial survey experiment with 1382 German employers recruiting for IT and business jobs, again including HE non-completers. A key focus of their study was to estimate the hiring chances of non-completers across different labour market segments and relative to typical competitors in each segment (high school graduates for apprenticeship positions; applicants with a vocational degree for skilled worker positions; HE graduates for graduate positions). They found that the employer-reported hiring probability for graduate positions is approximately 25 percentage points lower for HE non-completers than for graduates. Neugebauer and Daniel did not consider direct questions on employers' perceptions of graduates *vs*. non-completers or of the value of education in general.

While both these studies provide important evidence on sheepskin effects, they also have several limitations that we seek to address below. For example, the timing of dropout is unspecified in Di Stasio and Werfhorst's (2016) applicant profiles. A clear answer to *RQ1* (*How large are sheepskin effects in a survey experimental setting that minimizes biases due to unobserved confounding?*) requires that we separate study duration from program completion. More importantly, the two prior studies include no or only very limited measures of how employers view graduates vis-à-vis non-completers. We use a novel survey module on employer perceptions to contribute to theoretical debates about the nature of sheepskin effects and address *RQ2*: *What is it that degree completion signals to employers?* Is degree completion primarily seen as an indicator of (pre-existing) differences in general cognitive and/or non-cognitive skills, as suggested by signalling explanations? Or do employers view it as indicative of (occupation- and field-specific) skills taught in HE, which would seem to support human capital interpretations of sheepskin effects? In a further step, we will then link these survey questions to the experimental data. Our goal here is to better understand employer heterogeneity in the treatment of (hypothetical) applicants with and without a degree by providing an answer to *RQ3* (*Do employer perceptions predict hypothetical hiring behaviour in the survey experiment?*).

Study context: IT and business jobs in Germany

Our analysis focuses on entry-level jobs for university graduates in IT and business. Confining the experiment to two occupational fields facilitates the design of realistic vignettes that resemble real applicants (e.g. by mentioning specific programming skills rather than generic 'job-relevant skills' for IT applicants). We selected IT and business jobs because they are in high demand across a broad variety of firms, regions, and economic sectors. This allows us to survey a heterogeneous employer sample that should ensure good generalizability. In the concluding section, we discuss how sheepskin effects may manifest in other sectors.

The national context of our study is Germany, where school leavers need an HE entrance qualification to attend university. In most cases this is the so-called *Abitur*, typically obtained after 12 or 13 years of general schooling.

The standard period of study for a bachelor's degree is 6 semesters, but about one-third of all German HE entrants leave without graduating (OECD, 2013, p. 71). 47% of non-completers leave HE in the first two semesters but for 13% this takes seven semesters or longer (Heublein *et al*., 2017). While most HE non-completers subsequently pursue vocational education and training (VET) at the upper secondary level, direct entry into skilled employment is not uncommon, especially for those with prolonged study episodes (Heublein *et al*., 2017; Tieben, 2024). Moreover, it remains an open question to what extent enrolment in VET is a response to the (experienced or anticipated) labour market penalties for non-completion that are the focus of our study.

The most common way for non-completers to find a job is to respond to job advertisements (Heublein *et al*., 2017: 232), which is also reflected in a recent survey where a majority of employers indicated that they sometimes (66%) or often (22%) receive applications for graduate jobs from HE non-completers (Neugebauer *et al*., 2021, p. 19).

Germany is a country known for its strong linkages between educational credentials and occupations (DiPrete *et al*., 2017). This coupling is strongest for graduates of Germany's 'dual' (apprenticeship) system of vocational educational training, but also holds for HE graduates (Leuze, 2007; DiPrete *et al*., 2017). The forces underlying these strong education-to-work-linkages likely include non-binding and potentially implicit practices and conventions as well as binding regulations from general legislation, collective bargaining agreements, or organizational rules and codes.

These considerations suggest that sheepskin effects in Germany should, (a), be large by international standards and, (b), be partly attributable to the fact that many jobs are effectively restricted to individuals with specific degrees (i.e. to occupational closure). Regarding the first point, Berlingieri and Bolz (2020) indeed find Germany to have the largest wage gap between HE graduates and non-completers in their comparative study of 18 European countries. In our

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

analysis below, we will explore the role of occupational closure using employer self-reports of whether they would be permitted to hire non-completers and by restricting parts of the analysis to employers who say they could do so.

Data and methods

We study sheepskin effects using an online employer survey that combined a factorial survey experiment (Auspurg and Hinz, 2014) with a background questionnaire. The survey experiment presented employers with fictitious CVs of applicants (i.e. vignettes) and asked them to indicate how likely they would be to invite the applicant for a job interview. The experimental setup provides full control over the information available to employers. Unlike observational studies, we can therefore rule out unobserved confounding by design. Moreover, randomization ensures that CV characteristics such as academic performance and HE completion are uncorrelated across vignettes (see next section for details), while they are likely correlated in practice and partially unobserved in most employee data sets. The background questionnaire, which immediately followed the vignettes, provides crucial additional information on employers' perceptions of graduates and non-completers as well as additional information on the firm context.

Respondent sample

To increase external validity, our study was targeted at respondents with actual involvement in recruitment. We first constructed a list of nine entry-level job titles for the two occupational fields, such as 'computer scientist' in IT, or 'fund manager' in business (see Table A1 in the Online Appendix). Based on this list, we then collected all job postings advertised on the online job portal of the German Federal Employment Agency and on two leading private job platforms over a period of eight months (September 2017 to May 2018). After removing multiple employer entries from this list, we drew a random sample of 2000 job postings from each occupational field (4000 in total) and invited the contact persons to our survey by email. 335 respondents (65% female) participated in the survey (response rate 8.4%).

Table A2 in the Online Appendix shows that respondents and non-respondents are broadly comparable with respect to key characteristics that could be proxied from information provided in the job postings (e.g. company size, location, sector, or gender of the contact person). Only for company size do we find a substantial difference between responders and non-responders, with the former being somewhat more likely to work in small firms with < 50 employees and less likely to work in larger firms. We acknowledge that the list of observable characteristics is quite short and that non-respondents might differ in terms of further unobserved characteristics that shape their relative assessment of completers and non-completers. However, it is worth pointing out that the invitation email mentioned only a general interest in recruitment decisions and made no reference to comparisons of these two groups, so systematic self-selection into the survey based on respondents' views of these groups seems relatively unlikely.

Table 1 shows that respondents are distributed across firms of different sizes and in different economic sectors. They mainly work as general or human resource managers, and over 98% have been involved in personnel selection at some point. On average, respondents have 8.7 years of experience in employee recruitment. We refer to them as 'employers' in the remainder of the paper.

Experimental design

We first prompted employers with one of nine standardized hypothetical job offers, customized to match the specific job title they had recently advertised (for an example, see Figure A1 in the Online Appendix). The text of the job offer was based on the results of a qualitative content analysis of twenty randomly selected real-world vacancies for each job title.[5] We then showed each respondent a set of eight CVs (vignettes) of fictitious applicants, all of whom had the same level of general schooling (the general HE entrance qualification *Abitur*) and all of whom had attended HE (albeit sometimes without completing their program). In contrast to, for example, the United States, there are hardly any status differences between universities in Germany, which is why we did not specify the name of the university. In a preface to the vignettes, we informed respondents that all applicants were labour market entrants who had convincingly stated their interest in the job in an error-free cover letter.

After each vignette, we asked employers how likely they would invite the candidate to interview for the job we had described (answers were recorded on an 11-point scale from 0% to 100%). This question is the basis for our main outcome variable, the invitation probability. We also asked employers to estimate a candidate's starting salary in an open-ended format: 'Regardless of whether you would invite the applicant, what gross annual salary (in €) would you find appropriate for this applicant?'[6].

Table 2 lists the CV attributes that were randomly varied across vignettes, along with their respective levels (for further details, see Online Appendix B). These attributes capture the characteristics that are crucial in candidate selection at this stage, according to both

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

Table 1 Employer sample

***N* Employers**	**335**
N Vignette ratings	2680
Employer characteristics % for binary and M (SD) for continuous variables	
Female	65.37
Age	39.22 (10.58)
Function	
General management	17.66
Human resources	76.65
Operating department	4.79
Other	0.90
Managerial responsibility (y/n)	54.02
Involved in personnel selection (y/n)	98.01
Experience (years) in personnel selection	8.70 (7.66)
Firm characteristics	
Firm size	
0-49 employees	26.33
50-499 employees	38.81
≥ 500 employees	34.87
IT jobs (vs. business jobs)	54.03
Sector[a]	
Manufacturing (including Construction, and Utilities)	11.49
Market services (Trade, Transportation, Business services)	50.21
Non-market services (Public administration, Social services)	38.30
Candidate supply[b]	2.65 (0.70)
Closure[c]	32.54

Note: [a] Broad economic sector classification by the International Labour Organization (ILO). [b] 'How do you perceive the current applicant situation with respect to suitable candidates for [POSITION]?' (1-no suited applicants to 5-many suited applicants). [c] Is it possible in your organization to fill the position of a [JOB TITLE] with someone who does not meet the formal qualifications (completed tertiary degree)?' 32.54 % state it is not possible, i.e. the position is closed to non-completers. Descriptive statistics are based on M = 10 imputed datasets.

Table 2 Vignette attributes and levels

(1)	Education: 1 = Graduation after 6 semesters, 2 = non-completion after 6 semesters, 3 = non-completion after 2 semesters
(2)	Academic performance (GPA) in university: 0 = bad (3.3), 1 = good (1.7)
(3)	Job-relevant internship: 0 = no, 1 = 3 months internship
(4)	Informally acquired skills: 0 = no, 1 = programming/stock consulting
(5)	Matching study focus[a]: 0 = low, 1 = high
(6)	Abroad experience: 0 = no, 1 = studied abroad for 1 semester
(7)	English skills: 0 = intermediate (B2), 1 = proficient (C1)
(8)	Academic performance (GPA) in high school: 0 = bad (2.9), 1 = good (1.8)
(9)	First name [b]: 0 = lower class (e.g. Kevin), 1 = middle class (e.g. Julius)

Note: [a] Low match = unfitting study focus within a fitting field of study (e.g. computer sciences major with focus on system integration applies for a job as computer science expert in the area of software development as opposed to in the area of system integration). High match = high fit between study focus and job. [b] First names are connoted with social background, as one of our pre-studies showed, and may thus evoke reactions that could influence the evaluation of applicants. To keep those influences under control, we selected German first names which have been shown to be associated either with lower class or middle class. Lower class names comprised 'Justin', 'Kevin', 'Pascal' and 'Steven', whereas 'Jakob', 'Felix', 'Konstantin' and 'Julius' represented middle class names. Applicants' last name was randomly sampled from a list of the 100 most common family names in Germany.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

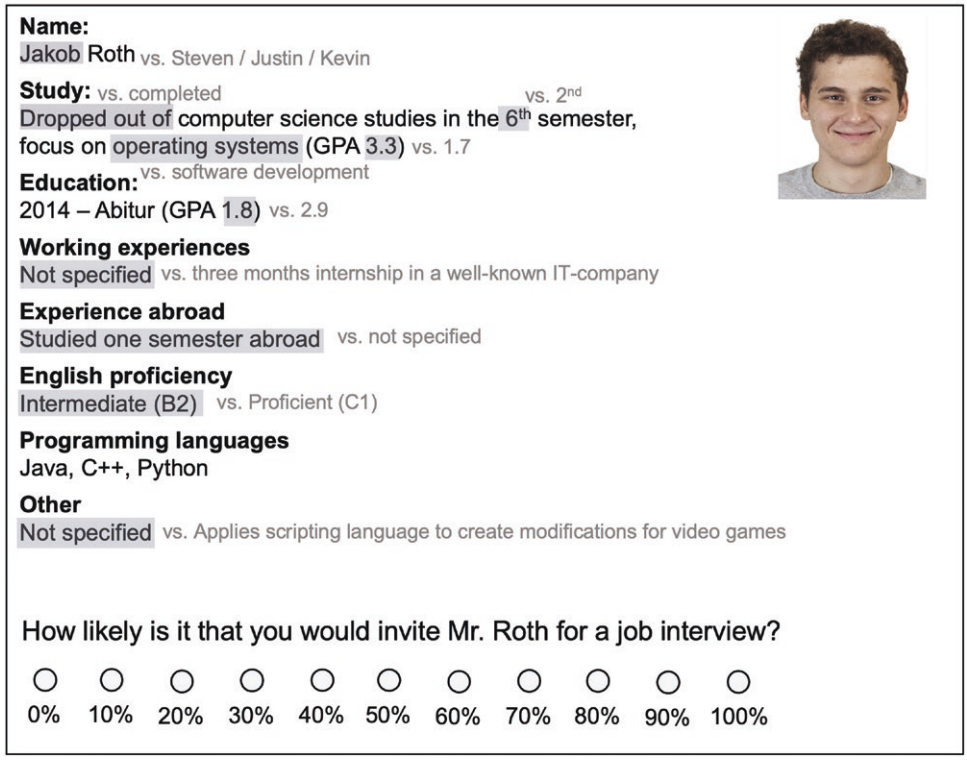

Figure 1 Example vignette (non-completer, IT)

Note: Attributes that were varied are shown with a grey background, and alternative levels are listed in lighter font (respondents did not see these alternatives).

previous research (Humburg and van der Velden, 2015; Piopiunik *et al*., 2020) and a qualitative pre-study with employers and job counsellors.[7] All CV attributes have two levels, except for education which takes three values: Graduation after 6 semesters, non-completion after 6 semesters (late non-completion), and non-completion after 2 semesters (early non-completion).

To create an orthogonal design, we proceeded as follows: We first drew a fraction of 64 vignettes from the universe of non-completer vignettes and allocated them to 16 sets of four vignettes, each containing two early and two late non-completion profiles. We then added four graduate vignettes to each set by mirroring the non-completion vignettes. That is, the graduate vignettes resembled the non-completer vignettes within each set, except in terms of education and the applicant's name and photograph (see next paragraph). Each vignette set thus comprised four non-completer vignettes (2 with an early and 2 with a late dropout) and four graduate vignettes. The sets were randomly assigned to employers, and vignettes were presented in random order to prevent primacy and systematic learning effects. Table A3 in the Online Appendix shows that the vignette dimensions are effectively orthogonal, and Figure A10 demonstrates that the effects of the (non-)completion show no strong trend from initial to later vignettes.

In order to not distract respondents from the attributes of primary interest, we used only male applicants with standard German-sounding names (see e.g. Quadlin, 2018 for evidence on gendered hiring standards). Each vignette was displayed graphically as a one-page CV resembling real-world examples. In Germany, CVs typically include a photograph of the applicant. Photos in our studies were selected from the Chicago Face Database (Ma *et al*., 2015) and were rated equally in terms of age and attractiveness by an independent rater sample. Figure 1 shows an example vignette for an IT position, translated into English and listing alternative attribute levels in a lighter font (respondents did not see these alternatives).

Two features of the vignettes warrant further discussion. First, while the ordering of vignette attributes is often randomized to avoid order effects (Auspurg and Hinz 2014), we decided not to do this in order to closely mimic real-life CVs, which tend to follow a relatively consistent structure. Second, as shown in Figure 1, our key treatment (study completion vs. non-completion)

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

not only appeared at the top of the vignettes, especially the non-completion conditions were also phrased more explicitly than they would typically be in real-life settings. Specifically, our fictitious applicants were characterized as 'dropouts', whereas real applicants might simply note that they studied a certain subject for a certain period of time, without explicitly describing themselves as non-completers. Does this explicit labelling of 'dropouts' compromise the external validity of our experiment? While it is impossible to give a definitive answer to this question, we believe that this concern should not be exaggerated for at least two reasons. First, employers in our qualitative pre-study confirmed that non-completers occasionally attempt to conceal their dropout status but noted that they would generally detect such cases anyway before extending an invitation for an interview (e.g. by taking a closer look at the application documents or requesting clarification). Second, our survey experiment is a low-stakes setting compared with the higher-stakes, real-life setting of screening actual applicants. It is therefore not obvious that mimicking the real-life CVs as closely as possible is the best strategy to ensure generalizability. Quite to the contrary, it might also be the case that subjects (i.e. employers) are less attentive in the low-stakes survey experiment and that the same piece of information therefore needs to be conveyed more directly to have the same effect.

Analytic strategy

To *identify sheepskin effects (RQ1)*, we compare the average invitation probability and starting salary of fictitious applicants with and without a degree. The crucial comparison for identifying the sheepskin premium is the one with applicants who left their program after six semesters (the standard time to degree) and thus attended HE for the same amount of time as graduates. The difference between those who terminated their studies after the sixth semester and those who did so after the second is informative about the value attached to HE attendance without graduation and provides a useful benchmark for the graduation effect. Importantly, all other candidate characteristics, including high school and university GPA are independent of the education treatment by design (we nevertheless include them in all regression models to improve statistical precision).

To *measure what degree completion signals* (RQ2), we rely on the following question from the background questionnaire: 'To what extent do you perceive non-completers to have an advantage or disadvantage over graduates on the following characteristics?' The question was followed by a list of 12 domains, again derived from our qualitative pre-study. Below we focus on the nine items that captured (cognitive and non-cognitive) skill dimensions: professional qualifications, theoretical knowledge, analytical thinking, verbal skills, resilience, perseverance, ability to work in a team, work motivation, and personal maturity. Results for the remaining three domains (labour costs, retention prospects, and age) are provided in the Online Appendix (Figure A2). Respondents provided their answers on a 5-point scale ranging from -2 (disadvantage of non-completers relative to graduates) to 2 (advantage), with 0 indicating that non-completers neither have a disadvantage nor an advantage relative to graduates on a given dimension. For the analysis below, we reversed the coding of these measures so higher values correspond to larger perceived advantages of graduates.

To *investigate if employer perceptions matter for hypothetical hiring behaviour* (*RQ3*), we examine cross-level interactions between respondent-level perceptions and degree status. Since employer perceptions turn out to be highly correlated across domains, we reduce dimensionality by averaging domain-specific skills. We consider both an *overall perceptions* variable that averages employer perceptions across all domains as well as more disaggregated versions that capture the following more coherent subsets of skills: *general cognitive skills* (analytical thinking, verbal skills), *occupation-specific skills* (professional qualifications, theoretical knowledge), *grit* (perseverance, resilience), and *other non-cognitive skills* (teamwork, work motivation, personal maturity). These skill subsets extracted from our qualitative pre-study are largely in line with what employers typically seek in new employees (e.g. Deming and Kahn, 2018; Jenkins and Wolf, 2018), even though there is little agreement in the literature about how non-cognitive skills should be defined and measured.

Estimation and multiple imputation procedure

We use linear multilevel (mixed) regression models to account for the nested data structure (vignette ratings nested in employers). Our basic specifications include all vignette attributes, and various employer-level covariates that may affect hiring chances, such as candidate supply, firm size, sector, and occupational field (see Table 1 for details).We further include vignette set and order as controls (Su and Steiner, 2020). We include random intercepts at the employer level. To ensure accurate statistical inference for the cross-level interactions between employer perceptions and degree completion we also include random slopes for the latter variable (Heisig and Schaeffer, 2019).[8] We report results in terms of predicted values, with all other covariates set to their respective means. Note that predicted invitation probabilities are obtained by converting the underlying 11-point (cf. Figure 1) scale

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

into probabilities and treating the resulting variables as continuous (and not from a probability model applied to a binary outcome).

We used multiple imputation via chained equations with 10 imputations to fill in missing values on both vignette ratings (4% for invitation probability, 11% for salary) and covariates (ranging from 6% for firm size to 22% for the graduate/non-completer comparison regarding verbal skills). The imputation models included all variables from the analysis and additional (i.e. auxiliary) employer characteristics collected in the questionnaire and inferred from the vacancies (e.g. experience in personnel selection or population at the firm location). We used predictive mean matching to impute continuous and ordered polytomous variables (e.g. invitation probability) and logistic regression for dichotomous variables.[9]

Results

RQ1: the importance of degrees for hypothetical hiring behaviour

Table 3 shows predicted invitation probabilities (in percent) and starting salaries (in EUR) for non-completers who left after 2 semesters (early non-completers), non-completers who left after 6 semesters (late non-completers), and graduates (after 6 semesters). We provide the full model results underlying these predictions, including coefficient estimates for all applicant characteristics, in Table A4 in the Online Appendix. Degree-holders enjoy a marked advantage over late non-completers, with a predicted invitation probability of 67.4% for the former and of only 38.5% for the latter. This is a striking difference in both absolute and relative terms: graduates' invitation probability exceeds that of late non-completers by approximately 29 percentage points or a factor of 1.75, despite the fact that both groups have attended HE for six semesters and are comparable in terms of key characteristics such as high school and university GPA.

The difference between the predicted invitation rates for late non-completers (38.5%) and early non-completers (28.4%) is smaller, with an absolute difference of about 10 per cent points and a probability ratio of 1.35. At the same time, the fact that we still find a meaningful difference between these two groups indicates that employers do value additional time spent at university. In some sense, this renders the much larger premium associated with graduation even more striking.

The completion advantage is also large relative to other applicant characteristics manipulated in the survey experiment. For example, university graduates with a good to very good GPA of 1.7 have an invitation probability of 73.3%, while graduates with a much lower GPA of 3.3 still have an invitation probability of 61.8%, setting all other applicant characteristics to their means (in the German system, passing GPA ranges from 1.0 to 4.0 with lower values indicating better performance).

Column 3 in Table 3 shows that graduates enjoy a large salary premium as well. According to our sample of employers, graduates would earn almost 6000 EUR more per year than late non-completers, whereas the latter could expect an additional 2000 EUR relative to early non-completers. This corresponds to advantages of, respectively, 17% and 6% relative to the lower-earning group. These estimates are broadly similar to benchmarks from observational studies, such as Flores-Lagunes and Light's (2010) analysis for the US, which finds starting wage gaps between 16 and 23 log points.[10]

As noted in the 'Study context' section, Germany is known for its tight coupling of formal qualifications and labour market attainment. Many occupations are difficult or even impossible to access without the appropriate certificates, in terms of both level of degree and field of study. This can be seen as a form of institutionalized credentialism, and it is natural to ask if and to what extent it drives the results presented so far.

Table 3. Predicted invitation probability (in %) and gross annual salary (in EUR)

	Invitation probability (in %)		Gross annual salary (in EUR)	
	Margin	(SE)	Margin	(SE)
Degree-holders (6 semesters)	67.4	(1.19)	40,542	(329)
Non-completers (6 semesters)	38.5	(1.35)	34,687	(331)
Non-completers (2 semesters)	28.4	(1.29)	32,655	(349)
N (employers)	335		335	
N (vignettes)	2680		2680	

Note: Predicted invitation probabilities and gross annual salaries from linear multilevel models with employer-level random intercept and random slopes for degree-holder/non-completer variables. All control variables were set to their respective means in calculating the predictions. Controls include other vignette dimensions, sector, occupational field, firm size, candidate supply, vignette set and order.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

We therefore asked employers whether they would be able to hire a non-completer for the advertised position in principle. Unlike for some professions such as physician or lawyer, there are no universally binding regulations that would make graduation a necessary requirement for practicing the jobs considered in our experiment. However, many employers are likely restricted by firm-level policies or collective agreements that mandate a completed degree. One third of our respondents actually stated that they could not hire a non-completer for the advertised position, whereas the remaining two thirds said that this was possible (cf. Table 1 above).

Table 4 shows predicted invitation probabilities and annual starting salaries based on the 226 employers in the latter group (i.e. those who say they could hire a non-completer). As one would expect, the difference in invitation probabilities between graduates and late non-completers is smaller in this 'no closure' sample than in the full employer sample (cf. Table 3 above), providing some evidence that formal degree requirements as a form of institutionalized credentialism play a role. The difference is not large, however. The invitation probability for degree-holders is virtually unchanged (67.6% in Table 4 vs. 67.4% in Table 3), while the one for late non-completers increases from 38.5% in the full (Table 3) to 43.5% (Table 4) in the reduced sample. This still amounts to an advantage of 24 per cent points or a factor of 1.56 in relative terms. The larger part of the sheepskin effect enjoyed by graduates thus *cannot* be attributed to rules or regulations that make graduation a mandatory prerequisite for employment. Interestingly, group-specific starting wages for the 'no closure' subsample are almost identical to those for the full sample, suggesting that employers who would not be able to hire non-completers abstracted from the fact they would not be able to do so when estimating the starting wages of the fictitious candidates.

RQ2: what does degree completion signal?

What explains the marked invitation and earnings premia enjoyed by graduates? What is it that makes them more attractive to employers than non-completers? Figure 2 provides an answer to these questions by summarizing the responses to our unique survey battery on employers' perceptions, with positive values indicating a perceived advantage and negative values a perceived disadvantage of graduates relative to non-completers.

Two domains, captured by two items each, stand out in terms of graduates having a clear advantage. The first is 'occupation-specific skills', measured by the items 'professional qualifications' and 'theoretical knowledge'. The second one is 'grit', captured by the items 'resilience' and 'perseverance'. The internal consistency of these two domains is high, as indicated by Cronbach's α of.83 and.90, respectively.

The remaining items actually show some advantages of non-completers over graduates, although most differences are quite small and statistically insignificant. The only domain where employers perceive non-completers to be clearly ahead of graduates is when it comes to being 'able to work in a team'. Attributed differences in 'general cognitive skills', captured by 'analytical thinking' and 'verbal skills', are rather small, particularly in the former case. The same holds for 'work motivation' and 'personal maturity' as further non-cognitive skill domains.

It is tempting to relate these assessments of graduates vis-à-vis non-completers to the theoretical explanations of sheepskin effects debated in the literature. Overall, both human capital and signalling accounts appear to receive some support. Graduates' clear perceived advantage with respect to occupation-specific skills seems more consistent with the human capital perspective, as these are precisely the types of skills that university programs are designed to develop (Bol and Heisig, 2021). The perceived advantage in terms of

Table 4 Outcomes for 'no closure' sub-sample

	Invitation probability (in %)		Gross annual salary (in EUR)	
	Margin	(SE)	Margin	(SE)
Degree-holders (6 semesters)	67.6	(1.42)	40,352	(386)
Non-completers (6 semesters)	43.5	(1.69)	34,792	(405)
Non-completers (2 semesters)	33.3	(1.57)	32,796	(422)
N (employers)	226		226	
N (vignettes)	1808		1808	

Note: Predicted invitation probabilities and gross annual salaries from linear multilevel models with employer-level random intercept and random slopes for degree-holder/non-completer variables. All control variables were set to their respective means in calculating the predictions. Controls include other vignette dimensions, sector, occupational field, firm size, candidate supply, vignette set and order.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

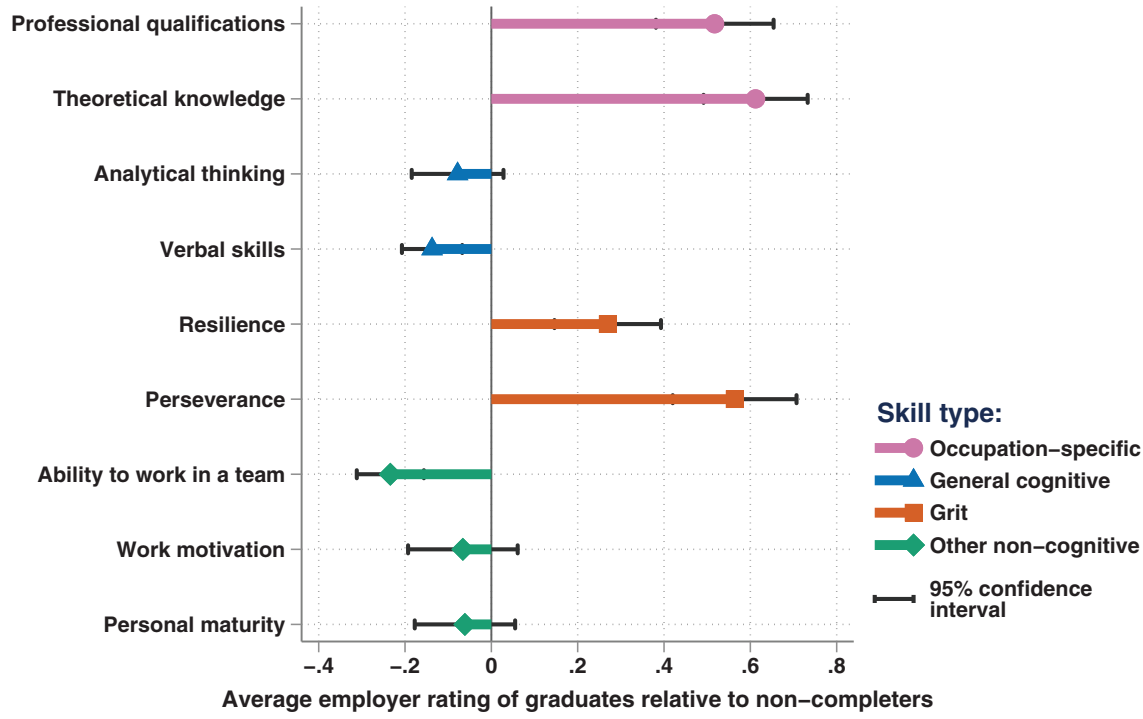

Figure 2 Perceived strengths and weaknesses of graduates relative to non-completers.

Note: Mean ratings and 95% confidence intervals based on five-point scale ranging from -2 (strong disadvantage of graduates) to 2 (strong advantage of graduates). *N* (employers) = 335.

grit, on the other hand, seems to better fit the idea that program completion signals pre-existing differences in non-cognitive. In that respect, it also interesting to see what kinds of (pre-existing) skills graduation does *not* seem to signal to our employers, namely a higher level of general *cognitive* skills.

These interpretations are tentative. For example, we cannot rule out that a trait such as perseverance is, at least to some extent, fostered and cultivated in the course of program completion and through the associated experiences that are not shared by non-completers—for example, the production of a thesis or some other kind of self-directed final output that marks the end of most programs of study. Relatedly, we cannot say anything about the extent to which graduates' perceived advantages in terms of occupation-specific skills originate from more efficient learning over their whole course of studies (Chiswick, 1973) or from their unique experiences during the completion phase.

RQ3: effect of employer perceptions on invitation probability

We now turn to the final step of our analysis. Do employer perceptions of graduates' strengths (and weaknesses) relative to non-completers matter for invitation probabilities? While Figure 2 displayed the average assessment of graduates vs. non-completers, we now focus on variation in these perceptions across employers and examine whether they predict the graduation premium in the experimental part of the survey. Technically, we estimate two models that include the scores for 'occupation-specific skills' and 'grit' (cf. Figure 2), obtained by averaging the underlying items, and interact them with the degree-holder/non-completer dummies, including the different skill domains one at a time.

Hainmueller *et al.* (2019) emphasize how conventional multiplicative interaction terms imply that the effect of the treatment variable (graduation) changes linearly over the full range of the moderating variable (employer perceptions). We assessed the plausibility of this assumption using Hainmueller *et al*'.s (2019) kernel smoothing estimator and found clear evidence that it does not hold (see Figures A3 and A4 in the Online Supplement). Specifically, the perception measures appear to be only weakly related to the magnitude of the completion vs. non-completion effect in the negative half of the employer perceptions scale. This negative half captures varying degrees of perceived advantages for non-completers over graduates. Approximately 20% of the employers in our sample express perceptions that fall into this part of the scale, with some

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

variation across skill domains. In the upper half of the perception scales, the graduation premium changes markedly and approximately linearly as employers ascribe increasingly strong advantages to graduates. In light of these results, we modelled the effect of the perceptions measures using linear splines with a single knot at 0, interacting both spline terms with the graduation vs. non-completion indicators. That is, we modelled the interaction as piecewise linear, allowing for a single change in its strength at 0—the scale point at which employer sees graduates and non-completers as roughly comparable in a given domain.

Figure 3 shows results for the invitation probability based on the full employer sample. The predicted invitation probability of graduates does not change much as their perceived advantage in terms of grit (left panel) or occupation-specific skills (right panel) increases. While increasing slightly, it falls between 65 and 70% over the full range of the employer perception measures. Employers' perceptions are much more important for the invitation rate of late non-completers. While showing no clear trend in the sparsely populated lower half of the scale, it declines sharply in the more heavily populated upper half—from over 40% for employers that see no advantage or disadvantage of graduates over non-completers to around 30% for those who see a very large advantage for graduates. This pattern is broadly similar for early non-completers, albeit on a lower overall level. Interestingly, the results in Figure 3 imply that even those employers who rate non-completers much more favourably than graduates (i.e. those located towards the left ends of the graphs) are considerably more likely to invite graduates. This could be read as evidence of strong credentialism in the German labour market. However, it should be kept in mind that these comparisons are based on relatively small subset of employers.

We provide additional results in the Online Appendix: Figure A3 shows predicted starting salaries for the full employer sample, while Figures A4 and

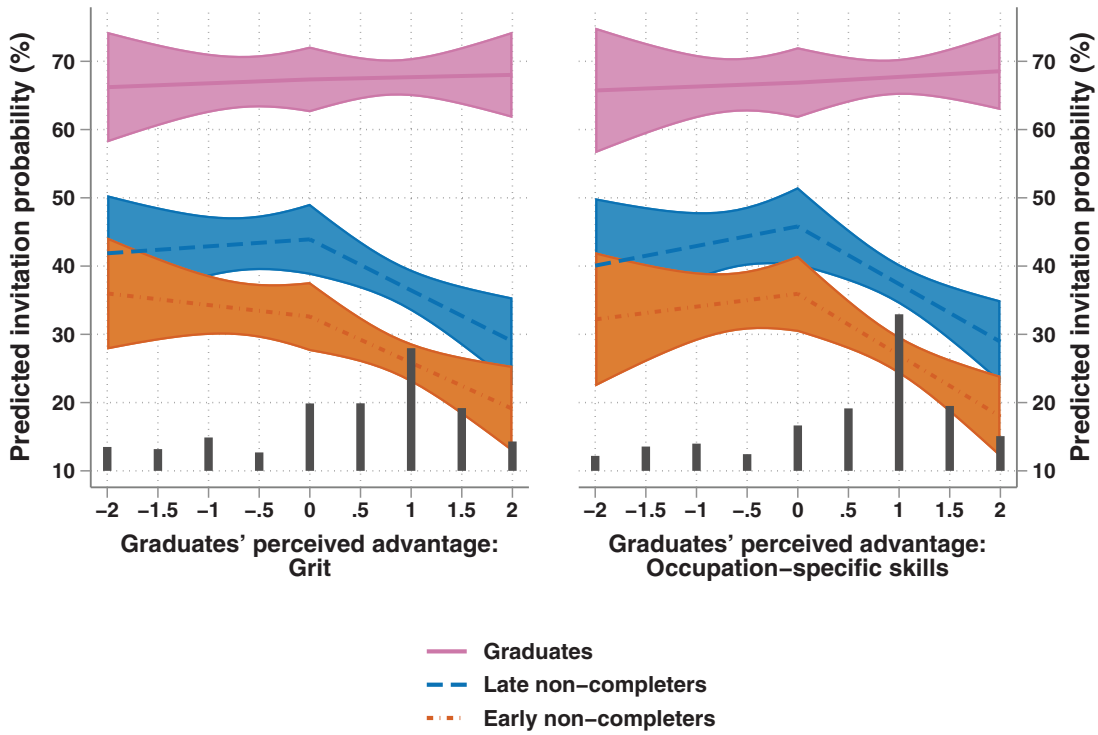

Figure 3 Employer perceptions (grit and occupation-specific skills) and invitation probability

Note: Predictive margins and 95% confidence intervals based on linear mixed effects models with interaction effects between graduates' perceived advantage in the grit domain (left graph) or occupation-specific skills domain (right graph) and the graduation status of the fictitious applicant. The perceived advantage measures are modelled using linear splines with a single knot at zero to capture changes in the magnitude of the interaction effect around this point (see Figures A3 and A4 in the Online Supplement). Perceptions are based on five-point scale ranging from -2 (strong disadvantage of graduates) to 2 (strong advantage of graduates). Models include an employer-level random intercept and random slopes for the graduation/dropout indicators. Controls: other vignette dimensions, sector, occupational field, firm size, candidate supply, vignette set and order. Vertical bars represent employer-level distributions of the perceptions measures (relative frequencies). *N* (employers) = 335; *N* (vignettes) = 2680.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

A5 show, respectively, predicted invitation probabilities and starting salaries for the no closure subsample. Results are qualitatively similar to those in Figure 3 in that the gap between graduates and non-completers grows as employer perceptions become more favourable towards the former. The most noteworthy difference to the invitation probability results in Figure 3 is that for starting wages this growing sheepskin effect is driven by increasing wages for graduates (rather than declining ones for non-completers).

While these are interesting results, we acknowledge that our ability to disentangle the relationships between *domain-specific* perceptions and invitation probabilities is very limited. Employer perceptions are highly correlated across items and, by implication, across domains (the Pearson correlation between the grit and occupation-specific skills measures used in Figure 3 is .72). The striking similarity of the relationships for grit and occupation-specific skills in Figure 3 already is an indication of this, and it becomes very clear when we run a (principle-component) factor analysis on all perception items displayed in Figure 2 above: This factor analysis yields a one-factor solution with an eigenvalue of 6.0 and factor loadings between .75 and .87. This means that some employers see relatively small differences between graduates and non-completers across all skill dimensions, while others perceive strong disadvantages for non-completers across all dimensions. Accordingly, the invitation premium for graduates is greater for these employers. Figures A6 to A9 in the Online Appendix accordingly depict predictions from models using an overall employer perceptions score obtained by averaging all nine skill domain items, both for invitation probabilities (A6, A8) as well as starting salaries (A7, A9) and for the full (A6, A7) as well as the 'no closure' employer sample (A8, A9). Results are very similar to those in the two panels of Figure 3.

In summary, this last step of the analysis shows that employer perceptions matter: those who view graduates more favourably relative to non-completers report a lower propensity to invite the latter to a job interview. When it comes to which types of perceptions play a role, our results suggest that it is difficult to single out a specific domain. High inter-item correlations suggest that those employers who see larger skills differentials between graduates and non-completers tend to do so more or less 'across the board'.

Conclusions

We have implemented and analysed a vignette survey experiment to shed light on the large labour market premium associated with successful (degree-conferring) completion of educational programs and of HE in particular, a phenomenon often referred to as a 'sheepskin effect'. Our analysis of German employers recruiting in IT and business occupations yields three main findings.

First, in line with observational studies (e.g. Jaeger and Page 1996; Flores-Lagunes and Light 2010), we find large sheepskin effects in a survey experimental setting that gives us full control over the information available to employers, both for reported invitation probabilities and for estimated starting wages. In our experiment, the annual earnings premium enjoyed by graduates is 17 % relative to late non-completers who left HE after 6 semesters and 24 % relative to early non-completers who left university after 2 semesters (cf. Table 3 above). These numbers are broadly comparable to observational estimates (see, e.g. Flores-Lagunes and Light 2010; Berlingieri and Bolz 2020), which could be read as evidence that observational studies tend to provide good approximations to the causal effect of program completion. Another interesting reference point is the study of DiStasio and van de Werfhorst (2016), who used a survey experiment broadly comparable to ours to study employer-reported *hiring* (rather than invitation) probabilities for fictitious graduates, including completers and non-completers. Back-of-the envelope calculations suggest that the effects on the log odds reported in this study translate into dropout penalties of ca. 22 per cent points in England and ca. 31 per cent points in the Netherlands, the latter being broadly comparable to the invitation penalty reported in Table 3 above. While interesting, these comparisons need to be taken with a grain of salt, because of differences in the countries and occupational fields studied, in the experimental treatments, and in the outcome variable.

Second, the average employer perceives degree-holders to strongly outperform non-completers in terms of occupation-specific skills and in terms of grit, captured by survey questions about 'perseverance' and 'resilience'.[11] Interestingly, we do not find a perceived advantage of graduates in terms of general cognitive skills, captured by 'analytical thinking' and 'verbal skills'. If anything, employers in our sample perceive non-completers to be slightly ahead of graduates in this respect. These findings seem broadly consistent with the fact that the final months of higher education usually focus on thesis writing and final examinations, activities that require perseverance and self-management and that tend to emphasize specific subject- and job-related skills and expertise.

Third, employers who view graduates more favourably relative to dropouts show a stronger preference for the former in the survey experiment. We cannot identify a subset of perceptions that would be decisive for understanding employer heterogeneity in the valuation of degree completion. It rather seems as if

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

some employers are broadly more favourable towards non-completers and others less so, viewing them as more—or less—competitive with graduates on most or all fronts.

On a theoretical level, our results provide support for all three mechanisms that have been argued to explain sheepskin effects. The large degree premium enjoyed by our fictitious graduates is partly driven by employers who state that they cannot hire a non-completer for the advertised position, a form of *credentialist* occupational closure (Weeden 2002). When it comes to adjudicating between *signalling* and *human capital* explanations, proponents of the former might rightly point to the fact that the payoff to four additional semesters of study (comparison between late and early non-completers) is dwarfed by the premium associated with graduating (comparison between degree-holders and late non-completers). At the same time, our finding that the average employer sees a clear advantage of graduates when it comes to occupation-specific skills does support human capital interpretations. That grit emerges as the second domain where graduates are perceived to outperform non-completers may again seem more consistent with a signaling perspective that sees credentials as mere markers of pre-existing (personality) differences between graduates and non-completers. However, it is also plausible that these skills are at least partly cultivated through the characteristic challenges that students face in the course of program completion.

While our findings shed new light on long-standing debates about the sources of sheepskin effects and labour market returns to education more broadly, they cannot settle these debates once and for all. In line with earlier research (Müller, 2005) , we would argue that, to a certain extent, all theoretical mechanisms play a role in understanding the educational pay-off in the labour market. A 'mechanism contest' seeking to single out a primary or even exclusive explanation for sheepskin effects thus seems not fruitful. Instead, future research should focus on the conditions under which some mechanisms become more or less prominent than others.

A first limitation of our study is that we measure self-reported behavioural intentions rather than actual employer behaviour. While this raises concerns about distortions due to hypothetical or desirability biases (Forster and Neugebauer, 2024), factorial surveys can still help to gain insights into the thought processes and criteria that employers use when forming hiring intentions. In contrast, a correspondence study, which measures actual invitations to job interviews, would have made it difficult to capture employer perceptions about what degree completion actually signals. In addition, we cannot rule out that our measures are subject to other sources of bias such as post-hoc rationalization. To minimize such risks, we deliberately placed the questions on the strengths and weaknesses of graduates vis-à-vis dropouts at the very end of the survey (whereas the survey experiment came at the very beginning). It remains possible, however, that employers' answers to these questions were partly motivated by an attempt to rationalize their 'behaviour' in the experimental part.

A second factor that might threaten the ecological validity of our study is the prominent placement and explicit phrasing of the education attribute. While few people would label themselves as 'dropouts' in real application situations, we have argued that such exaggeration may not affect, or potentially even enhance, the validity of the experiment. In particular, information may need to be communicated more explicitly in the low-stakes survey setting to have the same effect as less explicit information in the higher-stakes real-life setting. This argument is speculative, however, and we cannot rule out that our design choices led to inflated estimates of sheepskin effects.

Third and somewhat related, our experiment orthogonalizes observable applicant characteristics such as graduation and university GPA. This orthogonality, while useful for identification, implies that some applicant types (e.g. non-completers with very good GPA) may be much more common in our experimental population than they are among applicants in the real world. An indication that this leads to a partially atypical applicant pool can be found in a survey question that we asked after the experiment: 'In your opinion, to what extent do the applicant profiles from the example correspond to a typical applicant for the position of [JOB DESCRIPTION]?' 3% of employers answered 'not at all', 49% 'rather less', 45% 'quite', and 3% 'completely'. In the presence of effect heterogeneity—that is, if the magnitude of the sheepskin effect varied substantially depending on other candidate characteristics—such discrepancies between the experimental and real-world distributions of applicant profiles could lead to a situation where experimental estimates are very different from the corresponding real-world estimands, as discussed by De la Cuesta *et al.* (2022) for conjoint experiments.[12] To address this concern, we estimated a series of models that sequentially include all two-way interactions between the (non-)completion treatments and the remaining applicant characteristics—and reassuringly found that such effect heterogeneities are limited in the present case (for details, see Table A5 in the Online Appendix).

A fourth concern is that the hypothetical hiring situation, where graduates and non-completers apply to the same kind of job, may be relatively uncommon in the actual labour market. In particular, to the extent that dropout is a consequence of mismatches between field

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

of study and student interests, non-completers would not seem likely to apply for jobs that are closely related to their former field of study. However, Heublein *et al.*, (2017, 17ff.) report on German HE dropouts finds that performance problems (e.g. excessive demands and pressures, failed examinations) are a much more important reason for HE dropout than a 'lack of motivation'. Another reason why direct completer-dropout competition might be rare, especially in Germany, is that HE non-completers will often enrol in vocational education and training at the upper secondary level rather than seeking employment right away. Again, however, while Heublein *et al.*, (2017, Figure 9.2) estimate that, in 2014, 43% of German non-completers were in vocational training six months after leaving HE, they also find that roughly a third worked for pay. In addition, the decision to return to (vocational) education after leaving HE might well be a response to (actual or anticipated) unsuccessful job search efforts, that is, to the very penalties for non-completion identified in our experiment.

While our focus on two specific occupational fields in one particular country was key to designing a realistic and internally valid experiment, it inevitably raises questions about generalizability. As for the country setting, Germany is a major economy with a skill formation system that is characterized by strong education-to-work linkages and (related) institutional and legal rules tying access to specific occupations to qualifying credentials. The country thus seems a 'most likely case' for finding (large) sheepskin effects, as also suggested by the supply-side analysis of Berlingieri and Bolz (2020) who find that Germany has the largest graduate-dropout wage gap across a sample of 18 European countries. As for the occupational fields studied here, labour demand in IT and business tends to be high, which should increase employers' willingness to consider non-completers and reduce the magnitude of sheepskin effects. When it comes to the role of occupational closure, our best guess is that IT and business fall somewhere in the middle of the distribution. In our sample, roughly one in three employers indicated that they would not be able to hire a non-completer for the advertised position. In the absence of empirical data, we can only speculate about the importance of this mechanism in other fields. However, it seems very likely that it is (much) more important in heavily regulated fields such as health care and law, while it may well be less important in fields such as social science.

Future studies should implement similar experiments in other occupational fields and country contexts or, even better, employ explicit country-comparative designs to assess the generalizability of our results and to elucidate how institutional contexts shape the magnitude and sources of sheepskin effects. Another interesting avenue for future research is to assess what completion of VET signals to employers, as dropping out of vocational training also leads to a wage penalty (Patzina and Wydra-Somaggio, 2020).

Future work could also extend our analysis of employer perceptions by broadening the list of assessment dimensions and by asking employers to evaluate more specific applicants with respect to these domains (rather than asking for wholesale comparisons of graduates and non-completers like we did in our survey). It also seems worthwhile to explore the firm- or individual-level factors driving employer heterogeneity in perceptions of graduates and non-completers: Why do some recruiters attach greater importance to degree completion than others? Exploratory analysis of our own data did not provide clear insights into this issue.

Last but not least, our findings suggest some general lessons for interventions. As noted previously, the graduation premium is mirrored by labour market disadvantages on the part of non-completers. While the reasons for dropping out of higher education are diverse and likely require tailored countermeasures to prevent non-completion, our results suggest that employer perceptions—including potential stereotypes and negative biases—are important factors to address when it comes to giving non-completers a second chance in the labour market. Related, micro-credentials (European Commission, 2020) and similar instruments might improve the labour market prospects of non-completers by increasing the visibility and transparency of their skill sets.

Endnotes

1 The term 'sheepskin' is a reference to former times when diplomas were commonly printed on this material.

2 See Table 4, Model 1, and Table 5, Model 1, in Flores-Lagunes and Light (2010). The different estimates are due to different approaches to approximating years of education, which the authors refer to as 'time in school'.

3 A related literature studies the phenomenon of 'ability revelation' in (higher) education (Stinebrickner and Stinebrickner, 2014 ; Arcidiacono *et al.*, 2024), in the sense of both self-revelation—students finding out about their own learning capabilities as they progress through educational programs—and the revelation of abilities to others, most importantly employers, through performance in education and successful program completion (Arcidiacono, Bayer, and Hizmo, 2010).

4 We note that the distinction between demand-side and supply-side evidence is primarily analytic. While we frame employer decision-making as a demand-side phenomenon, it inevitably involves assessing supply-side actors like applicants and employees. Likewise, patterns in supply-side data (e.g. wage differentials) reflect the interplay between supply-side behaviours (e.g. job search) and demand-side decisions (e.g. job offers).

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

5 The prompting proved to be realistic. 70% of all employers stated in the subsequent survey that it corresponded 'quite well' or 'completely' to a typical job offer in the respective field.
6 To increase the response rate, we followed up with a categorical query for those who did not respond initially: Can you perhaps match the gross annual salary appropriate for this applicant to one of these salary categories (13 salary categories)? For these responses, we used the midpoint of the salary categories as best estimate for the metric salary variable.
7 In the qualitative pre-study, we conducted 13 semi-structured interviews of about one hour with job counsellors and recruiters from companies looking for computer scientists or business economists.
8 Conventional linear regression with employer-level cluster-robust standard errors produces very similar results (available upon request).
9 We addressed the nested data structure in two steps. First, we imputed missing employer characteristics including the average invitation probability to incorporate key information from the vignette part of the survey. Second, we imputed missing vignette ratings separately for graduate and non-completer profiles, accounting for differences in the relationship between employer perceptions and vignette ratings across the two groups. This ensures compatibility of the imputation with the analysis models that include cross-level interactions related to RQ3.
10 Berlingieri and Bolz (2020) report a substantially larger graduate-non-completer gap of almost 35 log points for Germany, but their estimate refers to adults aged 25 to 64 and is therefore less comparable with our or Flores-Lagunes and Light's (2010) estimates of starting salaries.
11 We acknowledge that our use of the term 'grit' differs slightly from Duckworth *et al*'.s (2007) in that we do not measure 'passion' which they consider another key component of this complex trait.
12 De la Cuesta *et al* (2022) suggest to address this issue by ensuring that the joint distribution of characteristics in the experimental 'population' closely matches its real-world counterpart (e.g. through weighting). However, this approach requires data to estimate the latter, which does not seem to be available for the case at hand.w

Supplementary Data

Supplementary data are available at *ESR* online.

Acknowledgements

Jan Paul Heisig and Martin Neugebauer contributed equally to this work. We would like to express our gratitude to Annabell Daniel for her significant contributions to the data collection process and valuable feedback provided on an earlier version of this paper. We also thank seminar participants at the European Consortium for Sociological Research 2021 Annual Conference (online) and WZB for their helpful comments and suggestions.

Funding

This work was supported by the German Federal Ministry of Education and Research (grant number 01PX16012 to MN).

Author Contributions

Martin Neugebauer (Conceptualization [Equal], Data curation [Lead], Formal analysis [Equal], Funding acquisition [Lead], Investigation [Lead], Writing - original draft [Equal], Writing - review & editing [Equal]), Jan Paul Heisig (Conceptualization [Equal], Formal analysis [Equal], Visualization [Lead], Writing - original draft [Equal], Writing - review & editing [Equal]), and Thijs Bol (Conceptualization [Equal], Writing - original draft [Supporting])

Data Availability Statement

The data underlying this article will become available via the Research Data Centre for Higher Education Research and Science Studies (FDZ-DZHW).

References

Arcidiacono, P. *et al.* (2010). Beyond signaling and human capital: education and the revelation of ability. *American Economic Journal: Applied Economics*, **2**, 76–104.
Arcidiacono, P. *et al.* (2024). College attrition and the dynamics of information revelation. *Journal of Political Economy*, **133**, 153–110.
Auspurg, K., and Hinz, T. (2014). *Factorial Survey Experiments*. Los Angeles: Sage.
Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: Columbia University Press.
Belman, D. and Heywood, J. S. (1997). Sheepskin Effects by Cohort: Implications of Job Matching in a Signaling Model. *Oxford Economic Papers*, **49**, 623–637.
Berlingieri, F., and Bolz, T. (2020). *Earnings of University Dropouts Across Europe. ZEW Discussion Paper No. 20-085. Mannheim*.
Bills, D. B. (2003). Credentials, signals, and screens: explaining the relationship between schooling and job assignment. *Review of Educational Research*, **73**, 441–469.
Bills, D. B., Di Stasio, V. and Gërxhani, K. (2017). The demand side of hiring: employers in the labor market. *Annual Review of Sociology*, **43**, 291–310.
Bol, T. and Heisig, J. P. (2021). Explaining wage differentials by field of study among higher education graduates: evidence from a large-scale survey of adult skills. *Social Science Research*, **99**, 102594.
Bol, T. and Van de Werfhorst, H. G. (2011). Signals and closure by degrees: the education effect across 15 European countries. *Research in Social Stratification and Mobility*, **29**, 119–132.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

Bol, T. and Weeden, K. A. (2014). Occupational closure and wage inequality in Germany and the United Kingdom. *European Sociological Review*, **31**, 354–369.

Chiswick, B. (1973). Schooling, screening, and income. In Solmon L. C. and Taubman P. J. (Eds.), *Does College Matter? Some Evidence on the Impacts of Higher Education*. New York: Academic Press, pp. 151–158.

Collins, R. (1979). *The Credential Society: A Historical Sociology of Education and Stratification*. New York: Academic Press.

de la Cuesta, B., Egami, N. and Imai, K. (2022). Improving the external validity of conjoint analysis: the essential role of profile distribution. *Political Analysis*, **30**, 19–45.

Deming, D. and Kahn, L. B. (2018). Skill requirements across firms and labor markets: evidence from job postings for professionals. *Journal of Labor Economics*, **36**, S337–S369. https://doi.org/10.1086/694106

Deterding, N. M. and Pedulla, D. S. (2016). Educational authority in the "Open Door" marketplace: labor market consequences of for-profit, nonprofit, and fictional educational credentials. *Sociology of Education*, **89**, 155–170.

DiPrete, T. A. *et al*. (2017). School-to-Work linkages in the United States, Germany, and France. *American Journal of Sociology*, **122**, 1869–1938.

Di Stasio, V. and Van De Werfhorst, H. G. (2016). Why does education matter to employers in different institutional contexts? a vignette study in England and the Netherlands. *Social Forces*, **95**, 77–106.

Duckworth, A. L. *et al*. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, **92**, 1087–1101.

European Commission. (2020). *A European approach to micro-credentials: Output of the micro credentials higher education consultation group: Final report*. Brussels: European Commission.

Flores-Lagunes, A. and Light, A. (2010). Interpreting degree effects in the returns to education. *The Journal of Human Resources*, **45**, 439–467.

Forster, A. and Neugebauer, M. (2024). Factorial survey experiments to predict real-world behavior: a cautionary tale from hiring studies. *Sociological Science*, **11**, 886–906.

Giani, M. S., Attewell, P. and Walling, D. (2020). The value of an incomplete degree: heterogeneity in the labor market benefits of college non-completion. *The Journal of Higher Education*, **91**, 514–539.

Hainmueller, J., Mummolo, J. and Xu, Y. (2019). How much should we trust estimates from multiplicative interaction models? simple tools to improve empirical practice. *Political Analysis*, **27**, 163–192.

Heisig, J. P. and Schaeffer, M. (2019). Why you should always include a random slope for the lower-level variable involved in a cross-level interaction. *European Sociological Review*, **35**, 258–279.

Heublein, U. *et al.* .(2017). Zwischen Studienerwartungen und Studienwirklichkeit: Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen. *DZHW Forum Hochschule*, **1-2017**, 1–298.

Humburg, M. and van der Velden, R. (2015). Skills and the graduate recruitment process: evidence from two discrete choice experiments. *Economics of Education Review*, **49**, 24–41.

Hungerford, T. and Solon, G. (1987). Sheepskin effects in the returns to education. *The Review of Economics and Statistics*, **69**, 175–177.

Huntington-Klein, N. (2020). Human capital versus signaling is empirically unresolvable. *Empirical Economics*, **60**, 2499–2531.

Jaeger, D. A. and Page, M. E. (1996). Degrees matter: new evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics*, **78**, 733.

Jenkins, A., and Wolf, A. (2018). Chapter eight: employers' selection decisions: the role of qualifications and tests. In *What's the Good of Education?* (pp. 147–168). Princeton: Princeton University Press.

Ketel, N. *et al*. (2016). The returns to medical school: evidence from admission lotteries. *American Economic Journal: Applied Economics*, **8**, 225–254.

Klein, D., Mishra, S., and Müller, L. (2021). Die langfristigen individuellen Konsequenzen des Studienabbruchs. In Neugebauer M., Daniel H.-D., and Wolter A. (Eds.), *Studienerfolg und Studienabbruch*. Wiesbaden: Springer VS, pp. 281–302.

Leuze, K. (2007). What makes for a good start? Consequences of occupation-specific higher education for career mobility: Germany and Great Britain compared. *International Journal of Sociology*, **37**, 29–53.

Ma, D. S., Correll, J. and Wittenbrink, B. (2015). The Chicago face database: a free stimulus set of faces and norming data. *Behavior Research Methods*, **47**, 1122–1135.

Matkovic, T. and Kogan, I. (2012). All or Nothing? The Consequences of Tertiary Education Non-Completion in Croatia and Serbia. *European Sociological Review*, **28**, 755–770.

Mincer, J. (1970). The Distribution of labor incomes: a survey with special reference to the human capital approach. *Journal of Economic Literature*, **8**, 1–26.

Mora, J. J. and Muro, J. (2008). Sheepskin effects by cohorts in Colombia. *International Journal of Manpower*, **29**, 111–121.

Müller, W. (2005). Education and youth integration into European labour markets. *International Journal of Comparative Sociology*, **46**, 461–485.

Neugebauer, M. and Daniel, A. (2022). Higher education non-completion, employers, and labor market integration: experimental evidence. *Social Science Research*, **105**, 102696.

Neugebauer, M., Bröder, C., and Daniel, A. (2021). *Studienausstieg – und dann? Der Übergang in Ausbildung und Beruf aus Sicht von Arbeitgeberinnen und Arbeitgebern*. Bielefeld: wbv.

OECD. (2013). *Education at a Glance 2013: OECD Indicators*. Paris: OECD.

Patzina, A. and Wydra-Somaggio, G. (2020). Early careers of dropouts from vocational training: signals, human capital formation, and training firms. *European Sociological Review*, **36**, 741–759.

Payne, S. S. C. (2023). Equalization or reproduction? 'some college' and the social function of higher education. *Sociology of Education*, **96**, 104–128.

Petzold, K. (2017). The role of international student mobility in hiring decisions. a vignette experiment among

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026

German employers. *Journal of Education and Work*, **30**, 893–911.
Piopiunik, M. *et al*. (2020). Skills, signals, and employability: an experimental investigation. *European Economic Review*, **123**, 103374.
Pons, E. and Blanco, J. M. (2005). Sheepskin effects in the Spanish labour market: a public–private sector analysis. *Education Economics*, **13**, 331–347.
Quadlin, N. (2018). The Mark of a Woman's Record: gender and academic performance in hiring. *American Sociological Review*, **83**, 331–360.
Scholten, M. and Tieben, N. (2017). Vocational qualification as safety-net? Education-to-work transitions of higher education dropouts in Germany. *Empirical Research in Vocational Education and Training*, **9**, 7.
Sørensen, A. B. (2000). Toward a sounder basis for class analysis. *American Journal of Sociology*, **105**, 1523–1558.
Stinebrickner, R. and Stinebrickner, T. (2014). Academic performance and college dropout: using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics*, **32**, 601–644.
Su, D. and Steiner, P. M. (2020). An evaluation of experimental designs for constructing vignette sets in factorial surveys. *Sociological Methods & Research*, **49**, 455–497.
Tieben, N. (2024). Destinations after higher education non-completion: the role of social background and pretertiary vocational qualifications. *Empirical Research in Vocational Education and Training*, **16**, 9.
Weeden, K. A. (2002). Why do some occupations pay more than others? Social closure and earnings inequality in the United States. *American Journal of Sociology*, **108**, 55–101.
Weiss, A. (1995). Human Capital vs. Signalling Explanations of Wages. *Journal of Economic Perspectives*, **9**, 133–154.

Martin Neugebauer is Professor of Sociology of Education at the Karlsruhe University of Education, Germany. His research interests encompass inequalities in education, returns to education, subjective well-being, and experimental methods. His recent work has been published in Sociological Science, European Sociological Review, and Social Science Research, among others.

Jan Paul Heisig is head of the research group 'Health and Social Inequality' at WZB Berlin Social Science Center and Professor of Sociology at Freie Universität Berlin. He works on a broad range of topics in the areas of social inequality, social stratification, and quantitative methods. Current projects focus on cross-national and regional variation in health inequalities, environmental inequalities, discrimination in health care, labour market returns to education, and multilevel modelling. Recent articles have appeared in journals such as American Sociological Review, British Journal of Political Science, European Sociological Review, and Socio-Economic Review.

Thijs Bol is Professor of Sociology at the University of Amsterdam and co-director of the Amsterdam Centre for Inequality Studies. His research focuses on inequality in education, the labour market, and science. His recent work has been featured in journals such as Proceedings of the National Academy of Sciences, American Sociological Review, American Journal of Sociology, Social Forces, and Sociological Science.

Downloaded from https://academic.oup.com/esr/article/42/2/165/8157227 by Universiteit van Amsterdam user on 12 May 2026