Research in Social Stratification and Mobility 93 (2024) 100959

Contents lists available at ScienceDirect

Research in Social Stratification and Mobility

journal homepage: www.elsevier.com/locate/rssm

Variation in unemployment scarring across labor markets. A comparative factorial survey experiment using real vacancies

Stefan Sacchi [a], Robin Samuel [b,*]

[a] Department of Social Sciences, University of Bern, Fabrikstrasse 8, CH-3012 Bern, Switzerland
[b] Department of Social Sciences, University of Luxembourg, 11, Porte des Sciences, LU-4366 Esch-sur-Alzette, Luxembourg

ARTICLE INFO

Keywords:
Unemployment scarring
Hiring
Stigmatization
Comparative analysis
Factorial survey experiment

ABSTRACT

Unemployment may severely impede access to (good) jobs. We focus on the effects of unemployment scarring on the chances of young workers to get hired and evaluate the extent to which they are affected in labor markets with different levels of unemployment. Drawing on Goffman's work on stigmatization and on queuing theory, we derive two potentially complementary micro-level explanations with opposing macro-level implication. We address the variation in unemployment scarring across 20 labor markets in four European countries based on factorial survey experiments embedded in real hiring situations. The results suggest that in labor markets with persistently low levels of unemployment, stigmatization, as proposed by Goffman, is the main source of unemployment scarring. We find no evidence that unemployment scarring is weaker when unemployment and the number of job seekers are low, as we inferred from queuing approaches. Our study contributes to expanding knowledge of context variability in unemployment scarring.

1. Introduction

Unemployment is associated with substantial and potentially lasting scarring effects on a wide range of career outcomes and life chances (Clark & Lepinteur, 2019; Strandh et al., 2014; Wanberg, 2012). Reported scars include reduced chances of being hired (Van Belle et al., 2018), wage loss (Gregg & Tominey, 2005), and well-being (Mousteri et al., 2018). An early spell of unemployment may be a *trigger event* (cf. Gangl, 2006) that gives rise to knock-on effects on subsequent labor market outcomes. Fragmented and protracted transitions into employment or precarious first jobs can further send negative signals to employers (Booth et al., 2002; Gebel, 2009; Pedulla, 2016). This may decrease the chances of finding good, permanent jobs when compared to labor market entrants with impeccable résumés. Overall, the research literature suggests that unemployment can lead to diverse and potentially long-lasting impairments within and beyond working life; we refer to these as scarring effects. In this article, we focus on the reduced chances of previously unemployed people finding new jobs, which may play a crucial role in explaining subsequent career-related scarring effects: Reduced hiring chances can lead to delayed re-employment with diminished job quality, lower income, etc. (Manzoni & Mooi-Reci, 2020; Stewart, 2007) and, mediated by this, to manifold disadvantages in the subsequent career. Reduced hiring chances can set the stage for various downstream scarring effects, making them a highly relevant subject of sociological research.

The evidence on the association of unemployment with hiring chances is mixed because of the challenge of inferring scarring and stigmatization from observational data (Ayllón et al., 2022). Furthermore, most studies focus on single contexts. Demand-side studies have increasingly turned to using experimental designs as they allow for a more direct examination of discrimination and stigmatization (for an overview on correspondence experiments, see Baert, 2018; on factorial survey experiments, McDonald, 2019). Concurrently, an increasing number of comparative experimental studies have addressed the need to better understand the role of context in unemployment scarring (see, e.g., Di Stasio & Lancee, 2020).

Our study adds to this emergent strand of experimental cross-national literature and proposes two, potentially complementary, micro mechanisms to explain the variability in the strength of unemployment scarring in hiring across diverse labor market contexts. First, drawing on Goffman (1963), we argue that the stigmatizing effect of unemployment is mitigated or fully suppressed when unemployment remains persistently high over long periods of time. Second, we build on queuing theory (Thurow, 1975) to argue that when hiring is difficult —

* Corresponding author.
E-mail address: robin.samuel@uni.lu (R. Samuel).

https://doi.org/10.1016/j.rssm.2024.100959
Received 13 October 2023; Received in revised form 17 May 2024; Accepted 15 July 2024
Available online 18 July 2024

due to current labor market tightness and small applicant pools, as is the case when unemployment is low — unemployment scarring will be alleviated. Our study offers a possible explanation for the inconsistent results of previous studies, covering a broad range of labor market conditions. We investigate the context variability of unemployment scarring based on 20 factorial survey experiments (FSEs) among recruiters responsible for filling real vacancies in five occupational fields across four European countries. The experimental, comparative approach and the integration of the factorial survey experiments into real hiring processes further enables a rigorous test of contextual variability in unemployment scarring. The focus of the empirical analysis is on the contrasting implications of the two proposed micro-mechanisms at the macro level of entire labor markets. While we also provide a fairly direct test of the proposed queueing-mechanism at the micro-level of individual recruiters, this is not feasible with the available data as regards the stigmatization mechanism. In Section 2, we proceed by providing an overview of the background pertaining to this study, explicate our theoretical arguments, and put forth our hypotheses. Following this, we elaborate on the methods employed and the analytical techniques utilized (Section 3). We then present our main results using graphs and reduced tables, concluding our analyses with several robustness checks in Section 4. Finally, we discuss our results and some implications in Section 5.

2. Background and theory

2.1. Variations in scarring across labor market contexts

There is a limited body of research addressing the impact of labor market conditions on unemployment scarring. A small number of experimental field studies (Birkelund et al., 2017; Farber et al., 2017; Kroft et al., 2013; Nunley et al., 2017; Nüß, 2018) address the *hiring chances* of previously unemployed job seekers by applying a correspondence audit design (Gaddis, 2018). To this end, fictitious job applications are submitted to real online job postings. This makes it possible to analyze whether and how the callback rates depend on the experimentally generated variation in selected applicant characteristics. Even if, with one exception (Kroft et al., 2013), this was not one of the main objectives of the studies mentioned, they can also assess the influence of the labor market context based on variations between local or occupational labor markets. Some of the mentioned studies (Birkelund et al., 2017; Kroft et al., 2013; Nüß, 2018) find greater unemployment scars in hiring (or stronger duration dependence) in favorable labor market conditions. In contrast, Farber et al. (2017) and Nunley et al. (2017) do not find evidence for an impact on hiring chances irrespective of labor market conditions. The evidence on the potentially moderating impact of labor market conditions on unemployment scarring in hiring is hence mixed. As long as the fictitious nature of the applications remains unnoticed, such studies enable reliable causal conclusions with high external validity. The mixed evidence is thus likely due to methodological differences in the covered labor market conditions, the selection of job ads and the composition of the fictitious applicant pool. Uncontrolled demand-side moderators on the level of jobs, companies, or recruiters, which are hardly observable in this type of experimental design, may also play a role.

There is additional evidence from nonexperimental studies on how employment-related scarring effects in a broader sense depend on labor market conditions. Note that this type of study is relevant here only insofar as such scarring effects are assumed to stem from the reduced hiring chances of the previously unemployed. Some of those studies provide clear evidence, while others offer limited evidence, that favorable labor market conditions may increase the scarring effects on wages (Ball, 2011; Lupi & Ordine, 2002; Mooi-Reci & Ganzeboom, 2015; Van Dijk & Folmer, 1999) and later unemployment risks (Biewen & Steffes, 2010; Omori, 1997). A comparative study by Ayllón et al. (2022) on the unemployment risks of previously unemployed individuals finds no evidence or even contrary effects of contextual variation in 11 of 12 examined European countries. In summary, consistent with the experimental evidence some, but not all, of the studies conclude that the disadvantages of the previously unemployed are amplified under favorable labor market conditions.

Institutional differences in education systems and labor market regulations may be expected to contribute further to the contextual variability in scarring effects (e.g., DiPrete, 2002; Gangl, 2004, 2006). Demand-side moderators of scarring effects at the job, firm, and recruiter levels likely also play a role in contextual variability. Factorial survey experiments on job applications in four European countries suggested that unemployment scarring was exclusively (Shi et al., 2018) — or, depending on the educational system, mostly (Imdorf et al., 2017) — limited to job applicants with skills that matched the corresponding job requirements. Experimental studies conducted in the US and Germany indicated that scarring effects were weaker in smaller companies and when vacancies had to be staffed urgently (Farber et al., 2017; Nüß, 2018). Considering the focus of our study, we restrict ourselves to this brief reference at this point but address the issue in Section 4.1.

2.2. The effect of unemployment experience on hiring chances at the individual level

Before we theorize the role of labor market conditions in moderating signaling and unemployment scars in hiring (Section 2.3), we present two lines of arguments on why unemployment experience may affect subsequent hiring chances at the individual level. The first centers on what unemployment may indicate to recruiters about the productivity or performance of an applicant (*productivity argument*). The second focuses on risk assessment (*risk argument*).

The *productivity argument* holds that recruiters use available information to select job applicants according to their expected productivity. From this perspective, unemployment may be related to skills depreciation (Becker, 1964) and unobserved personal characteristics that are considered negatively associated with productivity (Aigner & Cain, 1977; Spence, 1973). This mechanism is likely to play out in situations where a candidate's productivity is unknown and can only be inferred from application documents or readily observable characteristics, such as in the first stages of a hiring process. Oberholzer-Gee (2008) summarized four explanatory mechanisms: First, recruiters likely associate previous unemployment spells with the lack of motivation and undesirable personality traits such as low reliability and punctuality, and lack of social skills (Luijkx & Wolbers, 2009). Second, they may take periods of unemployment as a signal of interrupted on-the-job training and human capital depreciation (Mooi-Reci & Ganzeboom, 2015). Third, unemployment may be associated with low levels of trainability (Di Stasio, 2014; Thurow, 1975). Fourth, previous unemployment spells indicate that other recruiters have chosen not to employ the individual, suggesting they had access to more information and were thus better prepared to judge the applicant's expected productivity (Oberholzer-Gee, 2008).

According to the *risk argument*, recruiters focus on avoiding severe placement errors while selecting applicants (Bills et al., 2017; Fraser et al., 2010; Hendricks et al., 2003). Placement errors may necessitate replacement or additional hires (Hillmer et al., 2004; Karsan, 2007; O'Connell & Kung, 2007). Conflicts may arise, for example, if newly hired candidates do not conduct their supervisory tasks as planned, resulting in negative spillover effects on the productivity of other employees and generating high turnover costs. Recruiters may view previous unemployment spells as indicating that applicants may not stay for long or become well attuned to the prospective work environment, thus increasing the risk for placement errors (Bonoli & Hinrichs, 2012).

While the approaches presented above emphasize productivity and risk differently, they consider hiring an information or signal extraction problem. Ample evidence supports the assumption that spells of unemployment decrease hiring chances (Brand, 2015; Clark & Lepinteur,

2019; Ho et al., 2012; Luijkx & Wolbers, 2009; Shi & Di Stasio, 2022). These explanations primarily revolve around differences in characteristics of job candidates and what they indicate in terms of expected productivity and placement risks, rather than around the role of labor market contexts. However, labor market conditions may have a bearing on how recruiters assess applicants and may likely moderate individual-level mechanisms.

2.3. Role of labor market contexts in unemployment scarring

We elaborate on two strands of micro-level arguments, which may explain how labor market contexts moderate the strength of unemployment scarring in hiring. The first strand draws on Goffman's (1963) seminal work on social stigma and is partly based on Solga's (2005, 2008) arguments on the stigmatization of low-skilled workers. The second strand builds on the concept of a labor queue based on Thurow's (1975) job competition model (see also, e.g., Bills, 1990; Di Stasio, 2014; Van Belle et al., 2018). We will argue that both theoretical approaches may contribute to advance our understanding in a complementary way, that is, are not to be seen as mutually exclusive.

2.3.1. Labor market conditions and stigmatization

We provide three interrelated arguments for how the level of unemployment in a labor market can influence the stigma experienced by job applicants with a history of unemployment. The general premise here is, as we have argued in Section 2.2, that employers perceive unemployment as a signal of low productivity or a placement risk due to individual deficiencies such as a lack of work motivation, unreliability, absenteeism, or inability to cooperate with colleagues and superiors (Goffman, 1963, p. 11ff.; Solga, 2005, p. 190 f.). How much stigmatization occurs, then will likely depend on the following circumstances: First, it is crucial whether employers interpret unemployment as self-inflicted or as attributable to supra-individual economic causes. A negative evaluation of the previously unemployed, both in the wider population (Buffel & Van de Velde, 2019) and in the context of recruiting (Ho et al., 2012), is likely contingent upon an individual attribution of causes. At the core of this argument is the *causal locus* — the extent to which spells of unemployment are considered to have occurred owing to internal factors for which the individual is responsible (e.g., lack of work commitment) or external factors, which are beyond the control of the individual (e.g., plant closures, bad macroeconomic conditions). Where unemployment is widespread and easily attributable to external factors, it loses most of its signal value regarding individual deficiencies. Hence, stigmatization on the grounds outlined in Section 2.2 will likely be reduced.

Second, where levels of unemployment are low and unemployment is rare, it becomes more visible and conspicuous. Thus, the risk of social stigmatization increases (Goffman, 1963, pp. 64–67). According to Gesthuizen et al., (2011, p. 268) "the risk of attaching negative attributes to social groups is larger, when the relative group size is small" (quoting Jones et al., 1984,p. 92). Thus, *relative group size* co-determines the occurrence of unemployment stigmatization. With respect to hiring, persistently low unemployment levels imply that most job applicants' résumés show a seamless, straightforward employment record. Any unexplained gap in the résumé (or an openly declared phase of unemployment) thus becomes a conspicuous signal of possible individual deficiencies and is, potentially, reason to not consider an application further.

Third, in labor markets with unfavorable conditions and persistently high layoff risks, even the most productive and motivated workers may struggle to find a new job after a job loss (Gibbons & Katz, 1991). However, unemployment represents a strong and thus suspicious norm deviation in contexts, where immediately finding a new job is considered easy. Under such circumstances, the expected stigmatization is reinforced by the apparent *unambiguousness of the signal* of individual deficiencies (Goffman, 1963, pp. 64–67; Solga, 2005, p. 190).

To sum up, we assume that unemployment is more likely perceived as self-inflicted – i.e. as a signal of undesirable individual attributes or past behavior (as outlined in Section 2.2) – in favorable labor markets with low levels of unemployment. Under such conditions, stigmatization is further reinforced by the small relative group size and the apparent unambiguousness of the signal. Taken together, this suggests that unemployment stigmatization is mitigated in labor markets with persistently high levels of unemployment *(Hypothesis 1,* hereinafter H1*)*. The higher the levels of unemployment, the more likely it is for recruiters to place the causal locus for unemployment experience at the external end of the continuum.

2.3.2. Labor queues, hiring difficulties, and unemployment scarring

The concept of labor queues motivates a second line of argument to assume an association between labor market conditions and unemployment scarring. It is based on the vacancy competition model by Thurow (1975): companies rank job applicants according to easily observable characteristics such as gender, educational attainment, and job history. The job is then filled by the highest ranked person who does not receive any better job offers and can thus be recruited. In Thurow's original model framework, the ranking of job applicants in the labor queue ultimately aims at minimizing the expected training costs. Some applications of the queueing concept (e.g., Reskin, 1991) do without this assumption which may be seen as unnecessarily restrictive from a sociological perspective (Sørensen & Kalleberg, 1981, p. 72). In the context of this paper, it suffices to assume that the sought-after applicant characteristics such as motivations or abilities cannot be observed directly, which is why companies resort to corresponding proxy variables when selecting applicants.

How successful firms are in hiring the highest ranked applicants in the labor queue will also depend on current labor market conditions: During an economic upturn with a tight labor market, recruiters will find it more challenging to attract sufficient job applicants possessing all or most of the desired skills and characteristics. In such circumstances, recruiters may increasingly consider applicants who are further down the labor queue (Thurow, 1975, p. 94), such as those with the "wrong" gender (Reskin & Roos, 1990), a foreign background (Bursell et al., 2021), or an unemployment episode in their curriculum vitae (cf. Van Belle et al., 2018), who they would normally exclude during the initial screening. Even if this does not necessarily diminish existing concerns about hiring the formerly unemployed (cf. Section 2.2), their chances are likely to improve in small applicant pools where there are no competing applicants who are convincing in every respect.

The notion that a tight labor market contributes to a more positive evaluation and to improved hiring chances of applicants with likely harmful characteristics is in line with some experimental studies. For example, there is evidence that recruiters evaluate applications with given educational achievements (Protsch, 2021) or those from women (Kübler et al., 2018) more favorably when recruitment is difficult. Similarly, experimental findings indicate that in tight labor markets, hiring chances are less affected by age (Farber et al., 2018) or foreign origin (Baert et al., 2015). Accordingly, *Hypothesis 2* (hereinafter H2) postulates that recruiters evaluate previously unemployed job applicants more positively and are more likely to hire them if it proves difficult to fill a vacant position. It should be noted that such difficulties in filling vacancies will increase when the labor market as a whole is tight and unemployment is low.

2.4. Implications of a possible interplay of both micro-mechanisms

The micro-mechanisms derived from stigma and queuing arguments provide two explanations on how labor market conditions mitigate unemployment scarring which are possibly complementary on a theoretical level. If both mechanisms take effect, we expect the hiring chance disadvantages of the previously unemployed to be attenuated or even disappear altogether at both low and high levels of unemployment

which may be present in different occupational labor markets. At the macro level of entire labor markets, this will then lead either to a missing association between scarring and labor market conditions, or to an inverted u-shaped relationship with maximal scarring at intermediate levels of unemployment. Our literature review presents mixed evidence on the importance of labor market conditions for the hiring prospects of the previously unemployed. While some of the studies support the Goffmanian hypothesis (H1), other studies find no significant contextual differences across labor markets. The inconsistent results may be due to the limited variability in the levels of unemployment covered by previous studies. This suggests that any mitigating mechanisms specific to certain conditions may have gone unnoticed. To address this gap, we simultaneously test both mechanisms across a broad range of labor market conditions.

3. Data and methods

We used pooled data from 20 factorial survey experiments (Imdorf et al., 2020), a widely used design to study hiring intentions and related scarring effects (for an overview, see McDonald, 2019). Each FSE was integrated into a recruiter survey and conducted simultaneously in five narrowly defined occupational fields (mechanics, finance, information technology, catering, and healthcare) across four countries (Bulgaria, Greece, Norway, and Switzerland) in 2016. The sampling was based on job advertisements for entry-level positions and thus restricted to recruiters who, at the time of the survey, oversaw filling an advertised vacancy in one of those fields. Recruiters were shown fictional CVs of a small number of hypothetical job applicants (vignettes) in a randomized order and were asked to rate the likelihood of considering each one for the advertised job position. Some fictional applicants were designed to conform to virtually all crucial job requirements of field-typical vacancies, while others showed likely unfavorable characteristics. The survey experiment exploited the systematic variations in some applicant characteristics, such as the type of job experience or unemployment spell (explained below), to assess their impact on the recruiters' evaluations and, ultimately, hiring chances (Imdorf et al., 2020).

Carefully designed FSEs exhibit high *internal validity*, as any systematic variation in the outcome (in this case, hiring intentions of recruiters) can be attributed to manipulations of experimental stimuli, allowing for the causal interpretation of their average effects (Auspurg & Hinz, 2015). Implementing an FSE directly in a real hiring situation has two advantages (Gutfleisch et al., 2021). First, participants in the FSE are recruiters who are currently searching for and evaluating suitable job candidates to fill a position in the real world. Thus, they are acquainted with all job-specific requirements and selection criteria, likely strengthening the *reliability* and *internal validity* of their evaluation of fictitious job candidates in the FSE (Auspurg & Hinz, 2015, pp. 17–21, 81, 113–118). Second, it triggers a strong "psychological realism," which is probably the main precondition of high *external validity* (Auspurg & Hinz, 2015, pp. 113–118) — the transferability of findings to a broader range of real-world hiring situations and labor market contexts. External validity is further strengthened by choosing attributes (e.g., certificates) for the fictitious applicants that are familiar to recruiters in the field. However, the information about job applicants that can be provided in a vignette is inevitably limited and requires simplification.

FSE are a well-established method to measure behavioral intentions (Wallander, 2009) and also suited to assess their adaptation to varying external conditions, e.g. in different labor markets (Finger, 2016; Protsch, 2021). As long as the main interest is the direction and strength of treatment effects, methodological studies also indicate a mostly close correspondence between behavioral intentions as measured by FSE and actual behavior (Hainmueller, Hangartner, & Yamamoto, 2015; Petzold & Wolbring, 2019; Oesch, 2020). According to Petzold and Wolbring (2019, p. 26), such high "behavioral validity" of FSE is likely when either the normative expectations regarding the behavior in the experimental situation are weak, or when the costs of norm-compliant behavior in real world situations are low. Both prerequisites appear to be satisfied by our study: Neither can the inclusion of unemployment as a criterion to evaluate job applicants be deemed illegitimate, nor does a pre-hiring evaluation of job candidates entail substantial costs. This may be expected to contribute to the behavioral validity of our study.

A properly designed FSE provides an unbiased estimate of the average causal effect within the underlying study population, whereas a *comparative analysis* of the strength of average effects across different (sub)populations may be biased if unobserved moderators are involved (see Appendix A for a formal derivation of this claim and Auspurg & Hinz, 2015, pp. 40–41). In the context of FSEs *based on real vacancies*, three major sources of micro-level heterogeneity and possible moderator effects are at play: *job, recruiter,* and *firm characteristics*. The multitude of potentially relevant *moderators* is not a concern in the context of a thoroughly randomized FSE; in particular, the causal interpretation of the estimated average effects still holds for any single FSE. However, unless we consider relevant moderator effects (i.e., non-negligible heterogeneous treatment effects), any *comparative* analysis of causal effects taken from a series of parallel FSEs (or, from subgroups within one FSE) may be distorted by *moderator bias*. Comprehensive control over possible moderator bias bound to all kinds of demand-side variables is therefore crucial for a sound analytic strategy (see Section 3.5.). We first describe the experimental design of the 20 FSEs, the labor market contexts covered, the sampling of job vacancies, and the empirical measurement.

3.1. Experimental design

The recruiters were asked to evaluate several fictitious job applicants, each represented by a vignette showing a hypothetical CV (see Appendix B for an example vignette). The applicant characteristics on the vignettes (i.e., the timing and duration of unemployment, gender, and level and field-specificity of educational attainment and previous work experience) were experimentally varied, in addition to a country-specific experimental variable (see Table 1).

Nationality and time since leaving formal education were kept constant. All hypothetical applicants had national citizenship and five years (between 2010 and 2016) of labor market participation. All vignettes were presented in the prevalent national language. Aside from one country-specific experimental variable (bottom of Table 1), the *levels* of all experimental variables were the same across all countries and occupational fields. To maximize "psychological realism" and the validity of the FSEs (Auspurg & Hinz, 2015), the fictitious candidates' educational attainment levels and work experience were represented by concrete educational certificates and job titles that were familiar to recruiters in the field.[1] Upon the exclusion of implausible combinations (e.g., applicants without educational credentials working as trained nurses), the experimental design resulted in 252 different vignettes, that is, plausible combinations of applicant characteristics (Imdorf et al., 2020). Based on pretest response rates, D-Optimization (Dülmer, 2007) was used to select a suitable subset of vignettes to be fielded, while maximizing variance and orthogonality of the experimental dimensions. D-Optimized subsets of 162 vignettes in Norway and Switzerland and subsets of 90 vignettes in Bulgaria and Greece were fielded (Hyggen et al., 2016; Imdorf et al., 2020).[2] Blocks of nine vignettes were assigned to 18 decks in Norway and Switzerland and to 10 decks in Bulgaria and Greece, optimizing the blocks for maximum variance and orthogonality between each vignette dimension (see, e.g., Atzmüller & Steiner, 2010). Each recruiter was randomly assigned to one of these decks in the survey

[1] The selection of concrete credentials and job titles appearing on the vignettes relied on detailed standardized criteria aimed at maximizing comparability across occupational labor markets and countries (see Appendix C and Imdorf et al., 2020).

[2] For 162 and 90 vignettes, D-efficiency was 99.2 and 99.8, respectively, for a model containing all main effects (Imdorf et al., 2020).

Table 1
Experimental dimensions used to construct the vignettes.

Dimension 1: Timing and duration of unemployment (7 Levels)		
	Unemployment timing	*Unemployment duration*
1	Never	-
2	Immediately after graduation	10 months
3	Immediately after graduation	20 months
4	Between two previous jobs	10 months
5	Between two previous jobs	20 months
6	Ongoing (at the time of hypothetical application)	10 months
7	Ongoing (at the time of hypothetical application)	20 months

Dimension 2: Education and work experience (9 Levels)		
	Level and sector-specificity of education	*Sector specificity of work experience*
1	Lower-secondary education (ISCED 2)	Sector-specific job with low requirements
2	Sector-specific secondary education (ISCED 35)	Sector-specific job with medium requirements
3	Sector-specific tertiary education (ISCED 5)	Sector-specific job with high requirements
4	Lower-secondary education (ISCED 2)	Sales job with low requirements
5	Sales-specific secondary education (ISCED 35)	Sales job with medium requirements
6	Sales-specific tertiary education (ISCED 5)	Sales job with high requirements
7	Lower-secondary education (ISCED 2)	Call center job
8	Sector-specific secondary education (ISCED 35)	Call center job
9	Sector-specific tertiary education (ISCED 5)	Call center job

Dimension 3: Gender (2 Levels)		
1	Female	
2	Male	

Country-specific dimensions (2 Levels)				
	Bulgaria: International work experience	*Switzerland: Job-hopping*	*Greece: ALMP participation*	*Norway: ALMP participation*
1	Yes	Yes	Yes	Yes
2	No	No	No	No

Note: Table by authors based on Imdorf et al. (2020). ISCED: International Standard Classification of Education, Version 2011; ALMP: active labor market program.

experiment. The order in which the vignettes were presented was randomized across recruiters to avoid bias owing to ordering effects. For each vignette, the recruiters were asked to rate the likelihood of considering each applicant for a given vacancy on an 11-point scale (0 = "practically zero" to 10 = "excellent"). Each vignette contains information on some crucial applicant characteristics (see Table 1). If this information is insufficient to determine whether the fictitious candidate meets any additional job requirements, recruiters were instructed to assume she does. They were also informed that they would have the opportunity to declare any additional requirements later.

3.2. Labor market contexts

Given the large variation in unemployment levels, the four countries selected present an ideal set of cases to test our hypotheses. Switzerland and Norway had low unemployment levels — with average unemployment rates of 4.7 % and 3.7 %, respectively — for the years preceding the FSE, that is, 2010–2016 (Eurostat, 2020). Bulgaria represented a medium unemployment-rate context (10.7 %), and Greece a high unemployment-rate context (22.5 %) on average between 2010 and 2016 (Eurostat, 2020). In each country, our data covered the same five occupational fields (mechanics, finance, information technology, catering, and healthcare). Each field was further restricted to a few detailed occupational categories according to the International Standard Classification of Occupations (ILO, 2012; four-digit codes of ISCO 08; see Appendix C). The occupational categories were selected to maximize international comparability (see Imdorf et al., 2019, 2020).

3.3. Collection of vacancies through job adverts

Real-world vacancies were sampled in each country with the contact information of the person responsible for filling the respective position (Imdorf et al., 2020). Vacancies were sampled from a broad selection of online job portals and company websites in each country. We relied on the pre-defined ISCO categories (and the related index of occupational titles) to restrict sampling to advertised vacancies with a job title from one of the five occupational fields defined above (Imdorf et al., 2020). All selected vacancies refer to entry-level jobs, typically requiring an upper-secondary vocational or bachelor's degree.

The response rates for the 20 FSEs ranged from 10.0 % to 27.4 %, resulting in an average of 15.8 %. The response rates for employer surveys are generally much lower than in population surveys (Anseel et al., 2010). The reported response rates are within the range of similar recruiter surveys using FSEs to study hiring intentions, and are, in some cases, even higher (see, e.g., Liechti et al., 2017).

Our final sample included 18,001 applicant evaluations from 2088 recruiters. Appendix D presents the number of observations by country and occupational field. Appendix E displays the average number of ratings per vignette in each sample and the number of respondents per deck.

3.4. Empirical measurements

The *dependent variable* in our analyses is the vignette ratings — recruiters' evaluation of the hiring chances of a fictitious job candidate regarding the advertised vacancy. The 11-point rating scale (see Appendix B) ranged from "practically zero" (0) to "excellent" (10). To test our hypotheses, we focused on a *dichotomous indicator of unemployment*, indicating either absence (0) or presence (1) of an unemployment spell in the CV. Thus, all unemployment spells of variable duration (either 10 or 20 months) and timing (at the beginning, in the middle, or at the end of the five-year occupational trajectory) are collapsed (cf. Table 1), enabling an overall test of (average) scarring effects.[3]

To investigate H1, we measured the long-term average *level of unemployment* in each labor market segment by combining official unemployment data for the four countries (taken from Eurostat) with expert ratings (taken from Imdorf et al., 2020, pp. 61–70) on the relative level of unemployment in the occupational fields within each country (for a detailed description of the calculation, see Appendix F). Table 2 shows the resulting average field-specific unemployment rates in the period from 2010 to 2016, which is aligned with the time the fictitious job applicants spent in the labor market (cf. Section 3.1).

There are strong disparities between the labor market segments that span the entire range of variation observed among OECD countries over the same observation period (OECD, 2023). It is hardly relevant whether we rely on a five-year average (as in Table 2) or a single annual rate when measuring the segment-specific levels of unemployment because relative differences in unemployment levels tend to be remarkably stable over time across both countries (e.g., OECD, 2005) and occupational

[3] There are arguments and limited empirical evidence (Eriksson & Rooth, 2014; Luijkx & Wolbers, 2009; Raaum & Røed, 2006) suggesting that the strength of scarring effects may differ depending on the timing of unemployment in the employment history. However, it is not theoretically evident why the moderating role of timing should depend on the labor market context. Developing theoretical arguments on this topic is beyond the scope of this paper. However, in response to a reviewer's comment, we empirically explored the idea and found no evidence for a moderating role of the timing of unemployment in relation to the interaction effects implied by H1 and H2.

Table 2
Estimated average occupation-specific unemployment rates for the period 2010–2016.

	Occupation				
Country	*Mechanics*	*Finance*	*Health Care*	*Catering*	*IT*
Norway	4.1 %	2.5 %	2.2 %	5.2 %	3.2 %
Switzerland	3.6 %	3.2 %	2.6 %	7.4 %	2.6 %
Bulgaria	7.7 %	7.1 %	3.7 %	7.1 %	3.4 %
Greece	22.5 %	17.0 %	13.3 %	17.0 %	11.5 %

labor markets (Buchs & Buchmann, 2017; European Commission, 2019; Heidenreich, 2015).[4] If we further consider that perceptions of unemployment levels and trends appear to be fuzzy (Cardoso et al., 2016; Kunovich, 2012), a medium-term average as a measure of stable differences in the prevalence of unemployment is well in line with a Goffmanian perspective.[5]

A four-point rating scale included in the recruiter survey was used to test the impact of *hiring difficulty* (H2). With the original job ad on the screen, recruiters were asked: "How difficult is/was it to find a suitable candidate for this advertised position?" (1 = "very easy," 2 = "quite easy," 3 = "quite difficult," 4 = "very difficult"). This gave us a micro-level measure of the perceived difficulty of filling the vacant position, which showed a good construct validity, including the expected positive relationships with skill requirements and supervisor positions and negative relationships with firm size and level of unemployment.[6] For the analyses, responses were dichotomized to indicate either "very difficult" (1) or "very easy," "quite easy," or "quite difficult" (0), to describe hiring difficulty for the vacancy at hand.

The literature review suggested that the degree of fit between skill requirements of the job and applicant qualification was a strong moderator of unemployment scarring effects. To control this moderator effect, we relied on a binary indicator of *skill fit* that measured whether a fictitious job applicant was equipped with the required level of education, field of study, and work experience. The skill fit indicator was coded 1 if the candidate met all three criteria, and 0 otherwise. The data skill requirements were taken from the recruiter survey (see Appendix G).

We further controlled for all experimental vignette variables (according to Table 1) and their interactions with country. Appendix H provides the correlation matrices between all experimental variables for each FSE. The nine levels of the experimental dimension "Education and work experience" were divided into two separate variables (i.e., level of education and field-specificity of education and work experience) and an interaction term. The vector of control variables also included dummy variables for countries and occupations, the country-specific experimental variable (cf. bottom of Table 1), and a dummy variable to control for primacy effects (first and second vs. all subsequent vignette ratings).

The indicators for hiring difficulty and skill fit relied (partly) on data from the recruiter survey, which resulted in a small share of missing item values (1.3 % at the level of recruiters). The missing values were imputed using chained equation and considering the multilevel structure of the data (Grund et al., 2018: Equation 13). Following White, Royston, and Wood (2011, p. 388), we generated five imputed datasets for the analysis. Wherever applicable, we pooled the estimates based on Rubin's Rules. For other statistics, we reported the median value and cross-imputation range (see Marshall et al., 2009).

3.5. Analytical strategy

We modeled vignette evaluations clustered at the level of recruiters. To consider this hierarchical data structure, we estimated linear multilevel regression models with random intercepts. Owing to the random assignment of vignettes to recruiters, covariates are uncorrelated with the error term by design, thus meeting a core assumption of random effect models (Auspurg & Hinz, 2015). We estimated the following baseline model:

$$\ln(Y_{\nu i}) = \beta_0 + \underbrace{\beta(U_{\nu i}\cdot \ln[LU_\nu])}_{\text{Hypothesis 1}} + \underbrace{\beta(U_{\nu i}\cdot D_\nu)}_{\text{Hypothesis 2}} + \underbrace{\beta C + \beta O + \beta(C\cdot A_{\nu i}) + \beta P_{\nu i}}_{\text{Control Variables}} + u_\nu + \varepsilon \quad (1)$$

$$\text{with } \nu \in \{1, \ldots, 2088\},\ i \in \{1, \ldots, 9\}$$

$Y_{\nu i}$ Recruiter assessment of the hiring chances of application i for vacancy ν

β_0 / β
Intercept / Vectors of regression coefficients

$U_{\nu i}$ Unemployment spell in fictitious application (vignette) i for vacancy ν (dummy variable)

LU_ν Long-term average of unemployment rate in labor market segment of vacancy ν

D_ν Recruiter perception of the difficulty to staff vacancy ν

C, O
Dummy variables for four countries (C) and five occupations (O)

$A_{\nu i}$ Characteristics of applicant i applying for vacancy ν (see Table 1)

$P_{\nu i}$ Dummy for primacy effects in the rating of applicant i applying for vacancy ν

u_ν Error term at the level of vacancies (recruiters) with $N(0, \sigma_u)$

ε Residuals with $N(0, \sigma_\varepsilon)$

Model (1) does also include the main effects and all lower-level interaction effects implied by the interaction terms, although this is not explicitly stated in Eq. (1).

To test for stigma effects (H1), we estimated the interaction effect between unemployment at the level of job candidates and the long-term average unemployment rate. A significant positive interaction effect would support H1. The test is based on elasticities, as the model includes the natural logarithms of the vignette ratings and the long-term average of unemployment rates. The idea behind this is that *relative* changes in the latter are proportional to those in the former. This specification takes into account that an increase in unemployment of, for example, two percentage points is barely perceptible if the initial level is high, such as 20 %. However, at a low initial level of 2 %, it implies a doubling, which will likely spark intense media coverage, a marked growth in the numbers of job applicants, and possibly an increase in unemployed individuals within personal networks, making it hard for recruiters to

[4] Accordingly, country-level data from Eurostat (2020) shows an almost perfect correlation between average unemployment rates from 2010 to 2016 and the annual unemployment rate for 2016 (i.e., the survey year). For the 34 countries covered by Eurostat, the correlation between the average unemployment rates from 2010 to 2016 and the unemployment rate during the survey (2016) amounts to .95 (or .97 for the four countries considered here).

[5] For the same reason, there is little need for more differentiated measurements, such as age-specific ones. However, it would likely be beneficial to consider regional differences, which is not possible with the available data.

[6] The results are based on the multivariate regression model at the level of recruiters or vacancies (n = 2088) from hiring difficulty on a selection of correlates known from the literature. The model also reveals a significant negative partial effect of the average long-term level of unemployment, implying a reduction of 0.5 rating points in hiring difficulty over the range of observed unemployment levels.

overlook. The test of the queuing argument (H2) is based on the interaction term between the perceived difficulty to staff a given vacancy and the applicants' unemployment. A significant positive interaction effect would corroborate H2.

The model contains a comprehensive vector of control variables (see above Section 3.4.), which includes dummies for countries and occupations along with the full vector of applicant characteristics. The strong variations among countries (see, e.g., Imdorf et al., 2019) are reflected in the models using interaction terms to control for country-specific effects of education, work experience, and other applicant characteristics.

Whereas the baseline model from Eq. (1) covered many important determinants of individual hiring chances, it does not account for the heterogeneity in applicant-vacancy fit. Existing research using FSEs suggests that unemployment scarring effects are limited to applicants who have the required skills (Shi et al., 2018; Imdorf et al., 2017: Table A4). As outlined above, comprehensive control over moderator effects is crucial if a comparative assessment of the causal effects from a series of FSEs is intended. We employed two complementary methods to control for the supposed moderator role of job-applicant fit.

First, we controlled for *micro-level skill fit* directly in our models by including skill fit indicator F_{vi} (introduced in Section 3.4.), which measured whether the fictional application meets the skill-related job requirements. The relevant interaction terms in Eq. (1) extend to $U_{vi} \bullet \ln[LU_v] \bullet F_{vi}$ and $U_{vi} \bullet D_v \bullet F_{vi}$ to test Hypotheses 1 and 2, respectively, resulting in the following specification of our full model (including all main effects and lower-level terms implied by the interactions):

$$\ln(Y_{vi}) = \beta_0 + \underbrace{\beta(U_{vi} \cdot \ln[LU_v] \cdot F_{vi})}_{\text{Hypothesis 1}} + \underbrace{\beta(U_{vi} \cdot D_v \cdot F_{vi})}_{\text{Hypothesis 2}} + \underbrace{\beta C + \beta O + \beta(C \cdot A_{vi}) + \beta P_{vi}}_{\text{Control Variables}} + u_v + \varepsilon \quad (2)$$

Our general expectation is that the main effect of unemployment (U_{vi}) and the moderator effects of the macro-level unemployment ($U_{vi} \bullet \ln[LU_v]$) and hiring difficulty ($U_{vi} \bullet D_v$) are stronger for applicants who meet all skill requirements of the job (i.e., $F_{vi} = 1$).

Second, invariant characteristics of the vignettes (e.g., number of years in the labor market) may also, to a greater or lesser extent, be in line with the job requirements of any given real vacancy, resulting in a variable degree of what we define as *generic applicant-vacancy-fit*. As with skill fit, CVs with an insufficient generic fit may be screened out before an unemployment signal can take effect. To account for a possible moderator effect of generic fit in our models, we applied an ebalancing weight (see Hainmueller 2012) to align different sample distributions with respect to generic applicant-vacancy-fit across the 20 FSEs. The underlying vacancy-level measure of generic fit relies on an extra vignette (otherwise excluded from modeling), which was designed to be free from any potentially negative signals. Appendix I provides the details on the measurement of generic applicant-vacancy-fit and the reweighting procedure. In our robustness checks, we assessed how reweighting affected our results.

We present our main results using graphs and reduced tables. Estimates with p values below the 0.05 level were considered statistically significant. The full regression tables are provided in Appendix J. We conclude our analyses with a series of robustness checks, including an assessment of potential moderator bias owing to otherwise uncontrolled job and firm characteristics.

4. Results

First, we estimate a *reference model* including only unemployment and control variables. In line with previous research and the general expectation outlined in Section 2.2, we find a significant negative effect of unemployment on the evaluation of the fictitious job candidates ($b = -0.044$, $p = 0.001$; see Appendix P). In our *baseline model*, we add an interaction term of unemployment with the labor market level of unemployment, according to H1. Although the term shows the expected positive sign, it is not significant (Table 3, Baseline Model). This implies a tendency for higher levels of unemployment in a labor market segment to be linked to lower levels of unemployment stigmatization.

In the *full model*, which adjusts for micro-level skill fit (see Eq. 2 in Section 3.5), we find that the three-way interaction term among CVs with an unemployment spell, the long-term level of unemployment, and applicant-vacancy fit is positive and significant (Table 3, Full Model). Overall, the findings from the *full model* are consistent with H1, which proposes that unemployment stigmatization is mitigated in labor markets with persistently high levels of unemployment. Additional tests suggest that unemployment scarring is *gradually* reduced at higher levels

Table 3
Predicting recruiter assessments of job candidates (multi-level regression models).

Logged vignette ratings	Baseline Model	Full Model	Reduced Model
	Coef.[a]	Coef.[a]	Coef.[a]
Hypothesis 1 (H1)			
CV with unemployment (ref: CV without unemployment spell)	− 0.043** (0.016)	− 0.022 (0.017)	− 0.021 (0.015)
CV with unemployment × Skill fit	/	− 0.120* (0.050)	− 0.153*** (0.046)
CV with unemployment × Level of unemployment[b]	0.010 (0.009)	− 0.005 (0.010)	− 0.004 (0.010)
CV with unemployment × Level of unemployment[b] × Skill fit	/	0.087** (0.030)	0.079** (0.030)
Hypothesis 2 (H2)			
CV with unemployment × Hiring "very difficult"	− 0.018 (0.032)	0.004 (0.035)	/
CV with unemployment × Hiring "very difficult" × Skill fit	/	− 0.128 (0.090)	/
Level of unemployment (Labor market segment)[b]	0.054 (0.032)	0.089** (0.032)	0.090** (0.032)
Skill fit (ref: application does not meet all requirements)	/	0.643*** (0.064)	0.674*** (0.059)
Level of unemployment[b] × Skill fit	/	− 0.254*** (0.033)	− 0.247*** (0.034)
Hiring "very difficult" (ref: "very easy" to "quite difficult")	− 0.026 (0.043)	− 0.045 (0.045)	/
Hiring "very difficult" × Skill fit	/	0.120 (0.083)	/
Control variables[c]	YES	YES	YES
Constant	0.829*** (0.082)	0.749*** (0.081)	0.731*** (0.080)
Random effects parameters			
Level 2 (Recruiters): Standard deviation	0.494 (0.011)	0.481 (0.011)	0.481 (0.011)
Level 1 (Vignettes): Standard deviation	0.540 (0.007)	0.530 (0.007)	0.530 (0.007)
Number of Ratings (Vignettes)	18,001	18,001	18,001
Number of Recruiters	2088	2088	2088

Note: Significance levels (two-tailed tests):
*** $p < 0.001$.
** $p < 0.01$.
* $p < 0.05$.
[a] . Regression coefficients and cluster-robust standard errors (in parentheses).
[b] . Level of unemployment: Logged averaged rates for the 2010 to 2016 period (see Appendix F).
[c] . Model contains controls for all experimental variables including country interactions (full table in Appendix J).

of unemployment, indicating no discontinuities in its effects.[7] The results also suggest a weaker or non-existent stigma effect when a job applicant does not meet all skill requirements of the vacancy. In Table 3, the main effect of an unemployment spell in the CV and the interaction effect with the skill fit indicator support this argument. This suggests that unemployment exerts a strong and significant stigma effect that is limited to job applicants with all required qualifications. Thus, candidates with insufficient qualifications are screened out before stigma effects may unfold.

Testing H2 (Table 3, Full Model), we find no indication that unemployment scarring was reduced when recruiters considered hiring difficult. The effect is not significant, irrespective of whether a job candidate has the required skills or not. An omnibus test of all model terms including the indicator for hiring difficulty revealed that they are jointly insignificant. We checked the robustness of this result by testing several alternative specifications without finding any support for H2. First, we re-estimated the full model while excluding all interaction terms implied by H1, and then those tied to H2 (see Cinelli et al., 2022; Keele et al., 2020). This did not change any of the results, either for H1 or for H2 (see Appendix K). Second, we restricted the test of H2 to the main effect of hiring difficulty and its two-way interaction with unemployment by excluding all other terms related to H2 from the model. The main effect and interaction remained insignificant (individually and jointly). Third, the results remained unchanged when we reran the model with alternative dichotomizations or a continuous specification of our ordinal indicator for hiring difficulty. Finally, the Bayesian information criterion (BIC) was lower for the reduced model with all terms involving hiring difficulty removed than for the full model (see Table 3), indicating a superior fit of the former (see Raftery, 1995).[8] Thus, the lack of empirical support for H2 is a robust finding. The results remained consistent when we used an indicator for hiring urgency (instead of hiring difficulty), which might also reduce discriminatory behavior by recruiters.

In Fig. 1, we plotted the marginal effects of unemployment implied by the model (vertical axis) as a function of the average level of unemployment in the labor market segment (horizontal axis). We present separate plots for fictitious job candidates with (Panel A) and without (Panel B) a full set of suitable skills. For job candidates with suitable skills, the crucial interaction follows a convex curvilinear pattern, with effects leveling off as contextual unemployment rates increase (Fig. 1, Panel A). At the lowest observed unemployment rate (2.15 %), a substantial and significant stigma effect of minus 1.4 rating points emerges (on the original 11-point scale). This amounts to an effect of half a standard deviation in the untransformed ratings of hiring chances and is of similar size as the effects of education and job experience.[9] We further find that, as the long-term level of unemployment increases, the stigma effects weaken and approach zero. Above an unemployment level of 5.8 %, the effects are small ($<$ 0.23 rating points) and no longer significant at the 5 %-level (Fig. 1, Panel A). Given that we have no evidence for counter-balancing forces as suggested by the queuing arguments for H2, the pattern of effects shown in Panel A of Fig. 1 is fully consistent with the expectations derived from Goffman (H1). In contrast,

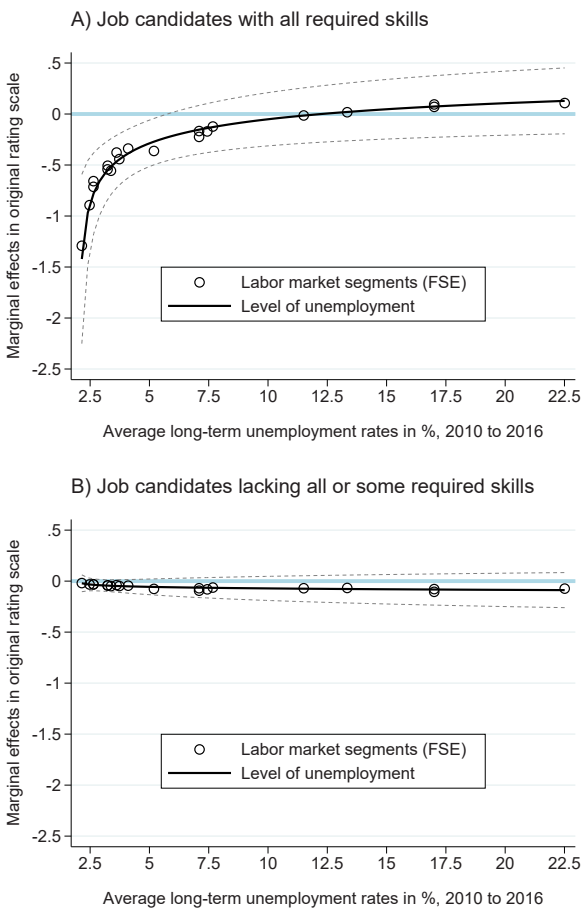

Fig. 1. Marginal effects of unemployment on hiring chances across labor market segments. Dashed lines are 95 % confidence intervals. *Panel A):* Effects for applications matching all vacancy skill requirements. *Panel B)*: Effects for applications not or only partially matching vacancy skill requirements. Full table of the model estimates is presented in Appendix J. Primacy effects and the country-specific experimental variables set to zero for calculation of marginal effects.

the marginal effects of unemployment for job applications without sufficient skills (Fig. 1, Panel B) are not associated with any effect owing to unemployment experience, irrespective of unemployment rates.

Finally, we highlight a few interesting implications of the models in Table 3. The results suggest a general tendency to assess job candidates more favorably when labor market conditions are harsh. This follows from the positive main effect of the level of unemployment in the labor market. Moreover, the strong main effect of skill fit indicates a huge advantage for job candidates with all the necessary qualifications and specializations. This advantage seems to dwindle in labor markets with high levels of unemployment, as there is a sizeable negative interaction effect between skill fit and high levels of unemployment.

4.1. Robustness checks

Our literature review suggests that unemployment scarring may depend on country-level moderators, including institutional variations in education and job-worker matching, as well as labor market regulations. Given the small number of countries in our study, it is hardly possible to directly assess such a moderator effect. Nevertheless, it is important to determine whether country differences alone can sufficiently explain our findings. Hence, we re-estimated the reduced model from Table 3, replacing the segmental unemployment rates with country dummies. The results shows that none of the relevant interactions between country and unemployment are significant (see Appendix L,

[7] Using a set of dummy variables for all possible dichotomizations of the continuous variable for long-term levels of unemployment (see Table 2), we tested whether the model fit (of the Reduced Model, see Table 3) would improve by substituting the continuous indicator (including its interactions) with one of its dichotomizations. The best-fitting of the 18 resulting models has a substantially higher median BIC of 33,220.6 across all imputations compared to the final model (see Table 3, Reduced Model with a median BIC of 33,162.8).

[8] Across all imputations, the BIC for the reduced model is invariantly lower. The median BICs for the reduced and the full model are 33,162.8, and 33,194.8, respectively.

[9] For example, the stigma effect is almost as strong as the total effect of suitable field-specific education and job experience and about twice as strong as that of the level of education (tertiary degree vs. compulsory schooling).

Country Model). Additionally, the BIC indicates an inferior fit for the model with country dummies compared to the reduced model from Table 3.[10] This finding indicates that *within-country variation* between occupational labor market segments contributed crucially to our overall result, implying that an explanation based on country-level arguments is insufficient.

Our indicator of the segment-specific levels of unemployment relies partly on expert ratings, which measure segment-specific deviations from the average national unemployment rate (known from official unemployment data). The calculation (see Appendix F) involves a rather arbitrary scaling parameter to transform the expert ratings into unemployment rates. We therefore re-estimated our models while choosing different scaling parameters that result in plausible segment-specific unemployment rates (see Appendix F). Our results proved robust against systematic variation of the scaling parameter (see Appendices J and L).

We used weights to balance different sample distributions in the 20 FSE with respect to generic applicant-vacancy fit (see Appendix I). To assess the impact of these weights, we reran our models on unweighted data. This induced minor differences in non-crucial model coefficients, but the substantive conclusions regarding both hypotheses did not change (see Appendix L and Appendix M). The most crucial result — the significant three-way interaction associated with H1 — becomes more apparent when we reweight the data to balance differences in generic fit. This is to be expected if the proposed reweighting for generic fit is successful in removing noise from the data (if H1 holds true) and indicates that balancing for generic fit may be a promising strategy for analyzing FSEs based on real-world vacancies. The weights to balance generic fit across the 20 FSEs are included in the supplemental materials to this article (see Appendix N; a detailed description of the procedure is provided in Appendix I).

The evidence addressing potential demand-side moderators of scarring effects on the level of firms, recruiters, and jobs is rather limited. There are some indications, however, pointing to the moderating effect of company size (Nüß, 2018), urgent staffing needs (Farber et al., 2017), and jobs with supervisor positions (Karren & Sherman, 2012). We hence extend our weights for generic applicant-vacancy-fit to further balance the 20 FSE-Samples regarding firm size (five or fewer employees), supervisor position (vs. employee/trainee positions), and the perceived importance of filling a vacant position ("very important" vs. "less important").[11] This procedure guarantees that the sample means of the potential moderators are the same in the 20 FSE and thus eliminates a possible moderator bias (see Appendix A). Running our main models using these extended weights did not substantially change our results (see Appendices J and L). The moderator bias owing to the additional demand-side variables thus seems negligible.

5. Discussion and conclusions

We used experimental methods to investigate how unemployment scarring in the hiring of young jobseekers depends on labor market conditions. We proposed two complementary micro-mechanisms that led us to assume opposing relationships at the labor market level. *First*, building on Goffman (1963), we argued that stigmatization effects owing to unemployment were weaker in labor markets with persistently high levels of unemployment (H1). *Second*, we inferred from queuing arguments that scarring effects owing to unemployment were weaker when jobs were hard to fill, that is, when unemployment and labor supply were low (H2). At the macro level of labor markets, our first argument suggested a negative relationship, whereas our second one suggested a positive relationship between the level of unemployment and scarring effects.

Our analysis used a set of standardized FSEs based on real job vacancies to assess the hiring chances of young workers in 20 different labor market contexts (five occupational fields in four European countries). The results supported H1, but we did not find evidence in favor of H2. The findings align with the assumption that the stigmatization of previously unemployed workers is reinforced in labor markets with persistently low unemployment levels. This stigmatizing effect is restricted to applicants who meet all skill requirements of the job, suggesting that unemployment is a downstream criterion in hiring when compared to adequate education and work experience.

The experimental design suggests the appropriateness of a causal interpretation of the observed stigma effects. The observed context variability across labor markets was robust to various tests for moderator bias. We further conducted various specification and robustness tests, which include the lack of evidence for H2, showing that the findings are robust throughout. The results regarding context variability are fully consistent with the Goffmanian perspective that focuses on how potential employers perceive the formerly unemployed. In summary, recruiters are more likely to perceive unemployment as self-inflicted (i.e., not caused by harsh labor market conditions or bad luck), when the level of unemployment is low (i.e., the group size is small). Simultaneously, unemployment represents a stronger and more conspicuous deviation from the social norm, making it a seemingly less ambiguous signal.

Given these results on unemployment stigmatization, it appears that other types of stigmatization may depend in a similar vein on the social context. This is particularly plausible for signals that, depending on the social context, are rather perceived as self-inflicted or fate, as is the case with unemployment. This potentially applies to a wide range of signals of success and failure in education and work. The best established example is the increasingly negative perception of the shrinking group of low-skilled workers in demanding labor markets (see Gesthuizen et al., 2011; Solga, 2005, 2008). These findings are well in line with Goffmanian reasoning and may serve as a more general model of context variation in the stigmatizing effects of potentially self-inflicted signals.

Our theoretical reasoning crucially relies on *causal locus* and presupposes that a given signal can be perceived as "self-inflicted." Therefore, it seems unlikely that the Goffmanian model of context variation also applies to social stigma arising from ascriptive characteristics like ethnicity. Research in social psychology suggests that stigmatization based on ascriptive characteristics depends in complex ways on various social context dimensions, including effective and/or perceived group size (see Samson & Bobo, 2014). Empirical research (ibid.; Mai, 2018, pp. 56–90, 141ff.) suggests that stigmatization *increases* with the group size of racial and ethnic minorities. There are also indications that the queuing approach may be more useful vis-à-vis context variation in stigmatization as evoked by ascriptive characteristics instead of unemployment (Baert et al., 2015; Kübler et al., 2018; Quillian & Midtbøen, 2021). These considerations let us assume that the Goffmanian model of context variation in stigmatization effects does not extend to ascriptive characteristics.

We focused on the stigmatization of young job seekers with extended spells of previous unemployment. The stigmatization we found is relevant beyond its immediate effect on hiring chances, because successful job applications are a bottleneck in the access to income, material security, and social participation. Hence, prolonged periods of unemployment may represent a paradigmatic "trigger event" (DiPrete, 2002; Gangl, 2006) that can induce a downward spiral whose consequences extend beyond the realm of work. Our findings show that, quite paradoxically, this risk is greatest in well-performing labor markets with invariably low levels of unemployment. In such environments, even well-skilled jobseekers face strong unemployment stigma in hiring, which may translate to long-lasting scars in and beyond employment.

[10] The median BIC over all imputations is higher for this model than for our main model (33,383 vs. 33,360).

[11] Based on the (imputed) variables *firmsize* (firm size), *jobpos* (supervisor position vs. employee/trainee positions), and *fillimpo* (perceived importance of filling a vacant position) from the recruiter survey.

Our study has some limitations. First, our data did not allow for a direct test of the micro mechanism behind H1 in recruiters' assessment of job candidates. The test of H1 relies on a comparative analysis of experimental data at the level of labor market segments. In contrast, we were able to subject our assumptions based on the queuing arguments (H2) to a direct test based on a micro-level measurement of the relevant recruiter perceptions at the individual vacancy level. Second, in labor markets with low levels of unemployment and a strong unemployment stigma, our analyses exploit between- and within-country variations from just two countries (Norway and Switzerland), both characterized by a strong dual vocational education and training system (see Imdorf et al., 2019, p. 97). In these countries, occupational specialization, training, and experience are of utmost importance for hiring chances (de Lange et al., 2014; Müller & Gangl, 2003; Shavit & Müller, 1998), which is also evident in our results (see also Appendix J). Future research can address the question of whether the association we found between unemployment levels and stigma also exists in countries with different education systems and a weaker role of occupations in the allocation of workers to jobs. Third, the experimental design is tailored to the first step in the multi-step hiring process, where job applications are screened to determine the ones that will be considered further (e.g., invited for a job interview). While the standardization required by our experimental design yields a high degree of internal and behavioral validity (Petzold & Wolbring, 2018), it limits the external validity of our findings to the first step in applicant selection — the initial screening of potentially many job applications. Moreover, the simplified and artificial situation implied by any FSE may pose a threat to external validity. We tried to counteract this by conducting the experiment with real vacancies and the responsible recruiters. This design also enabled a comprehensive control of possible moderator effects, for example, on the level of job characteristics.

With these limitations in mind, our study makes two contributions to sociological research on hiring and access to work. First, we add to the understanding of context variability in unemployment scarring by testing two potentially complementary micro-level explanations with opposing macro-level implications. Relying on 20 parallel FSEs, covering widely diverse labor market segments in four countries, we find no evidence that the interplay between current labor shortages and applicant queuing is a major driver of context-dependent unemployment scarring. Rather, stable long-term differences in labor market conditions prove decisive, which is consistent with the Goffmanian perspective on unemployment stigma. The study extends our theoretical understanding of how labor market conditions moderate unemployment stigmatization. The evidence presented supports the strong prevalence of unemployment stigmatization in labor markets with persistently low levels of unemployment, which is mitigated or fully suppressed at medium to high levels of unemployment. The sparse existing evidence suggests that the Goffmanian model is a valid general explanation of how stigmatization owing to potentially self-inflicted signals varies across social contexts. This covers a huge variety of potentially stigmatizing signals of failure and success in education, the labor market, and beyond.

Second, our study contributes to the literature by providing a rigorous test of the impact of unemployment scarring on hiring chances. We conducted a factorial survey experiment that was embedded in actual hiring processes, enabling a causal interpretation of the observed effects while strengthening external validity. Furthermore, we conducted our research across a comprehensive range of labor market conditions, which sets our study apart from previous studies.[12] Specifically, we were able to include low levels of unemployment (below 4 %), which provided a strong basis to exclude queueing mechanisms and evidence of scarring effects at low levels of unemployment, consistent with Birkelund et al. (2017) and aligning with the arguments derived from Goffman.

Our study highlights the benefits and challenges of cross-national experimental studies using real vacancies. The benefits include the potential for comparative causal inferences and understanding the role of context in shaping mechanisms at a micro level, whereas the challenges include harmonizing experimental design and controlling moderator bias. The latter can be addressed by comprehensively adjusting for possible moderators including job-applicant fit, which is particularly important when using real vacancies.

There is persuasive evidence from observational research (Filomena, 2023) that unemployment may have severe consequences for a wide range of outcomes in and beyond subsequent employment. However, the likely manifold causal mechanisms underlying all this are not sufficiently understood. Given that the access to jobs is key for subsequent chances in the labor market and beyond for many affected outcomes, our study focused on the impact of unemployment on the hiring chances at the beginning of a career. We found clear indications that, depending on labor market conditions, stigmatization of the unemployed is a main driver of unemployment scarring. This adds to an emergent strand of research addressing the role of employers and firms in generating inequalities in the labor market.

Funding

Collection of recruiter and expert data was funded by the Swiss State Secretariat for Education, Research and Innovation SERI (Grant Number 15.0089, 649395) as part of the Horizon 2020 project 'Negotiating early job-insecurity and labour market exclusion in Europe—NEGOTIATE' (Horizon 2020, Societal Challenge 6, H2020—YOUNG-SOCIETY-2014, YOUNG-1-2014).

CRediT authorship contribution statement

Robin Samuel: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Stefan Sacchi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Data Availability

A replication package is available at: https://osf.io/vd4gy/?view_only=01170e7a14a642d2a503aaee9019a80f.

Acknowledgements

We are grateful to Tamara Gutfleisch for her extensive comments. Additionally, we would like to thank the participants of the SEMILUX seminars, the Cathie Marsh Institute for Social Research (CMI) seminars, the Social Stratification Research Seminars, and various conferences where we presented our work. Any errors are our own.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.rssm.2024.100959.

References

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review, 30*(2), 175–187. https://doi.org/10.1177/001979397703000204

[12] For example, Farber et al. (2017) covered unemployment rates from 4.1 to 8.8 %, and Nunley et al. (2017) from 5.2 to 9.2 %.

Anseel, F., Lievens, F., Schollaert, E., & Choragwicka, B. (2010). Response rates in organizational science, 1995–2008: A meta-analytic review and guidelines for survey researchers. *Journal of Business and Psychology, 25*(3), 335–349. https://doi.org/10.1007/s10869-010-9157-6

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 6*(3), 128–138. https://doi.org/10.1027/1614-2241/a000014

Auspurg, K., & Hinz, T. (2015). *Factorial Survey Experiments*. Sage Publications,.

Ayllón, S., Valbuena, J., & Plum, A. (2022). Youth unemployment and stigmatization over the business cycle in Europe. *Oxford Bulletin of Economics and Statistics, 84*(1), 103–129. https://doi.org/10.1111/obes.12445

Baert, S. (2018). Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005. In S. M. Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance* (pp. 63–77). Springer International Publishing. https://doi.org/10.1007/978-3-319-71153-9_3.

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is There Less Discrimination in Occupations Where Recruitment Is Difficult?. In *ILR Review, 68* pp. 467–500). https://doi.org/10.1177/0019793915570873

Ball, P. (2011). *Mixed Signals: To what extent does male wage scarring vary with the characteristics of the local labour market in which unemployment was experienced?* (Discussion Paper 11/13). The University of Nottingham,.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Columbia University Press,.

Biewen, M., & Steffes, S. (2010). Unemployment persistence: Is there evidence for stigma effects. *Economics Letters, 106*(3), 188–190. https://doi.org/10.1016/j.econlet.2009.11.016

Bills, D. B. (1990). Employers' use of job history data for making hiring decisions. *The Sociological Quarterly, 31*(1), 23–35. https://doi.org/10.1111/j.1533-8525.1990.tb00315.x

Bills, D. B., Di Stasio, V., & Gërxhani, K. (2017). The demand side of hiring: employers in the labor market. *Annual Review of Sociology, 43*(1), 291–310. https://doi.org/10.1146/annurev-soc-081715-074255

Birkelund, G. E., Heggebø, K., & Rogstad, J. (2017). Additive or multiplicative disadvantage? the scarring effects of unemployment for ethnic minorities. *European Sociological Review, 33*(1), 17–29. https://doi.org/10.1093/esr/jcw030

Bonoli, G., & Hinrichs, K. (2012). Statistical discrimination and employers' recruitment. *European Societies, 14*(3), 338–361. https://doi.org/10.1080/14616696.2012.677050

Booth, A. L., Francesconi, M., & Frank, J. (2002). Temporary jobs: stepping stones or dead ends. *The Economic Journal, 112*(480), F189–F213. https://doi.org/10.1111/1468-0297.00043

Brand, J. E. (2015). The far-reaching impact of job loss and unemployment. *Annual Review of Sociology, 41*(1), 359–375. https://doi.org/10.1146/annurev-soc-071913-043237

Buchs, H., & Buchmann, M. (2017). *Job Vacancies and Unemployment in Switzerland 2006-2014: Labor Market Mismatch and the Significance of Labor Market Tightness for Unemployment Duration*. University of Zurich: Swiss Job Market Monitor,.

Buffel, V., & Van de Velde, S. (2019). Comparing negative attitudes toward the unemployed across european countries in 2008 and 2016: the role of the unemployment rate and job insecurity. *International Journal of Public Opinion Research, 31*(3), 419–440. https://doi.org/10.1093/ijpor/edy015

Bursell, M., Bygren, M., & Gähler, M. (2021). Does employer discrimination contribute to the subordinate labor market inclusion of individuals of a foreign background. *Social Science Research, 98*, Article 102582. https://doi.org/10.1016/j.ssresearch.2021.102582

Cardoso, A. R., Loviglio, A., & Piemontese, L. (2016). Misperceptions of unemployment and individual labor market outcomes. *IZA Journal of Labor Policy, 5*(1), 13. https://doi.org/10.1186/s40173-016-0069-6

Cinelli, C., Forney, A., & Pearl, J. (2022). A crash course in good and bad controls. *Sociological Methods & Research*. , Article 00491241221099552. https://doi.org/10.1177/00491241221099552

Clark, A. E., & Lepinteur, A. (2019). The causes and consequences of early-adult unemployment: evidence from cohort data. *Journal of Economic Behavior & Organization, 166*, 107–124. https://doi.org/10.1016/j.jebo.2019.08.020

de Lange, M., Gesthuizen, M., & Wolbers, M. H. J. (2014). Youth labour market integration across Europe. *European Societies, 16*(2), 194–212. https://doi.org/10.1080/14616696.2013.821621

Di Stasio, V. (2014). Education as a signal of trainability: Results from a vignette study with Italian employers. *European Sociological Review, 30*(6), 796–809. https://doi.org/10.1093/esr/jcu074

Di Stasio, V., & Lancee, B. (2020). Understanding why employers discriminate, where and against whom: The potential of cross-national, factorial and multi-group field experiments. *Research in Social Stratification and Mobility, 65*, Article 100463. https://doi.org/10.1016/j.rssm.2019.100463

DiPrete, T. A. (2002). Life course risks, mobility regimes, and mobility consequences: A comparison of Sweden, Germany, and the United States. *American Journal of Sociology, 108*(2), 267–309. https://doi.org/10.1086/344811

Dülmer, H. (2007). Experimental Plans in Factorial Surveys: Random or Quota Design. *Sociological Methods & Research, 35*(3), 382–409. https://doi.org/10.1177/0049124106292367

Eriksson, S., & Rooth, D.-O. (2014). Do employers use unemployment as a sorting criterion when hiring? evidence from a field experiment. *American Economic Review, 104*(3), 1014–1039. https://doi.org/10.1257/aer.104.3.1014

European Commission. (2019). *Labour Market and Wage Developments in Europe: Annual Review 2019*. Publications Office of the European Union,.

Eurostat. (2020). *Previous occupations of the unemployed*. ⟨http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_ugpis&lang=en⟩.

Farber, H. S., Silverman, D., & Wachter, T. M. von (2017). Factors determining callbacks to job applications by the unemployed: An audit study. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 3*(3), 168–201. https://doi.org/10.7758/rsf.2017.3.3.08

Filomena, M. (2023). Unemployment Scarring Effects: An Overview and Meta-analysis of Empirical Studies. *Italian Economic Journal*. https://doi.org/10.1007/s40797-023-00228-4

Finger, C. (2016). Institutional constraints and the translation of college aspirations into intentions—Evidence from a factorial survey. *Research in Social Stratification and Mobility, 46*, 112–128. https://doi.org/10.1016/j.rssm.2016.08.001

Fraser, R. T., Johnson, K., Hebert, J., Ajzen, I., Copeland, J., Brown, P., & Chan, F. (2010). Understanding employers' hiring intentions in relation to qualified workers with disabilities: preliminary findings. *Journal of Occupational Rehabilitation, 20*(4), 420–426. https://doi.org/10.1007/s10926-009-9220-1

Gaddis, S. M. (2018). An Introduction to Audit Studies in the Social Sciences. In S. M. Gaddis (Ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance* (pp. 3–44). Springer International Publishing. https://doi.org/10.1007/978-3-319-71153-9_1.

Gangl, M. (2004). Institutions and the Structure of Labour Market Matching in the United States and West Germany. *European Sociological Review, 20*(3), 171–187. https://doi.org/10.1093/esr/jch016

Gangl, M. (2006). Scar Effects of Unemployment: An Assessment of Institutional Complementarities. *American Sociological Review, 71*(6), 986–1013. https://doi.org/10.1177/000312240607100606

Gebel, M. (2009). Fixed-Term Contracts at Labour Market Entry in West Germany: Implications for Job Search and First Job Quality. *European Sociological Review, 25*(6), 661–675. https://doi.org/10.1093/esr/jcp005

Gesthuizen, M., Solga, H., & Künster, R. (2011). Context Matters: Economic Marginalization of Low-Educated Workers in Cross-National Perspective. *European Sociological Review, 27*(2), 264–280. https://doi.org/10.1093/esr/jcq006

Gibbons, R., & Katz, L. F. (1991). Layoffs and Lemons. *Journal of Labor Economics, 9*(4), 351–380. https://doi.org/10.1086/298273

Goffman, E. (1963). *Stigma: Notes on the Management of Spoiled Identity*. Prentice Hall,.

Gregg, P., & Tominey, E. (2005). The wage scar from male youth unemployment. *Labour Economics, 12*(4), 487–509. https://doi.org/10.1016/j.labeco.2005.05.004

Grund, S., Lüdtke, O., & Robitzsch, A. (2018). Multiple Imputation of Missing Data for Multilevel Models: Simulations and Recommendations. *Organizational Research Methods, 21*(1), 111–149. https://doi.org/10.1177/1094428117703686

Gutfleisch, T., Samuel, R., & Sacchi, S. (2021). The application of factorial surveys to study recruiters' hiring intentions: Comparing designs based on hypothetical and real vacancies. *Quality & Quantity, 55*(3), 775–804. https://doi.org/10.1007/s11135-020-01012-7

Heidenreich, M. (2015). The End of the Honeymoon: The Increasing Differentiation of (long-Term) Unemployment Risks in Europe. *Journal of European Social Policy, 25*(4), 393–413. https://doi.org/10.1177/0958928715594544

Hendricks, W., DeBrock, L., & Koenker, R. (2003). Uncertainty, Hiring, and Subsequent Performance: The NFL Draft. *Journal of Labor Economics, 21*(4), 857–886. https://doi.org/10.1086/377025

Hillmer, S., Hillmer, B., & McRoberts, G. (2004). The real costs of turnover: Lessons from a call center. *Human Resource Planning, 27*(3), 34 (+).

Ho, G. C., Shih, M., Walters, D. J., & Pittinsky, T. L. (2012). The Psychological Stigma of Unemployment: When joblessness leads to being jobless. *Academy of Management Proceedings, 2012*(1), 15371. https://doi.org/10.5465/AMB.15371abstract

Hyggen, C., Imdorf, C., Parsanaglou, D., Sacchi, S., Samuel, R., Stoilova, R., Shi, L.P., Yfanti, A., & Yordanova, G. (2016). *Understanding unemployment scars: A vignette Experiment of employers' decisions in Bulgaria, Greece, Norway and Switzerland* (7.2; NEGOTIATE Working Paper). ⟨http://orbilu.uni.lu/handle/10993/29309⟩.

ILO. (2012). *International Standard Classification of Occupations: ISCO-08*. International Labour Office,.

Imdorf, C., Shi, L.P., Sacchi, S., Samuel, R., Hyggen, C., Stoilova, R., Yordanova, G., Boyadjieva, P., Ilieva-Trichkova, P., & Parsanoglou, D. (2017). *Explaining Employers' Hiring Decisions: A Comparative Study of Employers' Risk Assessment* (7.3; NEGOTIATE Working Paper). University of Basel.

Imdorf, C., Shi, L. P., Sacchi, S., Samuel, R., Hyggen, C., Stoilova, R., Yordanova, G., Boyadjieva, P., Ilieva-Trichkova, P., & Parsanoglou, D. (2020). *NEGOTIATE Employer Survey. Scientific Use File. Data Documentation (NEGOTIATE Working Paper)*. University of Basel,. ⟨https://nsd.no/data/individ/publikasjoner/NSD2644/NY%20INFORMASJON%2013.05.2020%20NEGOTIATE_Employer_Survey_SUF_Documentation_v2_ci.pdf⟩.

Imdorf, C., Shi, L. P., Sacchi, S., Samuel, R., Hyggen, C., Stoilova, R., Yordanova, G., Boyadjieva, P., Ilieva-Trichkova, P., Parsanoglou, D., & Yfanti, A. (2019). Scars of early job insecurity across Europe: Insights from a multi-country employer study. In B. Hvinden, C. Hyggen, M. A. Schoyen, & T. Sirovátka (Eds.), *Youth Unemployment and Job Insecurity in Europe. Problems, Risk Factors and Policies* (pp. 93–116). Edward Elgar. https://doi.org/10.4337/9781788118897.00011.

Jones, E. E., Farina, A., Hastorf, A. H., Markus, H., Miller, D. T., & Scott, R. E. (1984). *Social Stigma: The Psychology of Marked Relationships*. Freeman and Company,.

Karren, R., & Sherman, K. (2012). Layoffs and unemployment discrimination: A new stigma. *Journal of Managerial Psychology, 27*(8), 848–863. https://doi.org/10.1108/02683941211280193

Karsan, R. (2007). Calculating the cost of turnover (https://doi.org/doi:) *Employment Relations Today, 34*(1), 33–36. https://doi.org/10.1002/ert.20139.

Keele, L., Stevenson, R. T., & Elwert, F. (2020). The causal interpretation of estimated associations in regression models. *Political Science Research and Methods, 8*(1), 1–13. https://doi.org/10.1017/psrm.2019.31

Kroft, K., Lange, F., & Notowidigdo, M. J. (2013). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*. *The Quarterly Journal of Economics, 128*(3), 1123–1167. https://doi.org/10.1093/qje/qjt015

Kübler, D., Schmid, J., & Stüber, R. (2018). Gender discrimination in hiring across occupations: A nationally-representative vignette study. *Labour Economics, 55*, 215–229. https://doi.org/10.1016/j.labeco.2018.10.002

Kunovich, R. M. (2012). Perceived Unemployment. *International Journal of Sociology, 42*(4), 100–123. https://doi.org/10.2753/IJS0020-7659420405

Liechti, F., Fossati, F., Bonoli, G., & Auer, D. (2017). The Signalling Value of Labour Market Programmes. *European Sociological Review, 33*(2), 257–274. https://doi.org/10.1093/esr/jcw061

Luijkx, R., & Wolbers, M. H. J. (2009). The Effects of Non-Employment in Early Work-Life on Subsequent Employment Chances of Individuals in The Netherlands. *European Sociological Review, 25*(6), 647–660. https://doi.org/10.1093/esr/jcp002

Lupi, C., & Ordine, P. (2002). Unemployment scarring in high unemployment regions. *Economics Bulletin, 10*(2), 1–8.

Mai, Q. D. H. (2018). *Precarious Work, Race, Space and Labor Market Stratification in the New Economy* (PhD dissertation). Vanderbilt University, (PhD dissertation) ⟨https://ir.vanderbilt.edu/handle/1803/15416⟩.

Manzoni, A., & Mooi-Reci, I. (2020). The cumulative disadvantage of unemployment: Longitudinal evidence across gender and age at first unemployment in Germany. *PLOS ONE, 15*(6), Article e0234786. https://doi.org/10.1371/journal.pone.0234786

Marshall, A., Altman, D. G., Holder, R. L., & Royston, P. (2009). Combining estimates of interest in prognostic modelling studies after multiple imputation: Current practice and guidelines. *BMC Medical Research Methodology, 9*(1), 57. https://doi.org/10.1186/1471-2288-9-57

McDonald, P. (2019). How Factorial Survey Analysis Improves Our Understanding of Employer Preferences. *Swiss Journal of Sociology, 45*(2), 237–260. https://doi.org/10.2478/sjs-2019-0011

Mooi-Reci, I., & Ganzeboom, H. B. (2015). Unemployment scarring by gender: Human capital depreciation or stigmatization? Longitudinal evidence from the Netherlands, 1980–2000. *Social Science Research, 52*, 642–658. https://doi.org/10.1016/j.ssresearch.2014.10.005

Mousteri, V., Daly, M., & Delaney, L. (2018). The scarring effect of unemployment on psychological well-being across Europe. *Social Science Research, 72*, 146–169. https://doi.org/10.1016/j.ssresearch.2018.01.007

Müller, W., & Gangl, M. (2003). The Transition from School to Work: A European Perspective. In W. Müller, & M. Gangl (Eds.), *Transitions from Education to Work in Europe: The Integration of Youth into EU Labour Markets* (pp. 1–22). Oxford University Press. https://doi.org/10.1093/0199252475.001.0001.

Nunley, J. M., Pugh, A., Romero, N., & Seals, R. A. (2017). The Effects of Unemployment and Underemployment on Employment Opportunities: Results from a Correspondence Audit of the Labor Market for College Graduates. *ILR Review, 70*(3), 642–669. https://doi.org/10.1177/0019793916654686

Nüß, P. (2018). *Duration dependence as an unemployment stigma: Evidence from a field experiment in Germany* (Working Paper 2018–06). Economics Working Paper. ⟨https://www.econstor.eu/handle/10419/179941⟩.

Oberholzer-Gee, F. (2008). Nonemployment stigma as rational herding: A field experiment. *Journal of Economic Behavior & Organization, 65*(1), 30–40. https://doi.org/10.1016/j.jebo.2004.05.008

O'Connell, M., & Kung, M.-C. (2007). The Cost of Employee Turnover. *Industrial Management, 49*(1), 14–19.

OECD. (2005). *How Persistent are Regional Disparities in Employment? The Role of Geographic Mobility*. OECD,.

OECD. (2023). *OECD Statistics*. ⟨https://stats.oecd.org/⟩.

Omori, Y. (1997). Stigma Effects of Nonemployment. *Economic Inquiry, 35*(2), 394–416. https://doi.org/10.1111/j.1465-7295.1997.tb01918.x

Pedulla, D. S. (2016). Penalized or Protected? Gender and the Consequences of Nonstandard and Mismatched Employment Histories. *American Sociological Review, 81*(2), 262–289. https://doi.org/10.1177/0003122416630982

Petzold, K., & Wolbring, T. (2018). What Can We Learn From Factorial Surveys About Human Behavior? *Methodology, 15*(1), 19–30. https://doi.org/10.1027/1614-2241/a000161

Protsch, P. (2021). Employers' recruitment contexts and hiring preferences in the German youth labor market. *Research in Social Stratification and Mobility, 73*, Article 100608. https://doi.org/10.1016/j.rssm.2021.100608

Quillian, L., & Midtbøen, A. H. (2021). Comparative Perspectives on Racial Discrimination in Hiring: The Rise of Field Experiments. *Annual Review of Sociology, 47*(1), 391–415. https://doi.org/10.1146/annurev-soc-090420-035144

Raaum, O., & Røed, K. (2006). Do Business Cycle Conditions at the Time of Labor Market Entry Affect Future Employment Prospects? (JSTOR) *The Review of Economics and Statistics, 88*(2), 193–210. https://doi.org/10.1162/rest.88.2.193.

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology, 25*, 111–163. https://doi.org/10.2307/271063

Reskin, B. F. (1991). Labour Markets as Queues. A Structural Approach to Changing Occupational Sex Segregation. In J. Huber (Ed.), *'Macro-Micro Linkages in Sociology* (pp. 170–192). Sage.

Samson, F. L., & Bobo, L. D. (2014). Ethno-Racial Attitudes and Social Inequality. In J. D. McLeod, E. J. Lawler, & M. Schwalbe (Eds.), *Handbook of the Social Psychology of Inequality* (pp. 515–545). Netherlands: Springer. https://doi.org/10.1007/978-94-017-9002-4_21.

Shavit, Y., & Müller, W. (1998). *From School to Work. A Comparative Study of Educational Qualifications and Occupational Destinations*. Clarendon Press,.

Shi, L. P., & Di Stasio, V. (2022). Finding a job after unemployment—Education as a moderator of unemployment scarring in Norway and German-speaking Switzerland. *Socio-Economic Review, 20*(3), 1125–1149. https://doi.org/10.1093/ser/mwaa056

Shi, L. P., Imdorf, C., Samuel, R., & Sacchi, S. (2018). How unemployment scarring affects skilled young workers: Evidence from a factorial survey of Swiss recruiters. *Journal for Labour Market Research, 52*(1), 7. https://doi.org/10.1186/s12651-018-0239-7

Solga, H. (2005). *Ohne Abschluss in die Bildungsgesellschaft*. Verlag Barbara Budrich,.

Solga, H. (2008). Lack of Training. Employment Opportunities for Low-Skilled Persons from a Sociological and Microeconomic Perspective. In K. U. Mayer, & H. Solga (Eds.), *Skill Formation. Interdisciplinary and Cross-National Perspectives* (pp. 173–204). Cambridge University Press.

Sørensen, A. B., & Kalleberg, A. L. (1981). An Outline of a Theory of Matching Persons to Jobs. In I. Berg (Ed.), *Sociological Perspectives on Labor Markets* (pp. 49–74). Academic Press.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. https://doi.org/10.2307/1882010

Stewart, M. B. (2007). The interrelated dynamics of unemployment and low-wage employment. *Journal of Applied Econometrics, 22*(3), 511–531. https://doi.org/10.1002/jae.922

Strandh, M., Winefield, A., Nilsson, K., & Hammarström, A. (2014). Unemployment and mental health scarring during the life course. *European Journal of Public Health, 24*(3), 440–445. https://doi.org/10.1093/eurpub/cku005

Thurow, L.C. (1975). *Generating inequality*. Basic Books. ⟨http://agris.fao.org/agris-search/search.do?recordID=US201300524426⟩.

Van Belle, E., Di Stasio, V., Caers, R., De Couck, M., & Baert, S. (2018). Why Are Employers Put Off by Long Spells of Unemployment? *European Sociological Review, 34*(6), 694–710. https://doi.org/10.1093/esr/jcy039

Van Dijk, J., & Folmer, H. (1999). Wage Effects of Unemployment Duration and Frequency. *Journal of Regional Science, 39*(2), 319–337. https://doi.org/10.1111/1467-9787.00136

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research, 38*(3), 505–520. https://doi.org/10.1016/j.ssresearch.2009.03.004

Wanberg, C. R. (2012). The Individual Experience of Unemployment. *Annual Review of Psychology, 63*(1), 369–396. https://doi.org/10.1146/annurev-psych-120710-100500

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377–399. https://doi.org/10.1002/sim.4067