Job Search, Overoptimism and Statistical Profiling: Can Information Provision Improve Job Search Outcomes?*

Nikolaj A. Harmon† **Robert Mahlstedt**‡
Mette Rasmussen§

February 15, 2026

Abstract

We estimate the causal effect of a large-scale information provision policy targeting unemployed workers in Denmark. The policy aimed to correct optimistic beliefs about reemployment prospects. Identification relies on a regression discontinuity design leveraging age discontinuities in the policy's statistical profiling tool. When unemployment insurance recipients are informed that they are at high risk of long-term unemployment, their likelihood of exiting unemployment increases significantly. These exits reflect two distinct mechanisms however: for some jobseekers the information treatment encourages faster job finding, while for others, it discourages job search altogether and shifts them from unemployment into passive support schemes.

Keywords: Job search, long-term unemployment, information provision, statistical profiling, regression discontinuity

JEL codes: J68, D83, C93

*This study was conducted as a part of the research initiative 'Behavioral Economics and Labor Market Performance'. Financial support by the Danish Ministry of Employment and the Independent Research Fund Denmark is gratefully acknowledged. We thank Gordon Dahl, Bas van der Klaauw, Søren Leth-Petersen, Sally Sadoff, and seminar and conference participants at the University of Copenhagen and the 2024 EALE meeting for valuable comments and suggestions. We are grateful to the Danish Agency for Labor Market and Recruitment for data access.

†University of Copenhagen; `nikolaj.harmon@econ.ku.dk`

‡University of Copenhagen and IZA; `robert.mahlstedt@econ.ku.dk`

§ROCKWOOL Foundation and University of Copenhagen; `mra@rff.dk`

1 Introduction

Unemployed workers often overestimate their reemployment prospects (Mueller et al., 2021; Mueller and Spinnewijn, 2023; Spinnewijn, 2015). Since these misperceptions can distort job search and prolong unemployment, a natural policy response is to implement information strategies designed to correct these beliefs. However, it is theoretically unclear whether such policies will improve employment outcomes. When jobseekers learn that job finding may be harder than expected, this could encourage them to intensify their search, thereby improving reemployment prospects. Alternatively, it could discourage them and lead them to give up on job search altogether.

We study the causal effects of a large-scale information provision policy that informs unemployment insurance (UI) recipients in Denmark about their job-finding prospects. Consistent with international evidence, Danish UI recipients typically overestimate their reemployment chances. Between 2015 and 2017, around 40% of new UI recipients were statistically profiled to predict their risk of long-term unemployment. Those assessed as high risk received a message highlighting their poor reemployment outlook. The policy sought to correct overly optimistic beliefs and prompt faster exits from unemployment by altering search behavior.

We analyze the causal effect of this information treatment on unemployment exits using a regression discontinuity (RD) design that exploits how the profiling tool assigns treatment. The tool relies on a decision tree in which age is a key splitting variable. This means that many jobseekers receive the treatment based solely on their exact age at profiling, with cutoffs at ages 28, 29, 54, and 56. Comparing individuals just above and below these thresholds provides quasi-random variation in treatment assignment.

Consistent with the policy's objective, the information treatment increases exits from unemployment: after six months, jobseekers just above the age cutoff—who receive the treatment—are seven percentage points less likely to receive unemployment benefits than those just below. These additional exits stem from two mechanisms: roughly 60% of these exits are driven by jobseekers encouraged to become self-supporting, typically by entering paid employment. The remaining 40% however reflect discouraged jobseekers who abandon search and transition into passive support programs that provide income without requiring active job search.

We interpret our findings using a simple job-search framework in which jobseekers choose their search effort or, potentially, exit unemployment by relying on passive support. The treatment (i) lowers expected job-finding probabilities for a given effort level and (ii) reduces the perceived returns to effort. These belief revisions generate heterogeneous behavioral responses: when the first effect dominates, some workers are *encouraged* to increase their search intensity, leading to higher job finding and earlier exits from unemployment. At the same time, the diminished reemployment prospects decrease the continuation value of receiving unemployment benefits and actively seeking employment. When passive support is relatively appealing, this can *discourage* job search altogether, leading jobseekers to exit unemployment and rely on passive support instead. In line with this notion, we find heterogeneous treatment effects along two dimensions. First, discouragement effects are particularly pronounced among workers for whom passive support schemes appear relatively attractive ex ante. Second, we document heterogeneity by baseline beliefs: effects are small and statistically insignificant for jobseekers already expecting a prolonged unemployment spell, consistent with the treatment providing little new information to this group.

Our study contributes to a growing literature on beliefs in job search. Prior research shows that unemployed workers are often overly optimistic about both their job-finding prospects (Spinnewijn, 2015; Mueller et al., 2021; Balleer et al., 2021) and their earnings potential (Caliendo et al., 2023; Conlon et al., 2018; Krueger and Mueller, 2016).[1] The key contribution of our paper is to causally estimate the effects of a policy aimed at correcting these beliefs on jobseekers' benefit receipt and employment outcomes.

Related field experiments have studied interventions that guide jobseekers toward specific jobs or occupations (Altmann et al., 2022; Le Barbanchon et al., 2023; Belot et al., 2019, 2022; Behaghel et al., 2022; Ben Dhia et al., 2022; Horton, 2017) or reduce search frictions and misperceptions more broadly through recruitment and mentoring programs (Abebe et al., 2025; Alfonsi et al., 2022; Bandiera et al., 2025), financial subsidies (Banerjee and Sequeira, 2023), and information brochures (Altmann et al., 2018). In contrast, our intervention directly targets a central variable in search models—beliefs about job-finding chances—providing novel evidence that downward belief revisions influence labor market outcomes through two distinct mechanisms: encouragement effects, where some jobseekers increase employment, and discouragement effects, where others withdraw from job search altogether.

Finally, our study adds to the literature on profiling tools in labor markets. Like in our setting, these tools use individuals' characteristics in statistical models to predict outcomes such as long-term unemployment risk. In most other contexts, however, profiling results are used to assign treatments—such as active labor market programs (ALMP) (see, e.g., Behncke et al., 2009; Black et al., 2007; Ernst et al., 2024; Frölich, 2008; Lechner and Smith, 2007; Staghøj et al., 2010)—without disclosing the profiling outcome to the individuals themselves.

[1] Other studies examine subjective beliefs in broader labor market contexts, including college graduates' earnings expectations and career choices (Cortés et al., 2023; Jiang and Zen, 2025), beliefs about skills in developing countries (Carranza et al., 2022; Kiss et al., 2023), and workers' perceptions of outside options (Jäger et al., 2024).

2 Empirical Setup

2.1 Unemployment and social security in Denmark

In Denmark, unemployment benefits operate under an opt-in system. Eligible jobseekers who have contributed for at least 12 months within the past three years can receive benefits for up to two years. The replacement rate is 90% of prior wages, capped at DKK18,866 (USD3,075) per month before taxes. Recipients must actively search for jobs and document their search efforts to maintain eligibility.

Unemployed individuals may also qualify for other public support programs. For instance, they can receive educational benefits when enrolling in secondary or tertiary education or vocational training. With a medical certificate, they may switch to sickness benefits for up to 22 weeks, while the parental leave system allows unemployed parents to care for children. Mothers receive 14 weeks of maternity leave after childbirth, and either parent can take up to 32 weeks of parental leave. Sickness and parental benefits are generally paid at the same level as unemployment benefits, and transitions between programs are common (Pedersen et al., 2012). Importantly, individuals on passive support programs are no longer subject to UI requirements, such as (1) applying for and documenting two jobs per week, (2) attending caseworker meetings, and (3) participating in ALMPs.

2.2 Profiling tool and information treatment

When becoming unemployed, all jobseekers register at the online portal of the Danish Employment Agency (*jobnet.dk*) to receive unemployment benefits. Upon registration, they are asked to answer a voluntary online survey.[2] The survey, advertised as preparation for the first caseworker meeting, includes 12 questions on education, search strategies, job preferences, and expectations. Around 40% of new unemployed respond, over 80% within the first two weeks.

During our study period (July 2015 - August 2017), the responses of jobseekers who completed the questionnaire were combined with administrative data and fed into a statistical profiling tool.[3] As detailed below, the profiling tool predicts whether individuals face a heightened risk of long-term unemployment, defined as being unemployed for 26 weeks or more. After completing the survey, jobseekers receive an on-screen message about their upcoming caseworker meeting, while those classified as "high risk" also see a note on their elevated risk:

> *Your characteristics indicate that reentering employment might be challenging for you. Our analysis reveals that persons with similar characteristics, who were previ-*

[2]For individuals without benefits in the past 180 days, the survey remains open twelve weeks after registration.

[3]We analyze the first profiling tool used from July 2015 to August 2017 (STAR, 2015); the model was later revised, and the risk assessment abolished in March 2022.

ously unemployed like you, encountered greater difficulties in securing new employment compared to other unemployment benefit recipients. We recommend discussing and planning the necessary steps with your caseworker to find a job quickly.

As we elaborate in Section 3, we expect this message to reduce jobseekers'—typically overoptimistic—beliefs about their reemployment prospects and thereby affect their search behavior. In our empirical analysis, we estimate the causal effect of being flagged as high risk and thus receiving this message. In addition, the profiling result is also shared with caseworkers. In principle, this could affect counseling, in which case our estimates would capture the combined effect of changes in caseworker behavior and jobseekers receiving the information. In practice, however, most caseworkers report that the assessment does not influence their advice (STAR, 2021). We return to this point and provide empirical evidence on caseworker responses in Section 4.5.

2.3 Discontinuities in the profiling tool

The profiling tool we study was developed independently by the Danish Employment Agency and is based on a machine-learning algorithm that predicts long-term unemployment. Using administrative data on unemployed workers from 2011–2013, the agency trained a single decision tree on observable characteristics such as age, ethnicity, employment history, industry, education, and prior benefit receipt. The tree predicted the probability of remaining unemployed for at least 26 weeks by splitting on variables that maximized differences in mean outcomes, thereby partitioning jobseekers into distinct nodes with associated risk scores. The 13 nodes with the highest predicted risk were classified as high risk and, during the study period, jobseekers in these high-risk nodes received the information treatment, while those in other nodes were classified as low risk and did not receive treatment. Depending on the time period, the model classifies about 15–20% of jobseekers as at high risk.

Although the aim of this paper is not to evaluate the performance of the profiling tool, we provide a basic assessment of its predictive performance. In Appendix Table A.1 we treat the output from the classification tool as a binary prediction for experiencing long-term unemployment and present a corresponding confusion matrix showing true and false positives and negatives, based on a period when no jobseekers received the information treatment and excluding the tool's training period. The tool exhibits predictive power: among individuals classified as high risk, 52% experience long-term unemployment, compared with 35% among those not classified as high risk. Viewed as a binary classifier, however, the tool's predictive performance is clearly limited. Only about one in four individuals who actually experience long-term unemployment are flagged as high risk by the tool; the true positive rate is 23%. Although the false positive rate is also relatively low at 12%, overall accuracy is only 64%. The effects estimated later in the paper should therefore be interpreted as reflecting information provision in the context of a profiling tool with limited discriminating power.

For our analysis, we exploit age discontinuities in treatment assignment created by the decision tree. Conditional on other characteristics, treatment often hinges on whether jobseekers' age exceeds a cutoff. Knowing the exact decision tree used by the agency allows us to identify the relevant cutoffs applicable to different jobseekers. For instance, a jobseeker without employment in the past year and with a health background is treated only if above age 29. Similarly, someone with recent employment, public administration experience, and no UI receipt in five years is treated only if above age 56. These discontinuities enable a regression discontinuity (RD) design.

2.4 Data and regression discontinuity design

Our data cover all newly registered UI recipients from July 2015 to August 2017 who completed the online survey. For these individuals, we observe survey responses, profiling outcomes (high-risk classification and treatment assignment), employment outcomes from administrative registers, and the input variables used to reconstruct the profiling tool.

If we ignore all age-based splits, the decision tree partitions jobseekers into 27 mutually exclusive subgroups based on the remaining input variables. The complete list of subgroups is presented in Appendix Table A.2.[4] Within eight of these subgroups, there is no variation in treatment assignment, meaning that we cannot leverage them in our RD analysis. In the remaining 19 subgroups, treatment assignment depends on whether a worker's age exceeds a specific age threshold. For the main analysis, we exclude 11 small subgroups with very few observations near the threshold, as well as two subgroups composed entirely of non-Danish individuals, who face distinct labor market challenges and language barriers that may limit treatment effects.[5] After these restrictions, six sizable subgroups remain where treatment is assigned based on age cutoffs. These constitute our main analysis sample, with thresholds at ages 28, 29, 54, and 56.[6]

For our RD specification, we define the running variable RV_i as the difference between the individuals' age at survey completion and the relevant age cutoff (28, 29, 54, or 56). Without measurement error, individuals with $RV_i \geq 0$ would receive the treatment, and those with

[4]In addition, Appendix Figure A.3 provides a graphical illustration of the partition into subgroups.

[5]As we show in Appendix Table A.5, including these subgroups has little impact on our results.

[6]The cutoffs likely mirror age patterns in Danish labor markets: long-term unemployment risk increases in the twenties when transitioning from education to stable employment, remains flat between 30 and 50, and rises again in the fifties when nearing labor force exit.

$RV_i < 0$ would not. In practice, predicted and actual treatment sometimes differ due to idiosyncratic measurement error in historical input variables used by the profiling tool. Updates to administrative data imply that variables we reconstruct ex post can slightly differ from those used in real time.[7] This leads us to assign a few jobseekers to the wrong subgroup and cutoff.

The presence of measurement error makes our RD design fuzzy. However, because the error is negligible and the first-stage relationship between crossing the age cutoff and receiving treatment exceeds 0.85, we focus on intention-to-treat (ITT) effects on labor market outcomes. We estimate these effects using the following specification on the weighted sample of individuals with RV_i close to zero:

$$Y_i = \beta_0 + \beta_1 RV_i + \beta_2 RV_i \times T_i + \tau T_i + \beta_3 X_i + \varepsilon_i, \tag{1}$$

where RV_i is the running variable (measured in weeks), T_i is an indicator for crossing the treatment threshold ($RV_i \geq 0$), and X_i is a set of predetermined controls. Our main specification uses observations within the optimal bandwidth that minimizes mean squared errors (Cattaneo et al., 2017) for our primary outcome: an indicator for receiving unemployment benefits 26 weeks after survey completion, the same horizon used in the profiling model. Observations are weighted with a triangular kernel. We exclude individuals completing the survey in week zero (the exact threshold week) because measurement error is likely highest in these cases.

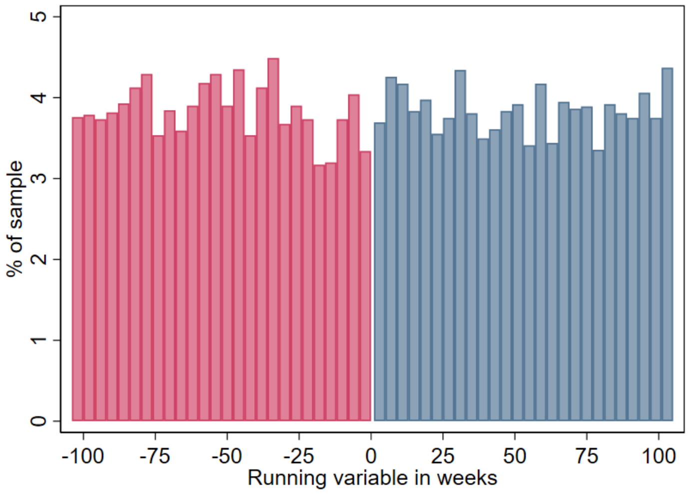
2.5 Validity of empirical approach and sample statistics

The RD approach identifies the causal effect of the information treatment if individuals on both sides of the cutoff are similar in all relevant characteristics except treatment status (i.e., no systematic sorting). This assumption is plausible here because there are no incentives or institutional factors that push individuals to enter UI or complete the survey right before or after a cutoff. Moreover, the survey invitation does not mention the profiling tool, and neither its existence nor the algorithm was common knowledge during our study period.

Empirically, Figure 1 shows that the density of the running variable is smooth around the cutoff (see McCrary, 2008). Only minor distributional differences appear, and a statistical test does not reject continuity ($p = 0.178$). Appendix Table A.3 further shows no meaningful discontinuities in over 30 predetermined characteristics, with only two significant at the 10% level. Overall, individuals just above and below the cutoff are highly comparable.

[7] Variables like the share of the past year spent on UI may be updated as new reports arrive. We use the most recent data to assign workers to subgroups and age cutoffs, which can occasionally misclassify their treatment status.

Figure 1: Distribution of running variable



Note: The figure shows the distribution of the running variable (measured in weeks) around the age cutoff defining whether individuals are predicted to be at low or high risk of long-term unemployment.

Additionally, we assess the robustness of our findings in several ways. First, we estimate alternative RD specifications, varying controls, using quadratic instead of linear trends, and applying a narrower, manually selected bandwidth (Table A.4). Second, we report robust confidence intervals (see Calonico et al., 2014; Cattaneo et al., 2019) alongside the main estimates. Third, we expand the sample to include subgroups with few observations and immigrants (Appendix Table A.5). Finally, we adjust for multiple hypothesis testing by reporting sharpened q-values controlling the false discovery rate (Anderson, 2008) (Appendix Table A.6). Results remain highly robust across all checks.

We also note that the RD design identifies treatment effects only for jobseekers near the age cutoff. In addition, only those completing the initial questionnaire are statistically profiled and included in the analysis. To assess representativeness, Appendix Table A.7 compares three groups: all unemployment spells during the study period, survey respondents, and our analysis sample within 150 weeks of the cutoff. Although some differences are expected, the analysis sample is broadly comparable to the overall population of jobseekers. Older workers—who also tend to have higher prior earnings and stronger family ties—are somewhat overrepresented, plausibly due to the RD cutoff.

2.6 Baseline beliefs and predictive power of information treatment

Before analyzing treatment effects, we present descriptive evidence on jobseekers' baseline beliefs about reemployment. The survey asked new UI recipients about their expected job finding duration. Using non-treated individuals near the cutoff (150 weeks), we compare these expectations to realized outcomes to assess belief accuracy without treatment.

Consistent with prior evidence, jobseekers are overly optimistic (Panels A and B, Figure 2). While 48.5% expect to find a job within three months, only 31% do so. Conversely, just 2% anticipate job finding to take more than six months, yet 40% remain unemployed after six months. Panel C shows that the actual median unemployment duration exceeds expectations by about two to three months, irrespective of whether jobseekers anticipate finding employment within one month, two to three months, or four to six months. Panel D further indicates that 47% of jobseekers are overly optimistic, compared to 9% who are pessimistic, with overoptimism most pronounced among those expecting rapid reemployment.

Figure 2: Characterization of jobseekers' baseline beliefs

Note: The figure shows the distribution of the perceived and realized job finding prospects among non-treated individuals within a bandwidth of 150 weeks below the age cutoff ($N = 5,189$).
The expected job-finding prospects (Panel A) is elicited during the survey before treated jobseekers were informed about their heightened risk of long-term unemployment using the following question: *How quickly do you think you will find a new job?*
The corresponding realized job finding (Panel B) refers to the duration between the completion of the survey and the first month in which the individual was employed (as observed in the administrative records). The category 'leave UI without job' accounts for all jobseekers taking up an education, sickness or parental benefits.
Panel C shows the median realized unemployment duration (blue dots) across groups of jobseekers with different baseline beliefs. The red capped lines indicate the range of durations consistent with accurate beliefs.
Panel D displays the share of jobseekers with optimistic, accurate, and pessimistic beliefs, based on a comparison between their expectations and realized job finding duration, for the full sample as well as for subgroups defined by baseline beliefs.

Given these optimistic baseline beliefs, we expect treated individuals to revise their beliefs downward in response to the message. This assumes that jobseekers view the information as both credible and novel. Appendix Table A.8 substantiate this by showing that the risk classification predicts unemployment durations even after controlling for baseline beliefs.[8] Under a Bayesian learner benchmark, this suggests that jobseekers *should* update beliefs downward when receiving the treatment.

3 Theoretical Considerations

The descriptive evidence suggests substantial scope for individuals to adjust their subjective beliefs and that it is reasonable for the average treated jobseeker to raise their perceived risk of long-term unemployment. In this section, we theoretically illustrate how updating jobseekers' beliefs may influence their behavior.

3.1 Job search framework

While receiving unemployment benefits, individuals search for jobs in continuous time and receive a flow utility of b. Individuals choose their level of search effort, $s \geq 0$, which affects the rate at which job offers arrive, according to the arrival rate function $\lambda(s)$. Without loss of generality, we assume that each additional unit of search effort entails a flow utility cost of one.[9]

Inspired by, e.g., Mueller and Spinnewijn (2023), jobseekers hold subjective beliefs about the job arrival rate function, $\widehat{\lambda}(s)$, and maximize utility as if $\widehat{\lambda}(s)$ represents the true function. We assume that $\widehat{\lambda}(s)$ is a strictly increasing and concave function. For illustrative purposes, we further assume that all jobs offer the same employment value denoted by V, which is high enough to be preferred over unemployment.[10]

Conditional on being on unemployment benefits and searching, we denote the perceived continuation value by $\bar{U}$. Assuming a time discount rate of $\rho > 0$, this satisfies a standard asset pricing equation, reflecting the flow benefits and costs, and the likelihood of transitioning to employment:

$$\rho\bar{U} = \max_{s} \quad b - s + \widehat{\lambda}(s)\left(V - \bar{U}\right) \tag{2}$$

[8]During the pre-profiling period (April 2014–June 2015), when no one received the treatment, individuals classified as high risk were more likely to remain unemployed after six months, conditional on their baseline beliefs.

[9]Let $\gamma(s)$ denote any invertible search cost function. Suppose a worker chooses search effort s to generate an arrival rate $\lambda(s)$ at total cost $\gamma(s)$. This setup is equivalent to one where the worker instead chooses effort $s' = \gamma(s)$, which yields an arrival rate $\lambda(\gamma^{-1}(s'))$ at total cost s'.

[10]This assumption abstracts from individuals' reservation wage, the minimum offer they accept. If jobseekers sometimes reject offers, they may lower this threshold when learning about reduced returns to search, which could further accelerate job finding and unemployment exit.

Additionally, we allow for the possibility that jobseekers exit unemployment benefits entirely for passive support, with the most attractive option offering an exogenously fixed continuation value of R. Depending on the individual, this could involve starting a new education and receiving educational support, receiving sickness or parental benefits, or even some form of self-support without a job.

Individuals' optimal decisions have a simple, useful characterization. Focusing on interior solutions, the optimal search effort, s^*, when receiving unemployment benefits, follows from the first-order condition:

$$1 = \underbrace{\frac{\partial\widehat{\lambda}}{\partial s}}_{\text{effect on job finding}} \cdot \underbrace{\left(V - \bar{U}\right)}_{\text{utility gain from job finding}} \tag{3}$$

On the left is the marginal effort cost, normalized to one. On the right is the marginal benefit of search: an additional unit of effort increases the job-finding rate by $\frac{\partial\widehat{\lambda}}{\partial s}$, and securing a job yields a utility gain of $\left(V - \bar{U}\right)$.

The choice between unemployment benefits and switching to passive support is straightforward, relying solely on a comparison of continuation values. The overall continuation value W is given by:

$$W = \max\left\{R, \bar{U}\right\}$$

If $\bar{U} \geq R$, it is optimal to receive unemployment benefits and continue searching. Otherwise, the individual will choose to exit into passive support.

3.2 Potential effects of the information treatment

The treatment informs jobseekers that they have a heightened risk of long-term unemployment. As job seekers typically hold optimistic beliefs about job finding (see Section 2.6), we anticipate this to *lower* their expectations regarding the likelihood of success when searching for employment. We therefore consider the comparative statics of transitioning from some initial perceived job offer arrival rate $\widehat{\lambda}_b(s)$ to a less optimistic one, denoted as $\widehat{\lambda}_u(s)$, where $\widehat{\lambda}_u(s) < \widehat{\lambda}_b(s)$ for all s.

Intensive margin of search: We first examine the effect of this change on optimal search effort conditional on remaining on unemployment benefits. A shift to less optimistic beliefs is likely to have two off-setting effects (Mueller and Spinnewijn, 2023), implying that the net effect on search is ambiguous. This follows from the first-order condition (Equation (3)), which shows that the marginal returns to search depend on two factors: $\frac{\partial\widehat{\lambda}}{\partial s}$ and $\left(V - \bar{U}\right)$.

First, if the perceived arrival rate function $\widehat{\lambda}$ shifts from the optimistic $\widehat{\lambda}_b(s)$ to the strictly lower $\widehat{\lambda}_u(s)$, the value of unemployment $\bar{U}$ declines (see Equation (2)). This raises the utility gain from job finding, $\left(V - \bar{U}\right)$, thereby increasing the marginal benefit of search and *encouraging* greater effort.

Second, changes in the perceived arrival rate function also affect the first-order condition by altering its slope, $\frac{\partial \widehat{\lambda}}{\partial s}$. The direction of this effect depends on the difference in shape between $\widehat{\lambda}_b(s)$ and $\widehat{\lambda}_u(s)$. In the context of our treatment, a plausible scenario is that less optimistic beliefs reduce the perceived effectiveness of one's own effort, i.e., $\frac{\partial \widehat{\lambda}_u}{\partial s} < \frac{\partial \widehat{\lambda}_b}{\partial s}$. This reduces the marginal utility of search and thereby *discourages* effort.

The net effect of changing beliefs depends on whether the *encouragement* or *discouragement* effect dominates. Appendix Section A.2 provides graphical illustrations of both scenarios. As shown below, however, our empirical results suggest that—at least in the aggregate—encouragement effects tend to dominate on the intensive margin in our setting.[11] A sufficient condition for this to hold is that the proportional decline in the slope of $\widehat{\lambda}$ is smaller than the decline in the level of $\widehat{\lambda}$ (see Appendix Section A.3 for a derivation):

$$\left(\frac{\partial \widehat{\lambda}_b}{\partial s} - \frac{\widehat{\partial \lambda}_u}{\partial s}\right) \Big/ \frac{\partial \widehat{\lambda}_b}{\partial s} < \widehat{\lambda}_b(s) - \widehat{\lambda}_u(s). \tag{4}$$

Extensive margin of search: Beyond the intensive margin of search, belief updates may also affect the extensive margin choice between remaining on unemployment benefits and switching to passive support. The prediction is clear: if the perceived arrival rate shifts from the optimistic $\widehat{\lambda}_b(s)$ to the less optimistic $\widehat{\lambda}_u(s)$, the value of searching, $\bar{U}$, declines. If $\bar{U}$ falls below R, the jobseeker optimally switches to passive support.

Along the extensive margin, less optimistic beliefs therefore unambiguously generate discouragement, potentially leading some individuals to abandon job search and enter passive support. This is particularly likely for jobseekers with access to relatively attractive passive support schemes, where the initial gap between $\bar{U}$ and R is small. Appendix A.2 provides detailed graphical illustrations.

4 Results

This section summarizes RD estimates of the treatment effects. Following the theoretical framework, we distinguish three mutually exclusive states: (1) receiving unemployment benefits with active job search requirements, (2) self-support without public benefits, typically reflecting em-

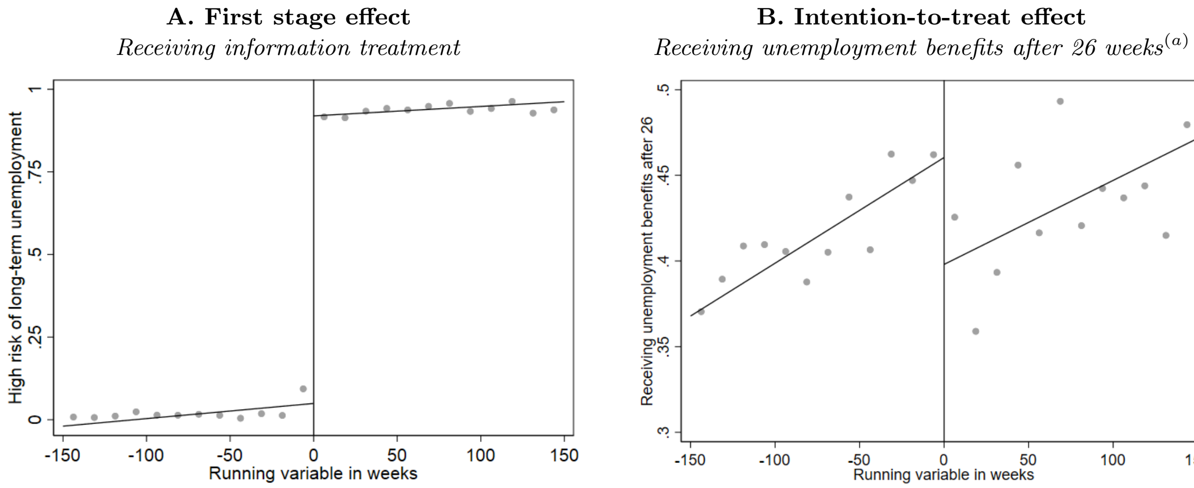
[11] Since our empirical analysis estimates average effects, we cannot rule out that some individuals experience discouragement effects on the intensive margin, which may be masked in the aggregate.

ployment, and (3) receiving other public benefits not tied to active search requirements. We first present results for the overall sample (Section 4.1) and then examine heterogeneity by the attractiveness of alternatives to job search (Section 4.2) and baseline beliefs (Section 4.3). Lastly, we examine transitions into paid employment (Section 4.4) and the role of caseworkers (Section 4.5).

4.1 Average effects of information treatment

Before turning to treatment effects on benefit status, we examine how the probability of receiving the information treatment varies with the running variable. Panel A of Figure 3 shows a sharp jump at the cutoff, from near zero to about 90 percent. As noted earlier, imperfect compliance reflects minor measurement errors in the profiling tool's input variables.

Figure 3: Graphical illustration of regression discontinuity



Note: The figure illustrates the regression discontinuity around the age cutoff. Panel A depicts the relationship between the running variable and the likelihood of actually receiving the information treatment after survey completion. Panel B depicts the relationship between the running variable and the likelihood of still receiving unemployment benefits 26 weeks later. Each dot represents the average for individuals within bins of the running variable, while the solid lines represent an RD-regression fitted to the underlying data.
[(a)]Includes public benefits subject to job search requirements, for example unemployment insurance (UI) benefits and social assistance.

Exits from unemployment: Having established a strong first stage, we next examine ITT effects on benefit status 26 weeks after survey completion—the horizon used in the profiling tool's risk assessment. Panel B of Figure 3 shows a sharp drop at the cutoff: treated individuals just above are less likely to receive unemployment benefits than those just below. Column (1) of Table 1 reports ITT estimates using the optimal bandwidth, indicating a significant reduction in long-term unemployment of 6.9 percentage points, or about 16% ($p = 0.003$). Additionally, Appendix Figure A.4 depicts bi-weekly treatment effects over one year. The reduction in unemployment emerges around week 12, diminishes after week 30—likely reflecting that the control group catches up—and fades by the end of the observation period. On average, the treatment shortens unemployment benefit receipt by about 1.4 weeks ($p = 0.058$)

Table 1: RD estimates: effect of information treatment

		Predicted attractiveness of passive support[(a)]	
	Full sample (1)	Low (≤median) (2)	High (>median) (3)
A. Dependent variable: unemployment benefits after 26 weeks[(b)]			
Intention-to-treat effect	-0.069***	-0.068**	-0.062*
	(0.023)	(0.031)	(0.034)
Robust 90% confidence intervals	[-0.128 ; -0.015]	[-0.152 ; 0.000]	[-0.160 ; 0.003]
Mean dependent variable	0.427	0.425	0.429
B. Dependent variable: self-support after 26 weeks[(c)]			
Intention-to-treat effect	0.042*	0.059*	0.012
	(0.023)	(0.031)	(0.033)
Robust 90% confidence intervals	[-0.029 ; 0.085]	[-0.022 ; 0.132]	[-0.07 ; 0.092]
Mean dependent variable	0.503	0.535	0.474
C. Dependent variable: other public benefits after 26 weeks[(d)]			
Intention-to-treat effect	0.026**	0.009	0.050**
	(0.012)	(0.014)	(0.020)
Robust 90% confidence intervals	[0.013 ; 0.074]	[-0.014 ; 0.055]	[0.016 ; 0.116]
Mean dependent variable	0.069	0.040	0.097
No. of effective observation	9,227	5,360	4,214
Bandwidth left (in weeks)	155	191	128
Bandwidth right (in weeks)	114	119	125
Control variables	Yes	Yes	Yes
Optimal bandwidth	Yes	Yes	Yes
Polynomial	1	1	1

Note: The table reports the effects of the information treatment on outcomes measured 26 weeks after survey completion for the full RD sample (Column 1) and separately by job seekers' predicted attractiveness of passive support (Columns 2-3). We apply the optimal bandwidth selector proposed by Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks". In all specifications, we account for a set of covariates including socio-demographic characteristics (gender, origin, marital status, number of children, living in capital region), level and field of education, and labor market histories (average monthly working hours and earnings in the year prior to job loss, UI fund association, labor market status six months prior to job loss). Standard errors are reported in parentheses. ***/**/* indicates statistical significance at the 1%/5%/10%-level, respectively.
[(a)] The attractiveness of passive support is calculated using out-of-sample predictions from the logit model reported in Table A.11.
[(b)] Indicates the receipt of unemployment benefits—both unemployment insurance and social assistance—that are conditional on active job search requirements.
[(c)] Indicates the absence of any public benefit receipt.
[(d)] Indicates the receipt of public benefits not subject to active job search requirements, including educational support, sickness benefits, and parental leave.

Destinations upon leaving unemployment: These results confirm that the information treatment accelerates unemployment exit, consistent with the policy's overall objective. Through the lens of our theoretical framework, these additional exits may arise through two distinct mechanisms: (i) encouragement effects on the intensive margin, as jobseekers increase their search effort, find employment, and become self-supporting; or (ii) discouragement effects on the extensive margin, as jobseekers abandon search and transition into passive support. In Panels B and Cof Table 1, we disentangle these mechanisms by separately estimating treatment effects on exits into self-support and into passive-support schemes. The results indicate that the overall pattern reflects encouragement effects for some jobseekers and discouragement for others. Panel B shows that the treatment increases self-support by 4.2 percentage points (+8.3%; $p = 0.068$), while Panel C shows a 2.6 percentage point increase (+38%; $p = 0.032$) in exits from UI to other public transfers not tied to job search requirements.

4.2 Attractiveness of passive support programs

As discussed in Section 3, the treatment most likely induces individuals to abandon job search when, absent the intervention, they already perceive certain passive support schemes as relatively attractive. We now test this prediction in our data.

As a first indirect test, we exploit the distinct age composition of our sample. Due to the RD age cutoffs, workers are either in their late twenties or their mid-fifties, and we expect corresponding differences in the relative attractiveness of passive support schemes. For older workers, who are likely in worse health, we expect sickness benefits to play a more important role. Moreover, as younger workers are in their prime childbearing years, we expect pronounced gender differences, with women taking the vast majority of parental leave during our study period. Accordingly, Appendix Figure A.5 examines heterogeneous effects by age and gender and shows results broadly consistent with our predictions: the treatment increases exits into self-support among young men, extends parental leave among young women, and raises transitions into sickness benefits among older men.

To provide a more direct test, we construct an individual-level proxy for the attractiveness of passive support. For each individual in our analysis sample, we use observable characteristics to predict the likelihood of entering passive support programs, based on a model estimated with pre-intervention administrative data (see Appendix A.1 for details). We then classify individuals as having low (below-median) or high (above-median) predicted attractiveness of passive support. Columns (2) and (3) of Table 1 show separate treatment effects for these groups. While overall exit rates from unemployment are similarly affected in both groups, the nature of these exits differs markedly. Among jobseekers with relatively unattractive alternatives (column 2), the treatment raises the likelihood of becoming self-supporting (+5.9pp; $p = 0.058$) but does not significantly impact transitions into passive support. By contrast, for those with more attractive alternatives (column 3), the treatment increases transitions into passive support (+5.0pp; $p = 0.013$) without a significant impact on self-support. Taken together, these patterns indicate that discouragement effects arise mainly among individuals with relatively attractive alternatives.

4.3 Heterogeneity by baseline beliefs

Jobseekers' responses to the information treatment may also vary with their prior beliefs, as the size of the information shock differs between individuals with higher and lower initial confidence. Table 2 therefore presents heterogeneous treatment effects for three groups: (1) jobseekers with low initial confidence, who either expect to become long-term unemployed or are uncertain about their job-finding prospects before receiving the treatment, (2) those with moderate confidence, expecting to find a job within four to six months, and (3) highly confident jobseekers who expect to find a job within three months.

Treatment effects are concentrated among jobseekers who do not initially expect long-term unemployment (columns 2 and 3), whereas effects for those who already anticipate a prolonged unemployment spell are small and statistically insignificant (column 1). This pattern is reassuring for the validity of our approach, since the treatment provides little new information to individuals who already expect prolonged unemployment, limiting its potential impact.

Additionally, we find notable differences in unemployment exits between jobseekers with high and moderate levels of baseline confidence. Positive treatment effects on self-support are sizable and statistically significant only among individuals with moderate confidence (+13.8pp; $p = 0.006$, column 2), whereas no such effect is observed for highly confident individuals. Conversely, the treatment increases the likelihood of exiting unemployment without starting a job—shifting into passive support—only among highly confident job seekers (+4.6pp; $p = 0.002$; column 3). In light of our theoretical framework, a possible explanation for this pattern is that among initially highly confident individuals, the treatment triggers the largest downward revision in both their perceived job-finding rate and the value of remaining on UI, causing transitions into passive support to dominate for this group. At the same time, baseline beliefs are likely correlated with other characteristics that also affect the attractiveness of passive support.[12]

4.4 Self-support, paid employment and job quality

Our main analysis treats self-support (i.e., leaving the public benefit system) as a proxy for job finding. However, beyond transitions into paid employment and self-employment, this measure may also capture some exits from the labor force. We therefore complement the analysis by

[12] Appendix Table A.10 shows that workers with different baseline beliefs also differ in other observable characteristics.

Table 2: RD estimates: effect of information treatment by baseline confidence

	Baseline confidence (exp. job finding duration)		
	Low confidence (≥7m./unc.)	Moderate confidence (4-6m.)	High confidence (≤3m.)
	(1)	(2)	(3)
A. Dependent variable: unemployment benefits after 26 weeks[(a)]			
Intention-to-treat effect	-0.025	-0.153***	-0.060**
	(0.039)	(0.051)	(0.028)
Robust 90% confidence intervals	[-0.127 ; 0.062]	[-0.371 ; -0.126]	[-0.129 ; 0.007]
Mean dependent variable	0.476	0.519	0.363
B. Dependent variable: self-support after 26 weeks[(b)]			
Intention-to-treat effect	0.030	0.138***	0.015
	(0.039)	(0.050)	(0.029)
Robust 90% confidence intervals	[-0.091 ; 0.099]	[0.093 ; 0.333]	[-0.063 ; 0.077]
Mean dependent variable	0.412	0.422	0.588
C. Dependent variable: other public benefits after 26 weeks[(c)]			
Intention-to-treat effect	-0.005	0.015	0.046***
	(0.024)	(0.019)	(0.015)
Robust 90% confidence intervals	[-0.030 ; 0.087]	[-0.014 ; 0.080]	[0.018 ; 0.092]
Mean dependent variable	0.112	0.058	0.048
No. of effective observation	3,166	1,889	5,883
Bandwidth left (in weeks)	190	118	195
Bandwidth right (in weeks)	109	174	150
Control variables	Yes	Yes	Yes
Optimal bandwidth	Yes	Yes	Yes
Polynomial	1	1	1

Note: The table reports the effects of the information treatment on outcomes measured 26 weeks after survey completion for jobseekers with varying baseline beliefs. We apply the optimal bandwidth selector proposed by Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks". In all specifications, we account for a set of covariates including socio demographics (gender, origin, marital status, number of children, living in capital region), level and field of education, and labor market histories (average monthly working hours and earnings in the year prior to job loss, UI fund association, employment 6 months and receipt of parental leave/sickness benefits 26 weeks prior to job loss). Standard errors are reported in parentheses. ***/**/* indicates statistical significance at the 1%/5%/10%-level, respectively.
[(a)] Indicates the receipt of unemployment benefits—both unemployment insurance and social assistance—that are conditional on active job search requirements.
[(b)] Indicates the absence of any public benefit receipt.
[(c)] Indicates the receipt of public benefits not subject to active job search requirements, such as educational support, sickness benefits, and parental leave.

examining transitions into paid employment, along with cumulative working hours and earnings. Panel A of Appendix Table A.9 reports the corresponding estimates for the full sample as well as separately by baseline belief group. Consistent with the main results, the treatment significantly increases the probability of paid employment at both three and six months and raises accumulated working hours and earnings relative to the control group among jobseekers with moderate baseline confidence (see Column 3).[13] Moreover, the estimates show comparable relative increases in employment, hours, and earnings, suggesting that the treatment primarily

[13] Overall, among those with moderate confidence, increased paid employment explains about 80% of the treatment effect on self-support.

operates on the extensive margin of employment rather than on job quality or hours worked. This aligns with our framework that assumes identical jobs, where higher job-finding rates reflect greater search effort. Although lower reservation wages or reduced hours could also generate such effects, the evidence does not support these interpretations.

4.5 Worker's beliefs or caseworker's responses?

So far, we have interpreted the estimated causal effects as arising from changes in jobseekers' beliefs. Yet the profiling tool's risk classification is shared not only with jobseekers but also with their caseworkers, who may adjust their behavior and thereby influence job search outcomes (see, e.g., Behncke et al., 2010; Schiprowski, 2020). To assess this possibility, we examine treatment effects on outcomes directly controlled by caseworkers: the frequency of meetings and participation in ALMPs (see, e.g., Humlum et al., 2023). The results, summarized in Panel B of Appendix Table A.9, support the interpretation that the observed treatment effects are not driven by caseworker responses. First, we find little evidence that meeting frequency differs by risk classification. For jobseekers with moderate baseline confidence, the effect on monthly meetings is close to zero (column 2), suggesting this channel plays at most a minor role in their higher job-finding rates.[14] Second, we examine treatment effects on ALMP participation after survey completion—again finding no evidence that this channel can account for the faster unemployment exits.[15]

These patterns are consistent with qualitative evidence showing that most caseworkers do not rely on the risk assessment in their counseling. Formal interviews by the employment agency and our own conversations with caseworkers indicate that (1) they do not regard the assessment as new information, since they already observe the underlying inputs, and (2) they find the prediction too uncertain, as screening occurs very early in the unemployment spell (STAR, 2021). This view aligns with recent evidence from Germany (van den Berg et al., 2023) and Belgium (Ernst et al., 2024) where caseworker risk assessments perform as well as or better than algorithmic ones. Although we cannot rule out caseworker responses entirely, the strong dependence of treatment effects on jobseekers' baseline beliefs makes it unlikely that caseworkers are the primary driver.

[14]For reference, Schiprowski (2020) find that one additional meeting shortens unemployment duration by about 11 days. To explain the roughly 30-day reduction observed for individuals with moderate baseline confidence, the information treatment would need to induce more than 2.5 extra meetings—an effect size that our estimates confidently rule out.

[15]If anything, the treatment reduces ALMP participation, likely because treated jobseekers leaving UI after the message are no longer eligible for program assignment.

5 Conclusion

Unemployed workers often overestimate their reemployment prospects. This paper estimates the causal effect of a large-scale information treatment designed to correct these misperceptions by informing jobseekers—based on statistical profiling—of their heightened risk of long-term unemployment. We find that the treatment increases exits from unemployment benefits but the nature of these exits varies across worker groups. For some jobseekers, the exits reflect encouragement effects, with successful transitions into employment. For others, they reflect discouragement effects, as individuals abandon job search and instead move into passive support schemes.

In sum, our results show that information policies aimed at correcting subjective beliefs can increase job finding in some cases. However, given the growing use of risk profiling to target interventions among unemployed workers (Desiere et al., 2019), our findings also serve as a cautionary note: for individuals with reasonably attractive alternatives to job search, revealing profiling outcomes may discourage active search and weaken their attachment to the labor market. Combined with the fact that profiling tools may misclassify individuals as high risk due to prediction errors, such policies may inadvertently discourage unemployed workers who, absent the intervention, would have had favorable employment prospects. Given the limited predictive performance of the profiling tool in our setting, this concern is particularly salient. Overall, policymakers should therefore carefully consider when and how profiling is applied, and which jobseekers are informed of their results. Identifying what forms of information provision improve outcomes—and for which groups of workers—remains an important direction for future research.

References

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, S. Quinn, and F. Shilpi (2025): "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment," *The Economic Journal.* 2

Alfonsi, L., M. Namubiru, and S. Spaziani (2022): "Meet Your Future: Experimental Evidence on the Labor Market Effects of Mentors," *G^{2}LM—LIC Working Paper No. 87.* 2

Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018): "Learning about job search: A field experiment with job seekers in Germany," *Journal of Public Economics*, 164, 33–49. 2

Altmann, S., A. Glenny, R. Mahlstedt, and A. Sebald (2022): "The direct and indirect effects of online job search advice," IZA Discussion Paper No. 15830. 2

Anderson, M. L. (2008): "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103, 1481–1495. 7, 20

Balleer, A., G. Duernecker, S. K. Forstner, and J. Goensch (2021): "The Effects of Biased Labor Market Expectations on Consumption, Wealth Inequality, and Welfare," *CESifo Working Paper*, 9326. 2

Bandiera, O., V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2025): "The search for good jobs: evidence from a six-year field experiment in Uganda," *Journal of Labor Economics*, 43, 000–000. 2

Banerjee, A. and S. Sequeira (2023): "Learning by Searching: Spatial Mismatches and Imperfect Information in Southern Labor Markets," *Journal of Development Economics*, 164, 103111. 2

Behaghel, L., S. Dromundo Mokrani, M. Gurgand, Y. Hazard, and T. Zuber (2022): "Encouraging and Directing Job Search: Direct and Spillover Effects in a Large Scale Experiment," *mimeo*. 2

Behncke, S., M. Frölich, and M. Lechner (2009): "Targeting labour market programmes—results from a randomized experiment," *Swiss Journal of Economics and Statistics*, 145, 221–268. 2

——— (2010): "A caseworker like me–does the similarity between the unemployed and their caseworkers increase job placements?" *The Economic Journal*, 120, 1430–1459. 17

Belot, M., P. Kircher, and P. Muller (2019): "Providing advice to jobseekers at low cost: An experimental study on online advice," *Review of Economic Studies*, 86, 1411–1447. 2

——— (2022): "Do the long-term unemployed benefit from automated occupational advice during online job search?" *IZA Discussion Paper No. 15452*. 2

Ben Dhia, A., B. Crépon, E. Mbih, L. Paul-Delvaux, B. Picard, and V. Pons (2022): "Can a Website Bring Unemployment Down? Experimental Evidence from France," *NBER Working Paper No. 29914*. 2

Black, D. A., J. Galdo, and J. A. Smith (2007): "Evaluating the worker profiling and reemployment services system using a regression discontinuity approach," *American Economic Review*, 97, 104–107. 2

Caliendo, M., R. Mahlstedt, A. Schmeisser, and S. Wagner (2023): "The accuracy of job seekers' wage expectations," Working Paper. 2

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295–2326. 7

Carranza, E., R. Garlick, K. Orkin, and N. Rankin (2022): "Job Search and Hiring with Limited Information about Workseekers' Skills," *American Economic Review*, 112, 3547–3583. 2

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2017): "A practical introduction to regression discontinuity designs," *Cambridge Elements: Quantitative and Computational Methods for Social Science-Cambridge University Press I*. 6, 13, 16, 12, 18, 19, 23

——— (2019): *A practical introduction to regression discontinuity designs: Foundations*, Cambridge University Press. 7

Conlon, J. J., L. Pilossoph, M. Wiswall, and B. Zafar (2018): "Labor market search with imperfect information and learning," *NBER Working Paper No. 24988*. 2

Cortés, P., J. Pan, L. Pilossoph, E. Reuben, and B. Zafar (2023): "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab," *The Quarterly Journal of Economics*, 138, 2069–2126. 2

Desiere, S., K. Langenbucher, and L. Struyven (2019): "Statistical profiling in public employment services. An international comparison," *OECD Working Paper 224*. 18

Ernst, S., A. I. Mueller, and J. Spinnewijn (2024): "Risk scores for long-term unemployment and the assignment to job search counseling," in *AEA Papers and Proceedings*, vol. 114, 572–576. 2, 17

Frölich, M. (2008): "Statistical treatment choice: an application to active labor market programs," *Journal of the American Statistical Association*, 103, 547–558. 2

Horton, J. J. (2017): "The effects of algorithmic labor market recommendations: Evidence from a field experiment," *Journal of Labor Economics*, 35, 345–385. 2

Humlum, A., J. Munch, and M. Rasmussen (2023): "What Works for the Unemployed? Evidence from Quasi-Random Caseworker Assignments," IZA Discussion Paper No. 16033. 17

Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2024): "Worker Beliefs About Outside Options," *The Quarterly Journal of Economics*, 139, 1505–1556. 2

Jiang, M. and K. Zen (2025): "Information Asymmetry in Job Search," *Unpublished Manuscript*. 2

Kiss, A., R. Garlick, K. Orkin, and L. Hensel (2023): "Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search," *IZA Discussion Paper No. 16522*. 2

Krueger, A. B. and A. I. Mueller (2016): "A contribution to the empirics of reservation wages," *American Economic Journal: Economic Policy*, 8, 142–79. 2

Le Barbanchon, T., L. Hensvik, and R. Rathelot (2023): "How Can AI Improve Search and Matching? Evidence From 59 Million Personalized Job Recommendations," Working Paper. 2

Lechner, M. and J. Smith (2007): "What is the value added by caseworkers?" *Labour economics*, 14, 135–151. 2

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of econometrics*, 142, 698–714. 6

Mueller, A. I. and J. Spinnewijn (2023): "Expectations Data, Labor Market and Job Search," *Handbook of Economic Expectations*, 677–713. 1, 9, 10

Mueller, A. I., J. Spinnewijn, and G. Topa (2021): "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias," *American Economic Review*, 111, 324–63. 1, 2

Pedersen, J., J. B. Bjorner, H. Burr, and K. B. Christensen (2012): "Transitions between sickness absence, work, unemployment, and disability in Denmark 2004-2008," *Scandinavian journal of work, environment & health*, 516–526. 3

Schiprowski, A. (2020): "The role of caseworkers in unemployment insurance: Evidence from unplanned absences," *Journal of Labor Economics*, 38, 1189–1225. 17

Spinnewijn, J. (2015): "Unemployed But Optimistic: Optimal Insurance Design With Biased Beliefs," *Journal of the European Economic Association*, 13, 130–167. 1, 2

Staghøj, J., M. Svarer, and M. Rosholm (2010): "Choosing the best training programme: is there a case for statistical treatment rules?" *Oxford Bulletin of Economics and Statistics*, 72, 172–201. 2

STAR (2015): "Beskrivelse af profilafklaringsværktøjet til dagpengemodtagere," Report, Danish Agency for Labor Market & Recruitment. 3

——— (2021): "STAR's praksisundersøgelse af profilafklaringsværktøjet," Report, Danish Agency for Labor Market & Recruitment. 4, 17

van den Berg, G. J., M. Kunaschk, J. Lang, G. Stephan, and A. Uhlendorff (2023): "Predicting Re-Employment: Machine Learning versus Assessments by Unemployed Workers and by Their Caseworkers," *IZA Discussion Paper No. 16426*. 17

Online Appendix

Job Search, Overoptimism and Statistical Profiling:
Can Information Provision Improve Job Search Outcomes?

Nikolaj A. Harmon **Robert Mahlstedt**
Mette Rasmussen

This Online Appendix provides supplementary information regarding the following aspects:

- Section A.1 provides details on our individual measurement for the attractiveness of passive support programs (analyzed in Section 4.2 of the main text).
- Section A.2 provides additional discussion and graphical illustrations for our theoretical framework.
- Section A.3 derives a sufficient condition for changes in beliefs leading to encouragement effects on search effort, in the context of our theoretical framework.
- Section A.4 presents additional figures and tables.

A.1 Predicting the attractiveness of passive support programs

To estimate each job seeker's probability of using passive support programs in the absence of the information treatment, we estimate a prediction based on unemployed workers from the pre-intervention period (April 2012-June 2015), during which no UI recipients received the treatment. The estimation sample is restricted to individuals aged 22–35 and 48–62-—i.e., within six years of the age cutoffs in the analysis sample—yielding $N = 8,863$ observations.

The dependent variable is an indicator for receiving other public benefits (educational support, parental leave, or sickness benefits) one year after job loss, with a mean of 7.9%. We estimate a logistic regression model that includes a rich set of covariates: demographic characteristics, education, industry dummies, and measures of labor market history (see Table A.11). Notably, age and gender—the variables used in the additional subgroup analysis in Section 4.2—as well as prior benefit receipt and past earnings, are among the strongest predictors.

We apply the estimated coefficients to the main analysis sample to predict the probability of entering passive support programs. Figure A.6 presents the resulting distribution, which shows substantial heterogeneity in predicted probabilities: moving from the 10th to the 90th percentile increases the likelihood of receiving alternative benefits from 1.9% to 8.2%. Finally, we divide the sample at the (pseudo) median—computed after restricting to the optimal bandwidth—and estimate separate RD regressions for subgroups with low and high predicted probabilities of entering passive support (see Table 1 and Section 4.2 for results and discussion).

A.2 Graphical illustrations of encouragement and discouragement effects in the theoretical framework

This section expands the discussion of the theoretical framework and provides graphical illustrations of the different possible effects of changing beliefs.

We begin by restating the theoretical framework from the main text, while also introducing some useful additional notation. Workers on unemployment benefits receive a flow utility of b and choose search effort s, at a utility cost of one per unit of effort. Search effort determines the perceived job offer arrival rate via the function $\widehat{\lambda}(s)$ with all jobs offering a continuation value of V that is sufficiently high to be preferred to unemployment. The time discount rate is ρ. Finally, we now define $U(s, \lambda)$ to be the value of unemployment *if* exerting a search effort of s while facing an arrival rate of λ. This implies the following asset pricing equation for $U(s, \lambda)$:

$$\rho U(s, \lambda) = b - s + \lambda \left(V - U(s, \lambda)\right). \tag{5}$$

The optimal search decision when on unemployment benefits can now be usefully characterized as job seekers jointly selecting a combination of search effort, s, and arrival rate, λ, subject to the perceived arrival function $\widehat{\lambda}(s)$. As in the main text, we let $\bar{U}$ denote the value of being unemployed when making this choice optimally:

$$\bar{U} = \max_{s,\lambda} \quad U(s, \lambda) \quad \text{s.t.} \quad \lambda = \widehat{\lambda}(s) \tag{6}$$

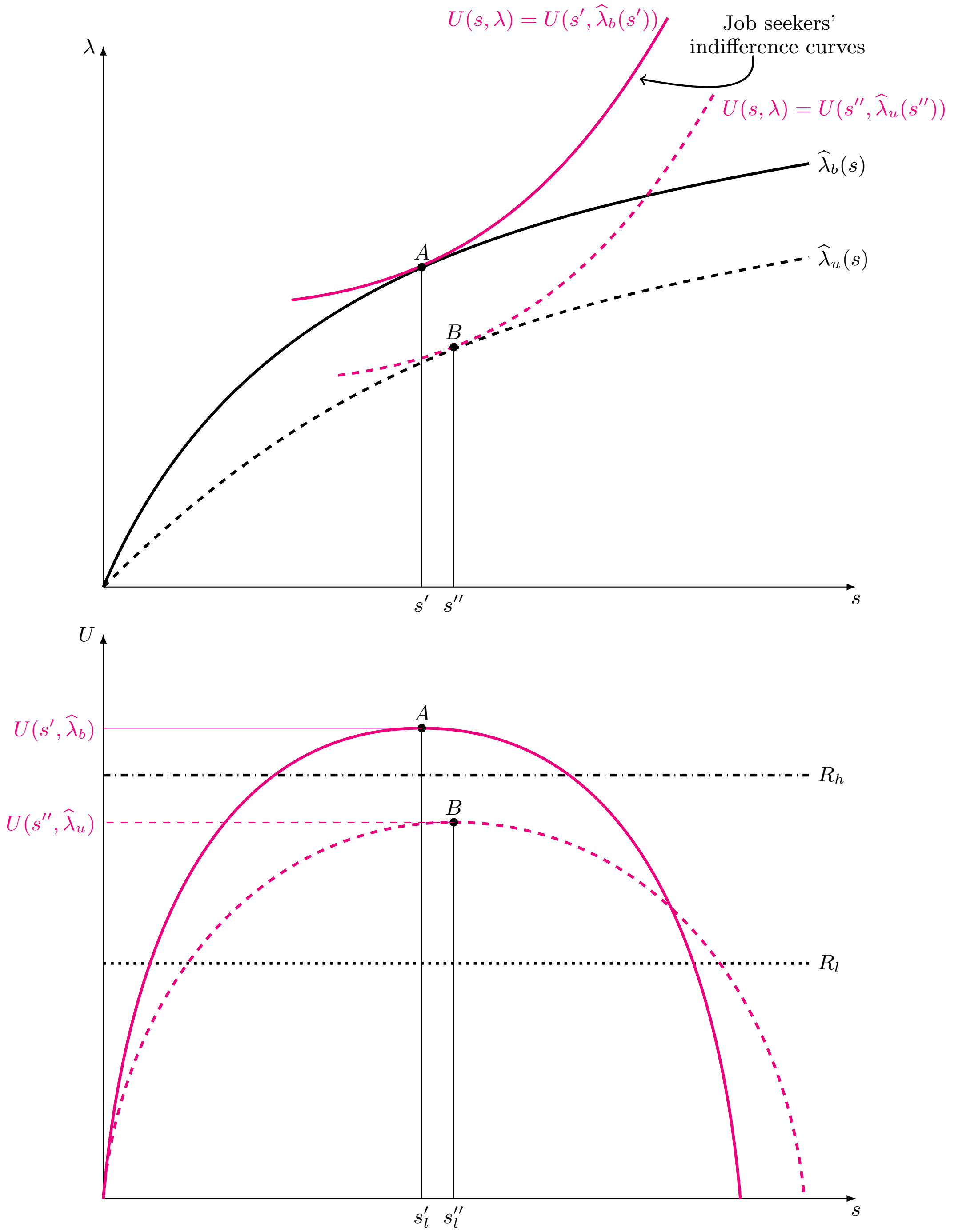
Besides the search effort decision, workers face a choice of leaving unemployment benefits for passive support which offers a continuation value of R. Completely analogous to the main text, this choice can be easily characterized by simply comparing continuation values. With W denoting the overall continuation value, we have:

$$W = \max\{R, \bar{U}\}$$

As in main text, we can then consider the comparative statics of transitioning from some initial perceived job offer arrival rate function $\widehat{\lambda} = \widehat{\lambda}_b$ to a less optimistic one, $\widehat{\lambda} = \widehat{\lambda}_u$, where $\widehat{\lambda}_u(s) < \widehat{\lambda}_b(s)$ for all s. Figures A.1 and A.2 illustrates the comparative statics for two different scenarios. The top panels of the figures depict the intensive-margin search effort decision while individuals receive unemployment benefits, shown in a coordinate system with s and λ on the axes. Following the characterization in (6) the optimal decision involves selecting a combination of s and λ along the curve $\widehat{\lambda}(s)$, with utility increasing as one moves upwards and to the left. When beliefs change, the curve shifts downwards from $\widehat{\lambda}_b(s)$ to $\widehat{\lambda}_u(s)$ resulting in a new optimum. Comparing the cases in Figures A.1 and A.2, however, we see that the effect on search effort is ambiguous; depending on the specific case, search effort may increase or decrease in the response to the changed beliefs.

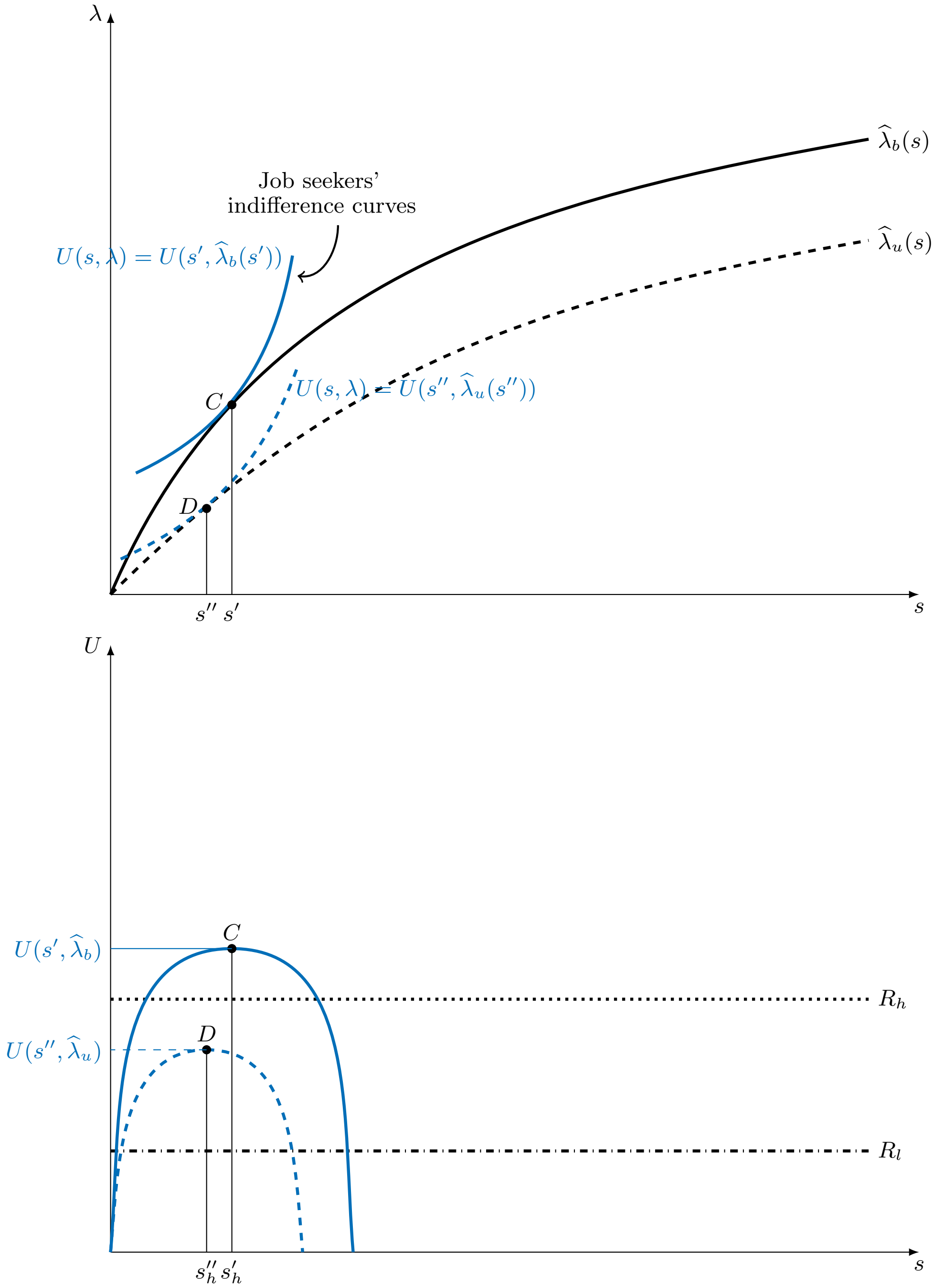
The lower panels of the figures show the corresponding comparative statics for the extensive-margin decision of whether to stop searching altogether and transition to passive benefits. The horizontal axis represents search effort s, and the vertical axis represents utility. In both figures, belief changes reduce the value of unemployment, implying the same potential effect: if this reduction is sufficiently large, the jobseeker will opt for passive benefits. As the figures further illustrate, this outcome becomes more likely when a relatively attractive passive support option is available (R_h rather than R_l).

Figure A.1: Graphical illustration of encouragement effect



Note: The figure illustrates the potential effects of the information treatment within the search framework outlined in Section 3.
The upper panel illustrates how the information treatment potentially encourages job seekers to increase their search effort, s. The treatment reduces the perceived job finding rate from $\widehat{\lambda}_b(s)$ (black solid line) to $\widehat{\lambda}_u(s)$ (black dashed line). For the job seeker characterized by the red indifference curves, this induces an increase in the optimal effort level from s' to s''. The overall value of being on UI is reduced from $U(s,\lambda) = U(s', \widehat{\lambda}_b(s'))$ to $U(s,\lambda) = U(s'', \widehat{\lambda}_u(s''))$.
The lower panel depicts the relationship between job seekers' search effort, s, and the perceived value of search, U, before (red solid line) and after (red dashed line) receiving the information treatment. The perceived value of search decreases from $U(s', \widehat{\lambda}_b)$ to $\rho U(s'', \widehat{\lambda}_u)$ due to the information treatment. Given that $U(s'', \widehat{\lambda}_u) < R_h$, individuals eligible for the high outside option, R_h (represented by the dash-dotted line), cease their search activities entirely. Given that $U(s'', \widehat{\lambda}_u) > R_l$, individuals eligible for the low outside option, R_l (dotted line) continue their search and adjust their search effort as illustrated in the upper panel.

Figure A.2: Graphical illustration of discouragement effect



Note: The figure illustrates the potential effects of the information treatment within the search framework outlined in Section 3.

The upper panel illustrates how the information treatment potentially encourages job seekers to reduce their search effort, s. The treatment reduces the perceived job finding rate from $\widehat{\lambda}_b(s)$ (black solid line) to $\widehat{\lambda}_u(s)$ (black dashed line). For the job seeker characterized by the blue indifference curves, this induces a reduction in the optimal effort level from s' to s''. The overall value of being on UI is reduced from $U(s,\lambda) = U(s', \widehat{\lambda}_b(s'))$ to $U(s,\lambda) = U(s'', \widehat{\lambda}_u(s''))$.

The lower panel depicts the relationship between job seekers' search effort, s, and the perceived value of search, U, before (blue solid line) and after (blue dashed line) receiving the information treatment. The perceived value of search decreases from $U(s', \widehat{\lambda}_b)$ to $\rho U(s'', \widehat{\lambda}_u)$ due to the information treatment. Given that $U(s'', \widehat{\lambda}_u) < R_h$, individuals eligible for the high outside option, R_h (represented by the dash-dotted line), cease their search activities entirely. Given that $U(s'', \widehat{\lambda}_u) > R_l$, individuals eligible for the low outside option, R_l (dotted line) continue their search and adjust their search effort as illustrated in the upper panel.

A.3 Sufficient condition for less optimistic beliefs to increase intensive-margin search effort

In this section, we derive a sufficient condition under which a shift to less optimistic beliefs increases search effort along the intensive margin, conditional on remaining on unemployment benefits.

We derive this sufficient condition by examining the following formulation of the intensive-margin worker problem from the main text:

$$\rho\bar{U}(\alpha) = \max_{s} \quad b - s + \widehat{\lambda}(s,\alpha)\left(V - \bar{U}(\alpha)\right) \tag{7}$$

$$\widehat{\lambda}(s,\alpha) = \alpha\widehat{\lambda}_b(s) + (1-\alpha)\widehat{\lambda}_u(s) \tag{8}$$

$$s^*(\alpha) = \underset{s}{\operatorname{argmax}} \quad b - s + \widehat{\lambda}(s,\alpha)\left(V - \bar{U}(\alpha)\right) \tag{9}$$

$$\rho > 0 \qquad \widehat{\lambda}_u(s) \qquad \widehat{\lambda}_b(s) \geq 0 \qquad s \geq 0 \qquad \alpha \in [0,1] \tag{10}$$

The problem above is identical to the general worker problem from the main text except for the fact that the perceived offer arrival rate $\widehat{\lambda}$ is now dependent on an additional parameter α, implying that the value of unemployment $\bar{U}$ also depends on α. In particular, equation (8) specifies that the perceived arrival rate is a convex combination of the optimistic arrival rate function $\widehat{\lambda_b}$ and the less optimistic arrival rate function $\widehat{\lambda_u}$ with weights governed by α. For $\alpha = 1$ the problem is equivalent to the worker problem from the main text where the worker believes the job offer arrival rate function to be the optimistic $\widehat{\lambda}_b(s)$, while for $\alpha = 0$ it is equivalent to the problem where the worker instead believes it to be the less optimistic $\widehat{\lambda}_u(s)$.

Equation (9) formally defines the optimal search effort choice s^* as a function of α. To establish the sufficient condition from the main text we will show that (for interior solutions, $s^*(\alpha) > 0$)) the condition implies that s^* is strictly decreasing in α (e.g. shifts towards the more optimistic beliefs always lead to lower search effort, and vice versa for shifts towards less optimistic beliefs).

As in the main text, we continue to maintain that the perceived job offer arrival rate is differentiable, strictly increasing and strictly concave for all s:

$$\frac{d\widehat{\lambda}_u}{ds}, \frac{\widehat{\lambda}_b}{ds} > 0 \qquad \frac{d^2\widehat{\lambda}_u}{ds^2}, \frac{d^2\widehat{\lambda}_b}{ds^2} < 0 \tag{11}$$

We also maintain that the optimistic job offer distribution involves strictly higher arrival rates and a (weakly) higher marginal returns to effort for all levels of s:

$$\lambda_u(s) < \lambda_b(s) \qquad \frac{\partial \widehat{\lambda}_u}{\partial s} \leq \frac{\partial \widehat{\lambda}_b}{\partial s} \tag{12}$$

Finally, as in the main text, we assume that employment is sufficiently attractive to always be preferable to unemployment benefits $(V - \bar{U}(\alpha))$. A primitive condition guaranteeing this is:

$$V > \frac{b}{\rho} \tag{13}$$

Deriving the sufficient condition

To show the desired results, we consider the first-order condition of Equation (7), which characterizes the optimal effort choice $s^*(\alpha)$ for interior solutions:

$$1 = \frac{\partial \widehat{\lambda}}{\partial s}\left(V - \bar{U}(\alpha)\right) \tag{14}$$

Let $F(s, \alpha) = \frac{\partial \widehat{\lambda}}{\partial s}\left(V - \bar{U}\right)$ denote the right-hand side of Equation (14). By the implicit function theorem, s^* is strictly decreasing in α if $\frac{\partial F}{\partial s} < 0$ and $\frac{\partial F}{\partial \alpha} < 0$. Below, we show that our sufficient condition ensures these inequalities hold at all points satisfying the first-order condition.

First, the fact that $\frac{\partial F}{\partial s} < 0$ follows immediately from computing the derivative since $\left(V - \bar{U}(\alpha)\right)$ is strictly positive and does not depend on s:

$$\frac{\partial F}{\partial s} = \frac{\partial^2 \widehat{\lambda}}{\partial s^2}\left(V - \bar{U}(\alpha)\right) < 0$$

To verify that our sufficient condition implies $\frac{\partial F}{\partial \alpha} < 0$, we differentiate the right-hand side of the first-order condition (14):

$$\frac{\partial F}{\partial \alpha} = \frac{\partial^2 \widehat{\lambda}}{\partial s \partial \alpha}\left(V - \bar{U}(\alpha)\right) - \frac{\partial \widehat{\lambda}}{\partial s}\frac{\partial \bar{U}}{\partial a}$$

Next, applying the envelope theorem to Equation (7), we obtain: $\frac{\partial \bar{U}}{\partial \alpha} = \frac{\partial \widehat{\lambda}}{\partial \alpha}\left(V - \bar{U}(\alpha)\right)$. Substituting this expression yields:

$$\frac{\partial F}{\partial \alpha} = \left(V - \bar{U}(\alpha)\right)\left(\frac{\partial^2 \widehat{\lambda}}{\partial s \partial \alpha} - \frac{\partial \widehat{\lambda}}{\partial s}\frac{\partial \widehat{\lambda}}{\partial \alpha}\right)$$

Since $\left(V - \bar{U}(\alpha)\right) > 0$, the sign of $\frac{\partial F}{\partial \alpha}$ will be determined by the sign of $\frac{\partial^2 \widehat{\lambda}}{\partial s \partial \alpha} - \frac{\partial \widehat{\lambda}}{\partial s} \frac{\partial \widehat{\lambda}}{\partial \alpha}$. Using Equation (8), we can relate this expression directly to the two belief parameters, $\widehat{\lambda}_u$ and $\widehat{\lambda}_b$:

$$
\begin{aligned}
\frac{\partial^2 \widehat{\lambda}}{\partial s \partial \alpha} - \frac{\partial \widehat{\lambda}}{\partial s} \frac{\partial \widehat{\lambda}}{\partial \alpha} &= \left(\frac{\partial \widehat{\lambda}_b}{\partial s} - \frac{\partial \widehat{\lambda}_u}{\partial s} \right) - \left(\alpha \frac{\partial \widehat{\lambda}_b}{\partial s} + (1-\alpha) \frac{\partial \widehat{\lambda}_u}{\partial s} \right) \left(\widehat{\lambda}_b(s) - \widehat{\lambda}_u(s) \right) \\
&\leq \left(\frac{\partial \widehat{\lambda}_b}{\partial s} - \frac{\partial \widehat{\lambda}_u}{\partial s} \right) - \frac{\partial \widehat{\lambda}_b}{\partial s} \left(\widehat{\lambda}_b(s) - \widehat{\lambda}_u(s) \right) \\
&= \frac{\partial \widehat{\lambda}_b}{\partial s} \left(\left(\frac{\frac{\partial \widehat{\lambda}_b}{\partial s} - \frac{\partial \widehat{\lambda}_u}{\partial s}}{\frac{\partial \widehat{\lambda}_b}{\partial s}} \right) - \left(\widehat{\lambda}_b(s) - \widehat{\lambda}_u(s) \right) \right) \qquad (15)
\end{aligned}
$$

The inequality in the second to last step follows from the fact that $\left(\alpha \frac{\partial \widehat{\lambda}_b}{\partial s} + (1-\alpha) \frac{\partial \widehat{\lambda}_u}{\partial s} \right)$ and $\left(\widehat{\lambda}_b(s) - \widehat{\lambda}_u(s) \right)$ are both positive and that $\left(\alpha \frac{\partial \widehat{\lambda}_b}{\partial s} + (1-\alpha) \frac{\partial \widehat{\lambda}_u}{\partial s} \right)$ attains its maximum for $\alpha = 1$. Finally, inspecting the last expression in Equation (15), we see that since $\frac{\partial \widehat{\lambda}_b}{\partial s} > 0$, a sufficient condition for $\frac{\partial F}{\partial \alpha} < 0$ is indeed:

$$
\frac{\frac{\partial \widehat{\lambda}_b}{\partial s} - \frac{\partial \widehat{\lambda}_u}{\partial s}}{\frac{\partial \widehat{\lambda}_b}{\partial s}} < \widehat{\lambda}_b(s) - \widehat{\lambda}_u(s)
$$

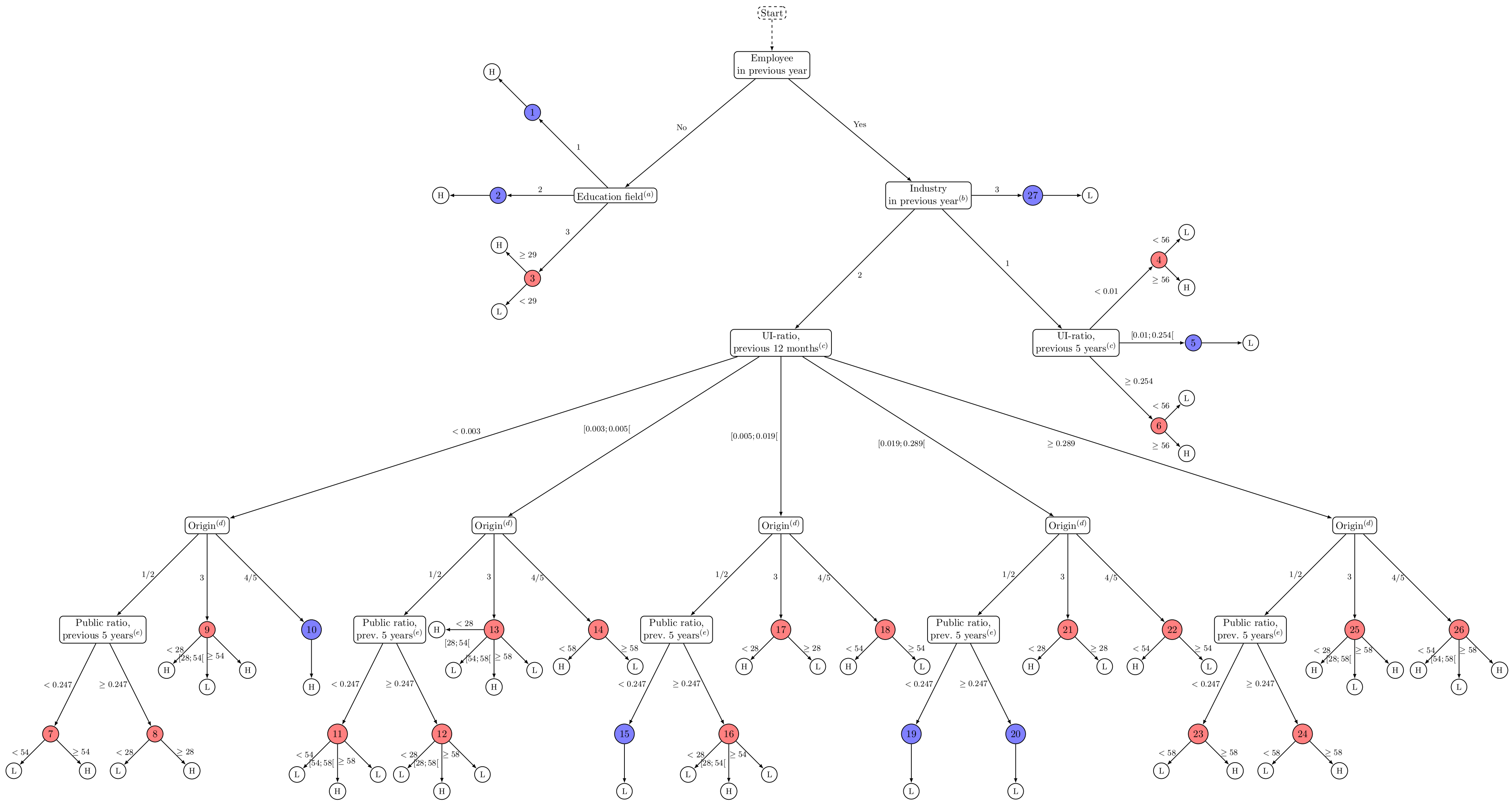
This is the sufficient condition given in the main text.

A.4 Additional Figures and Tables

This subsection presents the following additional figures and tables:

- Figure A.3 illustrates the decision tree underlying the profiling tool.
- Figure A.4 shows the bi-weekly evolution of treatment effects over a one-year horizon.
- Figure A.5 presents heterogeneous treatment effects across demographic groups.
- Figure A.6 depicts the distribution of the predicted attractiveness of alternatives to job search.
- Table A.1 shows a confusion matrix, depicting the share of jobseekers classified by the profiling algorithm's prediction of long-term unemployment alongside their actual outcome.
- Table A.2 shows details on the subgroups underlying the RD design.
- Table A.3 shows balance tests examining the validity of the RD design.
- Table A.4 presents estimation results using alternative RD specifications.
- Table A.5 presents estimation results for alternative samples.
- Table A.6 presents estimation results accounting for multiple hypothesis adjustment.
- Table A.7 reports summary statistics comparing the RD sample to all jobseekers completing the survey and the full population of unemployed workers.
- Table A.8 examines the predictive power of the profiling tool information for unemployment duration.
- Table A.9 presents treatment effects on labor market and caseworker-related outcomes.
- Table A.10 reports summary statistics comparing different belief groups.

Figure A.3: Illustration of Decision Tree and Partition into Subgroups



Note: The figure illustrates the partitioning of job seekers into subgroups generated by the decision tree underlying the profiling algorithm. Ignoring splits based on age, the tree partitions job seekers into 27 subgroups, each represented by a colored dot. Blue dots represent subgroups in which there is no variation in treatment (age does not affect treatment) and red dots represent subgroups with where least one age cutoff generates variation in treatment. For each of the red dots, the additional white dots shows how job seekers in these subgroups are assigned a predicted risk (e.g. treatment vs. no treatment) based on age cutoffs, where H denotes high-risk (treatment) and L denotes low-risk (no treatment). A further description for all the 27 subgroups is presented in Table A.2.

$^{(a)}$ *Education field:* 1 - humanities, religion, aesthetic or missing education; 2 - social work, office, non-commercial, or pedagogical training, fishery, agriculture or food, scientific education 3 - Manufacturing and crafts, health, transportation and communications

$^{(b)}$ *Industry categories:* 1 - public administration, health, teaching, employed in unknown activity, manufacturing, mining and quarrying, utilities, agriculture, forestry or fishery; 2 - trade, logistics, business services, culture, leisure or other services, real estate, information and communication, financial or insurance services. 3 - Construction

$^{(c)}$ *UI benefit ratio* is defined as the fraction of days receiving UI benefits in the previous 12 months and five years, respectively.

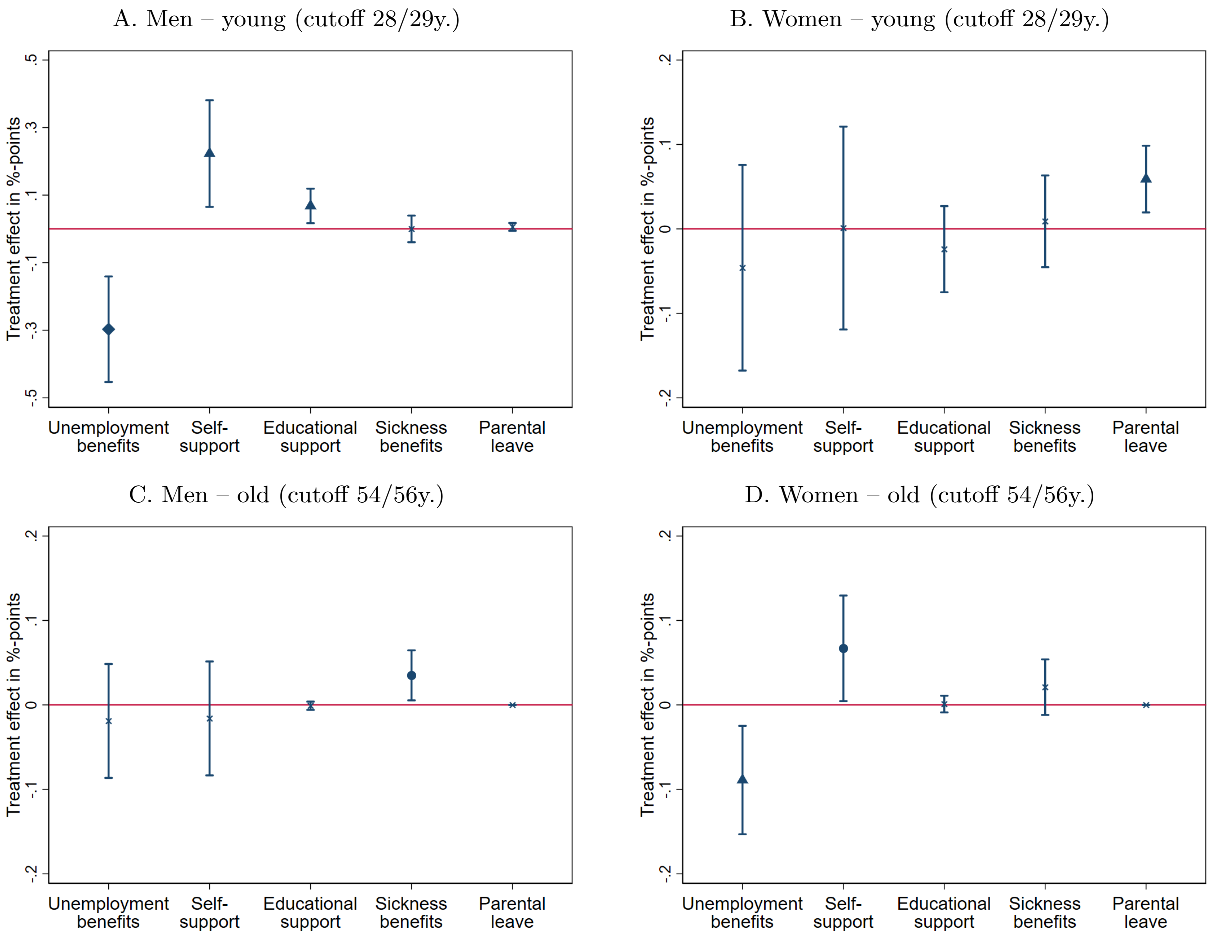
$^{(d)}$ *Origin categories:* 1 - Danish, 2 - descendent (Western country), 3 - immigrant (Western country), 4 - descendent (non-Western country), 5 - immigrant (non-Western country)

$^{(e)}$ *Public benefit ratio* is defined as the fraction of days receiving any public transfers in the previous five years.

Figure A.4: Time profile of treatment effects on benefit status

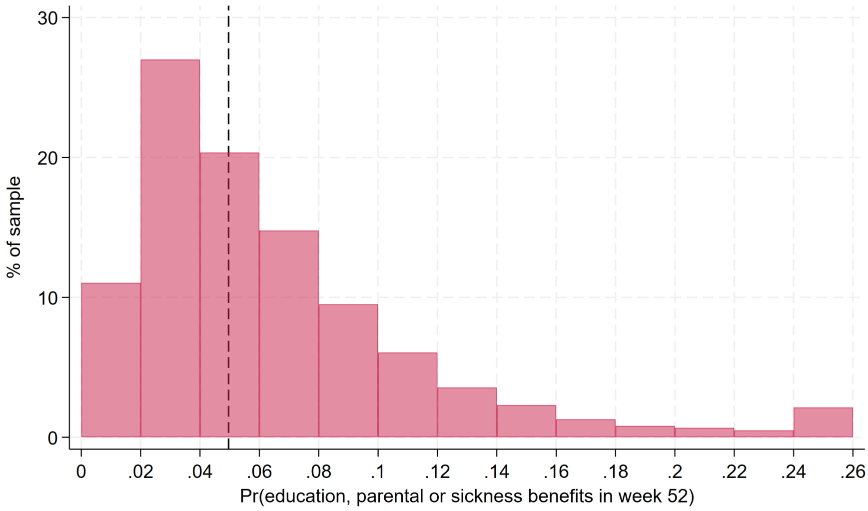
Note: The figure shows treatment effects on the likelihood of receiving unemployment benefits over time including 90% confidence intervals. We apply the optimal bandwidth selector proposed by Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks".

Figure A.5: Heterogeneity by demographic characteristics



Note: The figure shows intention-to-treat effects (including 90% confidence intervals) of the information treatment on the benefit status 26 weeks after survey completion based on RD regressions for various subgroups. The young sample includes men and women facing age cutoffs in their late twenties (age 28 or 29). The old sample includes men and women facing age cutoffs in their mid-fifties (age 54 or 56). We apply the optimal bandwidth selector proposed by Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks" and account for a set of covariates including socio demographics (gender, origin, marital status, number of children, living in capital region), level and field of education, and labor market histories. ♦/▲/● indicates statistical significance at the 1%/5%/10%-level, respectively.

Figure A.6: Predicted probability of using passive support



Note: The figure displays the predicted probability of receiving alternative benefits (education, parental, or sickness benefits) in week 52 relative to job loss for unemployed job seekers in the gross analysis sample (i.e., prior to applying the optimal bandwidth restriction, $N = 30,621$). Predictions are generated using a logistic regression model that includes all covariates listed in Table A.11 as predictors. The model is estimated on a pre-intervention sample of job seekers (April 2012–June 2015), during which no UI recipients were exposed to the information treatment. The sample is restricted to individuals aged 22–35 and 48–62. Predicted values are winsorized at the 98th percentile. The dashed line indicates the median (calculated in the main analysis after restricting to the optimal bandwidth).

Table A.1: Confusion matrix of decision tree

	Percent of job seekers	
	Actual long-term unemployment	
	No	Yes
Predicted high risk of long-term unemployment		
No	55.2	29.1
Yes	7.2	8.5

Note: The table shows the share of jobseekers classified by the profiling algorithm's as having a high vs. not-high risk of long-term unemployment, alongside their actual outcome, defined as remaining unemployed for at least 26 weeks. The sample comprises all unemployed workers in the pre-intervention period (April 2014–June 2015), when no information treatment was provided, and is not restricted to subgroups with age discontinuities ($N = 55,225$). Viewing the profiling tool as a binary classifier, the diagonal entries indicate the proportion of jobseekers correctly classified by the algorithm.

Table A.2: Subgroups in the Representation of the Decision Tree Isolating Age Cutoffs

Subgroup	Employee in previous year	Former industry$^{(a)}$	Education field$^{(b)}$	UI benefit ratio 5 years$^{(c)}$	UI benefit ratio 12 months$^{(c)}$	Public benefit ratio 5 years$^{(d)}$	Origin$^{(e)}$	Age-induced treatment	Observations (+/- 3 years)$^{(f)}$
1	No	0	1	-	-	-	-	-	-
2	No	0	2	-	-	-	-	-	-
3	No	0	3	-	-	-	-	Age >29	905
4	Yes	1	-	ratio <0.001	-	-	-	Age >56	2456
5	Yes	1	-	0.001 <ratio <0.254	-	-	-	-	-
6	Yes	1	-	ratio >0.254	-	-	-	Age >56	910
7	Yes	2	-	-	ratio <0.003	ratio <0.246	1	Age >54	2559
8	Yes	2	-	-	ratio <0.003	ratio >0.246	1	Age >28	552
9	Yes	2	-	-	ratio <0.003	-	2	Age <28 or Age >54	1072 ; 144
10	Yes	2	-	-	ratio <0.003	-	3	-	-
11	Yes	2	-	-	0.003 <ratio <0.004	ratio <0.246	1	54 <Age <58	36 ; 32
12	Yes	2	-	-	0.003 <ratio <0.004	ratio >0.246	1	28 <Age <58	10 ; 13
13	Yes	2	-	-	0.003 <ratio <0.004	-	2	Age <28 or 54 <Age <57	19 ; 0 ; 1
14	Yes	2	-	-	0.003 <ratio <0.004	-	3	Age <58	4
15	Yes	2	-	-	0.004 <ratio <0.018	ratio <0.246	1	-	-
16	Yes	2	-	-	0.004 <ratio <0.018	ratio >0.246	1	28 <Age <54	33 ; 27
17	Yes	2	-	-	0.004 <ratio <0.018	-	2	Age <28	60
18	Yes	2	-	-	0.004 <ratio <0.018	-	3	Age <54	14
19	Yes	2	-	-	0.018 <ratio <0.288	ratio <0.246	1	-	-
20	Yes	2	-	-	0.018 <ratio <0.288	ratio >0.246	1	-	-
21	Yes	2	-	-	0.018 <ratio <0.288	-	2	Age <28	192
22	Yes	2	-	-	0.018 <ratio <0.288	-	3	Age <54	76
23	Yes	2	-	-	ratio >0.288	ratio <0.246	1	Age >58	88
24	Yes	2	-	-	ratio >0.288	ratio >0.246	1	Age >58	247
25	Yes	2	-	-	ratio >0.288	-	2	Age <28 or Age >54	40 ; 12
26	Yes	2	-	-	ratio >0.288	-	3	Age <54 or Age >58	19 ; 15
27	Yes	3	-	-	-	-	-	-	-

Note: The table summarizes the 27 subgroups resulting from a partitioning of the full sample using all input variables and associated cutoffs *except age*.
$^{(a)}$*Industry categories:* 1 - public administration, health, teaching, employed in unknown activity, manufacturing, mining and quarrying, utilities, agriculture, forestry or fishery; 2 - trade, logistics, business services, culture, leisure or other services, real estate, information and communication, financial or insurance services. 3 - Construction
$^{(b)}$*Education categories:* 1 - humanities, religion, aesthetic or missing education; 2 - social work, office, non-commercial, or pedagogical training, fishery, agriculture or food, scientific education 3 - Manufacturing and crafts, health, transportation and communications
$^{(c)}$ *UI benefit ratio* is defined as the fraction of days receiving UI benefits in the previous 12 months and five years, respectively.
$^{(d)}$ *Public benefit ratio* is defined as the fraction of days receiving any public transfers in the previous five years.
$^{(e)}$*Origin categories:* 1 - Danish, 2 - descendent (Western country), 3 - immigrant (Western country), 4 - descendent (non-Western country), 5 - immigrant (non-Western country)
$^{(f)}$The last column reports the number of observations in our sample within a bandwidth of three years around the corresponding age cutoff.

Table A.3: Balance test: RD regression for pre-determined covariates

	Coef.	$P-$value
Demographics		
Age in years	0.654	[0.185]
Male	0.003	[0.906]
Danish citizen	-0.008	[0.574]
Migrant	0.008	[0.574]
Western origin	0.009	[0.271]
Married	-0.008	[0.725]
Any children	-0.022	[0.688]
Expecting a child	0.002	[0.735]
Living in Capital Region	-0.009	[0.641]
Level of education		
Primary or unknown	-0.006	[0.431]
Lower secondary	0.012	[0.488]
Upper secondary	0.013	[0.568]
Short-cycle tertiary	-0.021	[0.070]
Bachelor's degree (or equivalent)	-0.001	[0.940]
Master's degree (or equivalent)	0.003	[0.795]
Field of education		
Humanities	0.005	[0.510]
Manufacturing and craft	0.003	[0.813]
Natural science	-0.001	[0.886]
Pedagogical	0.008	[0.549]
Social science, administration and trade	-0.013	[0.399]
Health	-0.013	[0.333]
Transportation and communication	0.003	[0.719]
Agriculture	-0.002	[0.574]
Labor market status (26 weeks prior to survey)		
Employed	-0.008	[0.715]
Unemployment benefits	0.005	[0.339]
Parental benefits	-0.013	[0.119]
Sickness benefits	-0.012	[0.446]
Previous industry		
Business services	-0.001	[0.889]
Finance and insurance	-0.002	[0.866]
Trade and transportation	-0.008	[0.301]
Manufacturing	-0.022	[0.241]
Information and communication	0.001	[0.968]
Culture, leisure and other services	-0.004	[0.811]
Public administration, education and health	0.000	[0.946]
No. of observations	9,227	

Note: This table summarizes balance tests on pre-determined covariates observed in the administrative and survey data, measured at the time of the survey. We report regression coefficients using the covariates as the dependent variable in RD regressions. We include observations within the optimal bandwidth around the age cutoff (determined based on the main outcome, "unemployed at week 26 after survey") and weight observations using a triangular kernel. Separate tests are conducted for each variable. $P-$values are shown in square brackets.

Table A.4: Robustness: RD estimates for alternative specifications

	Main specification (1)	No controls (2)	Quadratic specification (3)	Manual bandwidth (4)
A. Dependent variable: unemployment benefits after 26 weeks$^{(a)}$				
Intention-to-treat effect	-0.069***	-0.069***	-0.079***	-0.068**
	(0.023)	(0.023)	(0.028)	(0.031)
Robust 90% confidence intervals	[-0.128 ; -0.015]	[-0.131 ; -0.016]	[-0.142 ; -0.019]	[-0.100 ; 0.054]
Mean dependent variable	0.427	0.427	0.427	0.427
B. Dependent variable: self-support after 26 weeks$^{(c)}$				
Intention-to-treat effect	0.042*	0.042*	0.045	0.029
	(0.023)	(0.023)	(0.028)	(0.031)
Robust 90% confidence intervals	[-0.029 ; 0.085]	[-0.027 ; 0.088]	[-0.032 ; 0.091]	[-0.102 ; 0.052]
Mean dependent variable	0.503	0.503	0.503	0.503
C. Dependent variable: other public benefits after 26 weeks$^{(b)}$				
Intention-to-treat effect	0.026**	0.027**	0.033**	0.038**
	(0.012)	(0.012)	(0.015)	(0.016)
Robust 90% confidence intervals	[0.013 ; 0.074]	[0.012 ; 0.074]	[0.017 ; 0.084]	[0.007 ; 0.089]
Mean dependent variable	0.069	0.069	0.069	0.069
No. of effective observations	9,227	8,933	13,999	5,024
Bandwidth left (in weeks)	155	144	226	75
Bandwidth right (in weeks)	114	116	179	75
Control variables	Yes	No	Yes	Yes
Optimal bandwidth	Yes	Yes	Yes	No
Polynomial	1	1	2	1

Note: The table presents a robustness check with alternative RD specifications, reporting the effects of the information treatment on outcomes measured 26 weeks after survey completion. In specification (2), we exclude control variables; all other specifications include covariates capturing sociodemographics (gender, origin, marital status, number of children, residence in the capital region), education level and field, and labor market history. Specification (3) uses a quadratic (second-order polynomial) instead of a linear specification. Specification (4) applies a manually selected bandwidth, while all other specifications use the optimal bandwidth selector of Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks". Standard errors are reported in parentheses. ***/**/* indicates statistical significance at the 1%/5%/10%-level, respectively.

$^{(a)}$Indicates the receipt of unemployment benefits—both unemployment insurance and social assistance—that are conditional on active job search requirements.

$^{(b)}$Indicates the absence of any public benefit receipt.

$^{(c)}$Indicates the receipt of public benefits not subject to active job search requirements, such as educational support, sickness benefits, and parental leave.

Table A.5: Robustness: RD estimates across different jobseeker samples

	Main analysis sample (1)	Adding subgroups with few obs. (2)	Adding immigrants (3)
Dependent variable: unemployment benefits after 26 weeks$^{(a)}$			
Intention-to-treat	-0.069***	-0.058***	-0.050***
	(0.023)	(0.022)	(0.019)
Robust 90% confidence intervals	[-0.128 ; -0.015]	[-0.117 ; -0.012]	[-0.108 ; -0.012]
Mean dependent variable	0.427	0.426	0.446
B. Dependent variable: self-support after 26 weeks$^{(b)}$			
Intention-to-treat	0.042*	0.030	0.024
	(0.023)	(0.022)	(0.020)
Robust 90% confidence intervals	[-0.029 ; 0.085]	[-0.028 ; 0.078]	[-0.024 ; 0.071]
Mean dependent variable	0.503	0.504	0.486
C. Dependent variable: other public benefits after 26 weeks$^{(c)}$			
Intention-to-treat	0.026**	0.028**	0.026**
	(0.012)	(0.012)	(0.011)
Robust 90% confidence intervals	[0.013 ; 0.074]	[0.01 ; 0.068]	[0.01 ; 0.063]
Mean dependent variable	0.069	0.069	0.068
No. of effective observations	9,227	10,726	13,440
Bandwidth left (in weeks)	155	174	201
Bandwidth right (in weeks)	114	113	113
Control variables	Yes	Yes	Yes
Optimal bandwidth	Yes	Yes	Yes
Polynomial	1	1	1

Note: The table presents a robustness check that includes additional jobseeker samples, reporting the effects of the information treatment on outcomes measured 26 weeks after survey completion. Column (1) presents results for our main study sample (including subgroups 3 and 8 around cut-off 28 and 29; as well as 4, 6, 7, and 9 around cut-off 54 and 56, see Table A.2). Column (2) expands the sample to include individuals from subgroups with few observations (subgroups 11–14, 16, 18, and 22–26), while Column (3) further adds subgroups consisting exclusively of young descendants of immigrants (subgroups 9, 17 and 21, around cut-off 28). We apply the optimal bandwidth selector proposed by Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks". All specifications control for a set of covariates, including sociodemographic characteristics (gender, origin, marital status, number of children, residence in the capital region), education level and field, and labor market history. Standard errors are reported in parentheses. ***/**/* indicates statistical significance at the 1%/5%/10%-level, respectively.

$^{(a)}$Includes public benefits subject to job search requirements, for example unemployment insurance (UI) benefits and social assistance.

$^{(b)}$Indicates no receipt of public benefits.

$^{(c)}$Includes educational support, sickness benefits and parental leave.

Table A.6: Robustness: multiple hypothesis adjustment

		Baseline confidence (exp. job finding duration)		
	Full sample	Low confidence (≥7m./unc.)	Moderate confidence (4-6m.)	High confidence (≤3m.)
	(1)	(2)	(3)	(4)
A. Dependent variable: unemployment benefits after 26 weeks[a]				
Intention-to-treat effect	-0.069	-0.025	-0.153	-0.060
	[0.003]	[0.522]	[0.003]	[0.032]
	{0.011}	{0.353}	{0.011}	{0.045}
Mean dependent variable	0.427	0.476	0.519	0.363
B. Dependent variable: self-support after 26 weeks[b]				
Intention-to-treat effect	0.042	0.030	0.138	0.015
	[0.068]	[0.442]	[0.006]	[0.605]
	{0.069}	{0.326}	{0.014}	{0.380}
Mean dependent variable	0.503	0.412	0.422	0.588
C. Dependent variable: other public benefits after 26 weeks[c]				
Intention-to-treat effect	0.027	-0.005	0.015	0.046
	[0.024]	[0.835]	[0.430]	[0.002]
	{0.041}	{0.534}	{0.326}	{0.011}
Mean dependent variable	0.069	0.112	0.058	0.048
No. of effective observation	9,227	3,166	1,889	5,883
Bandwidth left (in weeks)	155	190	118	195
Bandwidth right (in weeks)	114	109	174	150
Control variables	Yes	Yes	Yes	Yes
Optimal bandwidth	Yes	Yes	Yes	Yes
Polynomial	1	1	1	1

Note: The table presents a robustness check that adjusts for multiple hypothesis testing, reporting the effects of the information treatment on outcomes measured 26 weeks after survey completion for jobseekers with different baseline beliefs. For each coefficient, unadjusted p-values are shown in square brackets, and sharpened q-values controlling the false discovery rate (Anderson, 2008) are shown in curly brackets. All specifications control for sociodemographic characteristics (gender, origin, marital status, number of children, residence in the capital region), education level and field, and labor market history.

[a] Indicates the receipt of unemployment benefits—both unemployment insurance and social assistance—that are conditional on active job search requirements.

[b] Indicates the absence of any public benefit receipt.

[c] Indicates the receipt of public benefits not subject to active job search requirements, such as educational support, sickness benefits, and parental leave.

Table A.7: Summary statistics

	Full population (1)	Survey respondents (2)	Analysis sample (3)
Demographics			
Age in years	38.49	38.25	49.55
Male	0.44	0.41	0.40
Migrant	0.12	0.13	0.08
Married	0.36	0.34	0.50
No. of children	1.17	1.11	1.52
Living in Capital Regions	0.25	0.27	0.22
Level of education			
Primary or unknown	0.04	0.05	0.02
Lower secondary	0.14	0.13	0.17
Upper secondary	0.43	0.42	0.44
Short cycle tertiary	0.05	0.06	0.07
University degree	0.24	0.24	0.30
Labor market history in previous year			
Any employment	0.80	0.75	0.85
Any educational support	0.16	0.20	0.09
Any unemployment	0.66	0.60	0.55
Avg. monthly working hours	75.86	71.49	91.33
Avg. monthly earnings (1,000 DKK)	14.51	13.86	20.45
No. of observations	681,827	249,264	14,830

Note: This table presents summary statistics based on administrative data for three samples: (1) all unemployment spells initiated during the study period (July 2015 to August 2017), (2) the subset of jobseekers who completed the survey, and (3) our analysis sample of individuals within 150 weeks around the age cutoff. Percentage shares unless indicated otherwise.

Table A.8: Predictive power of profiling information in pre-intervention period (April 2014 to June 2015)

Dependent variable	Receiving unemployment benefits in week 26		
	(1)	(2)	(3)
Predicted to be at-risk	0.068***		0.054***
	(0.011)		(0.011)
Baseline belief about job finding			
within 1 month		-0.216***	-0.209***
		(0.022)	(0.022)
within 3 months		-0.151***	-0.143***
		(0.014)	(0.014)
within 6 months		-0.056***	-0.051***
		(0.016)	(0.016)
more than 6 months		0.026	0.021
		(0.035)	(0.035)
already found a job		-0.331***	-0.323***
		(0.042)	(0.042)
other[a]		-0.287***	-0.307***
		(0.049)	(0.049)
Constant	0.439***	0.560***	0.530***
	(0.008)	(0.010)	(0.011)
No. of observations	7,930	7,930	7,930
R2 (standard)	0.005	0.028	0.031
R2 (out-of-sample)	0.005	0.028	0.031

Note: The table illustrates the predictive power of the profiling information in a sample of jobseekers in the pre-intervention period (April 2014-June 2015) when no UI recipients received the information treatment. Otherwise, we apply the same sample restrictions as for main analysis sample including individuals who would be within six years of the relevant age cutoff if the information treatment had been in place. The columns show linear regressions where the outcome variable is a dummy for remaining unemployed after 26 weeks and where the explanatory variables are various dummies: *Predicted to be at-risk* represents a dummy for being predicted as being at "high risk" of long-term unemployment (>26 weeks) according to the profiling tool (e.g. this is the information contained in the information treatment). *Baseline belief*-dummies correspond to survey answers to the question "How quickly do you think you will get a job?", with "I don't know" as the omitted category. ***/**/* indicates statistical significance at the 1%/5%/10%-level, respectively. To provide a valid measure of whether adding additional variables improves predictive accuracy, the table includes an out-of-sample R^2 computed via 5-fold cross validation.

Table A.9: RD estimates: effect of information treatment on labor market and caseworker-related outcomes

		Baseline confidence (exp. job finding duration)		
	Full sample	Low confidence ($\geq$7m./unc.)	Moderate confidence (4-6m.)	High confidence ($\leq$3m.)
	(1)	(2)	(3)	(4)
A. Dependent variable: labor market outcomes				
Paid employment in month 3$^{(a)}$	0.010	0.027	0.077**	-0.022
	(0.020)	(0.029)	(0.038)	(0.027)
	[+3.9%]	[+15.4%]	[+45.6%]	[−6.5%]
Paid employment in month 6$^{(a)}$	0.010	-0.001	0.111**	-0.017
	(0.023)	(0.036)	(0.048)	(0.029)
	[+2.4%]	[+0.3%]	[+32.9%]	[−3.3%]
Total working hours within month 1-6$^{(b)}$	7.93	9.53	65.29**	-8.29
	(13.66)	(19.93)	(25.70)	(18.54)
	[+3.4%]	[+6.0%]	[+38.8%]	[−2.7%]
Total labor earnings (in DKK1,000) within month 1-6$^{(c)}$	2.05	-0.19	13.70**	-0.01
	(2.88)	(3.66)	(5.70)	(3.99)
	[+4.4%]	[−0.6%]	[+39.9%]	[−0.1%]
B. Dependent variable: caseworker-related outcomes				
Frequency of caseworker meetings/month$^{(d)}$	0.021	0.036	-0.003	0.006
	(0.020)	(0.031)	(0.036)	(0.027)
	[+3.2%]	[+5.6%]	[−0.4%]	[+0.9%]
ALMP participation within month 1-3$^{(e)}$	0.000	-0.025	-0.023	0.020
	(0.022)	(0.039)	(0.050)	(0.027)
	[+0.0%]	[−6.3%]	[−5.6%]	[+6.2%]
ALMP participation within month 4-6$^{(e)}$	-0.039*	-0.030	-0.082	-0.022
	(0.022)	(0.039)	(0.052)	(0.027)
	[−10.1%]	[−6.7%]	[−17.3%]	[−7.0%]
No. of effective observation	9,227	3,166	1,889	5,883
Bandwidth left (in weeks)	155	190	118	195
Bandwidth right (in weeks)	114	109	174	150
Control variables	Yes	Yes	Yes	Yes
Optimal bandwidth	Yes	Yes	Yes	Yes
Polynomial	1	1	1	1

Note: The table reports the effects of the information treatment on labor market outcomes (Panel A) and caseworker-related outcomes (Panel B) for jobseekers with varying baseline beliefs. We apply the optimal bandwidth selector proposed by Cattaneo et al. (2017), which minimizes the mean squared error above and below the cutoff based on the dependent variable "receiving unemployment benefits after 26 weeks". In all specifications, we account for a set of covariates including socio demographics (gender, origin, marital status, number of children, living in capital region), level and field of education, and labor market histories. Standard errors are reported in parentheses and relative effects compared to the control group mean are shown in square brackets. ***/**/* indicates statistical significance at the 1%/5%/10%-level, respectively.
$^{(a)}$Indicates any paid employment within the corresponding month after survey completion.
$^{(b)}$Refers to the overall number of working hours accumulated since survey completion.
$^{(c)}$Refers to the overall labor earnings accumulated since survey completion.
$^{(d)}$Refers to the number of caseworker meetings per month during the initial unemployment spell in which the jobseeker completed the survey.
$^{(e)}$Indicates participation in any active labor market program during the corresponding time period since survey completion.

Table A.10: Summary statistics by baseline beliefs

	Expected baseline job finding duration		
	$\leq$3 months	4-6 months	$\geq$7 months or uncertain
	(1)	(2)	(3)
No. of observation	5,883	1,889	3,166
Age in years	48.41	48.83	50.41
Married	0.473	0.505	0.520
Number of children	1.51	1.49	1.54
Living in Capital region	0.224	0.258	0.233
Education			
Primary or unknown education	0.018	0.019	0.033
Lower secondary education	0.149	0.136	0.188
Upper secondary education	0.460	0.415	0.447
Short cycle tertiary education	0.067	0.079	0.066
University education	0.305	0.350	0.266
Labor market outcomes in previous year			
Any employment	0.812	0.819	0.842
Avg. monthly working hours	104.06	105.64	97.81
Avg. monthly earnings (in DKK1,000)	24.87	27.85	20.98
Previous industry			
Manufacturing	0.214	0.203	0.177
Finance, insurance and business services	0.138	0.119	0.101
Trade and transportation	0.027	0.049	0.027
Information and communication	0.111	0.111	0.133
Culture, leisure and other services	0.145	0.139	0.094
Public administration, education and health	0.007	0.006	0.005
Other industry	0.075	0.084	0.080
No or missing previous industry	0.282	0.290	0.383

Note: The table reports summary statistics for subgroups with varying levels of baseline confidence. We include observations within the optimal bandwidth (based on the outcome "receiving unemployment benefits 26 weeks after survey completion") around the age cutoff. Percentage shares unless indicated otherwise.

Table A.11: Out-of-sample prediction for receiving passive support

Dependent variable	Other public benefits after 52 weeks	
	(1)	(2)
	Raw coef.	Std. coef.
Demographics		
Age in years	-0.03***	-0.24***
	(0.01)	(0.05)
Male	-0.24**	-0.12**
	(0.09)	(0.05)
Danish origin	-0.03	-0.01
	(0.23)	(0.05)
Western origin	-0.18	-0.03
	(0.33)	(0.05)
Married	-0.03	-0.01
	(0.09)	(0.04)
Number of children	-0.04	-0.05
	(0.04)	(0.04)
Capital region	0.07	0.03
	(0.10)	(0.04)
level of education		
Primary or unknown education	0.13	0.02
	(0.40)	(0.05)
Lower secondary education	0.33	0.13
	(0.27)	(0.10)
Upper secondary education	0.02	0.01
	(0.26)	(0.13)
Short cycle tertiary education	-0.10	-0.02
	(0.30)	(0.07)
University education	-0.17	-0.06
	(0.26)	(0.10)
Previous industry		
Business services	0.81	0.25
	(0.61)	(0.19)
Finance and Insurance	-0.05	-0.01
	(0.69)	(0.13)
Trade and transportation	0.69	0.29
	(0.60)	(0.25)
Manufacturing	0.83	0.29
	(0.60)	(0.21)
Information, communication	0.75	0.16
	(0.63)	(0.13)
Other industry	0.93	0.46
	(0.60)	(0.29)
Labor market history		
Any employment previous 12 months	0.21	0.05
	(0.18)	(0.04)
Avg. monthly working hours in previous 12 months	-0.00	-0.10
	(0.00)	(0.09)
Avg. monthly earnings (1,000 DKK) in previous 12 months	-0.02***	-0.37***
	(0.01)	(0.11)
Benefit ratio (share of weeks)		
Unemployment benefits in previous 12 months	-0.74	-0.05
	(0.72)	(0.05)
Unemployment benefits in previous 5 years	-1.46***	-0.21***
	(0.42)	(0.06)
Any public benefits in previous 5 years	1.71***	0.31***
	(0.37)	(0.07)
No. of observations	8,863	8,863
R^2	0.071	0.071

Note: The table reports both raw and standardized coefficients from a logistic regression predicting the receipt of other public benefits (educational support, parental leave, or sickness benefits) 52 weeks after unemployment entry. The estimation sample consists of unemployed workers from the pre-intervention period (April 2012–June 2015), during which no individuals received the information treatment. The sample is restricted to those aged 22–35 and 48–62. We additional control for individuals' membership in UI funds and their benefit status 26 weeks before the beginning of the unemployment spell.