WORKING PAPER

UNEMPLOYMENT, INACTIVITY, AND HIRING CHANCES: A SYSTEMATIC REVIEW AND META-ANALYSIS

Liam D’hert
Stijn Baert
Louis Lippens

July 2024
2024/1089

GHENT UNIVERSITY

Department of Economics

Faculty of Economics and Business Administration
Campus Tweekerken, St.-Pietersplein 5, 9000 Ghent - BELGIUM

Unemployment, inactivity, and hiring chances: A systematic review and meta-analysis

Liam D'hert,[i] Stijn Baert,[ii] and Louis Lippens[iii]

Abstract

Policymakers' push for higher employment rates requires the activation of long-term unemployed jobseekers and inactive persons. However, stigma related to unemployment or inactivity can hinder their hiring chances when applying for a job. This systematic literature review investigates whether, when, and why periods of not working are penalised in hiring. Our review confirms that employers generally treat the unemployed and inactive less favourably than their employed counterparts. A meta-regression analysis of transnational experimental data points to heterogeneity by the duration of being out of work: short-term unemployment of up to six months positively affects hiring prospects, while the adverse effects of unemployment scarring become noticeable after about twelve months. We highlight evidence for signalling mechanisms underlying this pattern: immediate availability offsets the negative signals in short spells, whereas expectations about reduced productivity primarily drive the negative impact of longer spells. The latter negative signal is more pronounced when unemployment rates are low.

Keywords: unemployment, inactivity, hiring chances, systematic review, meta-analysis

JEL Classification: E24, J24, J64

[i] Corresponding author. Ghent University, Sint-Pietersplein 6, 9000 Ghent, Belgium. Liam.Dhert@UGent.be. ORCID: 0009-0002-3542-5204.

[ii] Ghent University, University of Antwerp, Université catholique de Louvain and Institute of Labor Economics. ORCID: 0000-0002-1660-5165.

[iii] Ghent University. ORCID: 0000-0001-7840-2753.

1. Introduction

Increasing the employment rate is a core ambition for governments in many European countries, not least because an ageing population overstretches social security systems (ILO, 2016; OECD, 2019, 2021). The employment rate is the percentage of employed people within the so-called 'recruitment population', usually defined as those between the ages of 20 – or sometimes 25 – and 64 years old (Baert, 2021). As well as the employed, the recruitment population includes the unemployed and the inactive. Encouraging these subgroups to work is crucial in raising the employment rate. They are mutually exclusive groups: the unemployed are those not working but available for and seeking work, while the inactive are neither working nor looking for a job (Aysun et al., 2014; Baert, 2021). The inactive can be further divided into five subgroups: the discouraged unemployed who have given up looking for work, those who have retired early, homemakers and caregivers, the long-term sick and disabled, and full-time students (Baert, 2021; ILO, 2016).

However, stigma on the part of employers may be a barrier to engaging the (long-term) unemployed and inactive population in employment. From a theoretical perspective, first, human capital theory suggests that employers believe a period of non-employment diminishes an individual's job-related training (Acemoglu, 1995). As work provides an opportunity to gain and retain knowledge and skills, such periods likely hinder the accumulation of additional human capital and may even cause a depreciation of previously acquired skills (Acemoglu, 1995; Becker, 1964; Pissarides, 1992). Second, signalling theory, as proposed in the stigma effect model of Vishwanath (1989), posits that employers may interpret a period of non-employment as an indicator of reduced ability and productivity. Since employers have only limited and imperfect information about a candidate's productivity when making hiring decisions, they often rely on the candidate's employment history to evaluate and minimise the mis-hire risk.

Given these theoretical rationales, the duration of a period of unemployment is expected to be a critical moderator of the unfavourable treatment of (long-term) unemployed and (formerly)[1] inactive job candidates. The longer a person remains jobless,

[1] When starting to look for work after a long period of inactivity, a person is no longer inactive in the strict sense of the word. Nevertheless, the preceding period of inactivity may still contribute to unequal hiring chances.

the less productive they are perceived to be (Bonoli & Hinrichs, 2012; Lockwood, 1991). This perception stems from the assumption that highly productive individuals secure employment sooner (Arrow, 1973; Connelly et al., 2011; Spence, 1973). Rational herding theory supports this idea, suggesting that employers may interpret an applicant's unemployment as indicative of negative evaluations by other employers who have opted not to employ them (Anderson & Holt, 1997; Banerjee, 1992; Bonoli & Hinrichs, 2012). This perception could heighten employers' reluctance to hire unemployed individuals, prolonging their unemployment spell (Banerjee, 1992). Employers may perceive the unemployment or inactivity spell as less informative for individuals out of employment for a relatively short period because of little to no skill depreciation (Atkinson et al., 1996; Pissarides, 1992). They may acknowledge that matching workers to firms takes time and that, in this respect, a (short) period of unemployment is often inevitable (Pissarides, 2000).

Besides unemployment duration, labour market conditions affect the signal unemployment conveys. In a loose labour market marked by a high unemployment rate, employers are more prone to attribute unemployment to external factors (e.g. an economic downturn) than to individual shortcomings (Gibbons & Katz, 1991; van den Berg & van Ours, 1996). This shift in perception can mitigate the stigma associated with unemployment, potentially reducing the negative impact on future job prospects for the unemployed. Consequently, individuals who are unemployed during periods of high unemployment may find it easier to (re-)enter the labour market than when unemployment is low.

Over the past two decades, the empirical literature estimating the unfavourable treatment of unemployed and formerly inactive job candidates has expanded substantially. To inform scientists and policymakers on the mixed findings in this literature and its gaps, we conduct the first comprehensive systematic review.[2] First, we map whether the unemployed and inactive experience inferior treatment from employers throughout the job application process. We record each study's outcome, noting whether the unemployed or

[2] We acknowledge two recent reviews on the broader topic of unemployment scarring. Borland (2020) summarises this literature, albeit non-systematically, employing non-standard meta-analysis techniques, and with a focus on Australian studies. Similarly, Filomena (2023) provides an overview of empirical studies on unemployment scarring through a meta-regression analysis. However, this review does not explore the underlying demand-side mechanisms of unemployment scarring nor analyse the precise impact of different durations of unemployment on hiring chances. We focus on unemployment scarring from a demand-side perspective, emphasising employer perceptions and the impact of particular durations of unemployment.

inactive have a lower, equal or higher likelihood of being offered a job than employed applicants. When different outcomes depend on moderators, we also categorise the results at the sub-study level. Second, we provide a qualitative overview of the empirical support for the aforementioned theoretical channels for this treatment. Third, we investigate the point at which demand-side unemployment scarring[3] takes effect, tackling the lack of consensus around its onset in the existing literature. We address this through a state-of-the-art meta-analysis of all recent correspondence experiments (encompassing nearly 67,000 fictitious applicants across seven countries) on unemployment impact on hiring chances.

We show that the long-term unemployed with spells exceeding twelve months are at a significant hiring disadvantage vis-à-vis the employed. Hiring chances decline noticeably as unemployment continues. However, the short-term unemployed with spells shorter than six months are preferred over their employed counterparts, likely due to their immediate availability. Our review confirms that long-term unemployment typically signals lower motivation, competence, and productivity. However, certain applicant characteristics may mitigate (such as social skills) or even counterbalance (such as race) this negative signal. Similarly, labour market conditions may influence the penalties faced by the unemployed, who are at a greater disadvantage in periods, regions or sectors characterised by a tight labour market. Research on inactivity, although limited, suggests that a period of inactivity elicits a harsher penalty than one of unemployment.

The article proceeds as follows. In Section 2, we outline the data and methods used to systematically chart and synthesise the existing literature on unemployment and inactivity scarring from a demand-side perspective. We then present in Section 3 a systematic overview of the empirical findings on the effect of unemployment or inactivity on hiring chances alongside (i) a synopsis of frequently highlighted moderators, (ii) a meta-analysis of the impact of the duration of unemployment on the scarring effect, and (iii) an overview of evidence for the underlying mechanisms. Section 4 concludes with a synopsis and policy recommendations.

[3] We do not intend to investigate the supply-side explanations of unemployment and inactivity scarring. For instance, the unemployed and inactive may indeed experience skill loss, become less motivated or see their professional networks weaken while out of work, making it harder for them to find a job. Instead, we focus on unemployment scarring from a demand-side perspective, emphasising employers' perceptions of these issues. Henceforth, the term 'scarring' refers to scarring from a demand-side perspective.

2. Data and methods

In the following subsections, we describe the study identification strategy, the screening process, the data extraction, and the synthesis methods.

2.1. Study identification

We systematically searched for studies that examined the impact of unemployment or inactivity on hiring chances through the Web of Science Core Collection. We adopted the SPIDER framework (Sample, Phenomenon of Interest, Design, Evaluation, Research type) developed by Cooke et al. (2012) to identify and select studies according to predefined eligibility criteria. As can be seen from Table 1, (i) the 'Sample' criterion was limited to unemployed, formerly unemployed, and inactive individuals, including homemakers and the discouraged unemployed but excluding full-time students, the long-term sick and disabled, and early retirees;[4] (ii) the 'Phenomenon of Interest' criterion was restricted to unequal treatment in the hiring of unemployed and inactive individuals; (iii) the 'Design' criterion covered empirical studies including (quasi-)experiments, field experiments, and survey research; (iv) the 'Evaluation' criterion was limited to hiring decisions; and (v) the 'Research type' criterion was restricted to quantitative and mixed-methods research. As the unequal treatment of specific applicant groups can be more precisely and causally tested and disentangled through experimental research than through correlational research that relies on secondary data (due to potential unobserved heterogeneity; Gaddis, 2018; Neumark, 2018), we excluded studies that use the latter approach. Our search was not explicitly constrained by a start date, although the Web of Science Core Collection only archives articles from 1955 onwards. The articles included in the study were published no later than 2023, as this was the most recent complete calendar year at the time of data collection.

< Table 1 about here >

[4] Since education is an investment in later labour market performance, we exclude students from the inactive individuals considered (Blundell et al., 1999). Furthermore, the hiring chances of the long-term sick and disabled are expected to be affected primarily by stigma relating to health rather than not having worked for a longer period *per se*, thus placing them beyond the scope of this literature review. Last, early retirees have voluntarily opted to exit the workforce, making their return highly unlikely (Baert, 2021).

Our search strategy consisted of a manual search of peer-reviewed articles. As concepts are often inaccurately classified as a result of imprecise predefined terms (Lefebvre et al., 2023), we identified a set of keywords commonly used in the texts of existing papers and reviews on our topic.[5] First, we systematically searched the Web of Science Core Collection based on these keywords. Second, we conducted backward and forward citation searches using the full text of the identified articles to reduce the risk of overlooking papers. All searches were performed in January 2024.

2.2. Study screening

The search process initially resulted in 2,245 records. We then excluded Web of Science categories that were unlikely to yield relevant results, maintaining our focus on social science research (an overview of the excluded categories can be found in Table A2 in the appendix). This additional filtering process resulted in the retention of 1,677 articles, with five duplicate articles which we removed. We subsequently evaluated the titles and abstracts of these 1,672 articles against our eligibility criteria and identified 103 records that required a full-text review. From these records, we excluded (i) 27 studies (26%) because they focused on labour market outcomes other than unequal treatment within hiring and selection processes (e.g. career quality); (ii) 18 studies (17%) because they examined vulnerable groups other than unemployed or inactive individuals (e.g. applicants with criminal justice involvement); (iii) 17 studies (17%) because they indirectly investigated unequal hiring chances (e.g. through the use of existing datasets); and (iv) 15 studies (15%) because they focused on employee rather than employer perspectives (e.g. work-related barriers experienced by inactive individuals). Finally, we could not find the complete text for three studies (3%) because the Web of Science database indexed only the titles without providing any other details or information on the authors.

Thus, we identified 23 studies that fully matched the criteria. Subsequently, we conducted a backward and forward citation search within these articles, yielding 830 new unique records. We retained five additional articles after screening (excluding 819 records

[5] Our search used the following keywords: ‘unemployed’, ‘nonemployed’, ‘inactive’, ‘out of work’, ‘jobless’, ‘housewife’, ‘househusband’, ‘homemaker’, ‘discouraged’, ‘gap’, ‘break’, ‘interruption’, and ‘spell’ in combination with ‘hiring’, ‘hire’, ‘job offer’, ‘application’, ‘employment success’, and ‘callback’.

based on title and abstract review and excluding six articles based on a full-text evaluation). No further relevant records emerged from subsequent citation and reference searches, leaving 28 retained articles. Figure 1 provides a structured overview of the study selection process, which adheres to the PRISMA framework (Page et al., 2021).

< Figure 1 about here >

2.3. Data extraction

Following the study identification and screening, we registered the following general information from each study: bibliographic details (including authors, year of publication, period of data collection, and region), employment status (in terms of unemployment, former unemployment, or inactivity), hiring prospects (in terms of job interview invitation or job offer), type of data (i.e. field experiment, lab experiment, or survey), main findings, and underlying mechanisms. In addition, in light of our 'when'-question we recorded the exact unemployment spell lengths (in months) assigned to the treatment groups (i.e. fictitious unemployed applicants) in all field and lab experiments. We also documented the gender and age of the fictitious applicants to investigate whether the effect of unemployment or inactivity varied by these factors. Table 2 provides descriptive statistics of the extracted variables.

< Table 2 about here >

To meta-analytically pinpoint the precise impact of various durations of unemployment, we tracked the callback rate by unemployment duration for all field experiments, that is, correspondence experiments with fictitious job applications with diverging unemployment treatments. If the authors reported different outcomes for other sub-groups, we logged these as separate treatment effects (e.g. Pedulla, 2018, reports different callback rates for unemployed ethnic majority applicants versus unemployed ethnic minority applicants). Additionally, we noted the classification of the outcome variable: if a callback entailed an invitation to a job interview (or any broadly defined positive response, such as a request for information), we labelled it 'interview invitation' (or 'positive reaction'). To calculate the unemployed–employed callback ratios, we documented the number of fictitious job applications and the number of positive callbacks in both the treatment and control groups.

Overall, the experiments involved almost 67,000 fictitious job applicants in response to genuine job vacancies across a final meta-sample of 90 treatment effects.

2.4. Synthesis

To investigate whether, when, and why the long-term unemployed and formerly inactive face inferior treatment during the hiring process, we present (i) an overall qualitative synthesis and (ii) a quantitative meta-regression analysis to address the question of when this occurs. As the research on the hiring chances of the inactive appeared scarce (see Section 3.1), we only applied meta-analytic synthesis to the study-level data of literature concerning the hiring prospects of the unemployed (versus the employed).

2.4.1. Qualitative synthesis

To give the reader a concise overview of the literature on unemployment and inactivity scarring from a demand-side perspective, we assembled the key data from each study into a comprehensive table. This overview includes the authors and publication year, region and country, type of non-employment, type of outcome, type of data, a summary of the main findings, and the mechanisms underlying these findings. The results section discusses the convergences and divergences among the studies included.

2.4.2. Meta-analysis

A meta-analysis systematically combines and analyses data from multiple studies, allowing researchers to derive generalisable conclusions by aggregating findings across various contexts. For a meta-analysis to produce valid and meaningful insights, the studies selected must include similar research questions and comparable designs and methods. Below, we describe the meta-analytic synthesis methods we applied to the extracted data.

From the study-level data extracted from the experimental studies on unemployment and hiring chances (see Section 2.3), we estimated unemployed–employed callback ratios (CB_k), as specified in Equation 1. These ratios are determined by dividing two proportions: (i) the proportion of positive callbacks for the group of unemployed job seekers (u_k) (i.e. the treatment group) relative to the total number of sent applications for this group ($n_{k(unemployed)}$), and (ii) the proportion of positive callbacks for the group of employed job seekers (e_k) (i.e. the control group) relative to the total number of sent applications for this group ($n_{k(employed)}$). Some studies included multiple sub-studies characterised by different treated groups (e.g. short-term versus long-term unemployed). In turn, most (sub-)studies included multiple treatment effects, delineated by differences in unemployment status, unemployment duration, country, gender, or age. The resulting number of treatment effects is defined as *k*.

$$CB_k = \frac{u_k / n_{k(unemployed)}}{e_k / n_{k(employed)}} \tag{1}$$

Subsequently, we calculated pooled unemployed–employed callback ratios for the entire sample and the various subsamples delineated by different unemployment durations to pinpoint the onset of unemployment scarring. The general specification, following Harrer et al. (2021), is given in Equation 2, where CB_{pooled} is the pooled unemployed–employed callback ratio; CB_k represents the unemployed–employed callback ratio for effect *k*; ε_k and ζ_k are the sampling (within-study) and distributional (between-study) errors; and w_k^* is the variance-adjusted weight for effect *k*.

$$CB_{pooled} = \frac{\sum_{k=1}^{K}(CB_k + \varepsilon_k + \zeta_k)w_k^*}{\sum_{k=1}^{K} w_k^*} \tag{2}$$

Our calculations of the pooled callback ratios incorporated several estimation methods. First, we calculated between-study error variance (ζ_k) using the Restricted Maximum Likelihood (REML) estimator (Viechtbauer, 2005). The REML estimator provides more accurate and less biased estimates than alternative estimators when the effect sizes are heterogeneous, as we expected them to be based on the various contexts of the included studies (Langan et al., 2019). Second, we used the Mantel–Haenszel method for binary outcome data without continuity correction to estimate the base of the variance-adjusted weights. This approach relies on the underlying count data, where large-sample effects with a higher share of treated subjects are given greater weight. Third, the confidence intervals of the pooled callback ratios were adjusted using the Knapp–Hartung method. This adjustment assumes a *t*-distribution rather than a normal distribution (under Wald-type tests) of the pooled effect, typically leading to wider (more conservative) confidence intervals (Harrer et al., 2021; Langan et al., 2019). Fourth, we integrated three-level models in our calculations that explicitly account for the multilevel nature of the data, where treatment effects are clustered within studies and sub-studies (Van den Noortgate et al., 2013). This approach counters the violation of between-study independence of the individual treatment effects (Harrer et al., 2021).

2.4.3. Meta-regression

Building on the pooled meta-analytic estimates, we ran weighted least squares mixed-effects meta-regression analyses to control for and help explain within- and between-study heterogeneity. The general meta-regression specification is given by Equation 3, where CB_k is the unemployed–employed callback ratio for effect *k*; CB_{pooled} is the pooled unemployed–employed callback ratio operationalised as the intercept; UD_k is the duration of unemployment in months for effect *k*; X_k is a vector of control variables for effect *k*; β and B are a predictor coefficient and a vector of predictor coefficients, respectively; and ε_k and ζ_k are the sampling and distributional errors.

$$CB_k = CB_{pooled} + \beta UD_k + X_k B + \varepsilon_k + \zeta_k \tag{3}$$

The set of control variables (X_k) comprises the country–year unemployment rate; a dummy for the response type 'positive reaction' with 'interview invitation' as the omitted reference response type; region dummies for Eastern Asia, Eastern Europe, Northern Europe, and Western Europe with Northern America as the omitted reference region; the year in which the experiment was conducted (if the experiment ran over several years, this variable represents the average of those years); dummies for the female gender and mixed genders with male gender as the omitted reference gender category; and age dummies representing young (21–30 years) applicants, old (51+ years) applicants, and applicants of unknown age with prime-aged (31–50 years) applicants as the omitted reference age group. We selected control variables based on those deemed relevant in almost every study included. We also identified several additional variables, including the fictitious applicant's marital status and job skill level, as well as unemployment anti-discrimination law in the study region. However, these variables appeared in few studies and were thus unsuitable for heterogeneity analysis. Consequently, they were omitted as control variables.

Labour market tightness can moderate employers' perceptions of unemployed applicants. To assess this moderation effect, we ran an extension of the general meta-regression analysis, interacting the unemployment duration with the prevailing unemployment rate on the link between the strength of the unemployment spell (in terms of duration) and callback differences between the unemployed and the employed. This specification is shown in Equation 4. In addition to the terms used in Equation 3, UR_{tc} represents the unemployment rate for year *t* and country *c*.

$$CB_k = CB_{pooled} + \beta_1 UD_k + \beta_2 UR_k + \beta_3 UD_k UR_{tc} + X_k B + \varepsilon_k + \zeta_k \tag{4}$$

Similar to the pooling of the callback ratios, we estimated the meta-regression models using the REML estimator (Viechtbauer, 2005). Moreover, standard errors were clustered at the (sub-)study level and corrected using the bias-reduced linearisation small-sample adjustment proposed by Bell and McCaffrey (2002). This cluster-robust variance estimation produces less biased standard errors when the number of clusters is small compared to alternative adjustment methods (Pustejovsky and Tipton, 2018).

2.4.4. Publication bias and robustness

Publication bias can impact the results of a meta-analysis (Friese & Frankenbach, 2020), most commonly taking the form of outcome reporting bias and selection bias (Harrer et al., 2021). It may be that only favourable results in light of the identified hypotheses are published, even if the analyses produce imprecise estimates. Studies with small sample sizes, for which it is easier to identify large effects haphazardly, are particularly prone to this overreporting. Selection bias can result from questionable research practices related to *p*-hacking (i.e. fiddling with the analyses until the required *p*-value threshold is reached) or HARKing (i.e. formulating hypotheses after the results are known). These practices can lead to false positives and, thus, an under- or over-estimation of the effect of interest (Stefan & Schönbrodt, 2023).

We followed academic best practices to measure and control for the by-effects of publication bias (Harrer et al., 2021; Irsova et al., 2023). In addition, we tested the robustness of our results with different estimation methods. More specifically, we (i) computed bias statistics and assessed funnel plot asymmetry visually, (ii) identified outliers and calculated outlier-adjusted (meta-regression) estimates, (iii) estimated small-study-corrected effects through PET-PEESE and limit meta-analyses, (iv) drew *p*-curves and examined the results of three-parameter selection models to assess sensitivity to study or effect selection based on *p*-values, and (v) conducted hierarchical Bayesian meta-analysis as an alternative approach to the frequentist's method used to calculate the original meta-analytic estimates. We refer the reader to Appendix B for more details about these analyses and their results. Findings from these analyses are discussed alongside the meta-analytic results presented in Section 3.2 where appropriate.

3. Results

This section presents and discusses the findings of our systematic literature review and meta-analysis. We first offer a qualitative overview of the reviewed articles on unemployment and inactivity scars concerning hiring chances. Next, we present the results of the meta-analysis, focusing on the relationship between unemployment duration and hiring chances for the unemployed compared to the employed. We discuss these results in light of the findings from the included studies. Finally, we examine the evidence for the mechanisms underlying our observations.

3.1. Main treatment effect

A schematic overview of the studies on unemployment, inactivity, and hiring chances can be found in Table 3. The articles reviewed are ordered alphabetically by author's name in Column (1). Column (2) specifies the geographical region in which each study was conducted. Columns (3) and (4) list the independent and dependent variables, respectively, linking unemployment, former unemployment or inactivity to employers' hiring decisions. The data collection method is specified in Column (5). Column (6) summarises the primary findings regarding the relationship between the variables in Columns (3) and (4), clarifying whether there is a positive, negative, or neutral effect. Finally, Column (7) describes the evidence for the mechanisms underlying the observed unequal treatment, delineating whether the authors merely connected their results to a theoretical framework and, on this basis, proposed a mechanism or whether they empirically scrutinised the mechanism.

< Table 3 about here >

A glance at the literature overview in Table 3 reveals that the majority of studies on unemployment and inactivity scarring originate from the USA and Europe and that the effect of unemployment on hiring probability is most frequently studied. Notably, most studies employ experimental methodologies, yet their findings diverge. All 28 studies focus at least partially on current unemployment, seven studies (25%) also investigate former unemployment, and four studies (14%) also address inactivity – with a specific emphasis on homemakers. Our systematic search revealed no results regarding the discouraged unemployed, despite their significant prevalence within the inactive population (Gammarano, 2019). Thus, the literature on inactivity scarring is considerably less extensive than that on unemployment scarring.

Overall, the literature covered indicates that unemployment substantially reduces an individual's employment opportunities. Most studies (86%) detect at least some unfavourable treatment of unemployed applicants, although around half suggest that it may be inconsequential, contingent upon one or more moderators (see Section 3.2 for a discussion of the most critical moderators). Only a few indicate that unemployment may enhance employment chances, yet no study reports an overall positive association, since this effect is consistently contingent upon the duration of unemployment.

The comparisons made in all but three studies involve contrasting the unemployed or inactive with their employed counterparts. Baert and Verhaest (2019) differ by comparing unemployed individuals to recent graduates and overeducated applicants. They find that the unemployed are at a hiring disadvantage compared with these groups. Cahuc et al. (2021) examine young unemployed applicants, factoring in their previous work experience. They conclude that work experience does not influence employment opportunities. Finally, Van Belle et al. (2018) test the effect of varying unemployment durations (one to 36 months) without an employed control group. They reveal a clear correlation between the duration of unemployment and the likelihood of obtaining employment, with a longer duration resulting in diminishing employment opportunities.

Seven studies investigating the effect of former unemployment (i.e. a period of unemployment followed by a job) on hiring chances yield mixed evidence for its effect on employers' hiring decisions. Laboratory studies all demonstrate a negative impact of previous unemployment, whereas field experiments report null findings. The participants in laboratory experiments were presumably aware that previous unemployment was a focus of the study, while those in field experiments were unaware of their participation, leading them to place less importance on former unemployment.

As for those studies examining inactivity, particularly among homemakers, all four report a negative effect on employment opportunities. Weisshaar (2018, 2021) suggests that inactivity is penalised more severely than unemployment, while Kristal et al. (2023) and Tomlin (2022) report no additional disadvantages for the inactive.

3.2. Moderators examined

Numerous factors may affect the scarring effect of non-employment. Therefore, we conducted a meta-analysis to investigate the moderators considered by the literature to be most critical – unemployment duration and unemployment rate – juxtaposing our findings with those reported in the studies in this review. Next, we examine other moderators that lay outside the scope of our meta-analysis but are frequently highlighted in the literature.

3.2.1. Unemployment duration

Although most studies agree that unemployment reduces job opportunities, the literature lacks consensus on the specific length of unemployment at which employers start to penalise the unemployed.[6] Shi et al. (2018, p. 8) acknowledge that "there is no agreement in the literature on the point of time when the scarring effect of unemployment becomes significant". Nunley et al. (2017) attribute these duration dependence discrepancies in the unemployment audit literature to variations in experimental design, population of interest,

[6] Filomena (2023) recently conducted a meta-analysis of 65 studies to quantify the scarring effects of unemployment, integrating both experimental and non-experimental studies. However, this combination complicates the interpretation of the meta-analytical results. Moreover, Filomena (2023) calculated measures of correlation, which only indicate the strength of the association between unemployment scarring and hiring chances, not the magnitude of the effect of unemployment on hiring. Additionally, the author categorised unemployment durations only into short-term and medium-long-term and, presumably due to this strict categorisation, found no evidence for duration dependence.

sample periods, and the institutional structure of labour markets. In what follows, we examine the effect of unemployment duration through a meta-regression analysis of recent correspondence experiments included in our review that consider differences in positive callbacks between the unemployed and the employed. Our analyses control for study-level covariates related to design, period, region, and labour market conditions (see Section 2.4.4 for estimation details).

In line with the general observations in Section 3.1, we find meta-analytic evidence of a negative effect of unemployment on hiring chances (see Figure 2). The average unadjusted effect of unemployment on positive callbacks following resume screening amounts to −7.31% ($CI_{95\%}$ = [−14.40%, 0.36%]). Controlling for unemployment duration, country–year unemployment rate, response type, region, year, gender, and age, the netted average effect of unemployment on hiring chances is −7.89% ($CI_{95\%}$ = [−14.67%, −0.56%]). This result is robust to removing outliers ($CI_{95\%}$ = [−11.42%, −0.57%]). Thus, signalling unemployment scars the unemployed candidate in terms of lower hiring probability (vis-à-vis the employed), but the average effect appears small.

< Figure 2 about here >

However, the overall result hides heterogeneity based on the duration of unemployment. When assessing the differential impact of unemployment scarring on hiring chances by unemployment duration, we observe weak evidence for a positive effect of unemployment on hiring chances in the short term and a more pronounced (and increasing) negative effect of unemployment on hiring probability in the medium to long term (see Figure 2). Unemployment of one to six months appears to have a marginally positive impact on positive callbacks for the unemployed by 8.23% ($CI_{95\%}$ = [−5.31%, 23.69%]); adjusting for study-level covariates and outliers, positive callbacks are, on average, 16.72% higher for the short-term unemployed than the employed ($CI_{95\%}$ = [8.91%, 25.10%]).

In contrast, after 12 months of unemployment, positive callbacks for the medium- to long-term unemployed (i.e. those unemployed for 13–18 months) drop substantially to −21.38%, on average and covariate-adjusted ($CI_{95\%}$ = [−34.36%, −5.84%]). Similarly, the long-term unemployed (19–36 months) receive 27.02% fewer positive callbacks than the employed, on average and covariate-adjusted ($CI_{95\%}$ = [−41.04%, −9.66%]). Interestingly, this figure is comparable to the global meta-analytic average of 29% fewer positive callbacks for racial and ethnic minorities in hiring discrimination research (Lippens et al., 2023). Based on our weighted least-squares meta-regression estimates, unemployed–employed callback ratios decrease by 1.70% on average per additional month of unemployment ($CI_{95\%}$ = [−2.43%, −0.96%]; see Figure 3 and Table A4, Model 1, in Appendix A), or by 2.35% on average when controlling for covariates ($CI_{95\%}$ = [−3.56%, −1.15%]; see Table A4, Model 2 in Appendix A). Our three-parameter selection models (accounting for selection bias) and hierarchical Bayesian meta-analytic estimates corroborate these findings (see Section B.I and Tables B9 and B10 in Appendix B).

< Figure 3 about here >

The above findings align closely with those of Kroft et al. (2013), who provide clear evidence of a hiring advantage for the short-term unemployed and a hiring penalty for the long-term unemployed in the United States. Our meta-analysis, however, aggregates many treatment effects using cross-study data and controls for covariates influencing duration effects. This approach allows us to estimate the impact of different unemployment durations transnationally. We observe an unemployment benefit for spells of up to approximately six months, while Duguet et al. (2018) indicate a maximum of five months, Kroft et al. (2013) approximately three months, and Oberholzer-Gee (2008) up to twelve months. Additionally, we observe that unemployment scarring becomes significant after approximately twelve months. The existing literature has been unsuccessful in pinpointing the onset of this effect: Eriksson and Rooth (2014) identify a significant negative impact at nine months and Kroft et al. (2013) report a shift from advantage to disadvantage for the unemployed after three months, although this decline does not intensify with prolonged unemployment. We find that job opportunities decrease as unemployment persists, in line with Oberholzer-Gee (2008). Regarding the studies that lie outside the scope of our meta-analysis because they do not involve a correspondence experiment, our results resonate well with the observations of Van Belle et al.'s (2018) vignette experiment, which also report a duration effect.

3.2.2. Unemployment rate

As indicated in the introduction, many studies suggest that the tightness of the labour market – and the corresponding unemployment rate – can influence employers' hiring decisions towards the unemployed (Gibbons & Katz, 1991; van den Berg & van Ours, 1996). Therefore, we reran the meta-regression analysis using data on the regional labour market conditions in terms of unemployment rate at the time that each study was conducted. More specifically, we considered the interaction between how long the applicants had been unemployed and the prevailing unemployment rates, whereas we had previously only controlled for these rates in our analyses.

Making the distinction between (moderately) tight and loose labour markets, we observe that the effect of duration dependence is more pronounced when unemployment rates are low. On average, for each one-unit increase in both unemployment duration (in months) and unemployment rate (per cent), the negative duration dependence effect is offset by 0.59% in terms of callback ratio ($CI_{95\%}$ = [0.14%, 1.45%]; see Table A4, Model 3 in Appendix A). Nevertheless, when we remove outliers from the analysis, the mitigating effect of labour market tightness disappears ($CI_{95\%}$ = [−0.19%, 0.51%]; see Table B4, Model 3 in Appendix B). Figure 4 depicts this relationship based on a categorical distinction. In moderately loose labour markets (with unemployment rates between 6.0% and 10.4%, i.e. the maximum unemployment rate in the studies included),[7] the effect of duration dependence visually weakens with increasing unemployment length more than in relatively tight labour markets (with unemployment rates ranging from 3.1%, i.e. the minimum unemployment rate in the studies included to just under 6.0%). These results appear robust to varying the threshold value between 5.0% and 7.0% to distinguish between tight and loose labour markets (see Figures A5 and A6 in Appendix A).

< Figure 4 about here >

Our findings are echoed by Shi and Wang (2022), who observe that unemployment penalises unemployed job seekers in Switzerland, where unemployment is low, but not in Greece, where unemployment is high. Conversely, Nunley et al. (2017) and Farber et al. (2016) detect no hiring difference between regions with tight and loose labour markets, probably because they do not include unemployment spells in excess of twelve months, which is the threshold at which we find that unemployment scarring noticeably begins.

---

[7] We collected the unemployment rates from annual ILOSTAT data (ILO, 2024), noting the country's unemployment rate for the specific year in which the study was conducted. For studies spanning multiple years, we calculated the average unemployment rate over those years.

3.2.3. Other moderators

Although not considered in our meta-regression analysis, due to insufficient compatible data to test for heterogeneity, the scarring literature identifies additional moderators – beyond unemployment duration and rate – that influence the hiring chances of the unemployed and the inactive. In this subsection, we qualitatively explore the most studied moderators, i.e. (i) job skill level, (ii) race, (iii) gender, (iv) age.

First, almost half of the studies cover occupations with varying skill requirements, while some only include low- or medium-skilled jobs, and a few only consider high-skilled jobs.[8] Eriksson and Rooth (2014) find that long-term unemployment only affects employment chances for low- and medium-skilled jobs. Employers may view the duration of unemployment as less meaningful for high-skilled jobs, since high-productivity workers often seek premium positions or demand higher wages, which can prolong their unemployment (Pissarides, 2000). Furthermore, the higher education level of candidates for high-skilled positions reduces employers' reliance on indirect productivity signals (e.g. long-term unemployment) since education is a better productivity indicator (Baert et al., 2015; Taubman & Wales, 1974). However, Bonoli (2014) reports that long-term unemployment causes less concern in sectors with lower qualification requirements, possibly because lower-skilled jobs reduce the importance of the lower productivity and qualification signals sent by long-term unemployment (Mosthaf, 2014), as discussed in Section 3.3.

Second, about half of the studies include candidates from diverse ethnic backgrounds. Among these, while some studies (e.g. Weisshaar, 2021) report that race does not affect job opportunities for the unemployed and inactive, others do note an impact of race, but with varied findings. For example, Birkelund et al. (2017) and Pedulla (2018) posit that racial biases crowd out the typical negative implications of unemployment. This interpretation aligns with the findings of Eriksson and Rooth (2014), who observe a more severe unemployment penalty for ethnic majority applicants, possibly because unemployment is considered a less critical signal for ethnic minority applicants, who are more likely to experience unemployment. Conversely, Pierné (2018) concludes that ethnic minority

[8] We determined the skill level of the jobs included in the studies based on the O*NET job characteristics database.

applicants experience an amplified negative impact from unemployment. However, this study focused solely on the low-skilled construction sector, whereas the other studies included a mix of sectors. Pierné (2018) suggests that this outcome may be attributed to the lower negative effects of ethnic origin, as minority applicants are more prevalent in this sector and, consequently, face less discrimination. Therefore, race does not mute the impact of unemployment, and the unemployment signal regains significance.

Third, we discuss two factors that frequently appear in the studies reviewed but which appear to lack any moderating effect: gender (in 25 studies, i.e. 89%) and age (in 22 studies, i.e. 79%). Regarding gender, earlier literature reviews (e.g. Filomena, 2023) indicate that overall unemployment scarring effects are more pronounced for men due to lower unemployment rates among them: becoming unemployed thus deviates more from the norm. However, our review does not corroborate this observation (e.g. Weisshaar, 2021, finds no gender-based differences). This discrepancy may be explained by the relatively similar unemployment rates by gender across the countries examined in this review (ILO, 2021). Regarding age, studies that include different age categories report no moderation effect of age. For instance, experiments conducted by Farber et al. (2016) and Nunley et al. (2017) focus on different age groups among fictitious applicants, but both report null interaction effects with unemployment duration.

Last, in addition to the moderators discussed, we identify, albeit less highlighted in the literature, three other categories of moderators that influence unemployment scarring in hiring: (i) application stages, (ii) applicant characteristics, and (iii) applicant's job history. Regarding (i), Manning (2000) notes that employers rely primarily on negative signals when evaluating resumes and shortlisting for interviews. Once candidates progress through this initial selection stage, employers prioritise more objective criteria, such as unemployment. On applicant characteristics (ii), Maurer-Fazio and Wang (2018) find that unemployment carries a negative signal more for married women than single women as they are more prone to job challenges related to childcare, flexibility, and productivity. Additionally, Shi et al. (2018) observe that unemployed applicants without occupation-specific training do not encounter unemployment penalties because they inherently signal lower competence. When considering the applicant's job history (iii), Eriksson and Rooth (2014) argue that having had a job following a period of unemployment may negate the signal of reduced productivity. Similarly, Norlander et al. (2020) assert that the stigma of unemployment dissipates if an applicant becomes unemployed due to circumstances beyond their control, (e.g. following collective dismissal).

3.3. Mechanisms uncovered

The above results reveal that employers generally see the long-term unemployed and inactive as inferior workers. In this subsection, we review the evidence for the mechanisms underlying this outcome.

As can be seen from Column (7) of Table 3, the lion's share of studies provide (assumed) support for signalling theory, while only a limited number endorse human capital theory (e.g. Shi et al., 2018), and a few find no evidence in favour of any specific theory (e.g. Shi & Di Stasio, 2022). Despite the prevailing support for signalling theory, the specific signal conveyed by a period of unemployment varies. Some studies conclude that unemployment, in particular, signals a lower level of motivation to employers (e.g. Van Belle et al., 2018), while others claim that it reflects reduced competence (e.g. Weisshaar, 2018). Other signals referenced include lack of professionalism, qualifications, dedication, trainability, warmth, and soft skills. Together, these signals foster a pervasive perception of reduced expected productivity, consistent with the stigma effect model developed by Vishwanath (1989). Similarly, Oberholzer-Gee (2008) and Van Belle et al. (2018) offer support for the rational herding theory, illustrating employers' reluctance to consider long-term unemployed candidates on the assumption that their reduced productivity has likely led other employers to pass over them.

Furthermore, in line with the findings in Section 3.2.1, the length of a period of unemployment impacts employers' perceptions of an unemployed candidate. For instance, Duguet et al. (2018) assert that the negative unemployment signal applies only to the long-term unemployed, with the short-term unemployed benefiting from their immediate availability. Kroft et al. (2013) suspect that short-term unemployed vis-à-vis employed applicants are perceived as more loyal and less prone to job hopping. They further suggest that employers may perceive unemployed applicants as having less bargaining power than the employed, who can rely on their current job. Consequently, employers may offer lower starting salaries to unemployed applicants, with the cost savings outweighing the perceived lower quality of the short-term unemployed. These findings align with those of Lockwood (1991), who posits that perceived productivity diminishes as the duration of unemployment increases.

The two studies by Weisshaar (2018, 2021) discern that inactivity signals even less commitment than unemployment, appearing a more severe violation of ideal worker norms. Weisshaar (2021) adds that providing additional information on past professional performance and social skills can mitigate the reduced employment opportunities for unemployed applicants. However, inactive caregiver applicants consistently face disadvantages, regardless of the counter-stereotypical information provided. This greater penalty for the inactive appears, therefore, to stem from more ingrained cognitive perceptions: inactivity for family reasons signals a violation of ideal worker norms (Weisshaar, 2021).

Last, the selection of studies discussed in Section 3.2.2 that find an interaction with labour market tightness may be linked to a lower level of stigma in tighter markets. In other words, the signalling value of unemployment increases when the labour market flourishes and decreases when it slackens. For instance, Birkelund et al. (2017) argue that employers are more sceptical of applicants who have a prolonged unemployment history in tight labour markets. Similarly, Shi and Wang (2022) state that a high unemployment rate leads employers to explain unemployment as due to external factors rather than individual failure. Kroft et al. (2013) note that slack labour markets correlate with a less negative perception of unemployment, given the higher prevalence of unemployed jobseekers. This finding is echoed in studies focusing on specific applicant characteristics and particular sectors. For instance, Shi and Di Stasio (2022) report that unemployed applicants with no vocational education or training are penalised less since they are more likely to face unemployment. Similarly, Birkelund et al. (2017) note that there is no differential treatment of unemployed applicants in the communication industry, where unemployment is more common, given the prevalence of freelance and precarious workers.

4. Conclusion

We systematically reviewed the empirical literature on whether, when, and why employers penalise periods of unemployment and inactivity. Through a state-of-the-art meta-analysis of transnational data from a homogeneous and substantial set of correspondence experiments with fictitious job applications, we also quantified the evolution in the unemployment penalty over the spell duration, resolving disagreement on this evolution in the existing literature.

Our review shows that employers use information on unemployment as a negative filter in the resume screening process and that the duration of unemployment is inversely related to hiring chances. More concretely, our meta-analytic estimates indicate that the penalty visibly appears after a one-year unemployment spell. In contrast, periods of approximately six months or less lead to a marginally higher hiring probability, presumably due to employers valuing the immediate availability of the unemployed candidate and placing less emphasis on the negative stigma and lower perceived worker productivity commonly associated with unemployment. Our literature synthesis indicates that the long-term unemployed are primarily seen as less motivated, less competent, and, therefore, less productive vis-à-vis their employed counterparts.

Inactive individuals encounter comparable challenges, but the penalty in terms of hiring probability appears more severe. For them, the stigma of lower commitment seems to come to the fore as an underlying mechanism. However, our research identified only four studies that scrutinised the hiring prospects of inactive applicants, each attributing the period of inactivity to childcare alone. Hence, inactivity scarring literature is scarce, particularly for non-homemakers, and requires substantial expansion since activating the inactive is a crucial factor in boosting employment rates.

Following our in-depth synthesis of the literature, we propose three policy recommendations to alleviate the effects of unemployment scarring. First, our findings demonstrate that the recently unemployed should be guided towards jobs as soon as possible, as they are at a hiring advantage during the first six months of unemployment that wanes significantly thereafter. Job search assistance programmes stand out as the most effective labour market policy initiative in this regard, as evidenced by meta-analyses by

Card et al. (2010) and Kluve (2010).

Second, focusing a resume on the number of years of experience rather than specific dates can direct employers’ attention away from the negative signal of unemployment, as proposed by Kristal et al. (2023) and Okoroji et al. (2023). However, this intervention has only been explored in a limited number of studies to date and requires further investigation.

Third, applicants can include positive counter-information to temper or mute negative signalling in their resumes. Highlighting a genuine intent to secure stable employment and maintaining the requisite skills to prevent a deterioration in skills are key strategies, as Van Belle et al. (2018) and Weisshaar (2021) have suggested. Countering stereotypes has been shown to be effective for ethnic minority applicants (King & Ahmad, 2010; Sachs et al., 2024) and ethnic minority candidate tenants (Ewens et al., 2014). Neumark (2018), however, questions whether additional information consistently reduces bias, as studies do not always accurately reflect real-world behaviour. Here, too, further research is necessary to confirm the effectiveness of such statements in mitigating stigma.

References

Acemoglu, D. (1995). Public policy in a model of long-term unemployment. *Economica*, *62*(246), 161–178. https://doi.org/10.2307/2554901

Anderson, L. R., & Holt, C. A. (1997). Information cascades in the laboratory. *American Economic Review*, *87*(5), 847–862. http://www.jstor.org/stable/2951328

Arrow, K. J. (1973). The theory of discrimination. In Ashenfelter, O. and Rees, A. (Eds.), *Discrimination in labor markets* (pp. 1–33). Princeton University Press. https://doi.org/10.1515/9781400867066-003

Arulampalam, W., Booth, A. L., & Taylor, M. P. (2000). Unemployment persistence. *Oxford Economic Papers*, *52*(1), 24–50. https://doi.org/10.1093/oep/52.1.24

Atkinson, J., Lesley, G., & Meager, N. (1996). *Employers, recruitment and the unemployed*. The Institute for Employment Studies. https://www.employment-studies.co.uk/system/files/resources/files/325.pdf

Aysun, U., Bouvet, F., & Hofler, R. (2014). An alternative measure of structural unemployment. *Economic Modelling*, *38*, 592–603. https://doi.org/10.1016/j.econmod.2014.02.009

Baert, S. (2021). The iceberg decomposition: A parsimonious way to map the health of labour markets. *Economic Analysis and Policy*, *69*, 350–365. https://doi.org/10.1016/j.eap.2020.12.012

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is there less discrimination in occupations where recruitment is difficult? *ILR Review*, *68*(3), 467–500. https://doi.org/10.1177/0019793915570873

Baert, S., & Verhaest, D. (2019). Unemployment or overeducation: Which is a worse signal to employers? *De Economist*, *167*(1), 1–21. https://doi.org/10.1007/s10645-018-9330-2

Banerjee, A. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, *107*(3), 797–817. https://doi.org/10.2307/2118364

Bateson, R. (2023). Perceptions of pandemic resume gaps: Survey experimental evidence from the United States. *PLOS ONE*, *18*(3). https://doi.org/10.1371/journal.pone.0281449

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. University of Chicago Press.

Bell, R. M., & McCaffrey, D. F. (2002). Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology*, *28*(2), 169–181.

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

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. (2014). Employers' attitudes towards long-term unemployed people and the role of activation in Switzerland. *International Journal of Social Welfare*, *23*(4), 421–430. https://doi.org/10.1111/ijsw.12086

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

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). Publication bias. In M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein (Eds.) *Introduction to meta-analysis* (pp. 277–292). John Wiley & Sons. https://doi.org/10.1002/9780470743386.ch30

Borland, J. (2020). Scarring effects: A review of Australian and international literature. *Australian Journal of Labour Economics*, *23*(2), 173–188. https://search.informit.org/doi/10.3316/informit.469140968965412

Brandolini, A., Viviano, E. (2018). Measuring employment and unemployment. *IZA World of Labor*, *445*. https://doi.org/10.15185/izawol.445

Cahuc, P., Carcillo, S., & Minea, A. (2021). The difficult school-to-work transition of high school dropouts: Evidence from a field experiment. *Journal of Human Resources*, *56*(1), 159–183. https://doi.org/10.3368/jhr.56.1.0617-8894R2

Card, D., Kluve, J., & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, *120*(548), 452–477. https://doi.org/10.1111/j.1468-0297.2010.02387.x

Cockx, B., & Picchio, M. (2013). Scarring effects of remaining unemployed for long-term unemployed school-leavers. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *176*(4), 951–980. https://doi.org/10.1111/j.1467-985X.2012.01086.x

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signalling theory: A review and assessment. *Journal of Management*, *37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Dorsett, R., & Lucchino, P. (2018). Young people's labour market transitions: The role of early experiences. *Labour Economics*, *54*, 29–46. https://doi.org/10.1016/j.labeco.2018.06.002

Duguet, E., Le Gall, R., L'Horty, Y., & Petit, P. (2018). How does labour market history influence the access to hiring interviews? *International Journal of Manpower*, *39*(4), 519–533. https://doi.org/10.1108/IJM-09-2017-0231

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Eriksson, S., & Lagerström, J. (2006). Competition between employed and unemployed job applicants: Swedish evidence. *Scandinavian Journal of Economics*, *108*(3), 373–396. https://doi.org/10.1111/j.1467-9442.2006.00462.x

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

Ewens, M., Tomlin, B., & Wang, L. C. (2014). Statistical discrimination or prejudice? A large sample field experiment. *Review of Economics and Statistics*, *96*(1), 119–134. https://doi.org/10.1162/REST_a_00365

Farber, H. S., Herbst, C. M., Silverman, D., & Von Wachter, T. (2019). Whom do employers want? The role of recent employment and unemployment status and age. *Journal of Labor Economics*, *37*(2), 323–349. https://doi.org/10.1086/700184

Farber, H. S., Silverman, D., & Von Wachter, T. (2016). Determinants of callbacks to job applications: An audit study. *American Economic Review*, *106*(5), 314–318. https://doi.org/10.1257/aer.p20161010

Filomena, M. (2023). Unemployment scarring effects: An overview and meta-analysis of empirical studies. *Italian Economic Journal*, 1–60. https://doi.org/10.1007/s40797-023-00228-4

Friese, M., & Frankenbach, J. (2020). p-Hacking and publication bias interact to distort meta-analytic effect size estimates. *Psychological Methods*, *25*(4), 456–471. https://psycnet.apa.org/doi/10.1037/met0000246

Gaddis, S. M. (2018). An introduction to audit studies in the social sciences. In Gaddis, S. M. (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 3–44). Springer. https://doi.org/10.1007/978-3-319-71153-9_1

Gammarano, R. (2019, August 7). *Persons outside the labour force: How inactive are they really?* ILOSTAT. https://ilostat.ilo.org/persons-outside-the-labour-force-how-inactive-are-they-really/

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

Gregg, P. (2001). The impact of youth unemployment on adult unemployment in the NCDS. *The Economic Journal*, *111*(475), 626–653. https://doi.org/10.1111/1468-0297.00666

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003107347

ILO. (2016). *Key indicators of the labour market* (9th ed.). International Labour Organization. https://webapps.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/publication/wcms_498929.pdf

ILO. (2021). *Gender gap in unemployment rates.* International Labour Organization. https://ilo.org/infostories/en-GB/Stories/Employment/barriers-women#unemployed-vulnerable/gap-unemployment

ILO. (2024). *Unemployment rate by sex and age (annual).* International Labour Organization. https://rshiny.ilo.org/dataexplorer25/?lang=en&id=UNE_DEAP_SEX_AGE_RT_A

Irsova, Z., Doucouliagos, H., Havránek, T., & Stanley, T. D. (2023). Meta-analysis of social science research: A practitioner's guide. *Journal of Economic Surveys*, *00*, 1–20. https://doi.org/10.1111/joes.12595

Jarosch, G., & Pilossoph, L. (2019). Statistical discrimination and duration dependence in the job finding rate. *The Review of Economic Studies*, *86*(4), 1631–1665. https://doi.org/10.1093/restud/rdy055

King, E. B., & Ahmad, A. S. (2010). An experimental field study of interpersonal discrimination toward Muslim job applicants. *Personnel Psychology*, *63*(4), 881–906. https://doi.org/10.1111/j.1744-6570.2010.01199.x

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, *17*(6), 904–918. https://doi.org/10.1016/j.labeco.2010.02.004

Kristal, A. S., Nicks, L., Gloor, J. L., & Hauser, O. P. (2023). Reducing discrimination against job seekers with and without employment gaps. *Nature Human Behaviour*, *7*(2), 211–218. https://doi.org/10.1038/s41562-022-01485-6

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

Langan, D., Higgins, J. P., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., Viechtbauer, W., & Simmonds, M. (2019). A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods*, *10*(1), 83–98. https://doi.org/10.1002/jrsm.1316

Lefebvre, C., Glanville, J., Briscoe, S., Featherstone, R., Littlewood, A., Metzendorf, M-I., Noel-Storr, A., Paynter, R., Rader, T., Thomas, J., & Wieland, L. S. (2023). Chapter 4: Searching for and selecting studies. In Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., & Welch, V.A. (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.4). https://www.training.cochrane.org/handbook

Lippens, L., Vermeiren, S., & Baert, S. (2023). The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments. *European Economic Review*, *151*, 104315. https://doi.org/10.1016/j.euroecorev.2022.104315

Lockwood, B. (1991). Information externalities in the labour market and the duration of unemployment. *The Review of Economic Studies*, *58*(4), 733–753. https://doi.org/10.1111/1468-0297.00666

Manning, A. (2000). Pretty vacant: Recruitment in low-wage labour markets. *Oxford Bulletin of Economics and Statistics*, *62*, 747–770. https://doi.org/10.1111/1468-0084.0620s1747

Maurer-Fazio, M., & Wang, S. (2018). Does marital status affect how firms interpret job applicants' un/employment histories? *International Journal of Manpower*, *39*(4), 567–580. https://doi.org/10.1108/IJM-09-2017-0251

Mosthaf, A. (2014). Do scarring effects of low-wage employment and non-employment differ between levels of qualification? *Scottish Journal of Political Economy*, *61*, 154–177. https://doi.org/10.1080/14616696.2012.677050

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, *56*(3), 799–866. https://doi.org/10.1257/jel.20161309

Norlander, P., Ho, G. C., Shih, M., Walters, D. J., & Pittinsky, T. L. (2020). The role of psychological stigmatization in unemployment discrimination. *Basic and Applied Social Psychology*, *42*(1), 29–49. https://doi.org/10.1080/01973533.2019.1689363

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

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

Okoroji, C., Gleibs, I. H., & Howard, S. (2023). Inferring incompetence from employment status: An audit-like experiment. *PLOS ONE*, *18*(3). https://doi.org/10.1371/journal.pone.0280596

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

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T., Mulrow, C. D., Shamseer, L., Tetzlaff, J., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, 71. https://doi.org/10.1136/bmj.n71

Pedulla, D. S. (2018). How race and unemployment shape labor market opportunities: Additive, amplified, or muted effects? *Social Forces*, *96*(4), 1477–1506. https://doi.org/10.1093/sf/soy002

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *JAMA*, *295*(6), 676. https://doi.org/10.1001/jama.295.6.676

Pierné, G. (2018). Hiring discrimination, ethnic origin and employment status. *International Journal of Manpower*, *39*(1), 152–165. https://doi.org/10.1108/IJM-05-2016-0104

Pissarides, C. A. (1992). Loss of skill during unemployment and the persistence of employment shocks. *The Quarterly Journal of Economics*, *107*(4), 1371–1391. https://doi.org/10.2307/2118392

Pissarides, C. A. (2000). *Equilibrium unemployment theory* (2nd ed.). MIT Press.

Pustejovsky, J. E. & Tipton, E. (2018). Small sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business and Economic Statistics*, *36*(4), 672–683. https://doi.org/10.1080/07350015.2016.1247004

Rücker, G., Schwarzer, G., Carpenter, J. R., Binder, H., & Schumacher, M. (2011). Treatment-effect estimates adjusted for small-study effects via a limit meta-analysis. *Biostatistics*, *12*(1), 122–142. https://doi.org/10.1093/biostatistics/kxq046

Sachs, T., Homan, A. C., & Lancee, B. (2024). Impression formation of majority and minority applicants during resume screening—Does processing more information reduce prejudice? *Journal of Applied Social Psychology*, *54*, 387–404. https://doi.org/10.1111/jasp.13047

Shi, L. P., & Di Stasio, A. 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

Shi, L. P., & Wang, S. (2022). Demand-side consequences of unemployment and horizontal skill mismatches across national contexts: An employer-based factorial survey experiment. *Social Science Research*, 104. https://doi.org/10.1016/j.ssresearch.2021.102668

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve and effect size. *Perspectives on Psychological Science*, *9*(6), 666–681. https://doi.org/10.1177/1745691614553988

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better p-curves: Making p-curve analysis more robust to errors, fraud, and ambitious p-hacking, a reply to Ulrich and Miller (2015). *Journal of Experimental Psychology: General*, *144*(6), 1146–1152. https://doi.org/10.1037/xge0000104

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

Stanley, T. D. (2017). Limitations of PET-PEESE and other meta-analysis methods. *Social Psychological and Personality Science*, *8*(5), 581–591. https://doi.org/10.1177/1948550617693062

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

Stefan, A. M., & Schönbrodt, F. D. (2023). Big little lies: A compendium and simulation of p-hacking strategies. *Royal Society Open Science*, *10*(2), 220346. https://doi.org/10.1098/rsos.220346

Suomi, A., Schofield, T., & Butterworth, P. (2022). Does receipt of unemployment benefits change recruiter perceptions of candidates' personality, work relevant skills and employability? *Work*, *71*(4), 1029–1041. https://doi.org/10.3233/WOR-205048

Taubman, P. J., & Wales, T. (1974). *Higher education and earnings: College as an investment and screening device*. NBER Books. https://ideas.repec.org/b/nbr/nberbk/taub74-1.html

Tomlin, B. (2022). Maternity breaks: Unemployment spells or relevant experience? *Journal of Economic Behavior & Organization*, *198*, 673–681. https://doi.org/10.1016/j.jebo.2022.04.015

Trzebiatowski, T. M., Wanberg, C. R., & Dossinger, K. (2020). Unemployed needn't apply: Unemployment status, legislation, and interview requests. *Journal of Management*, *46*(8), 1380–1407. https://doi.org/10.1177/0149206318823952

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 den Berg, G. J., & Van Ours, J. C. (1996). Unemployment dynamics and duration dependence. *Journal of Labor Economics*, *14*(1), 100–125. https://doi.org/10.1086/209805

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*, 576–594. https://doi.org/10.3758/s13428-012-0261-6

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. https://doi.org/10.3102/10769986030003261

Vishwanath, T. (1989). Job search, stigma effect, and escape rate from unemployment. *Journal of Labor Economics*, *7*(4), 487–502. https://doi.org/10.1086/298218

Weisshaar, K. (2018). From opt out to blocked out: The challenges for labor market re-entry after family-related employment lapses. *American Sociological Review*, *83*(1), 34–60. https://doi.org/10.1177/0003122417752355

Weisshaar, K. (2021). Employment lapses and subsequent hiring disadvantages: An experimental approach examining types of discrimination and mechanisms. *Socius*, *7*. https://doi.org/10.1177/23780231211019861

Williams, D. R., Rast, P., & Bürkner, P.-C. (2018). *Bayesian meta-analysis with weakly informative prior distributions*. PsyArxiv. https://doi.org/10.31234/osf.io/7tbrm

Declarations

Ethics approval and consent to participate

Ethical approval or consent to participate is not applicable to this literature review.

Consent for publication

Consent for publication is not applicable to this literature review.

Data and code availability

Data and code are available at https://osf.io/rwna9.

Declaration of competing interest

The authors declare that they have no competing interests.

Funding

We acknowledge financial support from FWO (Research Foundation – Flanders) under grant number G073924N. In addition, Louis Lippens acknowledges funding from FWO under grant number 12AM824N.

CRediT authorship contribution statement

Liam D'hert: Conceptualisation, Methodology, Investigation, Data curation, Writing – original draft, Writing – review & editing; **Stijn Baert**: Conceptualisation, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition; **Louis Lippens**: Conceptualisation, Methodology, Formal analysis, Writing – original draft, Visualisation, Writing – review & editing.

Tables and figures

Table 1. Eligibility criteria for study inclusion

(1) Criterion		(2) Description
S	Sample	**I:** Unemployed, formerly unemployed, and inactive individuals (including homemakers and the discouraged unemployed). **E:** Other groups of (i) inactive individuals such as long-term sick and disabled individuals, students, and early retirees; and (ii) vulnerable individuals (such as older workers or migrants).
PI	Phenomenon of Interest	**I:** Unequal treatment of unemployed and inactive individuals in hiring and selection. **E:** Unequal treatment in other labour market outcomes (e.g. wages).
D	Design	**I:** Empirical studies (including (quasi-)experiments, field experiments, and survey research). **E:** Meta-analyses, (systematic) reviews, case studies, administrative data research, and theoretical papers.
E	Evaluation	**I:** Hiring chances, measured by employers' hiring intentions and perceptions. **E:** Other labour market outcomes (e.g. wages).
R	Research type	**I:** Primary, quantitative, empirical research (including mixed methods). **E:** Secondary and exclusively qualitative research.

Notes. The following abbreviations are used: 'I' denotes 'inclusion', 'E' denotes 'exclusion', and 'e.g.' signifies 'for example'. This table is based on the SPIDER framework established by Cooke et al. (2012).

Table 2. Descriptive statistics (N = 28)

Variable	N	%
Region		
North America	17	61%
Europe	9	32%
Other	2	7%
Year[a]		
1996–2000	2	7%
2001–2005	0	0%
2006–2010	2	7%
2011–2015	13	46%
2016–2021	11	39%
Treatment		
Unemployed	17	61%
Formerly unemployed or unemployed	7	25%
Unemployed or inactive	4	14%
Hiring outcome		
Job offer	3	11%
Job interview invitation	18	64%
Applicant assessment	1	4%
Job offer or job interview invitation	5	18%
Job offer or applicant assessment	1	4%
Method		
Experimental (field)	13	46%
Experimental (lab)	8	29%
Experimental (field and lab)	3	11%
Survey	2	7%
Experimental (field) and survey	2	7%
Unemployment duration		
Short-term unemployed (spell < 1 year)	2	7%
Long-term unemployed (spell ≥ 1 year)	10	36%
Short- or long-term unemployed	13	46%
N/A	3	11%
Gender		
Male	4	14%
Female	6	21%
Male or female	15	54%
N/A	3	11%
Age		
21–30	16	57%
31–50	5	18%
51+	0	0%
Various	1	4%
N/A	6	21%

Notes. Abbreviations used: N/A denotes 'not applicable'.

[a] We opted to include the period of data collection rather than the year of study publication. For studies where data were collected over multiple years, we recorded the first year of the data collection period.

Table 3. Summary of the literature

(1) Authors (year)	(2) Region (country)	(3) Independent variable(s)[a]	(4) Dependent variable(s)	(5) Data	(6) Main findings (effect of (3) on (4))	(7) Mechanisms
Baert and Verhaest (2019)	Western Europe (Belgium)	Unemployment (long-term)	Job interview invitation	Experimental (field)[b]	Unemployed applicants are less likely to receive job interview invitations than recent graduates or overeducated applicants. **(-)**	The authors suggest that unemployment may **signal** lower productivity to recruiters.
Bateson (2023)	North America (United States of America)	Unemployment and former unemployment	Job offer	Experimental (lab)[c]	Unemployed applicants or applicants with employment breaks during the COVID-19 pandemic are less likely to receive job offers than applicants who remained employed. **(-)**	The author provides evidence that unemployment acts as a negative **signal** to recruiters. They perceive applicants with a pandemic-related unemployment spell or break as lacking professionalism, qualifications, motivation, and dedication.
Birkelund et al. (2017)	Northern Europe (Norway)	Unemployment (long-term)	Job interview invitation	Experimental (field)	Unemployed applicants are less likely than employed applicants to receive job interview invitations.[s,t] **(-)**	The authors suggest that unemployment may **signal** lower worker quality and lower motivation to recruiters.
Bonoli (2014)	Central Europe (Switzerland)	Unemployment (long-term)	Employee assessment	Survey	Employers perceive the long-term unemployed as less employable than those with shorter or no periods of unemployment. The effect is stronger for larger companies. **(-)**	The author provides evidence that unemployment acts as a negative **signal** to recruiters, including indications of lack of motivation, lower productivity, and personality problems.
Cahuc et al. (2021)	Western Europe (France)	Unemployment (long-term)	Job interview invitation	Experimental (field)[d]	Applicants who have remained unemployed since leaving school are not less likely to receive job interview invitations than applicants with employment experience. **(0)**	N/A
Duguet et al. (2018)	Western Europe (France)	Unemployment (duration)	Job interview invitation	Experimental (field)[e]	Short-term unemployed applicants are more likely to receive job interview invitations than those with permanent employment or long-term unemployment. The latter have similar invitation chances to those with permanent employment.[u] **(+/0)**	The authors suggest that long-term unemployment may reveal a **human capital depreciation** to recruiters. Short-term unemployed individuals are favoured due to their immediate availability.
Eriksson and Rooth (2014)	Northern Europe (Sweden)	Unemployment (duration) and former unemployment	Job interview invitation	Experimental (field)	Unemployed applicants are less likely than employed applicants to receive job interview invitations when unemployed for over nine months.[v] There is no unequal treatment for applicants with shorter periods of	The authors suggest that long-term unemployment may **signal** lower productivity to recruiters. Work experience following an unemployment spell may eliminate this negative signal.

					unemployment. Contemporary unemployment is more damaging than past unemployment. **(-/0)**	
Farber et al. (2016)	North America (United States of America)	Unemployment (duration)	Job interview invitation	Experimental (field)[f]	Unemployed applicants are not less likely than employed applicants to receive job interview invitations. The duration of the unemployment spell has no impact. **(0)**	N/A
Farber et al. (2019)	North America (United States of America)	Unemployment (duration)	Job interview invitation	Experimental (field)[f]	Long-term unemployed applicants are less likely to receive job interview invitations than applicants with shorter spells of unemployment. Interim-employed applicants have even lower chances of receiving job interview invitations than unemployed applicants when applying to high-skilled positions. **(+/-)**	The authors suggest that long-term unemployment may **signal** lower productivity to recruiters.
Kristal et al. (2023)	Western Europe (United Kingdom)	Unemployment (long-term) and inactivity (homemaker)	Job interview invitation and job offer	Experimental (field and lab)	Unemployed and inactive applicants are less likely to receive job interview invitations and job offers than employed applicants, regardless of whether they explain their period of unemployment. However, applicants who structure their resumes based on years of experience rather than specific dates (thus concealing their unemployment) are not disadvantaged. **(-)**	The authors suggest that unemployment and inactivity may **signal** lower productivity to recruiters. They provide evidence that concealing unemployment spells leads recruiters to focus on work experience, thereby eliminating unemployment penalties.
Kroft et al. (2013)	North America (United States of America)	Unemployment (duration)	Job interview invitation	Experimental (field)[g]	Unemployed applicants are slightly more likely to receive job interview invitations than employed applicants during the initial months of unemployment (up to three months), substantially less likely in mid-length spells (up to eight months), but are no less likely to do so in more extended (more than eight months) spells. The effect is stronger in tight labour markets. **(+/-)**	The authors suggest that long-term unemployment may **signal** lower productivity to recruiters. They also indicate that short-term unemployed applicants are favoured over employed applicants because they are more loyal, less prone to job hopping, and immediately available.
Manning (2000)	Western Europe (United Kingdom)	Unemployment	Job interview invitation and job offer	Survey	Unemployed applicants are less likely to receive job interview invitations than employed applicants. However, once they reach the interview stage, they face no further hurdles in securing a job offer. **(-/0)**	The author suggests that unemployment may be a negative **signal** to recruiters for job interviews but not job offers. When making final hiring decisions, employers prioritise subjective (e.g. personality) factors over objective criteria (e.g. work history).

Maurer-Fazio and Wang (2018)	East Asia (China)	Unemployment (duration) and former unemployment	Job interview invitation	Experimental (field)[h]	Married women in short- or long-term unemployment are less likely than the employed to receive job interview invitations, but the opposite is true for single women in short- or long-term unemployment. Employment following a period of unemployment eliminates these unequal interview chances. **(+/-)**	The authors suggest that unemployment may **signal** lower productivity (for married women) to recruiters, while in other instances, they may view it as a signal of immediate availability (for single women).
Norlander et al. (2020)	North America (United States of America)	Unemployment (short-term)	Job interview invitation and job offer	Experimental (field and lab)[i]	Short-term unemployed applicants are less likely than employed applicants to receive job offers, except when the cause of their unemployment lies outside their control (e.g. the company goes out of business). Both field and laboratory experiments support this result. **(-)**	The authors provide evidence that unemployment acts as a negative **signal** to recruiters, including indications of lack of warmth and competence, except when the cause of the unemployment lies outside the unemployed applicant's control.
Nunley et al. (2017)	North America (United States of America)	Unemployment (duration) and former unemployment	Job interview invitation	Experimental (field)[j]	Unemployed applicants are not less likely than employed applicants to receive job interview invitations. The duration of the unemployment has no impact. Former unemployment also has no effect. **(0)**	N/A
Oberholzer-Gee (2008)	Central Europe (Switzerland)	Unemployment (duration)	Job interview invitation	Experimental (field)[k], survey	Unemployed applicants receive more job interview invitations than employed applicants during the initial months of unemployment (up to 12 months); there is no effect for mid-length spells (up to 24 months), and they are less likely to receive job interview invitations during more extended spells (over 24 months). **(+/0/-)**	The author provides evidence that long-term unemployment acts as a **signal** of lower productivity. Additionally, he provides proof of **rational herding** in that recruiters are reluctant to interview the long-term unemployed, believing that if they were productive, other recruiters would have already hired them.
Okoroji et al. (2023)	Western Europe (United Kingdom)	Unemployment (long-term)	Job interview invitation and job offer	Experimental (lab)[l]	Unemployed applicants are less likely than employed applicants to receive job interview invitations and job offers. **(-)**	The authors provide evidence that unemployment acts as a **signal** of lower competence to recruiters.
Pedulla (2018)	North America (United States of America)	Unemployment (long-term)	Job interview invitation	Experimental (field)	White unemployed applicants are less likely than employed applicants to receive job interview invitations, while this disparity is not observed for black unemployed (compared to employed) applicants.[w] **(-/0)**	The author suggests that long-term unemployment may **signal** lower quality to recruiters. Still, for black applicants, their race may already serve as a similar negative signal, limiting the additional adverse effects of unemployment for them.
Pierné (2018)	Western Europe (France)	Unemployment (short-term)	Job interview invitation	Experimental (field)[m]	Unemployed applicants of North African origin are less likely than employed applicants to receive job interview invitations, while this disparity is not (significantly)	The author suggests that unemployment may **signal** lower productivity to recruiters. He also suggests that employment status information is more relevant for

					observed for unemployed (compared to employed) applicants of French origin. **(-/0)**	applicants of North African origin than for applicants of French origin.
Shi and Di Stasio (2022)	Central and Northern Europe (Switzerland and Norway)	Unemployment (duration) and former unemployment	Job interview invitation	Experimental (lab)	Unemployed applicants and applicants with former unemployment are less likely to receive job interview invitations than (continuously) employed job seekers. The duration of the unemployment spell does not moderate this effect. Applicants with a vocational education and training (VET) background are even more penalised since unemployment is less likely for them. **(-)**	N/A
Shi and Wang (2022)	Central and Southern Europe (Switzerland and Greece)	Unemployment (duration) and former unemployment	Job offer	Experimental (lab)	Unemployed applicants and applicants with former unemployment are less likely to receive job offers in Switzerland. In contrast, in Greece, their chances are equal to those of employed applicants. The duration of the unemployment spell does not moderate this effect. **(-/0)**	The authors suggest that regional differences in the unemployment rate moderate this effect, with Greek recruiters attributing unemployment less to individual failure due to the high unemployment rate, whereas in Switzerland, characterised by a substantially lower unemployment rate, being unemployed is perceived as a personal failure.
Shi et al. (2018)	Central Europe (Switzerland)	Unemployment (duration) and former unemployment	Job interview invitation	Experimental (lab)	Unemployed applicants and applicants with former unemployment are less likely to receive job interview invitations than (continuously) employed job applicants, yet only when they have occupation-specific education and job experience. Long periods of unemployment are penalised more than shorter periods. **(-/0)**	The authors suggest that unemployment may reveal a **human capital depreciation**. For those without occupation-specific education and job experience, their background may already serve as a negative signal, eliminating the adverse effects of unemployment for them.
Suomi et al. (2022)	Oceania (Australia)	Unemployment	Employee assessment and job offer	Experimental (lab)	Recruiters perceive unemployed applicants as less employable than employed applicants. However, they are not less likely to receive job offers, even though this measure is highly correlated with employability.[x] **(-/0)**	The authors suggest that unemployment may **signal** less developed soft skills.
Tomlin (2022)	North America (United States of America)	Unemployment (long-term) and inactivity (homemaker)	Job interview invitation	Experimental (field)[n]	Unemployed and inactive applicants (all mothers) are less likely than employed applicants to receive job interview invitations, regardless of whether they explain their unemployment spell. **(-)**	The authors suggest that unemployment and former inactivity due to family reasons may reveal **human capital depreciation** to recruiters.
Trzebiatowski et al. (2020)	North America (United States of America)	Unemployment (duration)	Job interview invitation	Experimental (field)[o], survey	Long-term unemployed applicants are less likely to receive job interview invitations than the short-term unemployed, yet only in regions without unemployment	The authors provide evidence that long-term unemployment acts as a **signal** of lower competence to recruiters.

					status anti-discrimination legislation. There was no significant difference in job interview invitation rates between the long-term unemployed and the employed. **(-/0)**	
Van Belle et al. (2018)	Western Europe (Belgium)	Unemployment duration	Job interview invitation and job offer	Experimental (lab)[p]	The longer an applicant is unemployed, the lower the likelihood of receiving job interview invitations and job offers.[y] **(-)**	The authors provide evidence that long-term unemployment acts as a **signal** of lower motivation to recruiters and that recruiters are hesitant to interview the long-term unemployed, believing that if they were productive, other recruiters would have already hired them (i.e. **rational herding**).
Weisshaar (2018)	North America (United States of America)	Unemployment (long-term) and inactivity (homemaker)	Job interview invitation	Experimental (field and lab)[q]	Unemployed applicants are less likely than employed applicants to receive job interview invitations. Applicants who were previously inactive due to family reasons face an additional penalty, with inactive fathers facing the greatest disadvantage. **(-)**	The author provides evidence that unemployment acts as a **signal** of lower competence to recruiters. In contrast, inactivity due to family reasons signals less commitment, which is a more severe violation of ideal worker norms.
Weisshaar (2021)	North America (United States of America)	Unemployment (long-term) and inactivity (homemaker)	Job offer	Experimental (lab)[r]	Unemployed applicants are less likely than employed applicants to receive job offers. Applicants who were previously inactive due to family reasons face an additional penalty. Providing recruiters with information about an applicant's job performance and social skills mitigates the penalty for the unemployed but not for the inactive. **(-)**	The author suggests that the lower job offer chances of unemployed applicants may stem from informational biases, where employers base hiring decisions on assumptions drawn from job history. Penalties for homemakers seem to result from information-resistant cognitive biases: they remain disadvantaged even with multiple types of positive information. Inactivity due to family reasons violates ideal worker norms.

Notes. The following abbreviations are used: N/A (not applicable), i.e. (that is), e.g. (for example), '(+)' signifies a positive effect of the independent variable on the dependent variable, '(-)' denotes a negative effect, and '(0)' indicates that there is no discernible effect.

[a] When the term 'unemployment' is used, it refers specifically to the (fictional) applicant's current unemployment status during the job application process. It was consistently indicated whether the study considered only short-term (< 1 year), long-term (≥ 1 year), or various durations of unemployment. When the study also accounts for earlier breaks in employment history (i.e. an unemployment spell followed by employment), this is referred to as 'former unemployment'.

[b] Fictitious resumes were sent only for administrative and commercial jobs; fictitious applicants were all young men.

[c] Fictitious job applications were presented only for the hospitality industry.

[d] Fictitious resumes were sent only for gardener and receptionist vacancies; fictitious applicants were all young men without degrees.

[e] Fictitious resumes were sent only for accountant and sales assistant vacancies; fictitious applicants were all men.

[f] Fictitious resumes were sent only for administrative vacancies; fictitious applicants were all females with bachelor degrees.

[g] Fictitious resumes were sent only for administrative, customer service, and sales assistant vacancies; fictitious applicants were all young.

[h] Fictitious resumes were sent only for accountant and technical computer-support vacancies; fictitious applicants were 30-year-old females with university degrees and work experience.

[i] Fictitious resumes were sent only for accountant vacancies.

[j] Fictitious applicants were all recent college graduates with bachelor's degrees.

[k] The study was quasi-experimental, involving two genuine job seekers (two women in their late twenties) applying for real administrative vacancies.

[l] Fictitious job applications were only presented for assistant manager vacancies.

[m] Fictitious resumes were sent only for construction worker vacancies; fictitious applicants were all young men.

[n] Fictitious resumes were sent only for administrative vacancies; fictitious applicants were all mothers with bachelor degrees in psychology.

[o] Fictitious resumes were sent only for administrative, customer service, and sales assistant vacancies.

[p] Fictitious job applications were only presented for counter-assistant vacancies.

[q] Fictitious resumes were sent only for professional and managerial jobs; fictitious applicants were laid-off college-educated parents.

[r] Fictitious job applications were presented for marketing vacancies only. In contrast to the other studies, the experiment participants were not real recruiters but were drawn from a random sample of US adults.

[s] The study shows that the negative effect is absent in the information and communication industry. The authors suggest that this may be due to the high number of freelancers and precarious workers in this sector, which implies that unemployment is relatively common and, therefore, not penalised.

[t] The study shows that ethnic minority candidates face an additive disadvantage since they are also penalised for their race.

[u] The effect applies to accountant positions only.

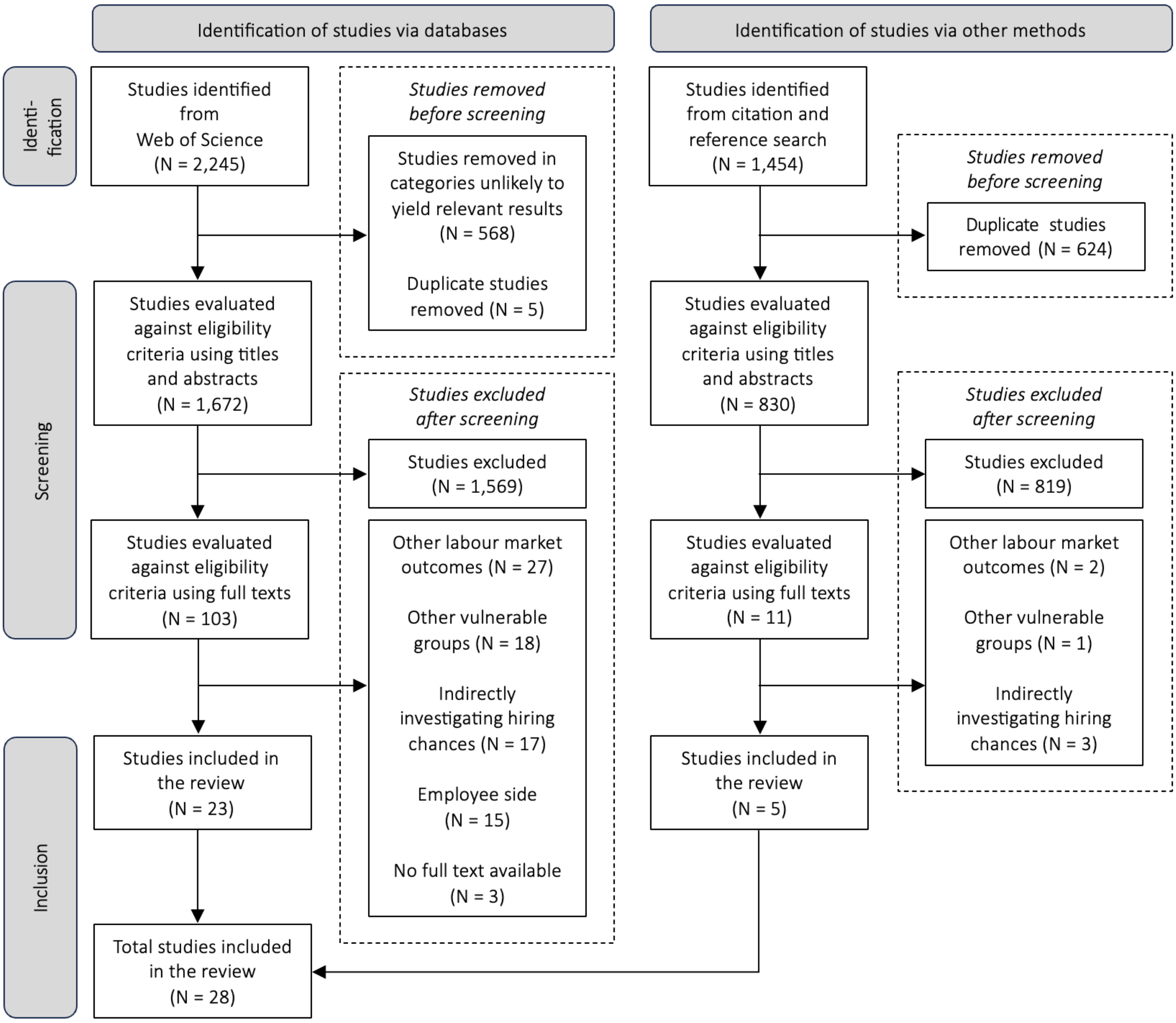
[v] The effect applies to medium- and low-skilled positions only.

[w] Given that the effect of unemployment status is not the primary focus of this study, the experience levels of the unemployed and employed applicants (reference group) were not equated. The unemployed applicants were assigned one year less experience than the employed applicants.

[x] Caution should be exercised when considering the results, as the authors acknowledge that the study was underpowered.

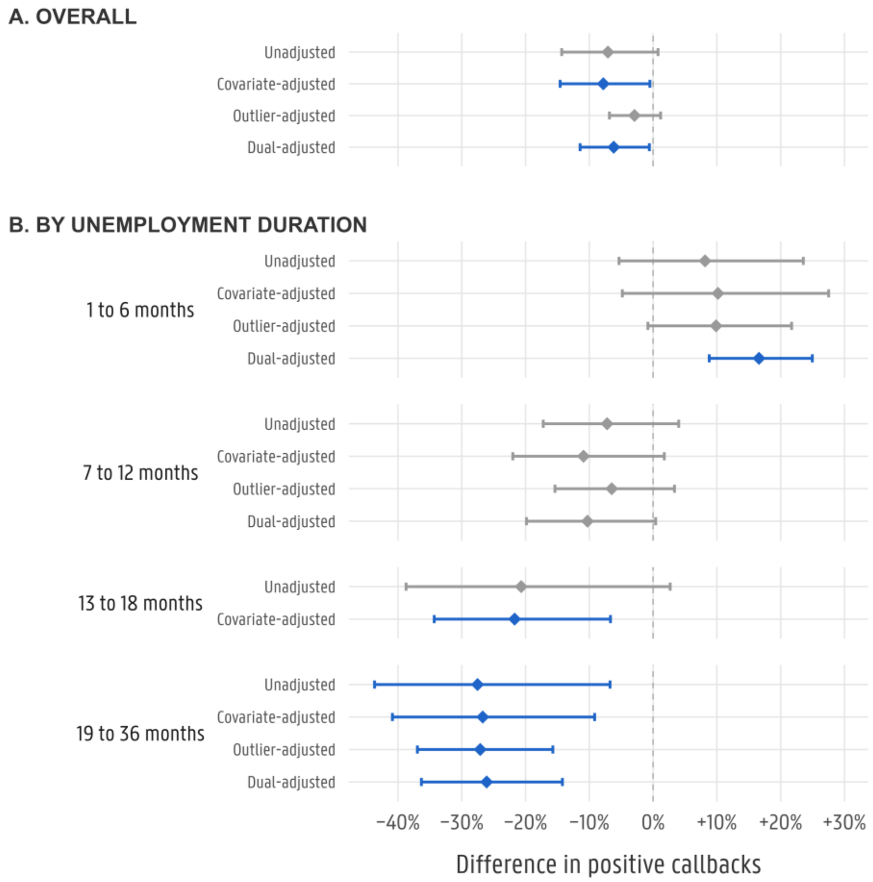
[y] This is the only study to examine different unemployment durations without comparing them to the standard benchmark of being employed.

Figure 1. Study selection flow diagram



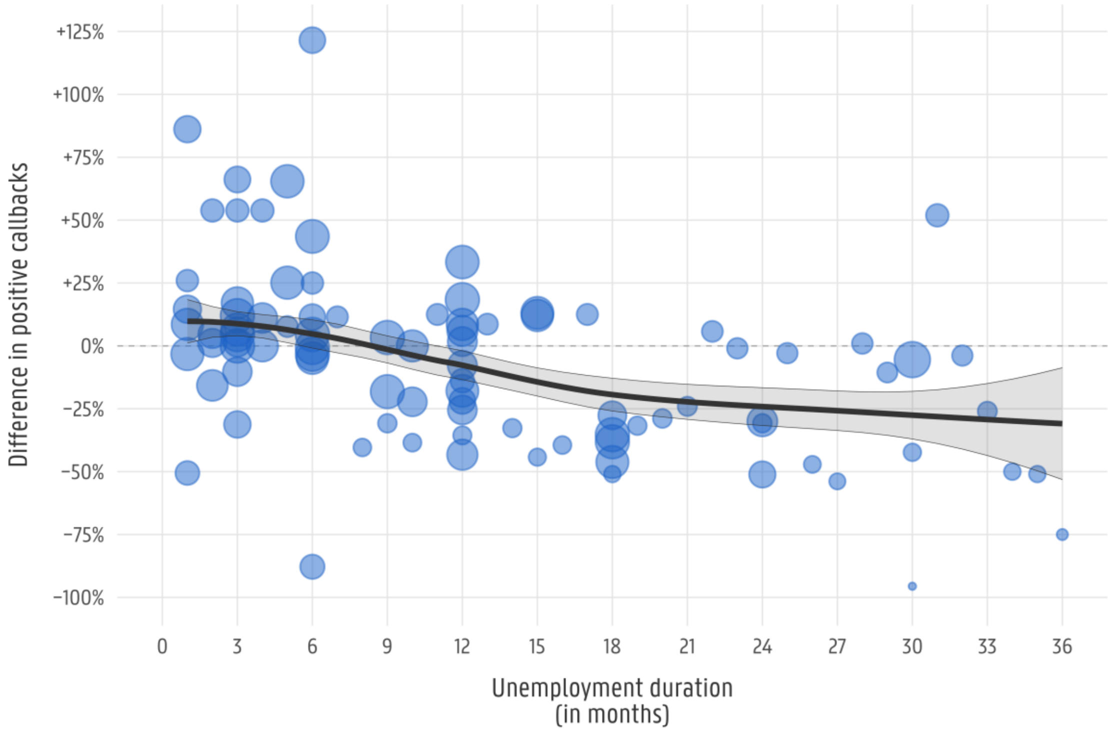
Notes. This figure is adapted from Page et al. (2021, p. 5).

Figure 2. Aggregated unemployed–employed differences in positive callbacks



Notes. Diamond shapes are aggregated differences in positive callbacks between the unemployed and employed by unemployment duration based on meta-analytic estimates (see Equation 2 and detailed results in Table A3 in Appendix A); negative estimates indicate a difference to the disadvantage of the unemployed. Error bars represent their 95% confidence intervals. Panel A shows callback differences for all unemployment durations (i.e. 1 to 36 months); Panel B shows callback differences by specific unemployment duration. Covariate-adjusted estimates are marginal means derived from a mixed effects meta-regression model with the following covariates: unemployment rate, response type, region, year, gender, and age (see Equation 3 and detailed regression results in Tables A4 and A5 in Appendix A). The covariate-adjusted estimate in Panel A also includes the duration variable as a covariate to deduce a netted estimate of the positive callback difference between the unemployed and employed across unemployment periods. Outlier-adjusted estimates exclude the unemployed–employed callback ratios of the studies for which the upper (lower) bound of the 95% confidence interval is lower (higher) than the lower (upper) bound of the confidence interval of the pooled random effects callback ratio (see detailed results in Table B3 in Appendix B). Dual-adjusted estimates are adjusted for outliers and covariates (see Equation 3 and detailed regression results in Tables B4 and B5 in Appendix B). In the absence of outliers for the models related to the 13 to 18 months specification, no outlier- or dual-adjusted estimates are reported. Blue-coloured (dark) estimates and error bars are statistically significantly different from zero at the 5% significance level; grey-coloured (light) estimates and error bars are not statistically significantly different from zero.

Figure 3. Employed–unemployed differences in positive callbacks by unemployment duration



Notes. Points are differences in positive callbacks between the unemployed and employed; positive estimates indicate a difference favouring the unemployed. The point size represents the respective meta-analytic random effect weight. The dark grey line shows the smoothed, locally weighted regression curve (LOESS) and represents the relationship between the unemployment duration and the estimated differences in positive callbacks derived from the general meta-analytic specification (see Equation 2). The specification of the weighted LOESS curve equates to $CB_k^* = \alpha + \beta UD_k + \varepsilon_k$, where CB_k^* is the predicted difference in positive callbacks for each effect *k*, α is the intercept, UD_k is the duration of unemployment in months for effect *k*, and β is the coefficient for the duration variable. The grey, semi-transparent ribbon shows its 95% confidence interval.

Figure 4. Employed–unemployed differences in positive callbacks by unemployment duration and unemployment rate

Notes. Points are differences in positive callbacks between the unemployed and employed by unemployment duration; negative estimates indicate a difference to the disadvantage of the unemployed. The point size represents the respective meta-analytic random effect weight. The lines show the weighted least squares (WLS) regression curves and define the relationship between the unemployment duration and the estimated differences in positive callbacks derived from the general meta-analytic specification (see Equation 2) by unemployment rate category (also see detailed regression results, including the interaction term between unemployment duration and the country–year unemployment rate in Table A4 in Appendix A). The specification of the WLS curves equates to $CB_k^* = \alpha + \beta UD_k + \varepsilon_k$, where CB_k^* is the predicted difference in positive callbacks for each effect k and for a given unemployment rate category, α is the intercept, UD_k is the duration of unemployment in months for effect k, and β is the coefficient for the duration variable. The semi-transparent ribbons show 95% confidence intervals.

Appendices

A. Additional tables and figures

Table A1. Reporting guidelines for meta-analyses in economics: a checklist

Criterion	Reporting
1. Research question and effects	
1.1. A statement of the specific theories, hypotheses, and effects studied.	We state the effects studied in Section 1. In addition, we detail these effects further in Sections 2.4.2 and 2.4.3. The expected theoretical mechanisms are clarified in Section 1.
1.2. A definition of how effects and their standard errors were measured.	We define how we estimated effects and standard errors in Sections 2.4.2 and 2.4.3.
1.3. A description of how measured effects are comparable.	We describe how the extracted data and the derived effects are comparable in Sections 2.1, 2.3, and 2.4.2.
2. Search, synthesis, and coding	
2.1. A description of how the research literature was searched.	We describe the study search process in Section 2.1.
2.2. A list of the rules for study inclusion or exclusion.	We list inclusion and exclusion criteria in Section 2.1. The study selection process is described in Section 2.2.
2.3. A statement addressing who searched, read, and coded the research literature.	We address the study identification, study screening, and data extraction processes in Sections 2.1., 2.2, and 2.3.
2.4. A list of the information coded for each study.	We list the coded information for each study in Section 2.3.
2.5. A description of the rule or method used to identify and omit outliers or influence points.	We describe the rule to identify outliers in the figure notes of Figure 2 and in Appendix B.
3. Modelling	
3.1. A table showing definitions of the coded variables along with their descriptive statistics.	We report aggregate descriptive statistics in Table 2. Variable descriptions can be retrieved from the supplementary data and code.
3.2. A description of the fitted meta-regression analysis and the strategy used to fit it.	We describe the estimation details of our meta-regression analysis and how we fit it in Section 2.4.3.
3.3. A report of publication, selection, or misspecification biases.	We briefly describe how we identified and controlled for publication bias in Section 2.4.4. Findings are concisely reported alongside the results of our main analysis in Section 3.2. A detailed report of the results of our bias-identifying and -correcting analyses can be found in Appendix B.
3.4. A description of the methods to accommodate heteroscedasticity and dependence across estimates.	We describe how we deal with dependence and clustering in Sections 2.4.2 and 2.4.3.
4. Reporting and interpretation	
4.1. Graph(s) of the effect sizes or other statistical displays of data.	We visualise estimated marginal means derived from our random effects meta-analytic and mixed effects meta-regression models in Figure 2. Figures 3 and 4 show the weighted treatment effects of each study (by unemployment duration and unemployment rate). Figures A1 to A4 in Appendix A show forest plots by unemployment duration. Figures B1 to B5 in Appendix B are funnel plots for the general specification and by unemployment duration.

(continued)

Criterion	Reporting
4.2. Robustness checks for meta-regression analyses and publication bias methods.	We extensively report on the robustness of our primary analyses and on publication bias in Appendix B. Findings of these robustness checks are concisely reported alongside the results of our main analysis in Section 3.2.
4.3. A discussion of the economic (or practical) significance of the main findings.	We discuss the economic and practical significance of our findings in Section 4.
4.4. A statement about sharing the data along with the codes of the main analyses.	See the data and code availability statement.

Notes. Criteria from this checklist were adapted from Havránek et al. (2020, p. 471–472).

Table A2. Excluded Web of Science categories

Category	Number of articles
Public Environmental Occupational Health	73
Law	53
Computer Science Information Systems	42
Environmental Studies	42
Rehabilitation	42
Computer Science Theory Methods	41
Engineering Electrical Electronic	41
Computer Science Interdisciplinary Applications	36
Computer Science Artificial Intelligence	32
Environmental Sciences	30
Urban Studies	27
Information Science Library Science	23
Computer Science Software Engineering	22
Health Care Sciences Services	22
Mathematics Interdisciplinary Application	22
Engineering Industrial	21
Green Sustainable Science Technology	20
Regional Urban Planning	20
Statistics Probability	20
Criminology Penology	19
Materials Science Multidisciplinary	19
Nursing	19
Geography	18
Engineering Multidisciplinary	17
Engineering Civil	16
Telecommunications	16
Medicine General Internal	14
Radiology Nuclear Medicine Medical Imaging	14
Astronomy Astrophysics	12
Computer Science Hardware Architecture	12
Engineering Manufacturing	12
Physics Applied	12
Physics Particles Fields	12
Transportation	12
Hospitality Leisure Sport Tourism	11
Medicine Research Experimental	11
Energy Fuels	10
Gerontology	10
Pharmacology Pharmacy	10
Surgery	10
Agricultural Economics Policy	9
Biochemistry Molecular Biology	9
Computer Science Cybernetics	9
Physics Nuclear	9
Area Studies	8
Biology	8

(continued)

Category	Number of articles
Physics Condensed Matter	8
Agriculture Multidisciplinary	7
Anthropology	7
Geosciences Multidisciplinary	7
Geriatrics Gerontology	7
Immunology	7
Oncology	7
Construction Building Technology	6
Emergency Medicine	6
Fisheries	6
Substance Abuse	6
Virology	6
Chemistry Physical	5
Clinical Neurology	5
Critical Care Medicine	5
Cultural Studies	5
Infectious Diseases	5
Linguistics	5
Neurosciences	5
Pediatrics	5
Physics Multidisciplinary	5
Zoology	5
Agricultural Engineering	4
Agronomy	4
Biophysics	4
Chemistry Multidisciplinary	4
Engineering Chemical	4
Food Science Technology	4
Forestry	4
Genetics Heredity	4
Nanoscience Nanotechnology	4
Nutrition Dietetics	4
Obstetrics Gynaecology	4
Optics	4
Remote Sensing	4
Respiratory System	4
Sport Sciences	4
Agriculture Dairy Animal Science	3
Automation Control Systems	3
Cell Biology	3
Engineering Environmental	3
Ergonomics	3
Evolutionary Biology	3
History Philosophy Of Science	3
Mathematics Applied	3
Mechanics	3

(continued)

Category	Number of articles
Medical Informatics	3
Nuclear Science Technology	3
Oceanography	3
Orthopedics	3
Physiology	3
Primary Health Care	3
Robotics	3
Water Resources	3
Biodiversity Conservation	2
Biotechnology Applied Microbiology	2
Chemistry Analytical	2
Chemistry Applied	2
Endocrinology Metabolism	2
Engineering Ocean	2
Geography Physical	2
Hematology	2
Imaging Science Photographic Technology	2
Language Linguistics	2
Mathematics	2
Medical Ethics	2
Medical Laboratory Technology	2
Metallurgy Metallurgical Engineering	2
Meteorology Atmospheric Sciences	2
Microbiology	2
Ornithology	2
Philosophy	2
Plant Sciences	2
Psychology Biological	2
Religion	2
Toxicology	2
Veterinary Sciences	2
Archaeology	1
Asian Studies	1
Cell Tissue Engineering	1
Chemistry Inorganic Nuclear	1
Dentistry Oral Surgery Medicine	1
Electrochemistry	1
Engineering Biomedical	1
Engineering Geological	1
Engineering Marine	1
Engineering Mechanical	1
Engineering Petroleum	1
Entomology	1
Film Radio Television	1
Geochemistry Geophysics	1
Geology	1

(continued)

Category	Number of articles
Horticulture	1
Instruments Instrumentation	1
Marine Freshwater Biology	1
Materials Science Coatings Films	1
Materials Science Paper Wood	1
Materials Science Textiles	1
Mathematical Computational Biology	1
Medieval Renaissance Studies	1
Music	1
Otorhinolaryngology	1
Paleontology	1
Pathology	1
Physics Atomic Molecular Chemical	1
Physics Fluids Plasmas	1
Physics Mathematical	1
Psychology Mathematical	1
Rheumatology	1
Theater	1
Transplantation	1
Urology Nephrology	1
Total	1,289

Notes. This table reports the number of excluded articles by Web of Science category unlikely to yield relevant results.

Table A3. Meta-analytic estimates of unemployed–employed differences in positive callbacks

Specification	Estimate	SE	$CI_{95\%}$	*t*	*p*	I^2
Overall	–0.0731	0.0371	[–0.1440, 0.0036]	–1.90	0.061	71.31%
1 to 6 months	0.0823	0.0711	[–0.0531, 0.2369]	1.20	0.237	71.45%
7 to 12 months	–0.0731	0.0511	[–0.1738, 0.0399]	–1.38	0.184	56.64%
13 to 18 months	–0.2078	0.0928	[–0.3878, 0.0251]	–1.99	0.072	68.95%
19 to 36 months	–0.2822	0.0911	[–0.4487, –0.0654]	–2.61	0.016	56.99%

Notes. Acronyms used: SE (standard error), CI (confidence interval). Estimates are differences in positive callbacks between the unemployed and employed. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. Standard errors were clustered at the (sub-)study level and corrected using a small sample adjustment. Marginal means are the effects of unemployment on positive callbacks, calculated at the average values of the continuous predictor variables and the reference categories of the categorical predictor variables. I^2-values around 25%, 50%, or 75% indicate a low, moderate, or high proportion of residual heterogeneity relative to the amount of unaccounted variability, respectively.

Table A4. Weighted least squares meta-regression of callback differences between the unemployed and employed on unemployment duration, unemployment rates, and other study-level covariates

Term	(1)	(2)	(3)
Intercept	0.1187$^{\times}$ (0.0587)	−264.0257 (108.8132)	−264.5789 (97.7941)
Unemployment duration *(in months)*	−0.0171* (0.0038)	−0.0238* (0.0063)	−0.0668* (0.0174)
Unemployment rate *(in per cent × 100)*	–	0.1937 (0.0848)	0.0977 (0.0949)
Response type: Interview invitation (ref.)			
Response type: Positive reaction	–	−0.1510 (0.1784)	−0.1100 (0.1479)
Region: Northern America (ref.)			
Region: Western Europe	–	−0.1354 (0.4215)	0.2932 (0.3326)
Region: Northern Europe	–	0.8258 (0.3915)	0.7897 (0.3984)
Region: Central Europe	–	2.4012 (1.0055)	2.5353 (0.9752)
Region: Asia	–	0.4661 (0.4269)	0.4942 (0.3969)
Year	–	0.1303 (0.0538)	0.1308 (0.0482)
Gender: Male (ref.)			
Gender: Female	–	0.5047 (0.5112)	0.6471 (0.4223)
Gender: Mixed	–	0.2017 (0.6141)	0.5081 (0.4847)
Age: Prime-aged (31–50) (ref.)			
Age: Young (21–30)	–	0.3135 (0.2277)	0.3081 (0.1695)
Age: Old (51+)	–	0.0629 (0.0800)	0.0040 (0.0724)
Unemployment duration × Unemployment rate	–	–	0.0059* (0.0023)
Marginal mean	−0.0947* (0.0384)	−0.0789* (0.0354)	−0.1043** (0.0350)
AIC	84.11	85.59	79.94
BIC	91.55	120.55	117.02
I^2	71.67%	60.23%	59.52%

Notes. Abbreviations and acronyms used: ref. (reference category), AIC (Akaike information criterion), BIC (Bayesian information criterion). Statistics are coefficient estimates with standard errors between parentheses. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. Standard errors were clustered at the (sub-)study level and corrected using a small sample adjustment. Marginal means are the effects of unemployment on positive callbacks, calculated at the average values of the continuous predictor variables and the reference categories of the categorical predictor variables. I^2-values around 25%, 50%, or 75% indicate a low, moderate, or high proportion of residual heterogeneity relative to the amount of unaccounted variability, respectively. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$, $^{\times}$ $p < 0.10$.

Table A5. Weighted least squares meta-regression of callback differences between the unemployed and employed on study-level covariates (by unemployment duration)

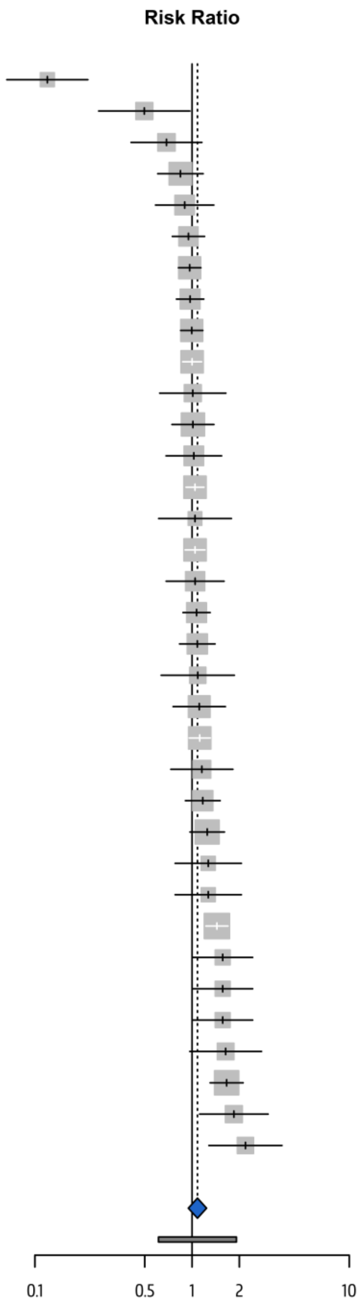
Term	1 to 6 months	7 to 12 months	13 to 18 months	19 to 36 months
Intercept	–97.5786 (127.9231)	76.0044 (132.5555)	8.5392 (17.5247)	–6.1658 (2.5709)
Unemployment rate *(in per cent × 100)*	–0.1057 (0.3809)	0.0191 (0.3304)	0.0705 (0.0192)	0.6950 (0.3024)
Response type: Interview invitation (ref.)				
Response type: Positive reaction	–0.6394 (0.5152)	0.1981 (0.3695)	–0.1317 (0.1134)	2.7396 (1.3925)
Region: America (ref.)				
Region: Europe	–0.2319 (1.0062)	0.0071 (0.9282)	0.2302 (0.1060)	3.2596 (1.3321)
Region: Asia	–0.8671 (0.5931)	0.0063 (0.4545)	–	–
Year	0.0492 (0.0627)	–0.0380 (0.0651)	–0.0047 (0.0087)	–
Gender: Male (ref.)				
Gender: Female	–0.4922 (2.8200)	0.1321 (1.2343)	–	–
Gender: Mixed	–0.7085 (1.6514)		–	–
Age: Prime-aged (31–50) (ref.)				
Age: Young (21–30)	0.5197 (0.3998)	–0.0388 (0.7452)	–	–
Age: Old (51+)	0.0558 (0.1294)	–0.1365 (0.3886)	–	–
Marginal mean	0.1022 (0.0786)	–0.1108ˣ (0.0545)	–0.2138*** (0.0600)	–0.2702*** (0.0741)
I^2	85.17%	60.80%	0.00%	42.38%

Notes. Abbreviations used: ref. (reference category). Statistics are coefficient estimates with standard errors between parentheses. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. Some variables were dropped from the regression models due to multicollinearity. Standard errors were clustered at the (sub-)study level and corrected using a small sample adjustment. I^2-values around 25%, 50%, or 75% indicate a low, moderate, or high proportion of residual heterogeneity relative to the amount of unaccounted variability, respectively. *** p < 0.001; ** p < 0.01; * p < 0.05, ˣ p < 0.10.

Figure A1. Forest plot: 1 to 6 months of unemployment

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95%-CI	Weight
Farber et al. (2016) 1-3	12	1,214	100	1,214	0.12	[0.07; 0.22]	1.5%
Norlander et al. (2020) 1-1	11	119	24	129	0.50	[0.25; 0.97]	2.2%
Pierné (2018) 1-2	22	301	32	301	0.69	[0.41; 1.15]	2.3%
Maurer-Fazio & Wang (2018) 2-1	51	720	86	1,024	0.84	[0.60; 1.18]	3.9%
Pierné (2018) 1-1	35	301	39	301	0.90	[0.59; 1.38]	2.9%
Farber et al. (2019) 2-3	87	747	200	1,629	0.95	[0.75; 1.20]	2.8%
Eriksson & Rooth (2014) 2-2	128	442	464	1,547	0.97	[0.82; 1.14]	3.7%
Farber et al. (2019) 2-1	145	1,215	200	1,629	0.97	[0.80; 1.19]	3.0%
Nunley et al. (2017) 1-2	182	1,158	378	2,394	1.00	[0.85; 1.17]	3.7%
Eriksson & Rooth (2014) 2-1	199	663	464	1,547	1.00	[0.87; 1.15]	3.8%
Farber et al. (2019) 3-2	25	257	36	374	1.01	[0.62; 1.64]	2.2%
Trzebiatowski et al. (2020) 1-1	71	588	72	604	1.01	[0.74; 1.38]	4.1%
Maurer-Fazio & Wang (2018) 1-3	29	367	81	1,052	1.03	[0.68; 1.54]	2.9%
Eriksson & Rooth (2014) 1-1	244	1,062	584	2,654	1.04	[0.92; 1.19]	3.8%
Kroft et al. (2013) 1-5	14	257	147	2,818	1.04	[0.61; 1.78]	1.4%
Eriksson & Rooth (2014) 1-2	183	796	584	2,654	1.04	[0.90; 1.21]	3.8%
Maurer-Fazio & Wang (2018) 1-1	27	367	74	1,052	1.05	[0.68; 1.60]	2.8%
Farber et al. (2019) 2-2	147	1,121	200	1,629	1.07	[0.88; 1.30]	3.0%
Farber et al. (2016) 1-1	108	1,214	100	1,214	1.08	[0.83; 1.40]	3.1%
Farber et al. (2019) 3-3	18	172	36	374	1.09	[0.64; 1.86]	2.0%
Trzebiatowski et al. (2020) 2-1	50	503	45	503	1.11	[0.76; 1.63]	3.6%
Nunley et al. (2017) 1-1	206	1,166	378	2,394	1.12	[0.96; 1.31]	3.8%
Farber et al. (2019) 3-1	31	279	36	374	1.15	[0.73; 1.82]	2.4%
Farber et al. (2016) 1-2	117	1,214	100	1,214	1.17	[0.91; 1.51]	3.1%
Duguet et al. (2018) 2-1	67	294	161	882	1.25	[0.97; 1.61]	4.4%
Kroft et al. (2013) 1-1	17	257	147	2,818	1.27	[0.78; 2.06]	1.5%
Kroft et al. (2013) 1-6	17	257	147	2,818	1.27	[0.78; 2.06]	1.5%
Oberholzer-Gee (2008) 1-1	53	68	170	314	1.44	[1.22; 1.69]	4.9%
Kroft et al. (2013) 1-2	21	257	147	2,818	1.57	[1.01; 2.43]	1.7%
Kroft et al. (2013) 1-3	21	257	147	2,818	1.57	[1.01; 2.43]	1.7%
Kroft et al. (2013) 1-4	21	257	147	2,818	1.57	[1.01; 2.43]	1.7%
Farber et al. (2019) 1-2	27	251	24	365	1.64	[0.97; 2.77]	2.1%
Duguet et al. (2018) 1-1	78	285	141	855	1.66	[1.30; 2.11]	4.5%
Farber et al. (2019) 1-1	33	272	24	365	1.85	[1.12; 3.05]	2.2%
Farber et al. (2019) 1-3	24	167	24	365	2.19	[1.28; 3.73]	2.0%
Random effects model		**18,865**		**47,861**	**1.08**	**[0.95; 1.24]**	**100.0%**
Prediction interval						**[0.61; 1.92]**	

Heterogeneity: $I^2 = 71\%$, $\tau^2 = 0.0806$, $p < 0.01$

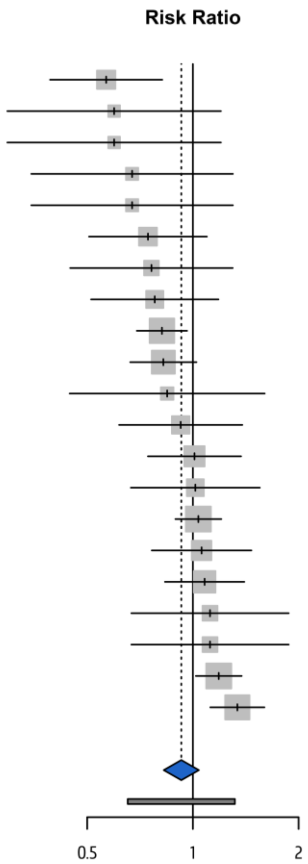


Notes. Acronyms used: RR (risk ratio), CI (confidence interval). Names of the effects are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset. Squares represent the callback ratios of the original effects. The square size represents the respective meta-analytic random effect weight. The horizontal lines depict the callback ratios' 95% confidence intervals. The (blue) diamond shape and vertical dotted line represent the unadjusted random effects pooled callback ratio. The (medium grey) rectangle represents the prediction interval.

Figure A2. Forest plot: 7 to 12 months of unemployment

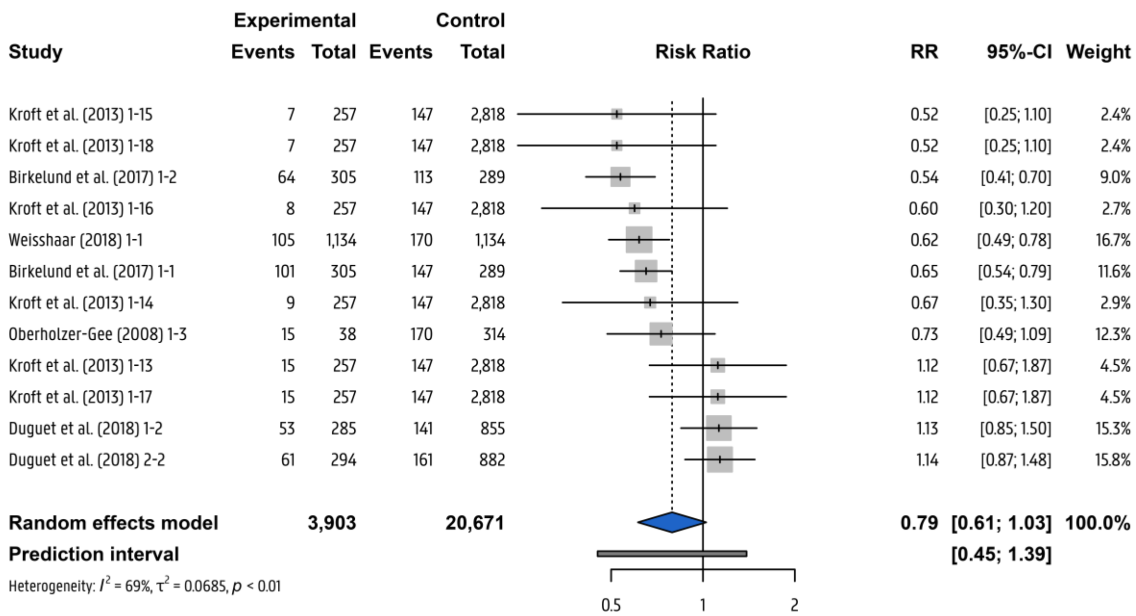
Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95%-CI	Weight
Pedulla (2018) 1-1	42	714	70	674	0.57	[0.39; 0.82]	4.7%
Kroft et al. (2013) 1-8	8	257	147	2,818	0.60	[0.30; 1.20]	1.9%
Kroft et al. (2013) 1-10	8	257	147	2,818	0.60	[0.30; 1.20]	1.9%
Kroft et al. (2013) 1-9	9	257	147	2,818	0.67	[0.35; 1.30]	2.1%
Kroft et al. (2013) 1-12	9	257	147	2,818	0.67	[0.35; 1.30]	2.1%
Maurer-Fazio & Wang (2018) 1-2	36	687	74	1,052	0.74	[0.51; 1.10]	4.5%
Farber et al. (2019) 3-4	19	259	36	374	0.76	[0.45; 1.30]	3.0%
Trzebiatowski et al. (2020) 2-2	37	531	45	503	0.78	[0.51; 1.18]	4.1%
Eriksson & Rooth (2014) 1-3	143	796	584	2,654	0.82	[0.69; 0.96]	8.4%
Farber et al. (2019) 2-4	114	1,127	200	1,629	0.82	[0.66; 1.02]	7.4%
Farber et al. (2019) 1-4	14	252	24	365	0.84	[0.45; 1.60]	2.2%
Maurer-Fazio & Wang (2018) 2-2	29	374	86	1,024	0.92	[0.62; 1.38]	4.3%
Trzebiatowski et al. (2020) 1-2	73	606	72	604	1.01	[0.74; 1.37]	5.7%
Pedulla (2018) 2-1	42	709	39	669	1.02	[0.67; 1.55]	4.0%
Eriksson & Rooth (2014) 2-3	157	505	464	1,547	1.04	[0.89; 1.20]	8.7%
Maurer-Fazio & Wang (2018) 1-4	56	687	81	1,052	1.06	[0.76; 1.47]	5.3%
Farber et al. (2016) 1-4	108	1,214	100	1,214	1.08	[0.83; 1.40]	6.5%
Kroft et al. (2013) 1-7	15	257	147	2,818	1.12	[0.67; 1.87]	3.1%
Kroft et al. (2013) 1-11	15	257	147	2,818	1.12	[0.67; 1.87]	3.1%
Nunley et al. (2017) 1-3	223	1,192	378	2,394	1.18	[1.02; 1.38]	8.7%
Oberholzer-Gee (2008) 1-2	50	69	170	314	1.34	[1.12; 1.60]	8.2%
Random effects model		**11,264**		**32,977**	**0.93**	**[0.83; 1.04]**	**100.0%**
Prediction interval						**[0.65; 1.32]**	

Heterogeneity: $I^2 = 57\%$, $\tau^2 = 0.0291$, $p < 0.01$



Notes. Acronyms used: RR (risk ratio), CI (confidence interval). Names of the effects are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset. Squares represent the callback ratios of the original effects. The square size represents the respective meta-analytic random effect weight. The horizontal lines depict the callback ratios' 95% confidence intervals. The (blue) diamond shape and vertical dotted line represent the unadjusted random effects pooled callback ratio. The (medium grey) rectangle represents the prediction interval.

Figure A3. Forest plot: 13 to 18 months of unemployment

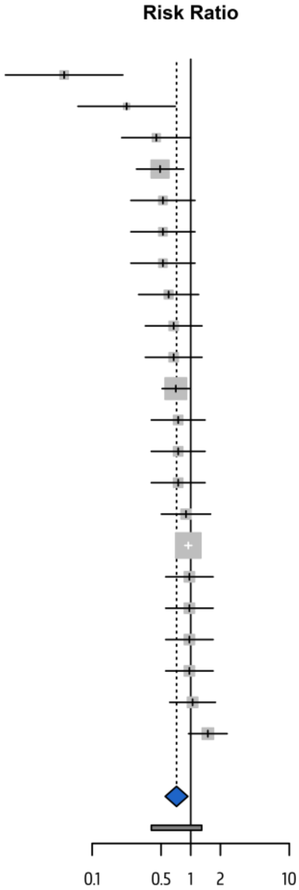


Notes. Acronyms used: RR (risk ratio), CI (confidence interval). Names of the effects are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset. Squares represent the callback ratios of the original effects. The square size represents the respective meta-analytic random effect weight. The horizontal lines depict the callback ratios' 95% confidence intervals. The (blue) diamond shape and vertical dotted line represent the unadjusted random effects pooled callback ratio. The (medium grey) rectangle represents the prediction interval.

Figure A4. Forest plot: 19 to 36 months of unemployment

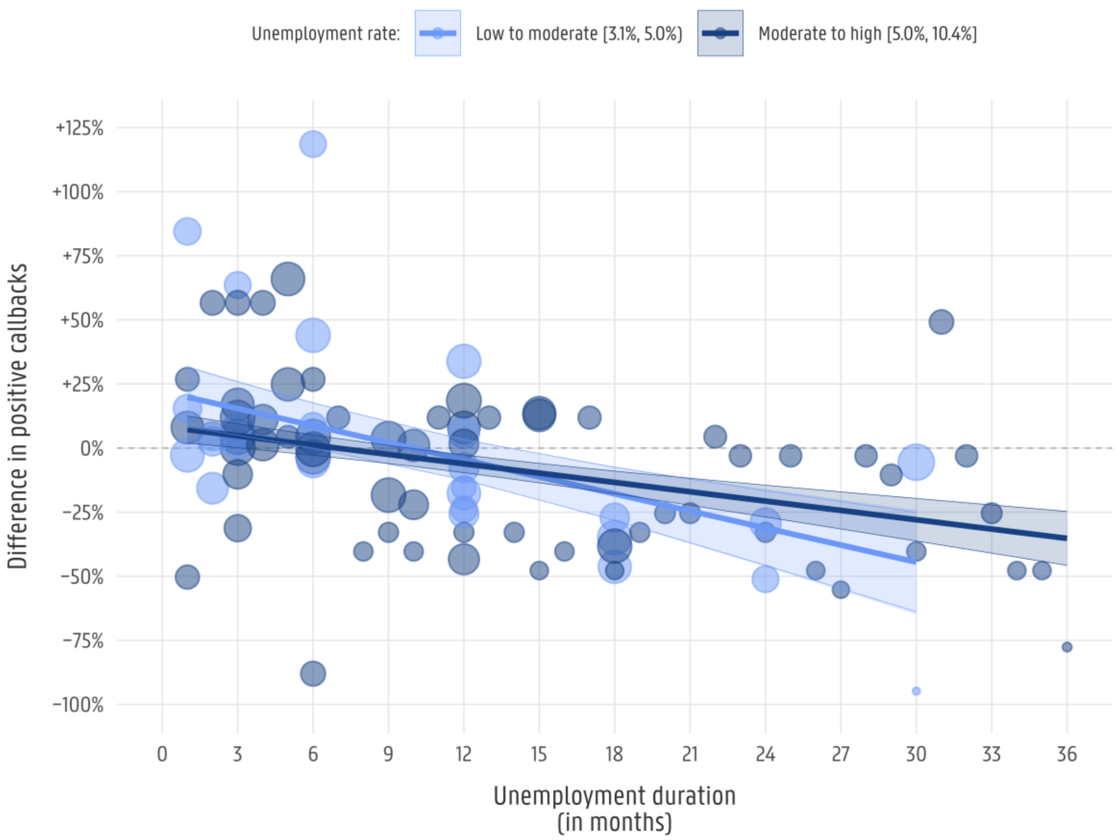
Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95%-CI	Weight
Oberholzer-Gee (2008) 1-5	2	71	170	314	0.05	[0.01; 0.20]	2.2%
Kroft et al. (2013) 1-36	3	257	147	2,818	0.22	[0.07; 0.70]	1.1%
Kroft et al. (2013) 1-27	6	257	147	2,818	0.45	[0.20; 1.00]	2.0%
Tomlin (2022) 1-1	15	98	31	99	0.49	[0.28; 0.85]	10.6%
Kroft et al. (2013) 1-26	7	257	147	2,818	0.52	[0.25; 1.10]	2.2%
Kroft et al. (2013) 1-34	7	257	147	2,818	0.52	[0.25; 1.10]	2.2%
Kroft et al. (2013) 1-35	7	257	147	2,818	0.52	[0.25; 1.10]	2.2%
Kroft et al. (2013) 1-30	8	257	147	2,818	0.60	[0.30; 1.20]	2.4%
Kroft et al. (2013) 1-19	9	257	147	2,818	0.67	[0.35; 1.30]	2.7%
Kroft et al. (2013) 1-24	9	257	147	2,818	0.67	[0.35; 1.30]	2.7%
Oberholzer-Gee (2008) 1-4	26	68	170	314	0.71	[0.51; 0.97]	15.5%
Kroft et al. (2013) 1-20	10	257	147	2,818	0.75	[0.40; 1.40]	2.8%
Kroft et al. (2013) 1-21	10	257	147	2,818	0.75	[0.40; 1.40]	2.8%
Kroft et al. (2013) 1-33	10	257	147	2,818	0.75	[0.40; 1.40]	2.8%
Kroft et al. (2013) 1-29	12	257	147	2,818	0.90	[0.50; 1.59]	3.2%
Kristal et al. (2023) 1-2	742	2,256	785	2,255	0.94	[0.87; 1.03]	21.3%
Kroft et al. (2013) 1-23	13	257	147	2,818	0.97	[0.56; 1.69]	3.4%
Kroft et al. (2013) 1-25	13	257	147	2,818	0.97	[0.56; 1.69]	3.4%
Kroft et al. (2013) 1-28	13	257	147	2,818	0.97	[0.56; 1.69]	3.4%
Kroft et al. (2013) 1-32	13	257	147	2,818	0.97	[0.56; 1.69]	3.4%
Kroft et al. (2013) 1-22	14	257	147	2,818	1.04	[0.61; 1.78]	3.5%
Kroft et al. (2013) 1-31	20	257	147	2,818	1.49	[0.95; 2.34]	4.2%
Random effects model		**7,119**		**53,706**	**0.72**	**[0.55; 0.93]**	**100.0%**
Prediction interval						**[0.40; 1.29]**	

Heterogeneity: $I^2 = 57\%$, $\tau^2 = 0.0739$, $p < 0.01$



Notes. Acronyms used: RR (risk ratio), CI (confidence interval). Names of the effects are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset. Squares represent the callback ratios of the original effects. The square size represents the respective meta-analytic random effect weight. The horizontal lines depict the callback ratios' 95% confidence intervals. The (blue) diamond shape and vertical dotted line represent the unadjusted random effects pooled callback ratio. The (medium grey) rectangle represents the prediction interval.

Figure A5. Employed–unemployed differences in positive callbacks by unemployment duration and unemployment rate: 5.0% threshold value



Notes. Points are differences in positive callbacks between the unemployed and employed by unemployment duration; negative estimates indicate a difference to the disadvantage of the unemployed. The point size represents the respective meta-analytic random effect weight. The lines show the weighted least squares (WLS) regression curves and define the relationship between the unemployment duration and the estimated differences in positive callbacks derived from the general meta-analytic specification by unemployment rate category. The specification of the WLS curves equates to $CB_k^* = \alpha + \beta UD_k + \varepsilon_k$, where CB_k^* is the predicted difference in positive callbacks for each effect *k* and for a given unemployment rate category, α is the intercept, UD_k is the duration of unemployment in months for effect *k*, and β is the coefficient for the duration variable. The semi-transparent ribbons show their 95% confidence intervals.

Figure A6. Employed–unemployed differences in positive callbacks by unemployment duration and unemployment rate: 7.0% threshold value

Notes. Points are differences in positive callbacks between the unemployed and employed by unemployment duration; negative estimates indicate a difference to the disadvantage of the unemployed. The point size represents the respective meta-analytic random effect weight. The lines show the weighted least squares (WLS) regression curves and define the relationship between the unemployment duration and the estimated differences in positive callbacks derived from the general meta-analytic specification by unemployment rate category. The specification of the WLS curves equates to $CB_k^* = \alpha + \beta UD_k + \varepsilon_k$, where CB_k^* is the predicted difference in positive callbacks for each effect *k* and for a given unemployment rate category, α is the intercept, UD_k is the duration of unemployment in months for effect *k*, and β is the coefficient for the duration variable. The semi-transparent ribbons show their 95% confidence intervals.

B. Publication bias and robustness checks

B.I. Overview of the results

Below, we provide an overview of the most important results of the analyses we performed to account for publication bias and verify the robustness of our specifications and methods. Where appropriate, they are contrasted with the findings from the main analyses.

First, we examined funnel plot asymmetry through bias statistics and visually inspected the funnel plots. We estimated two types of bias statistics based on Egger's test for asymmetry and Peters' binary effects adaptation of Egger's test (see Table B1; Egger et al., 1997; Peters et al., 2006). Egger's test suggests asymmetry in the overall specification and in the 7 to 12 months and 19 to 36 months unemployment specifications. Peters' adaptation only suggests funnel plot asymmetry in the 19 to 36 months unemployment specification. Visually, we notice that most values fall within or are close to the expected confidence intervals except for a few clear outliers (see Figures B1 to B5). These strongly deviating treatment effects were also detected (and controlled for) in our subsequent outlier analysis.

Second, we identified outliers defined as effects for which the upper (lower) bound of the 95% confidence interval of the unemployed–employed callback ratio is lower (higher) than the lower (upper) bound of the confidence interval of the pooled random effects callback ratio. Table B2 provides an overview of the outliers by estimation specification. Subsequently, we recalculated the meta-analytic weighted averages and meta-regression coefficients, leaving out the effects detected in the outlier analysis (see Tables B3 to B5).

The effect of unemployment (duration) on hiring chances remains statistically significant and substantial after adjusting for outliers. Based on our outlier-adjusted estimates, the short-term unemployed (up to six months) have marginally significantly better hiring chances, receiving about 10% more positive callbacks on average than the employed ($CI_{95\%}$ = [–0.67%, 21.84%]); see Table B3). In contrast, after one year of unemployment, the unemployed are significantly worse off in terms of hiring chances. Between 13 and 18 months of unemployment, the unemployed receive about 21% fewer positive callbacks on average ($CI_{95\%}$ = [–38.78%, 2.51%]); between 19 and 36 months, this statistic rises to about 27% fewer positive callbacks on average ($CI_{95\%}$ = [–37.03%, –15.86%], see Table B3). However, we no longer find a statistically significant moderation effect of the unemployment rate ($\beta = 0.0016$, $SE = 0.0018$, $p = 0.4457$), suggesting that the mitigating effect of a moderately loose labour market on the lower hiring chances for the long-term unemployed is not robust to our outlier-adjusted analysis (see Table B4).

Third, we estimated treatment effects that correct for the bias introduced by small-study effects. Typically, small studies are confronted with larger-than-normal statistical variance. Treatment effects generally need to be large to publish small sample studies in scientific outlets because only large effects tend to be statistically significant (Borenstein et al., 2011; Harrer et al., 2021). Therefore, small studies are more likely to show publication bias. We relied on two small-study effect methods to adjust for said bias, i.e. PET-PEESE and limit meta-analysis (Rücker et al., 2011; Stanley & Doucouliagos, 2014). PET-PEESE combines the precision-effect test (PET) and the precision-effect estimate with standard error (PEESE). Small-study bias-correction methods attempt to reduce the influence of studies with large standard errors by controlling for within-study variance. While these correction strategies should provide less biased estimates, they are known to perform worse if relatively few studies are included in the analysis or when between-study heterogeneity is high (Rücker et al., 2011; Stanley, 2017).

The PET-PEESE analyses suggest small effects drive the results of the 1 to 6 months and 19 to 36 months unemployment specifications (see Table B6). Following Harrer et al. (2011), we consider the results of the PEESE analysis if the *p*-value related to the PET estimate is smaller than 0.10. This rule holds for all estimates but the 19 to 36 months unemployment specification. On the one hand, the 1 to 6 months unemployment specification estimate in the PEESE analysis becomes statistically indistinguishable from zero ($CI_{95\%}$ = [–0.049, 0.0631]). On the other hand, the 19 to 36 months unemployment specification estimate in the PET analysis is heavily corrected and even appears to turn positive ($CI_{95\%}$ = [–0.0506, 0.2075]). The (direction of the) pooled effects concerning both specifications rely substantially on the treatment effects reported by Kroft et al. (2013), who computed unemployed–employed callback ratios for many different employment durations, resulting in various small-sample estimates. Because their overarching study is high-powered and their findings appear robust, we do not believe these overcorrections are justified.

Similarly, the limit meta-analysis appears to overcorrect the original pooled effects (see Table B7), presumably due to high between-study heterogeneity (see Table A3). Specifically for the 1 to 6 months and 19 to 36 months unemployment specifications, we also suspect the influence of small-sample estimates from Kroft et al. (2013) play a role in the overcorrection. Only the unemployed–employed callback ratio concerning the 13 to 18 months unemployment specification remains negative but is statistically insignificant from zero ($CI_{95\%}$ = [–0.3928, 0.1062]). For the above reasons, we place limited weight on the small-study corrected estimates from the PET-PEESE and limit meta-analyses.

Fourth, we evaluated the sensitivity of the pooled effects to effect selection based on *p*-values. In other words, we tested whether there is an overrepresentation of effects with just-significant or marginally significant *p*-values in the dataset that could indicate selection bias. To this end, we used *p*-curve and three-parameter selection models (Harrer et al., 2021). The *p*-curve method considers the distribution of *p*-values of the included treatment effects to identify bias (Simonsohn et al., 2014). When the null hypothesis is true, we expect the *p*-values to be uniformly distributed; when the null hypothesis can be rejected, we expect the *p*-value distribution to be right-skewed. The three-parameter selection models compare the likelihood of statistically significant versus non-significant effects to be selected for publication. An under-selection of non-significant effects would suggest selection bias.

The results of these sensitivity methods indicate little bias based on *p*-value selection. For each specification, the *p*-curve statistics show the presence of evidential value without indication for the absence or inadequacy of evidential value (see Table B8). All right-skewness tests produce statistically significant results, while the flatness tests indicate statistical insignificance, as expected under robustly significant results. Visually, however, we see increased reporting of *p*-values between 0.04 and 0.05 in the studies included in the 1 to 6 months unemployment specification (see Figure B6). Three out of four treatment effects included in this range are derived from Kroft et al. (2013), who reported multiple small-sample estimates (in an overall large-sample study) for a broad spectrum of unemployment durations. The pooled effects of the three-parameter selection models (see Table B9) are similar to those of the main meta-analysis (see Table A3). The likelihood ratio tests for the selection model parameters indicate that the results of our meta-analysis were not substantially influenced by a lower selection likelihood of non-significant results at the 5% or 10% statistical significance level thresholds.

Finally, fifth, we conducted a hierarchical Bayesian meta-analysis as a Bayesian alternative to the frequentist's analyses producing the original estimates. The Bayesian approach allows us to make claims about the probability of the bounds of the true effect given the observed data and (assumed) prior information about the distribution of the pooled effects and between-study heterogeneity. Following Harrer et al. (2021) and Irsova et al. (2023), we used weakly informative priors for the distributions of the pooled effects and between-study heterogeneity. We chose a normal distribution with mean 0 and variance 1 for the natural log of the callback ratios (i.e. $\ln CB \sim N(0,1)$), most log callback ratios are around zero with typical variations spanning about one log unit. Next, we chose a half-Cauchy distribution for the between-study variance, as between-study heterogeneity is non-negative and usually near zero (Harrer et al., 2021; Williams et al., 2018). The location parameter was set at 0, and the scaling parameter was set at 0.5 (i.e.). We fit Bayesian meta-analysis models with four Markov chains, 5000 iterations per chain, and the (sub-)study cluster variable as the random intercept. Our results are robust to using a lower, more precise scaling parameter of 0.3—as suggested by Williams et al. (2018)—for the between-study variance distribution.

The hierarchical Bayesian meta-analysis estimates are similar to our frequentist's analysis findings (see Table B10). We produced estimates for both the unadjusted and the outlier-adjusted specifications. The analysis affirms a likely positive impact of unemployment for the short-term unemployed (less than six months); the true effect falls in [−0.49%, 19.91%] (or [0.43%, 17.32%] for the outlier-adjusted specification) with a 90% probability. The long-term unemployed (more than 18 months) are the worst off; the callback penalty falls in [−47.34%, −2.04%] (or [−47.61%, −6.23%] for the outlier-adjusted specification) with a 90% probability. The observed density distributions and the posterior predictive density distributions of the hierarchical Bayesian meta-analysis models are visualised in Figure B7.

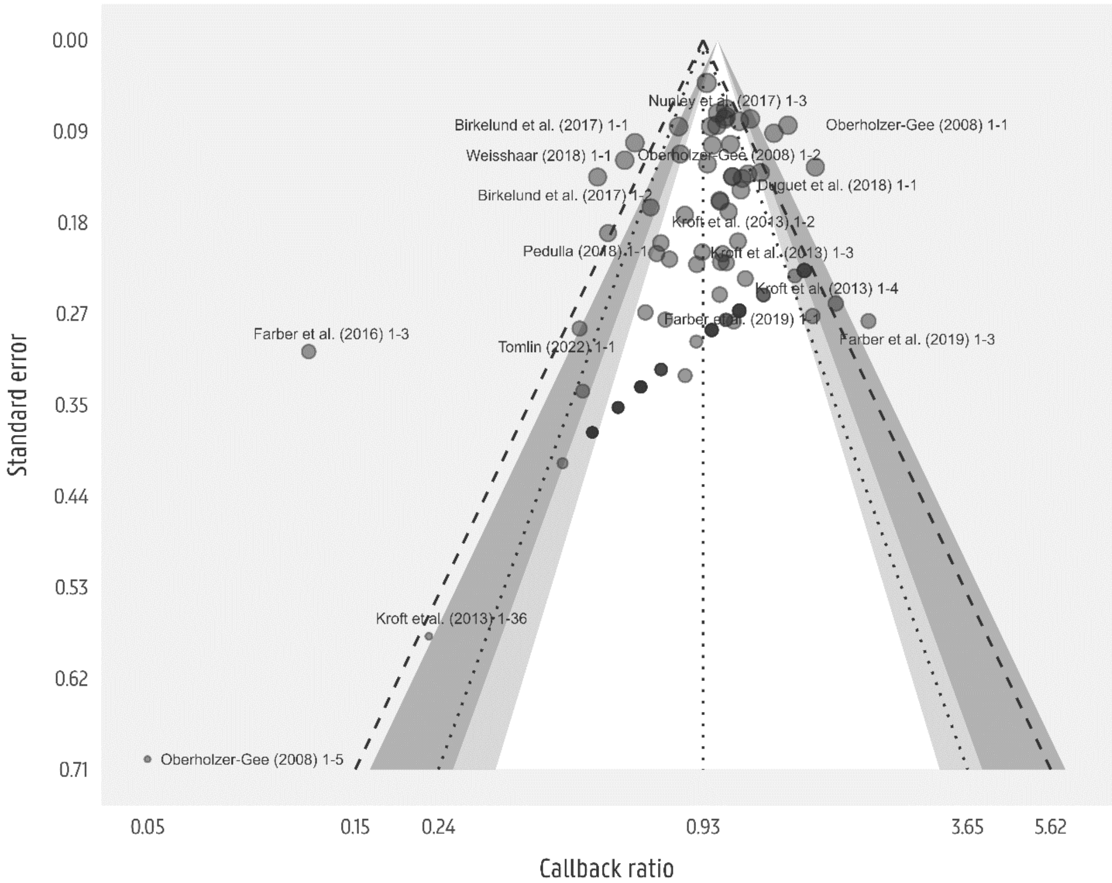
B.II. Funnel plot asymmetry

Table B1. Egger's and Peters' bias statistics for funnel plot asymmetry

Specification	Estimate	SE	$CI_{95\%}$	*t*	*p*
Egger					
Overall	–0.9559	0.3468	[–1.6356, –0.2762]	–2.76	0.007
1 to 6 months	–0.0480	0.6929	[–1.4061, 1.3100]	–0.07	0.945
7 to 12 months	–1.4928	0.6148	[–2.6977, –0.2878]	–2.43	0.025
13 to 18 months	0.2913	1.2701	[–2.1981, 2.7807]	0.23	0.823
19 to 36 months	–1.2219	0.3644	[–1.9362, –0.5076]	–3.35	0.003
Peters					
Overall	–112.3443	78.9136	[–267.0149, 42.3263]	–1.42	0.158
1 to 6 months	42.9592	118.0471	[–188.4130, 274.3315]	0.36	0.718
7 to 12 months	99.0796	125.2222	[–146.3560, 344.5151]	0.79	0.439
13 to 18 months	–23.0195	149.5654	[–316.1678, 270.1288]	–0.15	0.881
19 to 36 months	–382.2304	163.5733	[–702.8341, –61.6268]	–2.34	0.030

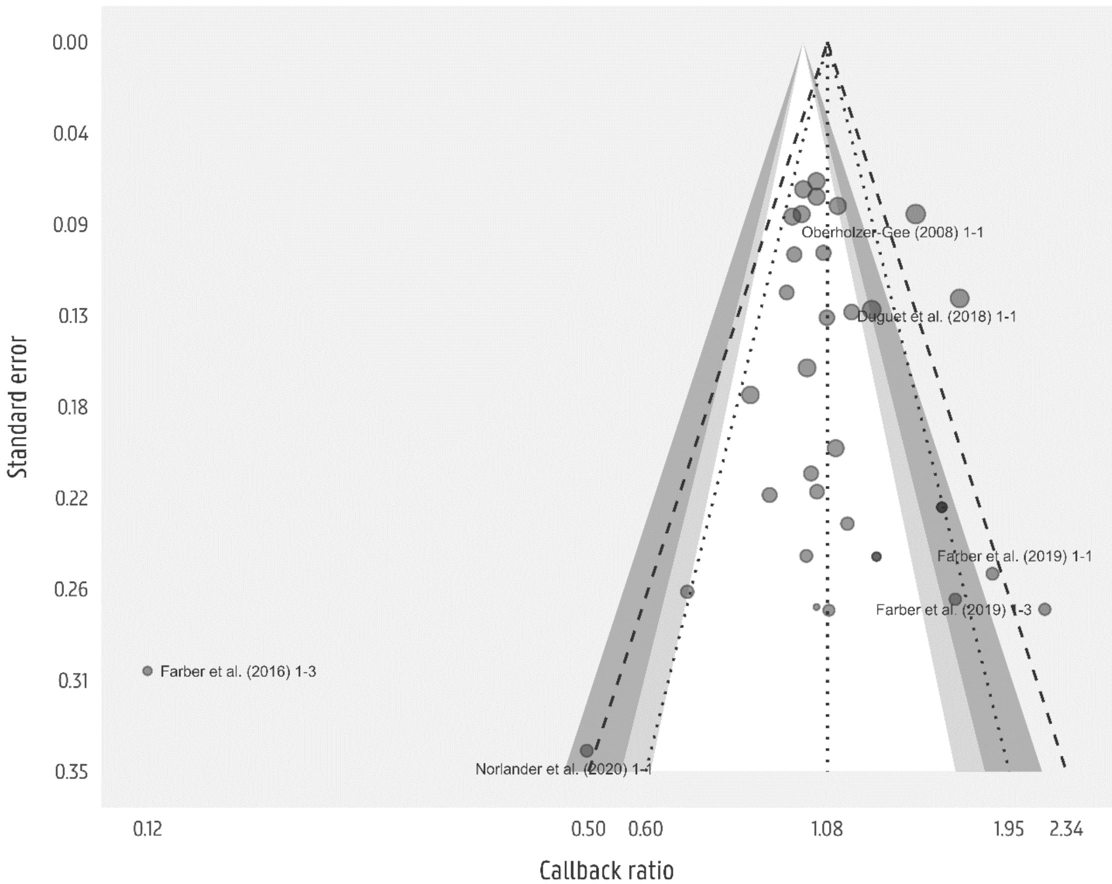
Notes. Bias statistics are derived from two asymmetry tests. Egger's bias statistic is based on the generic test for funnel plot asymmetry; Peters' bias statistic is based on a binary effects adaptation of Egger's test.

Figure B1. Contour-enhanced funnel plot: all treatment effects



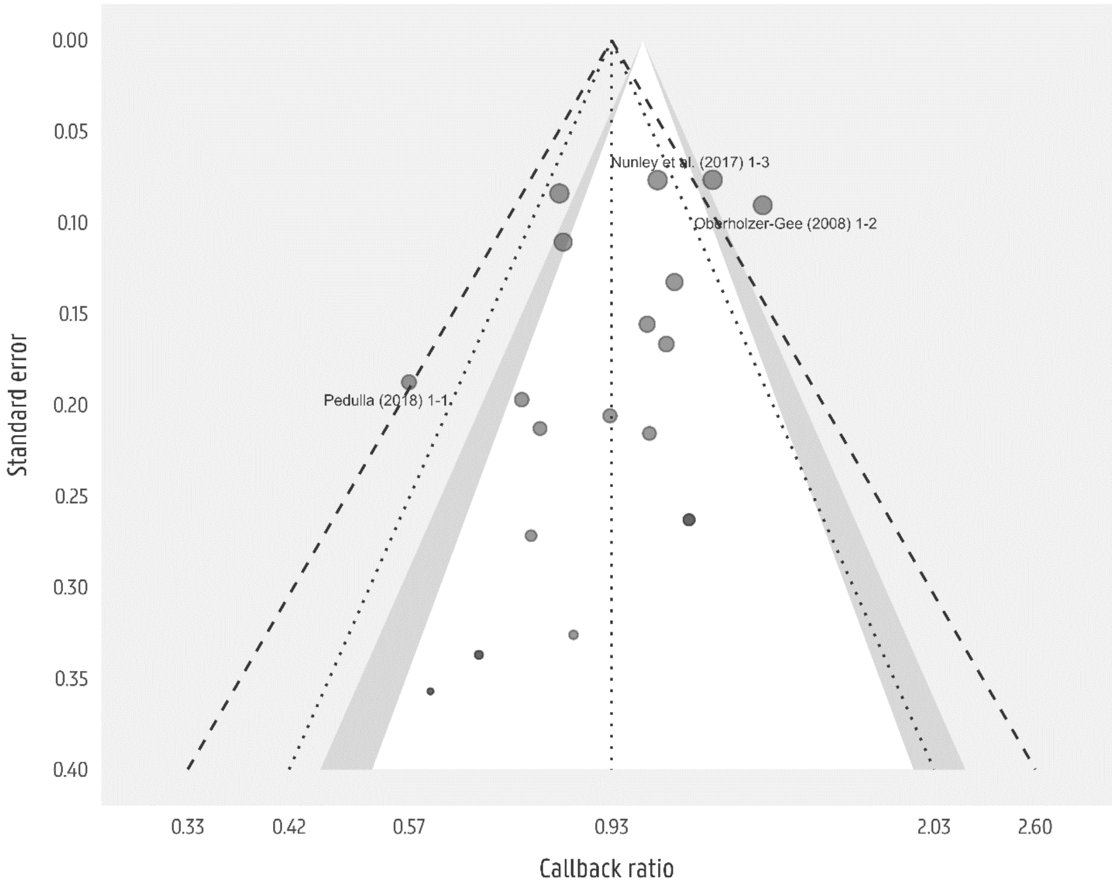
Notes. The x-axis is log-transformed. Semi-transparent points represent the callback ratios of the original effects. The point size represents the respective meta-analytic random effect weight. The vertical dotted line represents the unadjusted pooled callback ratio. Diagonal dotted (dashed) lines depict the 95% (99%) confidence intervals around this pooled callback ratio. The white (medium grey) ((dark grey)) triangular shapes depict the 90% (95%) ((99%)) confidence intervals around the null effect (i.e. a callback ratio of 1). Labelled points fall outside the 95% confidence intervals of the pooled callback ratio and the null effect (at a given standard error). Label names are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset.

Figure B2. Contour-enhanced funnel plot: 1 to 6 months of unemployment



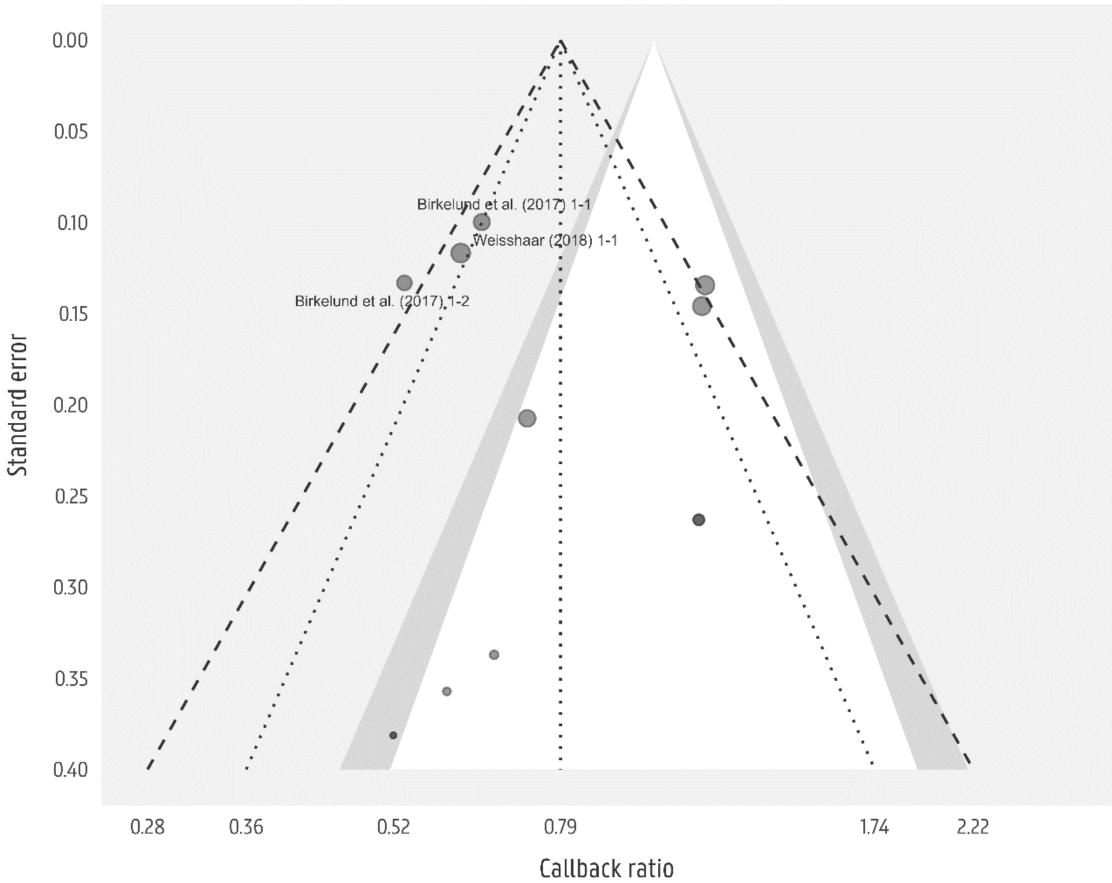
Notes. The x-axis is log-transformed. Semi-transparent points represent the callback ratios of the original effects. The point size represents the respective meta-analytic random effect weight. The vertical dotted line represents the unadjusted pooled callback ratio. Diagonal dotted (dashed) lines depict the 95% (99%) confidence intervals around this pooled callback ratio. The white (medium grey) ((dark grey)) triangular shapes depict the 90% (95%) ((99%)) confidence intervals around the null effect (i.e. a callback ratio of 1). Labelled points fall outside the 95% confidence intervals of the pooled callback ratio and the null effect (at a given standard error). Label names are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset.

Figure B3. Contour-enhanced funnel plot: 7 to 12 months of unemployment



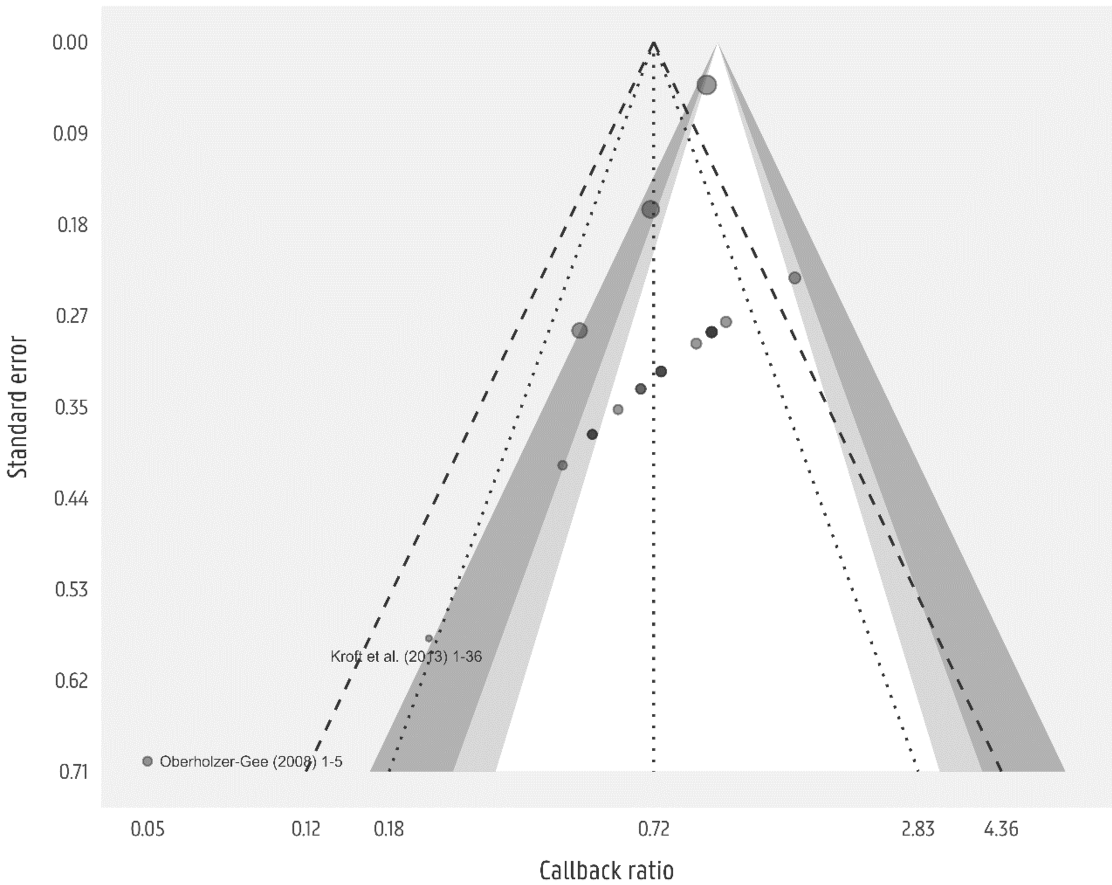
Notes. The x-axis is log-transformed. Semi-transparent points represent the callback ratios of the original effects. The point size represents the respective meta-analytic random effect weight. The vertical dotted line represents the unadjusted pooled callback ratio. Diagonal dotted (dashed) lines depict the 95% (99%) confidence intervals around this pooled callback ratio. The white (medium grey) ((dark grey)) triangular shapes depict the 90% (95%) ((99%)) confidence intervals around the null effect (i.e. a callback ratio of 1). Labelled points fall outside the 95% confidence intervals of the pooled callback ratio and the null effect (at a given standard error). Label names are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset.

Figure B4. Contour-enhanced funnel plot: 13 to 18 months of unemployment



Notes. The x-axis is log-transformed. Semi-transparent points represent the callback ratios of the original effects. The point size represents the respective meta-analytic random effect weight. The vertical dotted line represents the unadjusted pooled callback ratio. Diagonal dotted (dashed) lines depict the 95% (99%) confidence intervals around this pooled callback ratio. The white (medium grey) ((dark grey)) triangular shapes depict the 90% (95%) ((99%)) confidence intervals around the null effect (i.e. a callback ratio of 1). Labelled points fall outside the 95% confidence intervals of the pooled callback ratio and the null effect (at a given standard error). Label names are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset.

Figure B5. Contour-enhanced funnel plot: 19 to 36 months of unemployment



Notes. The x-axis is log-transformed. Semi-transparent points represent the callback ratios of the original effects. The point size represents the respective meta-analytic random effect weight. The vertical dotted line represents the unadjusted pooled callback ratio. Diagonal dotted (dashed) lines depict the 95% (99%) confidence intervals around this pooled callback ratio. The white (medium grey) ((dark grey)) triangular shapes depict the 90% (95%) ((99%)) confidence intervals around the null effect (i.e. a callback ratio of 1). Labelled points fall outside the 95% confidence intervals of the pooled callback ratio and the null effect (at a given standard error). Label names are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset.

B.III. Publication bias-adjusted estimates

Table B2. Outliers by estimation specification

Specification	Outliers
Overall	Birkelund et al. (2017) 1-1, Birkelund et al. (2017) 1-2, Duguet et al. (2018) 1-1, Farber et al. (2016) 1-3, Farber et al. (2019) 1-1, Farber et al. (2019) 1-3, Kroft et al. (2013) 1-36, Nunley et al. (2017) 1-3, Oberholzer-Gee (2008) 1-1, Oberholzer-Gee (2008) 1-2, Oberholzer-Gee (2008) 1-5, Pedulla 1-1, Tomlin 1-1, Weisshaar (2018) 1-1
1 to 6 months	Duguet et al. (2018) 1-1, Farber et al. (2016) 1-3, Oberholzer-Gee (2019) 1-3
7 to 12 months	Oberholzer-Gee (2008) 1-2, Pedulla (2018) 1-1
13 to 18 months	N/A
19 to 36 months	Kristal et al. (2023) 1-2, Kroft et al. (2013) 1-31, Oberholzer-Gee (2008) 1-5

Notes. Acronyms used: N/A (not applicable). Outliers are effects for which the upper (lower) bound of the 95% confidence interval of the unemployed–employed callback ratio is lower (higher) than the lower (upper) bound of the confidence interval of the pooled random effects callback ratio. Outlier names are defined by their unique identifier in the underlying dataset, which is composed of the authors' names, the year of publication, and the (sub-)study and effect sequence in the dataset.

Table B3. Meta-analytic estimates of unemployed–employed differences in positive callbacks (adjusted for outliers)

Specification	Estimate	SE	$CI_{95\%}$	t	p	I^2
Overall	–0.0290	0.0203	[–0.0686, 0.0122]	–1.41	0.163	31.67%
1 to 6 months	0.1001	0.0552	[–0.0067, 0.2184]	1.90	0.066	41.14%
7 to 12 months	–0.0656	0.0447	[–0.1549, 0.0332]	–1.42	0.173	29.31%
13 to 18 months	–0.2078	0.0928	[–0.3878, 0.0251]	–1.99	0.072	68.95%
19 to 36 months	–0.2721	0.0502	[–0.3703, –0.1586]	–4.60	<0.001	0.00%

Notes. Acronyms used: SE (standard error), CI (confidence interval). Estimates are differences in positive callbacks between the unemployed and employed. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. The clustering of the effects across (sub-)studies was modelled through three-level models. I^2-values around 25%, 50%, or 75% indicate a low, moderate, or high proportion of residual heterogeneity relative to the amount of unaccounted variability, respectively.

Table B4. Weighted least squares meta-regression of callback differences between the unemployed and employed on unemployment duration, unemployment rates, and other study-level covariates (adjusted for outliers)

Term	(1)	(2)	(3)
Intercept	0.0579 (0.0356)	−170.8460 (116.4623)	−180.0373 (108.6374)
Unemployment duration *(in months)*	−0.0089ˣ (0.0038)	−0.0164* (0.0028)	−0.0286 (0.0148)
Unemployment rate *(in per cent × 100)*	–	0.2328 (0.1139)	0.2303 (0.1020)
Response type: Interview invitation (ref.)			
Response type: Positive reaction	–	0.0239 (0.1405)	0.0421 (0.1446)
Region: Northern America (ref.)			
Region: Western Europe	–	0.2062 (0.3933)	0.3082 (0.3712)
Region: Northern Europe	–	0.8002 (0.5160)	0.8542 (0.4877)
Region: Central Europe	–	1.7394 (0.9979)	1.9212 (0.9862)
Region: Asia	–	0.5197 (0.3077)	0.5353 (0.3014)
Year	–	0.0837 (0.0574)	0.0882 (0.0535)
Gender: Male (ref.)			
Gender: Female	–	1.1002 (0.3598)	1.2416ˣ (0.4026)
Gender: Mixed	–	0.7124 (0.4187)	0.8099 (0.4073)
Age: Prime-aged (31–50) (ref.)			
Age: Young (21–30)	–	0.0790 (0.3005)	0.1213 (0.2863)
Age: Old (51+)	–	0.0529 (0.0486)	0.0512 (0.0418)
Unemployment duration × Unemployment rate	–	–	0.0016 (0.0018)
Marginal mean	−0.0529* (0.0224)	−0.0615* (0.0271)	−0.0683* (0.0284)
AIC	2.79	0.64	1.97
BIC	9.78	33.02	36.26
I^2	19.33%	0.00%	0.00%

Notes. Abbreviations and acronyms used: ref. (reference category), AIC (Akaike information criterion), BIC (Bayesian information criterion). Statistics are coefficient estimates with standard errors between parentheses. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. Standard errors were clustered at the (sub-)study level and corrected using a small sample adjustment. Marginal means are the effects of unemployment on positive callbacks, calculated at the average values of the continuous predictor variables and the reference categories of the categorical predictor variables. I^2-values around 25%, 50%, or 75% indicate a low, moderate, or high proportion of residual heterogeneity relative to the amount of unaccounted variability, respectively. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$, ˣ $p < 0.10$.

Table B5. Weighted least squares meta-regression of callback differences between the unemployed and employed on study-level covariates (by unemployment duration; adjusted for outliers)

Term	1 to 6 months	7 to 12 months	13 to 18 months	19 to 36 months
Intercept	−30.7184 (86.2813)	−1.1414ˣ (0.2337)	N/A	−0.3875ˣ (0.0407)
Unemployment rate *(in per cent × 100)*	0.0943 (0.1895)	0.1169 (0.0288)	N/A	0.0128 (0.0131)
Response type: Interview invitation (ref.)			N/A	
Response type: Positive reaction	−0.2638 (0.2677)	0.4600ˣ (0.1090)	N/A	−0.3782* (0.0105)
Region: America (ref.)			N/A	
Region: Europe	−0.0436 (0.5453)	0.3879 (0.1531)	N/A	–
Region: Asia	−0.7227 (0.3288)	0.5744ˣ (0.1351)	N/A	–
Year	0.0148 (0.0423)	–	N/A	–
Gender: Male (ref.)			N/A	
Gender: Female	0.7355 (1.4757)	–	N/A	–
Gender: Mixed	0.0192 (0.8346)	–	N/A	–
Age: Prime-aged (31–50) (ref.)			N/A	
Age: Young (21–30)	0.4725 (0.4121)	−0.0503 (0.0773)	N/A	–
Age: Old (51+)	0.0562 (0.1153)	−0.1047 (0.0472)	N/A	–
Marginal mean	0.1673*** (0.0390)	−0.1046ˣ (0.0461)	N/A	−0.2631*** (0.0509)
I^2	0.00%	23.12%	N/A	0.00%

Notes. Abbreviations and acronyms used: ref. (reference category), N/A (not applicable). Statistics are coefficient estimates with standard errors between parentheses. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. Some variables were dropped from the regression models due to multicollinearity. Standard errors were clustered at the (sub-)study level and corrected using a small sample adjustment. Marginal means are the effects of unemployment on positive callbacks, calculated at the average values of the continuous predictor variables and the reference categories of the categorical predictor variables. I^2-values around 25%, 50%, or 75% indicate a low, moderate, or high proportion of residual heterogeneity relative to the amount of unaccounted variability, respectively. In the absence of outliers for the models related to the 13 to 18 months employment specification, no outlier- or dual-adjusted estimates are reported. *** p < 0.001; ** p < 0.01; * p < 0.05, ˣ p < 0.10.

Table B6. PET-PEESE meta-analytic estimates of unemployed–employed differences in positive callbacks

Specification	Estimate	SE	$CI_{95\%}$	*z*	*p*
PET					
Overall	–0.0650	0.0331	[–0.1298, –0.0002]	–1.97	0.049
1 to 6 months	–0.1144	0.0505	[–0.2133, –0.0155]	–2.27	0.023
7 to 12 months	–0.1807	0.0684	[–0.3148, –0.0467]	–2.64	0.008
13 to 18 months	–0.2525	0.1407	[–0.5282, 0.0232]	–1.80	0.073
19 to 36 months	0.0784	0.0658	[–0.0506, 0.2075]	1.19	0.233
PEESE					
Overall	–0.0374	0.0195	[–0.0757, 0.0008]	–1.92	0.055
1 to 6 months	0.0071	0.0286	[–0.0490, 0.0631]	0.25	0.805
7 to 12 months	–0.0951	0.0404	[–0.1744, –0.0159]	–2.35	0.019
13 to 18 months	–0.2733	0.0713	[–0.4131, –0.1334]	–3.83	<0.001
19 to 36 months	–0.0423	0.0417	[–0.1240, 0.0394]	–1.02	0.310

Notes. Acronyms used: SE (standard error), CI (confidence interval), PET (precision-effect test), PEESE (precision-effect estimate with standard error). Estimates are differences in positive callbacks between the unemployed and employed. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. The clustering of the effects across (sub-)studies was modelled through three-level models.

Table B7. Limit meta-analytic estimates of unemployed–employed differences in positive callbacks

Specification	Estimate	SE	$CI_{95\%}$	*z*	*p*
Overall	0.1660	0.0676	[0.0408, 0.3062]	2.65	0.008
1 to 6 months	0.1850	0.1136	[–0.0181, 0.4300]	1.77	0.077
7 to 12 months	0.0755	0.0932	[–0.0925, 0.2745]	0.84	0.401
13 to 18 months	–0.1804	0.1254	[–0.3928, 0.1062]	–1.30	0.193
19 to 36 months	0.2068	0.3265	[–0.2899, 1.0508]	0.69	0.487

Notes. Acronyms used: SE (standard error), CI (confidence interval). Estimates are differences in positive callbacks between the unemployed and employed. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. The clustering of the effects across (sub-)studies was modelled through three-level models.

Table B8. *P*-curve statistics: right-skewness and flatness test statistics

	Right-skewness test					Flatness test					Evidential value	
Specification	$p_{binomial}$	z_{full}	p_{full}	z_{half}	p_{half}	$p_{binomial}$	z_{full}	p_{full}	z_{half}	p_{half}	Present?	Absent?
Overall	0.058	–6.52	<0.001	–7.79	<0.001	0.528	3.18	0.999	7.55	1.000	Yes	No
1 to 6 months	0.500	–3.40	<0.001	–5.88	<0.001	0.239	1.34	0.910	5.54	1.000	Yes	No
7 to 12 months	0.312	–1.99	0.023	–1.49	0.068	0.740	0.56	0.713	2.01	0.978	Yes	No
13 to 18 months	0.125	–6.02	<0.001	–5.69	<0.001	1.000	4.27	1.000	4.73	1.000	Yes	No
19 to 36 months	0.312	–2.29	0.011	–2.06	0.020	0.740	0.84	0.800	2.43	0.992	Yes	No

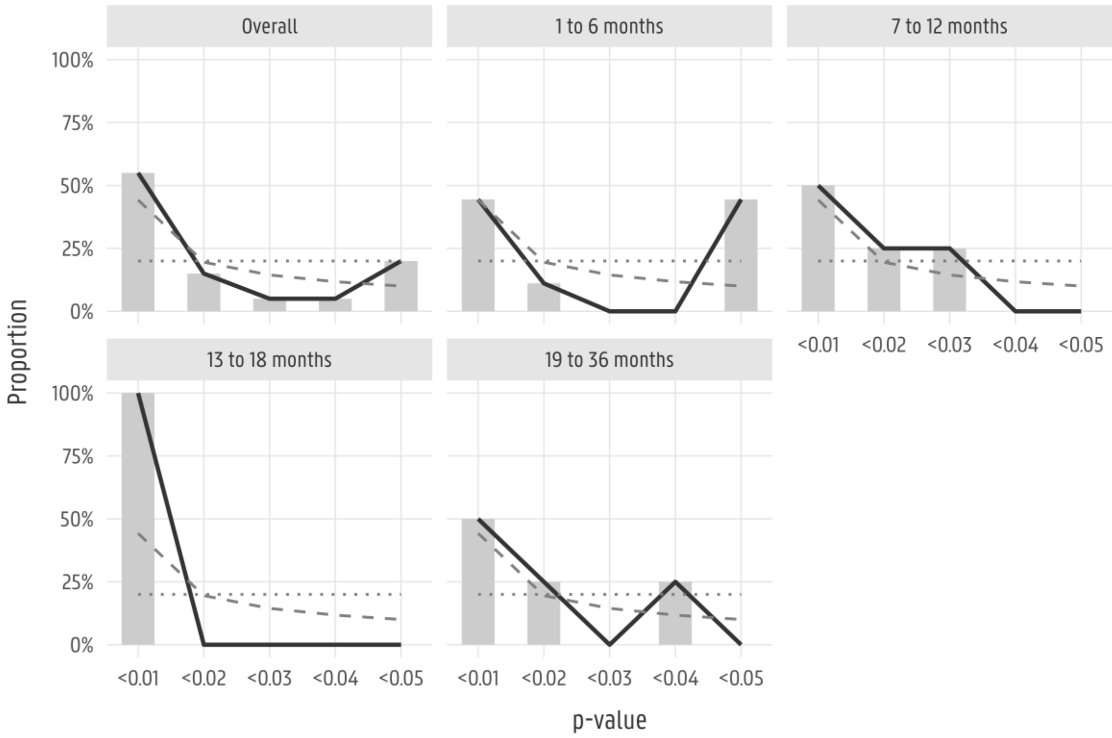
Notes. Estimates are differences in positive callbacks between the unemployed and employed. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. The clustering of the effects across (sub-)studies was modelled through three-level models. Following Simonsohn et al. (2015), we assume evidence is present if the *p*-value for the half curve in the right-skewness test is less than 0.05 or if the *p*-values for the full and half curves are less than 0.10. In addition, we assume evidence is absent or inadequate if the *p*-value for the full curve in the flatness test is less than 0.05 or if the *p*-values for the binomial test, full curve, and half curve are less than 0.10.

Table B9. Three-parameter selection model meta-analytic estimates of unemployed–employed differences in positive callbacks

	Selection model				Likelihood ratio test	
Specification	Estimate	SE	$CI_{95\%}$	*p*	χ^2	*p*
Two-sided alpha: 0.05						
Overall	–0.0866	0.0455	[–0.1757, 0.0026]	0.057	0.07	0.787
1 to 6 months	0.1451	0.1091	[–0.0687, 0.3589]	0.183	0.69	0.406
7 to 12 months	–0.1034	0.0573	[–0.2158, 0.0089]	0.071	1.43	0.232
13 to 18 months	–0.2446	0.1133	[–0.4665, –0.0226]	0.031	0.60	0.437
19 to 36 months	–0.2767	0.0899	[–0.4530, –0.1004]	0.002	0.56	0.453
Two-sided alpha: 0.10						
Overall	–0.0919	0.0455	[–0.1837, –0.0001]	0.050	0.18	0.673
1 to 6 months	0.1183	0.1091	[–0.0772, 0.3138]	0.236	0.25	0.615
7 to 12 months	–0.0895	0.0573	[–0.2103, 0.0312]	0.146	0.28	0.597
13 to 18 months	–0.2274	0.1133	[–0.4748, 0.0201]	0.072	0.99	0.319
19 to 36 months	–0.2953	0.0899	[–0.4642, –0.1264]	0.001	0.76	0.385

Notes. Acronyms used: SE (standard error), CI (confidence interval). Estimates are differences in positive callbacks between the unemployed and employed. Negative (positive) coefficients signify less (more) positive callbacks for the unemployed. The clustering of the effects across (sub-)studies was modelled through three-level models. χ^2-values are likelihood ratio test statistics indicative of the lower selection likelihood of non-significant results.

Figure B6. *P*-curves



Notes. The x-axis groups *p*-values (less than or equal to 0.05) of the included treatment effects in bins. Bars and dark grey lines represent the observed proportions of *p*-values (amongst those less than or equal to 0.05) per *p*-value bin. Dashed (dotted) lines depict the expected *p*-value distribution assuming 33% power (no effect).

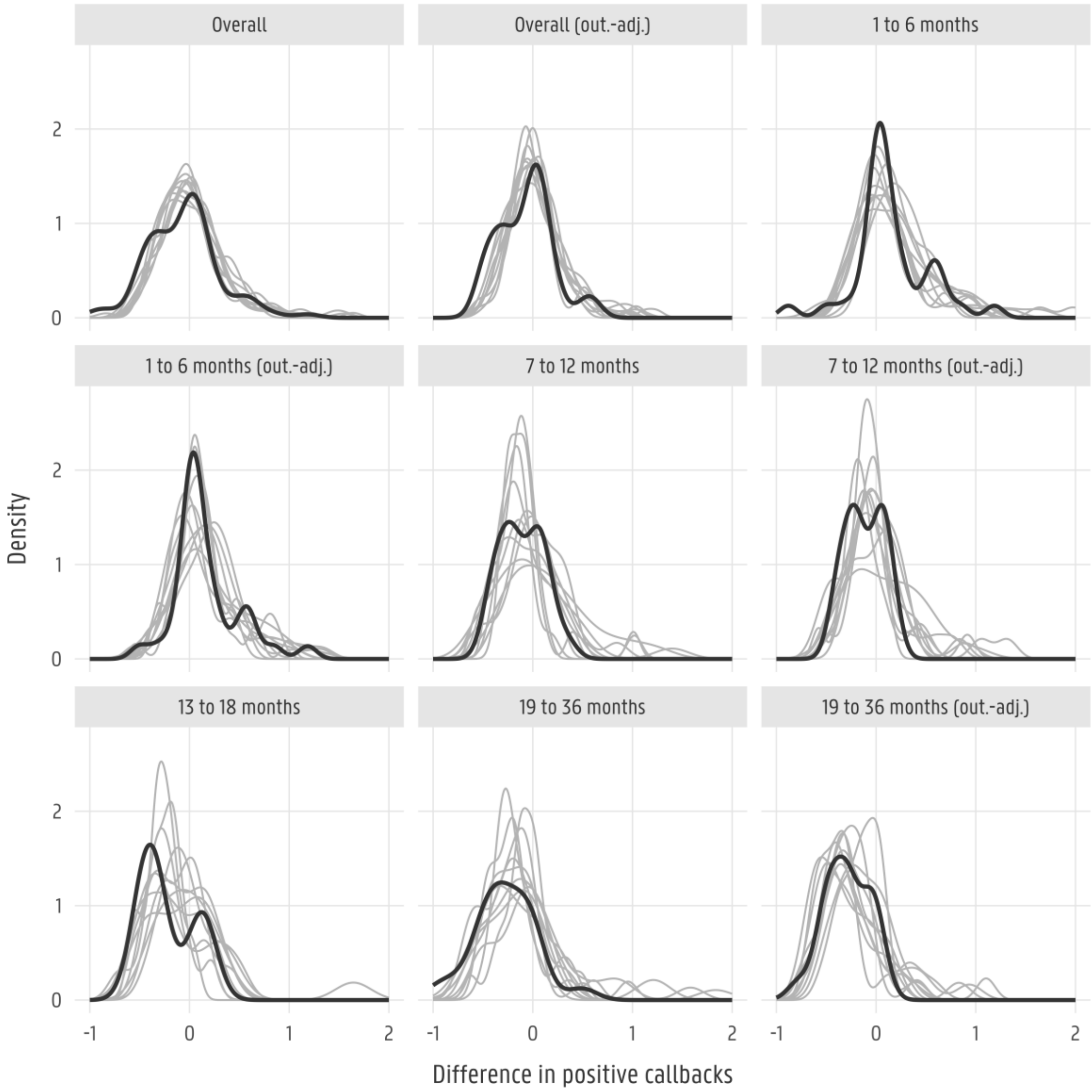
B.IV. Hierarchical Bayesian meta-analysis

Table B10. Hierarchical Bayesian meta-analytic estimates of unemployed–employed differences in positive callbacks

Specification	Estimate	EE	$CI_{90\%}$	$CI_{95\%}$
Unadjusted statistics				
Overall	−0.0616	0.0492	[−0.1397, 0.0198]	[−0.1543, 0.0355]
1 to 6 months	0.0920	0.0621	[−0.0049, 0.1991]	[−0.0260, 0.2231]
7 to 12 months	−0.0692	0.0539	[−0.1557, 0.0202]	[−0.1739, 0.0439]
13 to 18 months	−0.1946	0.1249	[−0.3613, 0.0472]	[−0.3958, 0.1293]
19 to 36 months	−0.2469	0.1413	[−0.4734, −0.0204]	[−0.5090, 0.0656]
Outlier-adjusted statistics				
Overall	−0.0276	0.0245	[−0.0680, 0.0126]	[−0.0762, 0.0212]
1 to 6 months	0.0839	0.0518	[0.0043, 0.1732]	[−0.0150, 0.1962]
7 to 12 months	−0.0590	0.0509	[−0.1405, 0.0272]	[−0.1576, 0.0470]
13 to 18 months	N/A	N/A	N/A	N/A
19 to 36 months	−0.2889	0.1240	[−0.4761, −0.0623]	[−0.5100, 0.0081]

Notes. Acronyms used: EE (estimation error), CI (credible interval), N/A (not applicable). The prior distribution of the log callback ratios follows a normal distribution; the prior distribution of the between-study heterogeneity follows a half-Cauchy distribution. Estimates are back-transformed differences in positive callbacks ($\Delta_{Callback} = e^{\hat{\beta}} - 1$). Estimation errors were also back-transformed using the delta method ($EE = e^{\hat{\beta}} * EE_{ln}$).

Figure B7. Observed and posterior predictive distributions of differences in positive callbacks



Notes. Abbreviations used: out.-adj. (outlier-adjusted). Dark grey lines represent the observed distributions in the underlying data. Light grey lines are 95% posterior predictive distributions.