Algorithms, Incentives, and Effort Adjustment: A Field Experiment with Caseworkers*

Patrick Arni
ZHAW, University of Bristol, IZA, CES-ifo

John Körtner
Collegio Carlo Alberto

May 28, 2026

Abstract

Algorithms are increasingly deployed to improve policy, yet little is known about how their predictions are factored into subsequent decisions in which humans retain final authority. We present evidence from a field experiment in which caseworkers in public employment agencies received algorithmic predictions of job seekers' unemployment durations, intended to shift effort toward those at risk of long-term unemployment. We find that caseworkers updated their beliefs and changed their behavior, but not in the way intended. Caseworkers increased effort among job seekers with predicted short unemployment durations, further accelerating unemployment exit rates among this group. Analyzing elicited prior beliefs, we show that this effort adjustment is inconsistent with standard incentive distortions, including cream-skimming: in the absence of predictions, we do not find more effort for beliefs of short durations. To understand the results, we provide a model of effort allocation under reputational concerns with imperfect attribution. In our model, the algorithm provides a public signal of case difficulty, leading caseworkers to shift effort to avoid the blame associated with durations exceeding predictions.

JEL: C53, D73, D83, J65

*We thank Rafael Lalive, Amelie Schiprowski, and Tara Slough for helpful comments and suggestions. We also thank seminar audiences at the CEPA in Potsdam, HEC Labor Group, IDHEAP, New York University, University of Bristol, University of Konstanz, ZHAW, and OECD ELS, and conference audiences at EALE (2023), SIOE (2024), ESPE (2025), IZA Workshop on Matching Workers and Jobs Online (2025), Frontiers of ALMP Workshop at the University of Bologna (2025), and SNoPE (2026). John Körtner acknowledges financial support from the Horowitz Foundation for Social Policy.

1 Introduction

The increasing popularity of predictive algorithms is driven by high expectations about their potential to improve decisions across policy fields (e.g. Kleinberg et al., 2015, 2018; Goel et al., 2016; Mastrobuoni, 2020; Mullainathan and Obermeyer, 2022). Leveraging data from past cases, such algorithms[1] are trained to predict decision-relevant outcomes for cases where those outcomes are not yet known at the time of the decision. The implicit assumption is that better predictions lead to better decisions. In the public sector, however, algorithms are not eliminating discretion but are used as supplemental decision aids with humans retaining final decision-making power, either by norm or by explicit legal rule.[2] As a result, the effects of algorithms depend on how humans update beliefs and adjust behavior in response to algorithmic information. To understand the impact of algorithms, we have to study the "humans in the loop" (e.g. De-Arteaga et al., 2020; Stevenson and Doleac, 2024). Yet evidence on behavioral responses to algorithms is still scant.

To help fill this gap, we provide evidence from a unique field experiment in the Swiss canton of Fribourg. The regional public employment service temporarily implemented an algorithm to predict the duration of unemployment for each newly registered job seeker. This approach, known as profiling, is a widespread strategy in public employment services (see Desiere et al., 2019, for an overview across OECD countries). The policy goal of profiling is to identify job seekers at risk of long-term unemployment and target them early with support. In Fribourg, caseworkers were asked a series of questions after the first meeting to collect supplemental information on the behavioral characteristics of the job seekers, which were then combined with administrative socio-demographic data. On this basis, the algorithm created a prediction of the expected duration a job seeker would need to find a job. Before

[1]Throughout the paper, we use the term "algorithm" to refer to the final model of a learning system that can be queried for predictions (in line with Ludwig and Mullainathan, 2021). Another common term in recent studies for such algorithms are "prediction machines" and their predictions are consequently called "machine predictions" (Agrawal et al., 2018; Kleinberg et al., 2018; Ribers and Ullrich, 2023; Harris and Yellen, 2024).

[2]Article 14 of the EU AI act requires that algorithmic systems must be designed such that they can be "effectively overseen by natural persons" (European Commission, 2021).

receiving the algorithmic prediction as a second opinion in a randomly selected half of the cases, the caseworker was additionally asked about her or his own prediction of the duration of unemployment. Matching job seekers and caseworkers to administrative databases, we can observe their decisions, actions and outcomes throughout the unemployment spell and beyond.

We find that caseworkers learn from the predictions, but change their behavior in unintended ways. While the intervention aimed at identifying job seekers at risk of long-term unemployment to increase support for this group, caseworkers instead shifted effort toward job seekers who already had good prospects. Among job seekers with short predicted durations, caseworkers were significantly more likely to schedule a second meeting within a month, indicating higher meeting intensity. These cases also exhibit lower reservation wages and higher unemployment exit rates, while post-unemployment wages remain unchanged. Analyzing results conditional on elicited prior beliefs, we find that the effect is largest where caseworkers also expected a short unemployment duration.

We interpret the evidence through a model of effort allocation under reputational concerns with imperfect attribution. Caseworkers are subject to evaluation: outcomes are monitored and supervisors discuss individual cases, creating implicit reputational pressure. Without algorithmic predictions, supervisors observe outcomes but cannot easily distinguish whether poor outcomes reflect inadequate effort or difficult circumstances. Algorithmic predictions introduce a public benchmark for expected case difficulty. When predictions are observable, supervisors can compare realized outcomes to what would have been expected given the prediction. Poor outcomes following favorable predictions become more difficult to attribute to circumstances alone and are therefore more likely to be attributed to insufficient effort. We model this as a benchmark-avoidance problem: caseworkers choose effort to avoid exceeding a prediction-based performance threshold. The model implies that effort increases when algorithmic predictions indicate shorter unemployment durations. In addition, the model implies that effort responses are strongest when the public benchmark is demanding but attainable. This occurs when the algorithm predicts a short unemployment duration and the caseworker's own prior belief also indicates good employment prospects. The predictions match our empirical findings.

The mechanism differs from cream-skimming (Heckman et al., 1997, 2002; Courty and Marschke, 2004; Heinrich and Marschke, 2010; Koning and Heinrich, 2013). In cream-skimming, distortions arise from background incentives. Our mechanism operates through observability: algorithms change what supervisors observe about case difficulty, altering how they attribute outcomes to effort versus circumstances. This builds on insights from Gibbons and Murphy (1992), Holmström (1999), and Ely and Välimäki (2003) that career concerns depend on how informatively outcomes reflect effort. By signaling case difficulty, predictions sharpen attribution and create asymmetric incentives to invest extra effort in easy cases to avoid blame.[3] The mechanism also differs from Albright (2023) and McLaughlin and Spiess (2025), who study settings with algorithmic action recommendations where decision-makers can deflect blame by deferring to the algorithm's advice. Our setting features outcome predictions rather than action recommendations, with supervisors evaluating outcomes rather than adherence to algorithmic guidance.

We contribute to several insightful strands of literature. First, we contribute to the literature on algorithms for policy decisions. A growing body of work shows algorithms can improve outcomes. Kleinberg et al. (2018) show algorithmic predictions can improve bail decisions by reducing prediction error. Ribers and Ullrich (2023) show that combining algorithmic predictions with physician diagnostic skill can substantially reduce antibiotic overprescribing. Belot et al. (2019), Mastrobuoni (2020), Grimon and Mills (2022), Sadka et al. (2024), and Harris and Yellen (2024) show algorithmic information improves outcomes in employment services, policing, child protective services, litigation, and equipment maintenance. However, as Stevenson and Doleac (2024) show in the context of judicial discretion,

[3]The psychology literature documents behavioral distortions under "defensive decision-making" among risk-averse decision-makers in organizations (e.g. Argyris, 1990; Ashforth and Lee, 1990; Artinger et al., 2019). We formalize the attribution mechanism and show how algorithmic predictions lead to defensive decision-making by altering the observability of case difficulty.

the impact depends on how algorithms interact with decision-makers. Albright (2023) shows algorithmic recommendations provide reputational cover in bail decisions: judges follow recommendations because doing so shields them from blame if defendants reoffend. Relatedly, McLaughlin and Spiess (2025) show how algorithmic recommendations can distort behavior by making it costly to deviate from prescribed actions. We instead study a mechanism that operates through outcome evaluation rather than action compliance (cf. Albright, 2023; McLaughlin and Spiess, 2025). When predictions make case difficulty observable to supervisors, they establish a public benchmark against which performance is evaluated. This removes cover rather than providing it.

Second, we contribute to research on performance incentives in organizations. A central challenge in organizational design is that performance measurement systems can distort behavior when workers have discretion and outcomes depend on unobservable circumstances (e.g. Gibbons and Murphy, 1992; Holmström and Milgrom, 1991; Baker, 1992; Gibbons, 1998; Prendergast, 1999). This problem is particularly acute in public sector organizations where formal incentives are weak, missions are complex, and workers retain substantial discretion (Dixit, 2002; Burgess and Ratto, 2003). In public employment services, caseworkers enforce search monitoring, facilitate program access, and have substantial effects on employment outcomes (e.g. Maibom et al., 2017; Schiprowski, 2020; Cederlöf et al., 2021; Böheim et al., 2022; Le Barbanchon et al., 2024). In this context, Heckman et al. (2002) formalize cream-skimming: performance evaluation induces workers to focus on cases where returns are the highest. Courty and Marschke (2004), Heinrich and Marschke (2010), and Koning and Heinrich (2013) document such responses empirically for job training programs. These studies identify distortions arising from the productivity structure. We identify a distinct mechanism: algorithms change what supervisors observe about case difficulty, altering how they attribute outcomes to effort. The mechanism applies beyond employment services to organizational contexts where algorithmic predictions make performance more observable.

Third, we contribute to the emerging literature using information provision experiments in different areas of economics. Information provision experiments cleanly identify the effect of one specific element in the information set of decision-makers by affecting their beliefs through provided real-world information. Haaland et al. (2023) provide a recent overview on this growing literature, with a focus on measuring beliefs and belief updating, which are the first-stage outcomes of the induced information shocks. We contribute to this literature with a novel information provision experiment that is, to our knowledge, the first to directly affect the beliefs and belief updating of caseworkers in the public employment service. Altmann et al. (2018) document an experiment in which job seekers were provided with additional information on success factors in the job search process, aimed at affecting their beliefs and search actions. In our experiment we follow one of the main design strategies by eliciting prior and posterior quantitative beliefs about the targeted outcome, as in Roth and Wohlfart (2020) on expectations about the macroeconomy.[4]

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and experimental design. Section 3 outlines theoretical mechanisms through which algorithmic predictions can affect caseworker behavior. Sections 4 and 5 present the empirical strategy and results. Section 6 develops a formal model of effort distortion under reputational concerns to explain the main results. Section 7 discusses implications and concludes.

2 The Field Experiment

2.1 Context

The context of our field experiment is the application of predictive algorithms in the Swiss unemployment insurance (UI) system in the form of profiling job seekers. The setting is well suited to study behavioral responses to algorithmic predictions because caseworkers retain substantial discretion in their interactions with job seekers throughout the unemployment

[4] An alternative design strategy is to apply probabilistic information treatments, as done by Zimmermann (2020) or Möbius et al. (2022) in different behavioral contexts.

spell. At the same time, the institutional setting generates detailed data on both caseworker actions and subsequent outcomes.

Switzerland operates a relatively generous UI system to provide income support. Individuals who paid contributions for at least eighteen months within the prior two years receive benefits of 70 to 80 percent of their previous earnings. The most common potential benefit duration is 400 working days, for individuals aged 25 to 55[5] (see also Arni et al., 2013, 2022; Lalive et al., 2005, for further descriptions of the Swiss UI system). This generosity is coupled with a comparably strict system of job search requirements, monitoring of obligations, and sanctioning in case of non-compliance.[6] At the same time, public employment services provide substantial job search assistance and access to active labor market programs.

Caseworkers. The implementation of enforcement measures and support programs is in the hands of public employment services (PES) and their caseworkers. Newly registered job seekers are assigned to a caseworker and usually attend a first counseling meeting within two weeks of registration. Caseworkers monitor search requirements, impose sanctions in cases of non-compliance, provide job search assistance, and assign labor market programs. They exercise substantial discretion in setting application requirements, scheduling meetings, assigning referrals, and deciding when to sanction.

Caseworkers address two central frictions in unemployment insurance systems. First, monitoring and sanctions mitigate moral hazard problems associated with benefit receipt (Moffitt, 1985; Chetty, 2008). Evidence from Switzerland shows that stricter monitoring and sanctions reduce unemployment duration but can also lower post-unemployment earnings (Arni and Schiprowski, 2019; Arni et al., 2013). Second, counseling, referrals, and labor market programs can reduce information frictions in job search and improve matching quality (Card et al., 2018). By providing information and referrals, caseworkers may also affect job

[5]Individuals below age 25 without children are entitled to 200 days, whereas entitlement above age 55 extends to 520 days and above age 60 to 640 days.

[6]The Swiss UI operates a monitoring system with the highest sanctioning rate in Europe, and these restrictive regimes have a significant threat effect on reducing unemployment duration (Arni et al., 2022).

seekers' beliefs about employment prospects and offered wages (Spinnewijn, 2015; Conlon et al., 2018; Mueller et al., 2021).

Caseworkers' Incentives. Given their considerable discretion, caseworkers are subject to oversight (Bundesrat, 2013). Internal guidelines recommend different intervention strategies depending on assessed employability (Table A.1). In addition, PES offices are evaluated using a performance index based on four outcome measures: rapid reintegration, long-term unemployment, benefit exhaustion, and re-registration rates (Table A.2).

Supervisors regularly monitor these indicators and discuss performance with caseworkers. Survey evidence shows that supervisors consider the first three indicators particularly important and that most supervisors check the index at least monthly (Kaltenborn and Kaps, 2013, p. 88). The consequences are typically indirect and include audits and learning visits across offices (Kaltenborn and Kaps, 2013, p. 33). At the individual level, performance evaluation is often informal: 23 percent of caseworkers have written performance targets linked to the index, while for 57 percent supervisors discuss performance measures without formalized targets (Kaltenborn and Kaps, 2013, p. 130).

Profiling. The algorithm was in the form of a profiling system. The goal of profiling is to identify job seekers at risk of long-term unemployment early in the spell and to allocate support to them (e.g. Black et al., 2007; Ernst et al., 2024). Algorithms can help by leveraging administrative data to estimate each job seeker's expected unemployment duration,[7] exploiting the observable heterogeneity in long-term unemployment risk across individuals (Mueller and Spinnewijn, 2025).

The profiling system used in Fribourg was based on proportional hazard models (see Arni and Schiprowski, 2015, for background on the project). The models relied on both administrative records and additional job seeker information collected by caseworkers during the first counseling meeting, including assessments of motivation, self-confidence, job search

[7] Körtner and Bonoli (2023) provide an overview of profiling designs and outcomes across countries.

efficiency, salary expectations, and application quality (see A.2 in Arni and Schiprowski, 2015). To construct a training sample, the canton conducted a five-month data collection phase during which caseworkers recorded these additional job seeker characteristics without yet receiving predictions.[8] These data were combined with administrative records and used to estimate the prediction models. The system consisted of 12 separate sub-models stratified by age, gender, and language region (Table A.3). Each model included approximately 250 variables from administrative records and caseworker assessments. Model performance was evaluated using the average absolute prediction error and a classification-based "signal rate." Across sub-models, the average absolute deviation between predicted and realized unemployment duration ranged from 58 to 93 days, while signal rates ranged from 69 to 90 percent in the training data (Arni and Schiprowski, 2015). Predictions were displayed to caseworkers through a visual "barometer" interface that grouped job seekers into risk categories (Figure A.1).

2.2 Experimental Design

The algorithm was designed to provide caseworkers with a second opinion on the job seeker's employment prospects and to support early intervention strategies for individuals at risk of long-term unemployment. The algorithmic prediction was supposed to help the caseworkers adjust their expectations on the job seeker's chances, as caseworkers often underestimate the individual risks of longer unemployment durations (see Figure A.3 and Arni and Schiprowski, 2015).

During the first meeting, typically held within two to three weeks of registration, caseworkers assessed each job seeker and recorded the supplemental information. These responses, combined with administrative variables, formed the input for the algorithm, which

[8] A central challenge with this approach was that many unemployment spells in the collected sample were still ongoing at the end of this period, leaving durations right-censored. To address this, a propensity score matching approach was used to impute missing durations by drawing on past UI data from Fribourg. This method allowed for the inclusion of long unemployment durations in the training sample. However, it also meant that imputations relied only on administrative variables, thus, the additional predictive value of the recorded caseworker inputs could not be exploited for long durations.

then produced a prediction of the unemployment duration. Although the algorithm generated predictions for all job seekers, these were shown to caseworkers in only a randomly selected half of the cases. Figure 1 illustrates the process. Following registration, each job seeker met with their assigned caseworker, who conducted the assessment and recorded the supplemental information in a designated form in the register data system. These data were merged with administrative records and processed by the statistical model. If the case was assigned to the treatment group, the caseworker received the algorithmic prediction before proceeding with the rest of the counseling process; otherwise, the consultation continued without the prediction.

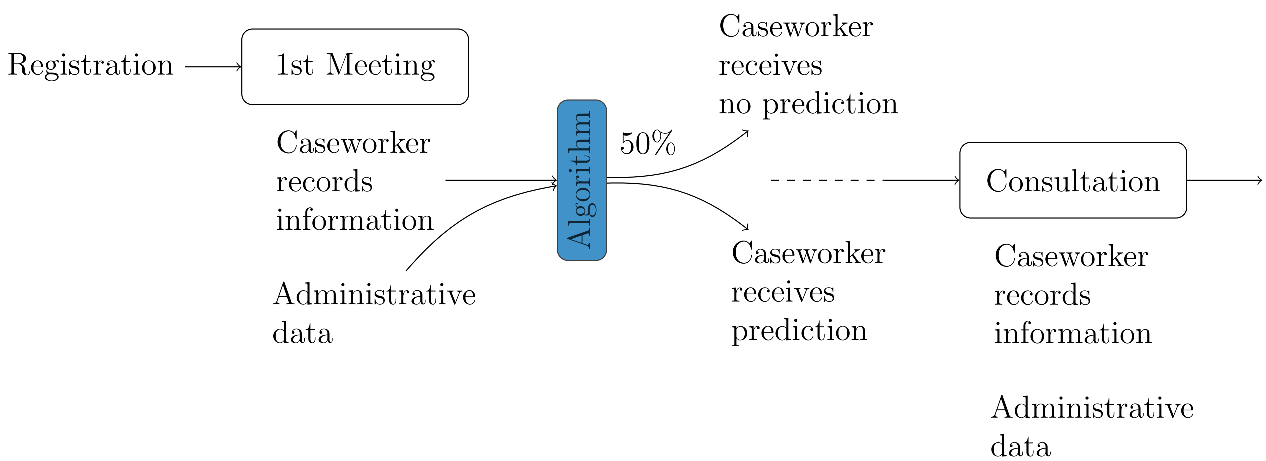

Figure 1: Description of the Process.

We observe both caseworker behavior and job seeker outcomes through linked administrative data. In addition, approximately three months after the initial consultation, caseworkers were asked to record updated assessments of the job seekers. A key feature of the design is that randomization occurred at the caseworker level: within each caseworker's caseload, algorithmic predictions were shown for half of the job seekers. This ensures that treatment variation is orthogonal to both job seeker characteristics and caseworker fixed effects, enabling a clean estimation of the algorithm's causal impact on individual decision-making.

2.3 Descriptive Statistics

Table 1 reports descriptive statistics for the sample. Figure A.2 in the Appendix documents the sample selection. We focus on the period October 2013 to February 2014, when caseworker participation in the experiment had stabilized following a training intervention in September 2013. The final sample includes 3,183 job seekers served by 86 caseworkers who treated both algorithm and control group job seekers. The treatment and control groups are well-balanced across the shown characteristics.

	Algorithm		**Control**	
	Mean	SD	Mean	SD
Prediction	158.26	(99.42)	153.45	(97.66)
Belief	139.25	(90.29)	140.60	(95.55)
Age	37.25	(11.73)	36.85	(11.55)
Gender (F)	0.41	–	0.40	–
Employment Rate (pre-UE)	91.46	(17.49)	92.20	(16.24)
Earnings (pre-UE)	3601.91	(2938.38)	3508.49	(2819.50)
Sector:				
Agriculture	0.01	–	0.01	–
Manufacturing, Industry	0.25	–	0.28	–
Services	0.74	–	0.71	–
Job Seekers	1572		1611	
Caseworkers	86		86	

Table 1: **Descriptive Statistics.** Descriptive statistics for the sample (see Figure A.2 for more details). Prediction and Belief are the algorithmic forecast and caseworker's prior belief of time to reemployment (days). Allowances are daily UI benefits (CHF). Employment Rate is pre-unemployment employment intensity (percent). Earnings are average monthly pre-unemployment earnings in the three months prior. Sectors follow the classification in FSO (2008).

Figure A.3 in the Appendix compares the distributions of caseworker beliefs, algorithmic predictions, and realized unemployment durations. Both caseworkers and the algorithm are overly optimistic, underestimating unemployment durations on average. The algorithm predicts long unemployment spells more frequently than caseworkers do, though both distributions remain skewed toward shorter durations relative to realized outcomes, underestimating unemployment durations on average by 82 and 65 days, respectively.

2.4 Signal

Before examining how caseworkers respond to algorithmic predictions, we characterize the information treatment itself. The prediction algorithm performed worse in the field experiment phase than anticipated in the calibration phase: the algorithm explains less than one percent of the variance in realized durations, while caseworker beliefs explain 3.3 percent. Calibration regressions of actual durations on predictions yield slopes of 0.04 for the algorithm and 0.33 for caseworker beliefs, indicating both severely underpredict the variance in unemployment outcomes. While this low precision reduced the value for the policy evaluation with regards to the main effects, it creates additional variation along the belief and prediction dimensions. This additional variation allows for a more detailed assessment of the treatment effects on the caseworker's beliefs and behavioral reactions. Figure 2 provides additional evidence on the informativeness of the signal.

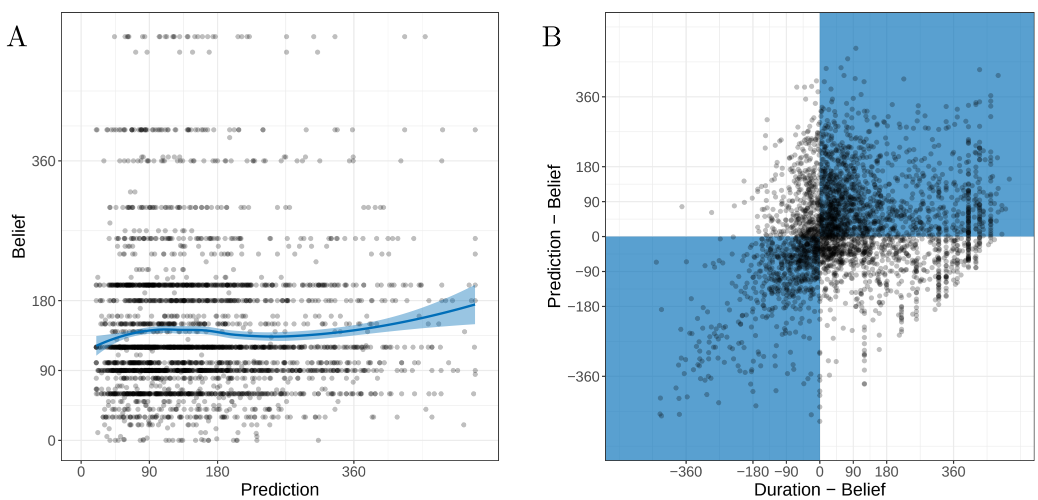

Figure 2: **Information Treatment.** The signal in the algorithmic predictions. Panel A plots caseworker expectations against algorithmic predictions. Panel B shows helpful signals: (i) if the prediction is shorter than the caseworker belief in the case of the realized duration being shorter than the belief (bottom left); (ii) if the prediction is longer than the caseworker belief in the case of the realized duration being longer than the belief (top right).

Panel A of Figure 2 plots caseworker expectations against algorithmic predictions. We see that both dimensions are barely correlated, which alleviates identification concerns in the empirical strategy. Panel B illustrates in which cases the prediction created a "helpful" signal, i.e. a signal that corrects the caseworker belief in the right direction towards the realized duration. About 60 percent of the signals are helpful, indicating that the algorithm provided additional information relative to caseworker beliefs.

3 Conceptual Framework

Before we lay out the results, we motivate what could change. While we use employment terminology, the framework applies to any decision-maker allocating effort under uncertainty (in line with e.g. Chan et al., 2022; Kleinberg et al., 2018; Ribers and Ullrich, 2023). Consider a caseworker choosing effort $e \in [0, 1]$ for a job seeker. Each job seeker has latent risk $\theta \in [0, 1]$, where higher values correspond to higher expected unemployment duration absent additional effort. The realized unemployment duration is $y = \mu(\theta, e, \varepsilon)$, increasing in θ and decreasing in e. The caseworker observes job seeker characteristics and forms a prior belief b about expected unemployment duration absent additional effort. In the treatment group, the caseworker also receives an algorithmic prediction p of unemployment duration. Both b and p are noisy signals of the underlying expected duration implied by θ.

3.1 Belief Updating

Given prior belief $b = \mathbb{E}[y]$ about expected unemployment duration and algorithmic prediction p, caseworkers form a posterior expectation $\tilde{b}$:

$$\tilde{b} = \begin{cases} \lambda b + (1 - \lambda)p & \text{if treatment group} \\ b & \text{if control group} \end{cases},$$

where $\lambda \in [0, 1]$ reflects how much weight the caseworker places on their prior relative to the algorithmic signal. In the control group, no additional information is received, such that posterior expectations equal prior beliefs.

Benchmark. The policy intention follows standard Bayesian updating, where decision-makers optimally combine prior beliefs with new information based on relative precision. Under Bayesian updating, the weight λ is determined by relative precision. If the prior is very precise, λ is close to 1 and the caseworker relies mainly on their own belief. If the algorithm is very precise, λ is close to 0 and the caseworker relies mainly on the prediction. The empirical prediction is that posterior beliefs $\tilde{b}$ should shift toward algorithmic predictions in treatment, with the magnitude depending on perceived relative precision.

Alternatives. In practice, caseworkers do not know the algorithm's precision. Thus, caseworkers may misspecify λ, placing suboptimal weight on the algorithmic signal. Algorithm aversion corresponds to the extreme case $\lambda = 1$, where caseworkers completely ignore predictions (Dietvorst et al., 2015; Dargnies et al., 2024). Other deviations include confirmation bias (Rabin and Schrag, 1999), ambiguity aversion (Epstein and Schneider, 2007), and limited attention to conflicting signals (Caplin and Dean, 2015). As we do not observe caseworkers' certainty about their beliefs, our ability to measure updating mechanisms is limited. Empirically, we focus on testing whether any belief updating occurs, or if we see absolute algorithm aversion, where no belief updating occurs.

3.2 Behavior Adjustment

The caseworker's effort decision hinges on a payoff function that depends on the expected unemployment duration and the effort provided. We consider a general payoff function

$$\pi(e) = -\alpha \mathbb{E}[y \mid e] - \frac{\beta}{2} e^2.$$

The first term captures the cost of poor outcomes. The second term is the direct cost of effort.[9] Since unemployment duration is not known at the time of the decision, the case-

[9] We assume quadratic effort costs to obtain interior solutions and tractable comparative statics (cf. Chan et al., 2022; Kleinberg et al., 2018; Ribers and Ullrich, 2023).

worker evaluates expected outcomes using their posterior belief. Let $\tilde{b}$ denote the posterior expectation of unemployment duration absent additional effort. Substituting $\mathbb{E}[y \mid e]$ with $\tilde{b}(1-e)$, we get the payoff function at the time of the decision: $-\alpha\tilde{b}(1-e) - (\beta/2)e^2$, where $\alpha, \beta > 0$. This specification captures that the marginal value of effort is higher for cases expected to have longer unemployment durations.

Benchmark. The policy intention was that caseworkers place positive weight on improved outcomes and face positive effort costs, such that $\alpha, \beta > 0$. Assuming an interior solution, the first-order condition yields optimal effort: $e^* = \alpha\tilde{b}/\beta$. Effort therefore increases in the posterior belief. Substituting the posterior belief and differentiating with respect to the prediction gives

$$\frac{\partial e^*}{\partial p} = \frac{\alpha(1-\lambda)}{\beta} > 0.$$

The benchmark prediction is therefore that treatment effects are positive for high p: caseworkers increase effort when predictions indicate higher unemployment risk.

Alternatives. Several mechanisms can overturn this benchmark. Bureaucrats may have misaligned intrinsic preferences (e.g. Prendergast, 2003; Dixit, 2002). For example, the caseworkers might place weight on avoiding poor outcomes for low-risk cases rather than high-risk cases, because of a preference to work with cooperative, motivated clients and to avoid the psychological costs of dealing with difficult cases. Beyond intrinsic preferences, caseworkers may be misaligned for instrumental reasons. Heckman et al. (1997, 2002) formalize "cream-skimming," where workers focus on easy cases when performance contracts reward successful placements regardless of case difficulty, or workers believe effort is more productive for easy cases (see also Courty and Marschke, 2004; Heinrich and Marschke, 2010; Koning and Heinrich, 2013). As the Swiss performance incentives do not only measure placement success but days in unemployment, the background incentives are not necessarily misaligned. However, caseworkers can still have misaligned productivity beliefs that enable cream-skimming. To model this, we allow the (perceived) effort effectiveness to vary across client types: $\mathbb{E}[y \mid e] = \tilde{b}(1 - \tau(\tilde{b})e)$, where $\tau(\tilde{b})$ captures how effective effort is for cases with posterior expected duration $\tilde{b}$. Substituting, we get the payoff function $-\alpha\tilde{b}(1 - \tau(\tilde{b})e) - (\beta/2)e^2$. The first-order condition yields: $e^* = \alpha\tilde{b}\tau(\tilde{b})/\beta$. With posterior updating, the effect of the prediction is:

$$\frac{\partial e^*}{\partial p} = \frac{\alpha(1-\lambda)}{\beta}\left[\tau(\tilde{b}) + \tilde{b}\tau'(\tilde{b})\right],$$

where the right-hand side is evaluated at the posterior. Under beliefs where $\tau'(\tilde{b}) < 0$ is sufficiently negative, effort would decrease in p: caseworkers would provide more effort to low-risk cases.

Another potential group of mechanisms arises from the algorithm. Albright (2023) shows that algorithmic recommendations provide reputational cover in bail decisions: judges face lower blame when following algorithmic advice, even if outcomes are poor. We can introduce a very general penalty term: $-\gamma R(\cdot)$, where $\gamma > 0$ captures reputational concerns. Albright (2023) motivates a mechanism involving recommendations that follow from the predictions and effort: $R(e, r(p))$, where $r(p)$ is the recommendation. Deviating increases reputational exposure. McLaughlin and Spiess (2025) examine a related mechanism where algorithmic recommendations alter preferences by imposing a cost on deviations. This motivates a penalty function of the form $R(|e - r(p)|)$, where the reputational cost increases in the distance between the caseworker's action and the algorithmic recommendation. In McLaughlin and Spiess (2025), institutional factors such as audit risk or fear of liability create psychological or economic costs for overriding algorithmic guidance. In our setting, predictions do not prescribe what to do but establish a public benchmark for what outcomes to expect. The theoretical and empirical effects of algorithms in settings with outcome-based performance evaluation are unexplored.

4 Estimation Strategy

To estimate the effect of access to the algorithm on various outcomes, we begin with the following specification:

$$y_i = \beta_0 + \beta_1 \text{Algorithm}_i + \gamma_j + \varepsilon_i$$

Here, y_i denotes the outcome of interest for job seeker i, and Algorithm_i is a treatment indicator equal to one if the caseworker received the algorithmic prediction. We include caseworker fixed effects γ_j, where j indexes caseworkers. Given random assignment of the prediction within caseworkers, $\hat{\beta}_1$ can be interpreted as the causal effect of access to the algorithm. Fixed effects additionally absorb differences in caseworker-level compliance with the experimental protocol, as caseworkers could choose not to complete the supplemental questionnaire required for inclusion in the experiment. Standard errors are clustered at the caseworker level to account for within-caseworker correlation in the error term.

We extend this baseline model to study heterogeneity in the treatment effect by interacting access with the algorithmic prediction:

$$\begin{aligned} y_i =& \beta_0 + \beta_1 \text{Algorithm}_i \\ &+ \beta_2 \text{Prediction}_i + \beta_3 [\text{Algorithm}_i \times \text{Prediction}_i] + \gamma_j + \varepsilon_i \end{aligned}$$

This specification allows treatment effects to vary with the signal provided by the algorithm. Under the benchmark framework in Section 3, caseworkers should increase effort for job seekers with longer predicted unemployment durations. By contrast, negative interaction effects for longer predicted durations are consistent with mechanisms such as cream-skimming or other forms of misaligned effort allocation. To capture potential nonlinear responses, we divide predictions into three categories. Specifically, we distinguish between predictions of at most 90 days, predictions between 90 and 180 days, and predictions above 180 days.

The thresholds reflect the color transitions in the barometer interface of the algorithm (see Figure A.1 in the Appendix). To examine whether the baseline findings depend on these exact thresholds, we additionally estimate continuous robustness specifications using either a linear prediction measure or natural spline functions with two and three degrees of freedom.

Beyond the prediction itself, our setting also allows us to incorporate caseworkers' own prior beliefs directly into the treatment analysis. We therefore estimate:

$$\begin{aligned} y_i =& \beta_0 + \beta_1 \text{Algorithm}_i + \beta_2 \text{Prediction}_i \\ &+ \beta_3 \text{Belief}_i + \beta_4 [\text{Algorithm}_i \times \text{Prediction}_i] + \beta_5 [\text{Algorithm}_i \times \text{Belief}_i] \\ &+ \beta_6 [\text{Algorithm}_i \times \text{Belief}_i \times \text{Prediction}_i] + \gamma_j + \varepsilon_i \end{aligned}$$

We define beliefs analogously to predictions by distinguishing between expected unemployment durations of at most 90 days, between 90 and 180 days, and above 180 days. This specification allows us to examine whether responses to the algorithm depend not only on the prediction itself, but also on the caseworker's prior assessment of the job seeker. In particular, it allows us to distinguish between cases in which the algorithm confirms prior beliefs and cases in which it contradicts them. To distinguish treatment-induced responses from pre-existing effort allocation patterns even more explicitly, for behavioral outcomes, we also estimate the relationship between prior beliefs and outcomes in the control group:

$$y_i = \beta_0 + \beta_1 \text{Belief}_i + \gamma_j + \varepsilon_i \quad \text{if Algorithm}_i = 0$$

This specification allows us to examine how caseworkers allocate effort to job seekers in the absence of algorithmic predictions. In particular, it allows us to assess whether caseworkers allocate more effort to job seekers they expect to have shorter unemployment durations, as implied by cream-skimming explanations.

5 Results

In this section, we report empirical results organized around the conceptual framework in Section 3. We first examine whether algorithmic predictions affect caseworkers' beliefs and behavior, and whether these responses are consistent with standard updating or alternative explanations. We then study whether behavioral responses translate into changes in job seekers' behavior and labor market outcomes. For each set of outcomes, we briefly describe the variables and their operationalization before presenting the empirical results and their implications. We present results graphically, with regression tables reported in Tables A.4–A.14. We focus on the interaction specifications that are central to understand mechanisms.

Belief Updating. We first examine whether algorithmic predictions affected caseworkers' beliefs. Figure 3 shows belief updating between the first and second information elicitation, separately by prediction and by the interaction of prediction and prior belief. We measure belief updating as the change between caseworkers' initial expectations elicited during the first counseling meeting and their updated expectations elicited approximately three months later. The analysis sample is restricted to the 1,174 cases (37 percent of the experimental sample) where both the first and the second survey and thus both belief elicitations were completed.

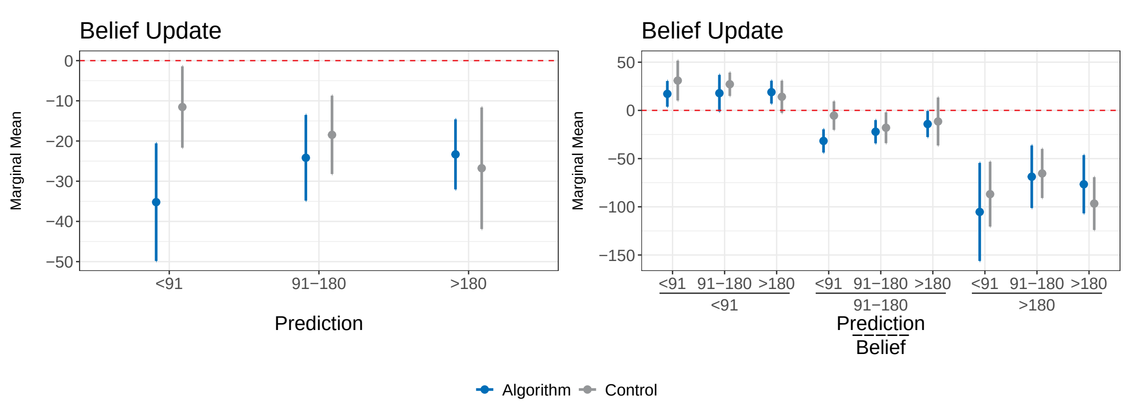

Figure 3: **Belief Updating.** Change in caseworkers' beliefs between the first and second information elicitation by treatment status, algorithmic prediction, and prior belief. Table A.4 reports estimates.

We observe belief updating in both treatment and control groups, indicating that caseworkers revised their expectations over the unemployment spell even without access to the algorithm. Relative to the control group, the algorithm leads only to modest additional updating, concentrated primarily among job seekers with short predicted unemployment durations. Across most prediction and belief cells, differences between treatment and control groups remain small. While caseworkers do not completely ignore the algorithmic signal, the prediction does not appear to substantially shift posterior beliefs beyond the information caseworkers acquire during the spell itself. A caveat is that the second elicitation often occurred relatively late in the unemployment spell. As a result, posterior beliefs incorporate substantial additional learning from job seeker behavior and labor market outcomes over time. To assess whether the results depend on the 90- and 180-day thresholds, we additionally estimate continuous specifications using linear and spline functions of the prediction measure, reported in Tables A.15–A.20. The results are qualitatively similar to the categorical specification and do not suggest that the findings are driven by the threshold choices.

Behavior Adjustment. To examine how caseworkers adjusted their behavior in response to algorithmic predictions, we examine four measures of caseworker behavior. Active labor market program (ALMP) assignment is a binary indicator equal to one if the job seeker was assigned to any labor market program within the first 90 days of unemployment. Application requirements measure the average number of applications caseworkers required job seekers to submit per month, computed across the first three months with positive monitoring data. Sanctions is a binary indicator equal to one if the job seeker received any executed sanction within the first 90 days. Meeting intensity measures whether a second meeting occurred within 30 days of the initial consultation. Shortly before the experiment, official guidelines shifted from fixed meeting schedules toward schedules conditional on job seeker employability (SECO, 2012).

Figure 4 reports treatment effects by prediction categories. Panel A shows that the main behavioral response operates through meeting intensity. Caseworkers increased follow-up meetings for job seekers with short predicted unemployment durations, while sanctions, ALMP assignments, and application requirements remained largely unchanged. The increase in meetings is concentrated among job seekers predicted to exit unemployment quickly and weakens as predicted unemployment duration increases.

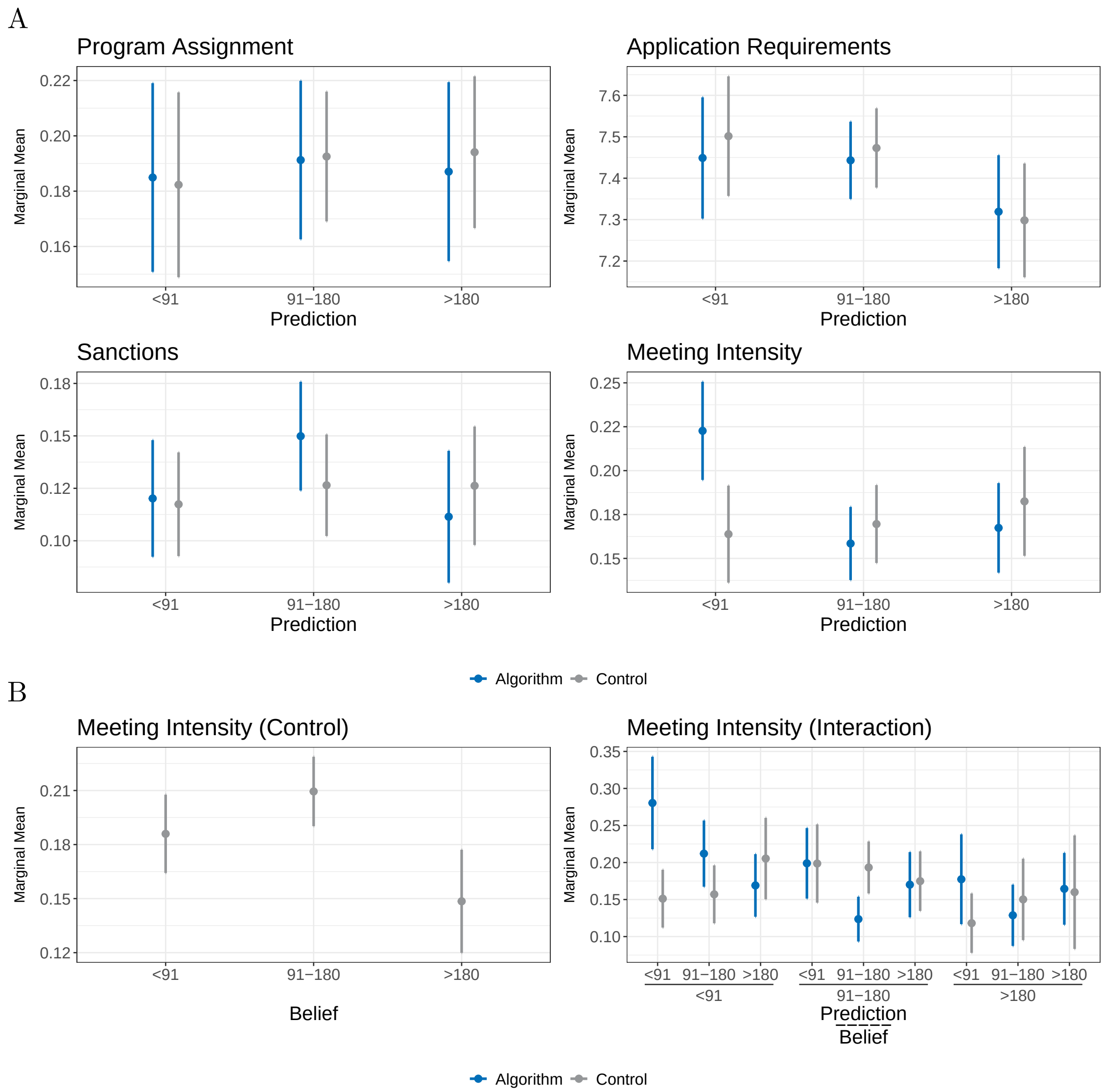

Figure 4: **Behavior Adjustment.** Change in caseworkers' behavior by treatment status, algorithmic prediction, and prior belief. Error bars show 95% confidence intervals with standard errors clustered at the caseworker level. Tables A.5–A.8 report estimates.

This pattern is opposite to the benchmark prediction in Section 3, under which higher predicted unemployment durations should induce greater effort. Instead, caseworkers increased effort for job seekers with comparatively favorable predicted prospects.

Panel B incorporates prior beliefs in addition to algorithmic predictions. Because meeting intensity is the only behavioral outcome displaying substantial heterogeneity across prediction groups, results are only shown for this outcome. Corresponding interaction estimates for the remaining behavioral outcomes are reported in Appendix Tables A.5–A.8. Panel B shows that the increase in meeting intensity for short predicted durations is not mirrored by prior beliefs alone. In the control group, meeting intensity does not systematically increase for job seekers believed to have short unemployment durations. Instead, the behavior is in line with institutional guides (Table A.1). The allocation pattern instead emerges conditional on the algorithmic prediction. This weakens standard cream-skimming explanations in which caseworkers systematically allocate greater effort to job seekers with better expected prospects independent of the algorithm. The meeting-intensity results are also robust to replacing the prediction categories with continuous prediction specifications, reported in Tables A.15–A.20. In both the linear and spline models, the treatment effect declines with predicted unemployment duration, consistent with the categorical specification. The continuous specifications therefore suggest that the findings are not mechanically driven by the 90- and 180-day thresholds.

Job Seekers' Behavior Adjustment. We next examine whether algorithmic predictions indirectly affected job seekers' behavior through their impact on caseworkers. We focus on two outcomes: the number of job applications submitted per month and the reservation wage reported by job seekers through the two caseworker information elicitations. Applications submitted measures the average number of job applications actually submitted per month, computed the same way as application requirements. Reservation wage adjustment measures the change in stated reservation wages (in CHF per month) between the initial and follow-up elicitations. The analysis sample for both measures is restricted to cases with complete elicitation data (N=1,174 for wage adjustment, N=3,013 for applications).

Figure 5 reports treatment effects by prediction and belief categories. We find no evidence that access to algorithmic predictions changed job seekers' application behavior. Across prediction groups, the number of submitted applications is similar between treatment and control groups. Thus, the behavioral responses documented above do not appear to operate through changes in observable search intensity. Reservation wages show a different pattern. Job seekers in the treatment group report somewhat lower reservation wages, particularly for short predicted unemployment durations. This suggests that the increased interaction documented in the previous subsection may have affected job seekers' expectations or willingness to accept available jobs.

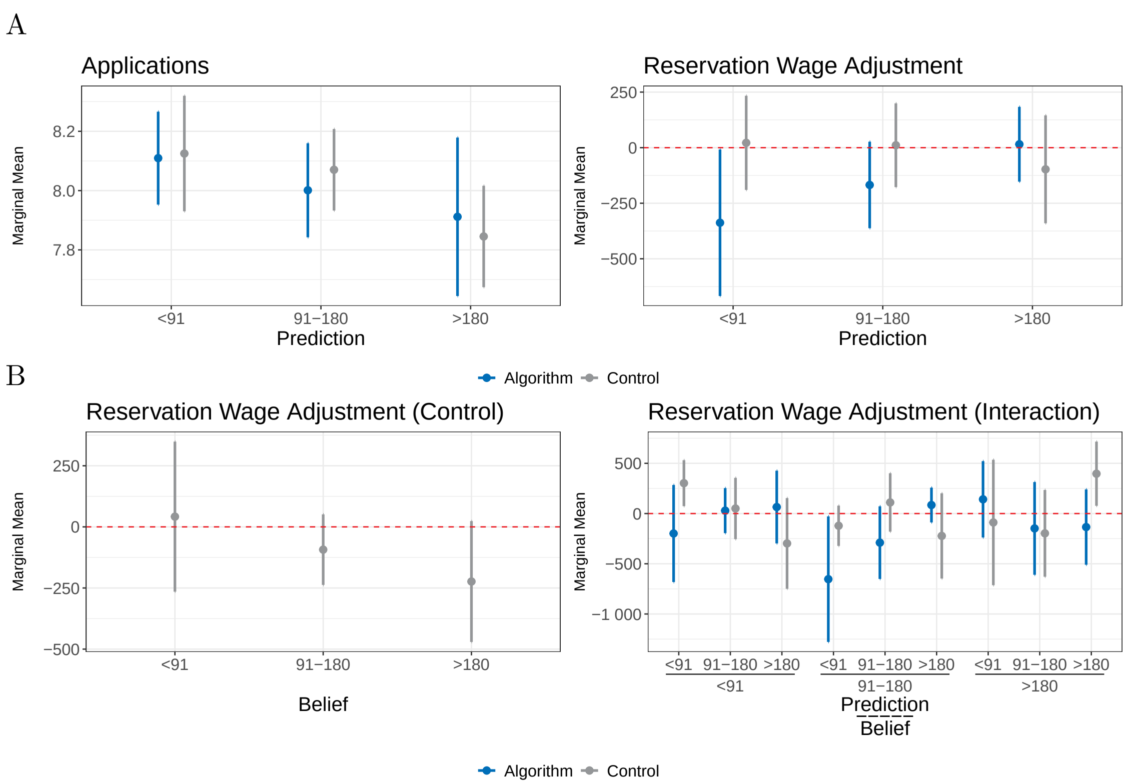

Figure 5: **Job Seekers' Behavior Adjustment.** Change in job seekers' behavior by treatment status, algorithmic prediction, and prior belief. Error bars show 95% confidence intervals with standard errors clustered at the caseworker level. Tables A.9 and A.10 report estimates.

Because reservation wages display heterogeneity across prediction groups, Panel B additionally incorporates prior beliefs and predictions jointly. The interaction patterns indicate that reservation wage responses vary across both dimensions, with the largest adjustments concentrated among job seekers with short predicted unemployment durations. Corresponding interaction estimates for application behavior are reported in Tables A.15–A.20. The reservation-wage results are also robust to continuous prediction specifications. Linear and spline models produce qualitatively similar patterns, indicating that the findings are not driven by the categorical thresholds.

Labor Market Outcomes. We next examine whether the behavioral responses documented above translated into changes in labor market outcomes. We define unemployment exits as de-registrations with transition into employment.

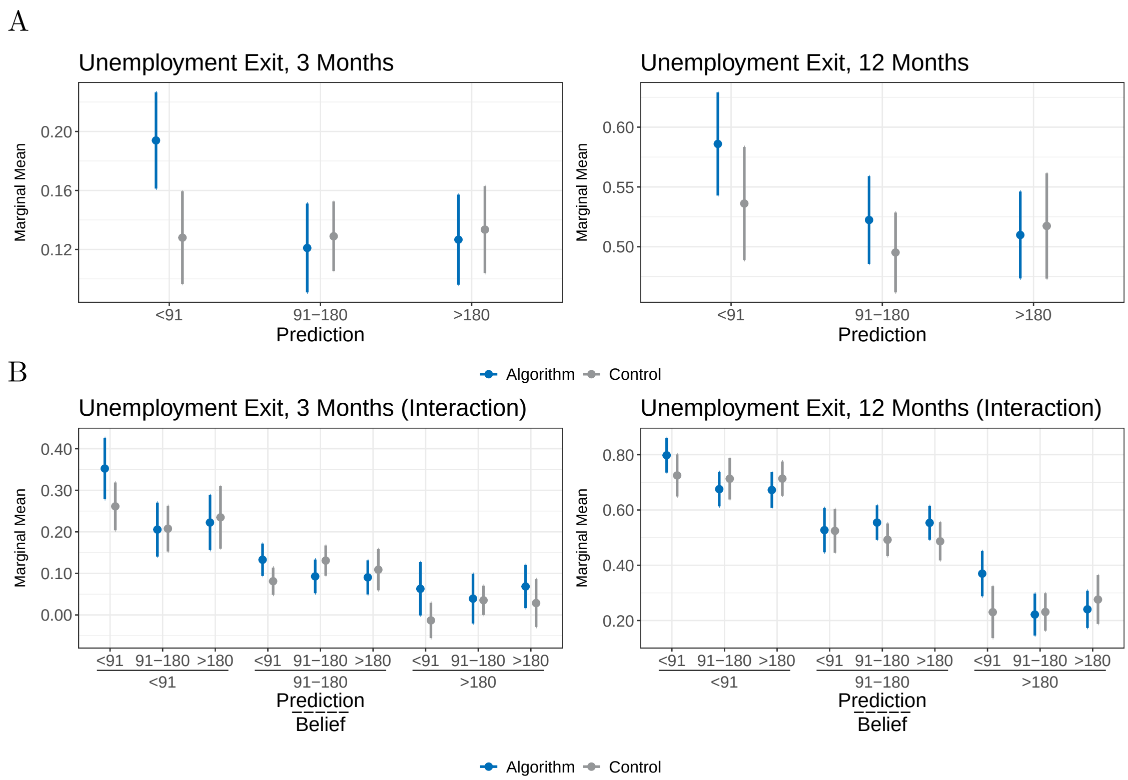

Figure 6: **Unemployment Duration.** Change in job seekers' exits to employment by treatment status, algorithmic prediction, and prior belief. Error bars show 95% confidence intervals with standard errors clustered at the caseworker level. Tables A.11 and A.12 report estimates.

Figure 6 shows that the intervention increased unemployment exits within 3 and 12 months primarily among job seekers with short predicted unemployment durations. Exit rates are higher in the treatment group for these job seekers, particularly at shorter horizons, while effects for medium- and long-duration predictions remain small throughout. The pattern mirrors the behavioral responses documented above: caseworkers increased meeting intensity for job seekers with favorable predicted prospects, and these job seekers subsequently exited unemployment more quickly. Panel B further shows that the exit effects are concentrated in cells with short predicted unemployment durations rather than short prior beliefs alone. This again weakens standard cream-skimming explanations based solely on caseworkers prioritizing job seekers they already believed to have favorable prospects. The unemployment-exit results are also robust to continuous prediction specifications. Linear and spline models yield qualitatively similar patterns, with treatment effects declining as predicted unemployment duration increases.

A remaining concern is whether faster exits came at the expense of job quality. To assess this, we examine post-unemployment job quality using average monthly earnings over the 24 months following unemployment exit and re-registration into unemployment within 24 months of exit. Both measures are derived from linked social security records. The sample size is slightly smaller due to right-censoring and incomplete data linkage.

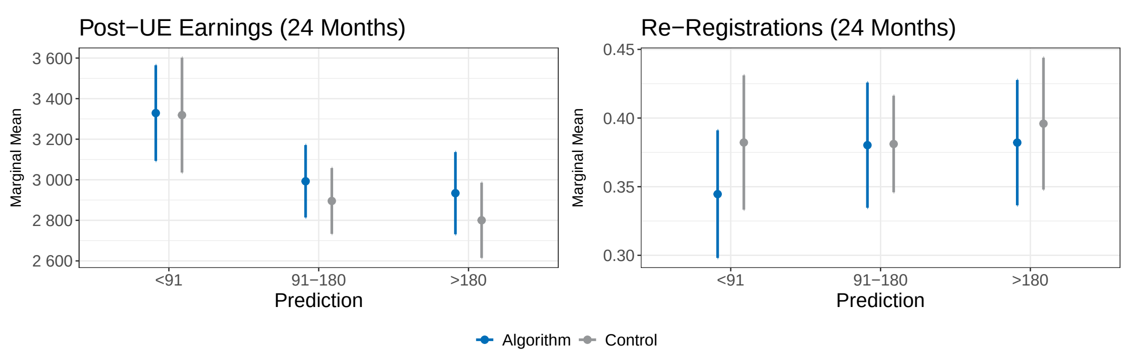

Figure 7: **Job Quality.** Change in job seekers' post-unemployment employment quality by treatment status and algorithmic prediction. Error bars show 95% confidence intervals with standard errors clustered at the caseworker level. Tables A.13 and A.14 report estimates.

Figure 7 shows no evidence that the intervention reduced job quality. Post-unemployment earnings and re-registration rates remain similar between treatment and control groups across prediction categories. Thus, while the algorithm changed caseworker behavior and accelerated exits among job seekers with short predicted unemployment durations, these effects do not appear to come at the expense of lower-quality employment matches.

6 Effort Distortion

We now develop a model showing how algorithmic predictions can distort effort allocation through reputational incentives. We return to the framework in Section 3. A caseworker forms a prior belief b about expected unemployment duration, receives an algorithmic prediction p in treatment, and chooses effort $e \in [0, 1]$ before duration y is realized. A key assumption is that supervisors observe predictions when evaluating performance based on realized outcomes.

Without predictions, supervisors observe outcomes but have limited information about case difficulty. Public predictions provide a benchmark for expected performance. When a case predicted to have a short duration instead lasts long, poor performance is harder to attribute to circumstances alone, changing caseworkers' reputational incentives.

Model. We model the caseworker response through a prediction-based performance threshold. Supervisors expect cases to resolve within a benchmark duration $t(p) = \kappa p$, with $\kappa > 0$, so that predicted difficult cases are allowed longer benchmark durations. A short prediction therefore creates a stricter benchmark.

We use two closely related versions of the model. The first applies the benchmark to expected duration. This isolates the basic benchmark-avoidance force: shorter predictions create stricter benchmarks and raise the effort threshold required to avoid blame. The second applies the benchmark to realized duration. This turns benchmark avoidance into a probabilistic problem: effort lowers the probability of failure rather than mechanically eliminating it. This second version characterizes attainability through the interior solution, showing explicitly that effort has little reputational value when the benchmark is already likely to be met or seems out of reach, and the highest value when expected duration is close to the benchmark. This links the public prediction to the caseworker's own prior belief.

We now formalize these two versions in turn, starting with the sharp case in which reputational evaluation is based on expected benchmark failure. Using the expected duration after effort from Section 3, $\mathbb{E}[y \mid e] = \tilde{b}(1-e)$, the caseworker's payoff is

$$\pi(e) = -\alpha\tilde{b}(1-e) - \frac{\beta}{2}e^2 - \gamma\mathbb{1}\{\tilde{b}(1-e) > \kappa p\},$$

where $\alpha > 0$ captures the cost of poor outcomes, $\beta > 0$ governs effort costs, and $\gamma > 0$ is the reputational cost of failing the benchmark. Absent the reputational penalty, optimal effort is $e^0(\tilde{b}) = \alpha\tilde{b}/\beta$. The reputational term creates a benchmark-avoidance motive. If the penalty is sufficiently large, the caseworker wants to be in the no-penalty region whenever this is feasible. The benchmark therefore imposes a lower bound on effort. The minimum effort that satisfies the benchmark is

$$e(\tilde{b}, p) = 1 - \frac{\kappa p}{\tilde{b}}.$$

This expression applies in the interior case in which the benchmark binds and is attainable, $0 < \kappa p < \tilde{b}$.[10]

We next replace expected benchmark failure with realized benchmark failure, which turns benchmark avoidance into a probabilistic problem. The previous version treats expected duration as the object evaluated against the benchmark. In practice, however, supervisors observe realized duration, while effort affects expected duration. We therefore write realized

[10] If $\kappa p \geq \tilde{b}$, the benchmark is already satisfied at zero effort. If the required effort exceeds one, the benchmark cannot be met with feasible effort. In the binding region, $e(\tilde{b}, p)$ is the benchmark-avoidance threshold, not necessarily the full optimum. If the unconstrained optimum $e^0(\tilde{b})$ already exceeds this threshold, the caseworker chooses $e^0(\tilde{b})$ while still avoiding the penalty.

duration as the conditional mean plus an idiosyncratic component, $y = \tilde{b}(1-e) + \varepsilon$, where $\mathbb{E}[\varepsilon \mid e, b, p] = 0$. The reputational cost γ is unchanged, but effort now reduces the probability of benchmark failure rather than deterministically eliminating it. The payoff is

$$\pi(e) = -\alpha\tilde{b}(1-e) - \frac{\beta}{2}e^2 - \gamma\left[1 - F\left(\kappa p - \tilde{b}(1-e)\right)\right],$$

where F is the cdf of ε. The first-order condition for an interior optimum is

$$\alpha\tilde{b} - \beta e + \gamma\tilde{b}f\left(\kappa p - \tilde{b}(1-e)\right) = 0.$$

The final term is the marginal reputational return to effort. It captures how much additional effort lowers the probability that the realized duration exceeds the benchmark.[11]

This expression formalizes perceived attainability. The marginal reputational return to effort, $\gamma\tilde{b}f(\kappa p - \tilde{b}(1-e))$, is maximized when the argument of f is near zero; that is, when expected duration $\tilde{b}(1-e)$ is close to the benchmark κp. If expected duration is far below the benchmark, failure is already unlikely and additional effort has little reputational value. If expected duration is far above the benchmark, failure remains likely even with additional effort, making marginal effort similarly unproductive. Effort therefore responds most strongly when the benchmark is demanding but attainable.

Empirical Predictions. The first implication is that predictions can reverse the intended allocation of effort. The policy objective was to direct additional effort toward job seekers with long predicted unemployment durations. The benchmark mechanism instead creates reputational pressure around short predictions, because short predictions imply stricter public benchmarks. In the deterministic model, this logic is direct. In the region where the benchmark binds and is attainable, holding posterior expected duration fixed, the benchmark-

[11] The first-order condition is used to characterize the local marginal incentive created by the benchmark. A unique global optimum would require additional regularity conditions on f.

avoidance effort threshold decreases in the algorithmic prediction:

$$\frac{\partial e}{\partial p} = -\frac{\kappa}{\tilde{b}} < 0.$$

If posterior beliefs also depend on the prediction, with $\tilde{b} = \lambda b + (1-\lambda)p$, the total derivative is

$$\frac{de}{dp} = -\frac{\kappa \lambda b}{[\lambda b + (1-\lambda)p]^2} \leq 0.$$

Thus, within the binding region, short algorithmic predictions raise the effort needed to avoid benchmark failure because they make the performance benchmark harder to satisfy.

The second implication concerns the interaction between predictions and prior beliefs. A short prediction creates a strict benchmark, but effort responds most strongly when the caseworker believes that this benchmark is still within reach. The second version captures this by making the marginal reputational return to effort largest near the benchmark: $\tilde{b}(1-e) \approx \kappa p$. Substituting $\tilde{b} = \lambda b + (1-\lambda)p$, the benchmark-margin condition becomes

$$[\lambda b + (1-\lambda)p]\,(1-e) \approx \kappa p.$$

This condition links the public prediction to the caseworker's own belief. When the algorithm predicts a short duration and the caseworker also expects a short duration, the benchmark is demanding but attainable. Additional effort can then meaningfully reduce the probability of benchmark failure. When the algorithm predicts a short duration but the caseworker expects a long duration, the benchmark is demanding but distant from the caseworker's expected outcome. In that case, the caseworker may think that failure is likely regardless of marginal effort, so the reputational return to effort is weaker.

The model therefore predicts that effort should increase for job seekers with short predicted unemployment durations, especially when short predictions confirm caseworkers' own prior beliefs. This is the pattern we observe in the data. Caseworkers increase meeting intensity for job seekers with short predicted unemployment durations, and this response is concentrated among cases in which short predictions confirm short prior beliefs (Figure 4, Panel B, right).

Contribution. Cream-skimming in Heckman et al. (1997, 2002) operates through the baseline payoff function: incentives favor easy cases, or effort is believed to be more productive for them. Such mechanisms should affect effort allocation in both treatment and control groups because they reflect the underlying incentive structure or production technology. Our mechanism operates through observability. The algorithm does not change formal incentives or effort productivity; it changes what supervisors observe about case difficulty. The benchmark $t(p) = \kappa p$ exists only when predictions are public, creating an attribution standard absent in the control group. Consistent with this distinction, we observe the effort-risk relationship only in treatment. In the control group, caseworker beliefs are not related to effort allocation (Figure 4, Panel B, left), which weighs against productivity-based cream-skimming as the primary explanation.

The mechanism also differs from recommendation-based mechanisms in Albright (2023) and McLaughlin and Spiess (2025). In those settings, algorithmic recommendations can change behavior by altering the cost of following or deviating from a prescribed action; recommendations can provide reputational cover or make deviation costly. In our setting, predictions do not prescribe actions but establish public outcome benchmarks. They therefore affect accountability by changing how realized outcomes are attributed to worker effort versus case circumstances. Recommendation-based mechanisms operate through the cost of following or deviating from prescribed actions, whereas prediction-based benchmarking operates through the attribution of outcomes relative to a public performance benchmark.

7 Conclusion

Algorithms are increasingly deployed to guide resource allocation in settings where decision-makers retain discretion over implementation. We study how caseworkers in public employment services respond to algorithmic profiling predictions in a field experiment conducted in the Swiss canton of Fribourg. We find that caseworkers respond to algorithmic signals, but not in ways aligned with policy objectives. Rather than shifting effort toward job seekers at risk of long-term unemployment, caseworkers increased effort for job seekers predicted to exit unemployment quickly. To interpret these findings, we develop a model in which caseworkers allocate effort under reputational concerns. When algorithmic predictions make case difficulty publicly observable, supervisors can more easily attribute poor outcomes to inadequate effort rather than difficult circumstances. Predictions therefore create performance benchmarks that concentrate reputational pressure on cases expected to leave unemployment quickly, inducing caseworkers to shift effort toward those cases.

Our findings contribute to a broader understanding of how algorithms interact with organizational incentives. First, they show that algorithmic tools can distort behavior even when humans retain full discretion and algorithms provide only information rather than recommendations. Second, they demonstrate that the effects of algorithmic predictions depend critically on institutional context. The mechanism we identify extends beyond unemployment services. Predictions may distort effort allocation whenever supervisors evaluate workers based on outcomes, case difficulty varies independently of worker effort, and algorithms make that difficulty observable. These conditions characterize many domains in which algorithmic decision support is increasingly used, including healthcare, education, criminal justice, and social services. In each setting, predictions can alter how supervisors attribute outcomes to effort versus circumstances, thereby reshaping effective incentives even when formal rules remain unchanged.

Several features of our implementation merit discussion. First, the algorithm's predictive accuracy was substantially lower in the field than anticipated during calibration. The model explained less than one percent of the variance in realized unemployment durations. While this limited the intervention's effectiveness as originally conceived, it created valuable variation for studying behavioral responses to algorithmic signals when predictions diverge from prior beliefs. Still, future implementations would require substantially more accurate prediction models to realize the intended policy benefits of early identification and targeting. Second, we do not observe caseworkers' beliefs about heterogeneous treatment effects. Caseworkers may have informal priors about which job seekers are most responsive to interventions, independent of risk predictions. Without data on these beliefs, we cannot fully assess whether targeting decisions reflected rational expectations about differential treatment returns across job seeker types. Third, caseworker resistance to the algorithm was substantial. As discussed in more detail in the background report (see Arni and Schiprowski, 2015), caseworkers found the supplemental questionnaire cumbersome, time-consuming, and duplicative of existing documentation obligations. This resulted in delayed entries and, in some cases, missing data that further diluted the intervention's impact. Early in the implementation, caseworkers largely refused to complete the questionnaire until supervisors issued additional instructions making its use quasi-mandatory.

Taken together, our findings highlight that the effects of algorithms depend on how predictive information interacts with organizational incentives and accountability structures. As algorithmic tools become increasingly embedded into public sector decision-making, understanding these institutional interactions will be central to designing systems that improve outcomes rather than distort effort allocation.

References

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Press.

Albright, Alex. 2023. "The Hidden Effects of Algorithmic Recommendations." *Working Paper.*

Altmann, Steffen, Armin Falk, Simon Jäger, and Florian Zimmermann. 2018. "Learning About Job Search: A Field Experiment With Job Seekers in Germany." *Journal of Public Economics* 164 33–49.

Argyris, Chris. 1990. *Overcoming Organizational Defenses: Facilitating Organizational Learning*. Boston: Allyn and Bacon.

Arni, Patrick, Gerard J. van den Berg, and Rafael Lalive. 2022. "Treatment Versus Regime Effects of Carrots and Sticks." *Journal of Business & Economic Statistics* 40 (1): 111–127.

Arni, Patrick, Rafael Lalive, and Jan C. Van Ours. 2013. "How Effective Are Unemployment Benefit Sanctions? Looking Beyond Unemployment Exit." *Journal of Applied Econometrics* 28 (7): 1153–1178.

Arni, Patrick, and Amelie Schiprowski. 2015. "Die Rolle von Erwartungshaltungen in der Stellensuche und der RAV-Beratung. Teilprojekt 2: Pilotprojekt Jobchancen-Barometer." Publikation Arbeitsmarktpolitik 43, State Secretariat for Economic Affairs (SECO).

Arni, Patrick, and Amelie Schiprowski. 2019. "Job Search Requirements, Effort Provision and Labor Market Outcomes." *Journal of Public Economics* 169 65–88.

Artinger, Florian M., Sabrina Artinger, and Gerd Gigerenzer. 2019. "C.Y.A.: Frequency and Causes of Defensive Decisions in Public Administration." *Business Research* 12 (1): 9–25.

Ashforth, Blake E., and Raymond T. Lee. 1990. "Defensive Behavior in Organizations: A Preliminary Model." *Human Relations* 43 (7): 621–648.

Baker, George P. 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy* 100 (3): 598–614.

Belot, Michele, Philipp Kircher, and Paul Muller. 2019. "Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice." *The Review of Economic Studies* 86 (4): 1411–1447.

Black, Dan A., Jose Galdo, and Jeffrey A. Smith. 2007. "Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach." *American Economic Review* 97 (2): 104–107.

Böheim, René, Rainer Eppel, Helmut Mahringer et al. 2022. "More Caseworkers Shorten Unemployment Durations and Save Costs. Results from a Field Experiment in an Austrian Public Employment Office." *WIFO*.

Bundesrat. 2013. "Wirksamkeit und Effizienz der öffentlichen Arbeitsvermittlung." *Schweizerische Eidgenossenschaft*.

Burgess, Simon, and Marisa Ratto. 2003. "The Role of Incentives in the Public Sector: Issues and Evidence." *Oxford Review of Economic Policy* 19 (2): 285–300.

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review* 105 (7): 2183–2203.

Card, David, Jochen Kluve, and Andrea Weber. 2018. "What works? A meta analysis of recent active labor market program evaluations." *Journal of the European Economic Association* 16 (3): 894–931.

Cederlöf, Jonas, Martin Söderström, and Johan Vikström. 2021. "What Makes a Good Caseworker?" *Working Paper*.

Chan, David C, Matthew Gentzkow, and Chuan Yu. 2022. "Selection with Variation in Diagnostic Skill: Evidence from Radiologists." *Quarterly Journal of Economics* 137 (2): 729–783.

Chetty, Raj. 2008. "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy* 116 (2): 173–234.

Conlon, John J., Laura Pilossoph, Matthew Wiswall, and Basit Zafar. 2018. "Labor Market Search with Imperfect Information and Learning." Working Paper 24988, National Bureau of Economic Research.

Courty, Pascal, and Gerald Marschke. 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives." *Journal of Labor Economics* 22 (1): 23–56.

Dargnies, Marie-Pierre, Rustamdjan Hakimov, and Dorothea Kübler. 2024. "Aversion to Hiring Algorithms: Transparency, Gender Profiling, and Self-Confidence." *Management Science*.

De-Arteaga, Maria, Riccardo Fogliato, and Alexandra Chouldechova. 2020. "A Case for Humans-In-The-Loop: Decisions in the Presence of Erroneous Algorithmic Scores." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12.

Desiere, Sam, Kristine Langenbucher, and Ludo Struyven. 2019. "Statistical Profiling in Public Employment Services: An International Comparison." *OECD Social, Employment and Migration Working Papers* No. 224.

Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey. 2015. "Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err." *Journal of Experimental Psychology: General* 144 (1): 114.

Dixit, Avinash. 2002. "Incentives and Organizations in the Public Sector: An Interpretative Review." *Journal of Human Resources* 696–727.

Ely, Jeffrey C., and Juuso Välimäki. 2003. "Bad Reputation." *Quarterly Journal of Economics* 118 (3): 785–814.

Epstein, Larry G., and Martin Schneider. 2007. "Learning under Ambiguity." *The Review of Economic Studies* 74 (4): 1275–1303.

Ernst, Sebstian, Andreas I. Mueller, and Johannes Spinnewijn. 2024. "Risk Scores for Long-Term Unemployment and the Assignment to Job Search Counseling." *AEA Papers and Proceedings*.

European Commission. 2021. "Artificial Intelligence Act." *European Parliament*.

FSO. 2008. *NOGA 2008: General Classification of Economic Activities*. Neuchâtel: Federal Statistical Office, Switzerland.

Gibbons, Robert. 1998. "Incentives in Organizations." *Journal of Economic Perspectives* 12 (4): 115–132.

Gibbons, Robert, and Kevin J. Murphy. 1992. "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." *Journal of Political Economy* 100 (3): 468–505.

Goel, Sharad, Justin M. Rao, and Ravi Shroff. 2016. "Personalized Risk Assessments in the Criminal Justice System." *American Economic Review* 106 (5): 119–123.

Grimon, Marie-Pascale, and Christopher Mills. 2022. "The Impact of Algorithmic Tools on Child Protection: Evidence from a Randomized Controlled Trial." *Working Paper.*

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. "Designing Information Provision Experiments." *Journal of Economic Literature* 61 (1): 3–40.

Harris, Adam, and Maggie Yellen. 2024. "Decision-Making with Machine Prediction: Evidence from Predictive Maintenance in Trucking." *Working Paper.*

Heckman, James, Carolyn Heinrich, and Jeffrey Smith. 1997. "Assessing the Performance of Performance Standards in Public Bureaucracies." *American Economic Review* 87 (2): 389–395.

Heckman, James J., Carolyn Heinrich, and Jeffrey Smith. 2002. "The Performance of Performance Standards.." *Journal of Human Resources* 37 (4): 778–811.

Heinrich, Carolyn J., and Gerald Marschke. 2010. "Incentives and Their Dynamics in Public Sector Performance Management Systems." *Journal of Policy Analysis and Management* 29 (1): 183–208.

Holmström, Bengt. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *The Review of Economic Studies* 66 (1): 169–182.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization* 7 24–52.

Kaltenborn, Bruno, and Petra Kaps. 2013. "Steuerung der öffentlichen Arbeitsvermittlung in der Schweiz." Publikation Arbeitsmarktpolitik 35, State Secretariat for Economic Affairs (SECO).

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human Decisions and Machine Predictions." *Quarterly Journal of Economics* 133 (1): 237–293.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer. 2015. "Prediction Policy Problems." *American Economic Review* 105 (5): 491–495.

Koning, Pierre, and Carolyn J. Heinrich. 2013. "Cream-Skimming, Parking and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts." *Journal of Policy Analysis and Management* 32 (3): 461–483.

Körtner, John, and Giuliano Bonoli. 2023. "Predictive Algorithms in the Delivery of Public Employment Services." In *Handbook of Labour Market Policy in Advanced Democracies*, 387–398, Edward Elgar Publishing.

Lalive, Rafael, Jan C. Van Ours, and Josef Zweimüller. 2005. "The Effect of Benefit Sanctions on the Duration of Unemployment." *Journal of the European Economic Association* 3 (6): 1386–1417.

Le Barbanchon, Thomas, Johannes Schmieder, and Andrea Weber. 2024. "Job Search, Unemployment Insurance, and Active Labor Market Policies." In *Handbook of Labor Economics*, Volume 5. 435–580, Elsevier.

Ludwig, Jens, and Sendhil Mullainathan. 2021. "Fragile Algorithms and Fallible Decision-Makers: Lessons from the Justice System." *Journal of Economic Perspectives* 35 (4): 71–96.

Maibom, Jonas, Michael Rosholm, and Michael Svarer. 2017. "Experimental Evidence on the Effects of Early Meetings and Activation." *The Scandinavian Journal of Economics* 119 (3): 541–570.

Mastrobuoni, Giovanni. 2020. "Crime Is Terribly Revealing: Information Technology and Police Productivity." *The Review of Economic Studies* 87 (6): 2727–2753.

McLaughlin, Bryce, and Jann Spiess. 2025. "Algorithmic Assistance with Recommendation-Dependent Preferences." *Working Paper: arXiv:2208.07626.*

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat. 2022. "Managing Self-Confidence: Theory and Experimental Evidence." *Management Science* 68 (11): 7793–7817.

Moffitt, Robert. 1985. "Unemployment Insurance and the Distribution of Unemployment Spells." *Journal of Econometrics* 28 (1): 85–101.

Mueller, Andreas I., and Johannes Spinnewijn. 2025. "The Nature of Long-Term Unemployment: Predictability, Heterogeneity, and Selection." *Journal of Political Economy* 133 (12): 3846–3902.

Mueller, Andreas I., Johannes Spinnewijn, and Giorgio Topa. 2021. "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias." *American Economic Review* 111 (1): 324–363.

Mullainathan, Sendhil, and Ziad Obermeyer. 2022. "Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care." *The Quarterly Journal of Economics* 137 (2): 679–727.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37 (1): 7–63.

Prendergast, Canice. 2003. "The Limits of Bureaucratic Efficiency." *Journal of Political Economy* 111 (5): 929–958.

Rabin, Matthew, and Joel L Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *Quarterly Journal of Economics* 114 (1): 37–82.

Ribers, Michael Allan, and Hannes Ullrich. 2023. "Machine Predictions and Human Decisions with Variation in Payoff and Skills: The Case of Antibiotic Prescribing." *Working Paper* October, 2023.

Roth, Christopher, and Johannes Wohlfart. 2020. "How Do Expectations About the Macroeconomy Affect Personal Expectations and Behavior?" *The Review of Economics and Statistics* 102 (4): 731–748.

Sadka, Joyce, Enrique Seira, and Christopher Woodruff. 2024. "Information and Bargaining through Agents: Experimental Evidence from Mexico's Labour Courts." *The Review of Economic Studies* rdae003.

Schiprowski, Amelie. 2020. "The Role of Caseworkers in Unemployment Insurance: Evidence from Unplanned Absences." *Journal of Labor Economics* 38 (4): 1189–1225.

SECO. 2012. "Mitteilung RAV/LAM/KAST 2012/10." *Staatssekretariat für Wirtschaft, Schweizerische Eidgenossenschaft.*

Spinnewijn, Johannes. 2015. "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs." *Journal of the European Economic Association* 13 (1): 130–167.

Stevenson, Megan T, and Jennifer L Doleac. 2024. "Algorithmic Risk Assessment in the Hands of Humans." *American Economic Journal: Economic Policy* 16 (4): 382–414.

Zimmermann, Florian. 2020. "The Dynamics of Motivated Beliefs." *American Economic Review* 110 (2): 337–363.

Appendix

Context

Employability	job seeker Group	Strategy	Activity / Intensity
High	Employable without restrictions	Integration	Meeting at least once every two months
Medium	Employable with slight obstacles	Support	Meeting once or twice every month and job search assistance or individual course program
	Not adequately employable with development potential	Qualification	Meeting once or twice every month and internship or training subsidy
	Low initiative and work motivation	Verification of cooperation	Meeting once or twice every month and temporary employment program
Low	High risk of long-term unemployment	Clarification	Meeting once per month and temporary employment program
	High risk of UI benefits expiry	Coordination	Meeting once every two months and support through interinstitutional cooperation (IIZ)

Table A.1: **Guidelines.** Triage in the Swiss UI system. Translated extracts from Mitteilung RAV/LAM/ KAST 2012/10, June 26, 2012 (SECO, 2012).

	Indicator	Operationalization	Weight
(1)	Rapid Reintegration	Number of daily benefit payments	50%
(2)	Long-Term Unemployment	Number of job seekers unemployed after 12 months	20%
(3)	Expiry of Unemployment Benefits	Number of job seekers reaching the max. of benefit payments	20%
(4)	Reregistrations	Number of job seekers who re-register	10%

Table A.2: **Evaluation.** Performance Indicators in the Swiss UI system.

Jobchancen-Barometer

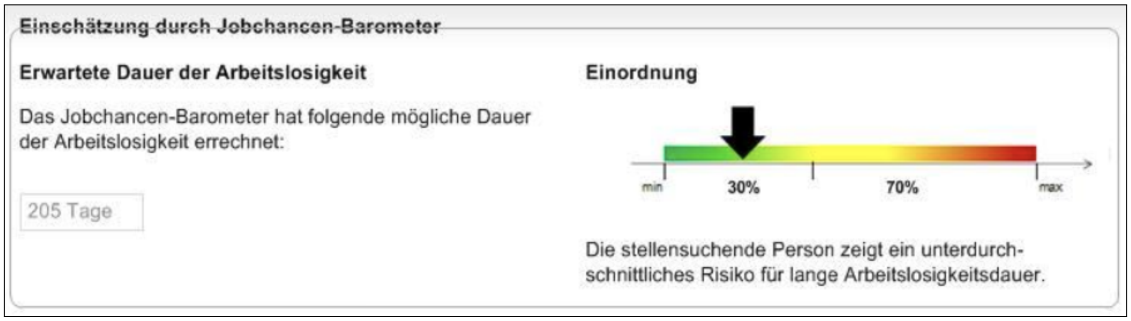

Figure A.1: **Barometer.** The "barometer" in the interface of the algorithm. Screenshot from Arni and Schiprowski (2015).

#	Risk Group	Range (in days)			
1	F, <30, German	0 – 66	67 – 87	88 – 196	197 –
2	M, <30, German	0 – 62	63 – 109	110 – 186	187 –
3	F, 30-50, German	0 – 68	69 – 100	101 – 184	185 –
4	M, 30-50, German	0 – 76	77 – 120	121 – 166	167 –
5	F, >50, German	0 – 94	95 – 159	160 – 211	212 –
6	M, >50, German	0 – 75	76 – 104	105 – 187	188 –
7	F, <30, French	0 – 74	75 – 92	93 – 186	187 –
8	M, <30, French	0 – 90	91 – 133	134 – 184	185 –
9	F, 30-50, French	0 – 71	72 – 95	96 – 181	182 –
10	M, 30-50, French	0 – 77	78 – 109	110 – 187	188 –
11	F, >50, French	0 – 65	66 – 84	85 – 202	203 –
12	M, >50, French	0 – 108	109 – 144	145 – 199	200 –

Table A.3: **"Barometer" Categories.** The 12 groups were defined by job seekers' interacted characteristics: gender, age groups (below 30, 30-50, above 50), and language region: German-speaking vs French-speaking areas. The shape of the baseline hazards was flexibly estimated using a piecewise-constant model (split into 16 pieces) with each of the 12 sub-groups.

Descriptives

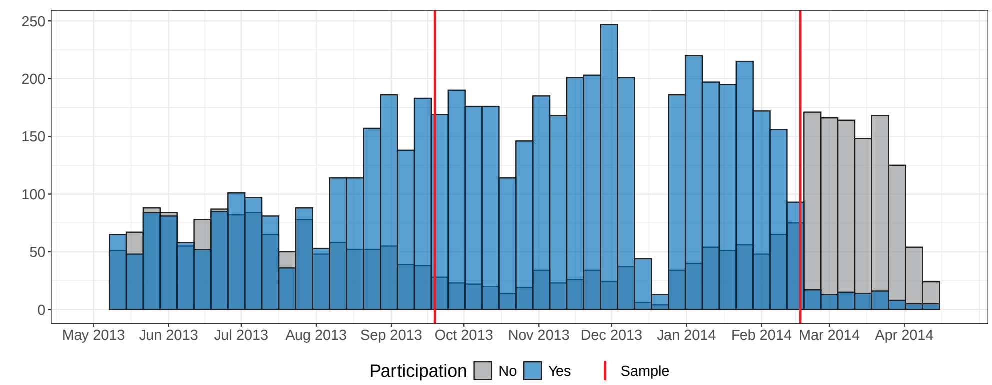

Figure A.2: **Sample Selection.** We focus on job seekers registered between October 2013 and February 2014. Prior to October 2013, caseworker participation rates were low. A training intervention in September 2013 reminded caseworkers of the importance of the pilot project, after which participation stabilized at approximately 90% (Arni and Schiprowski, 2015, Chapter 3). The final sample size is 3,192, slightly smaller than the 3,276 reported in the background report. The difference is due to updated administrative data, excluded job seekers with multiple participations in the same unemployment spell, and excluded caseworkers who only appear in control or treatment group

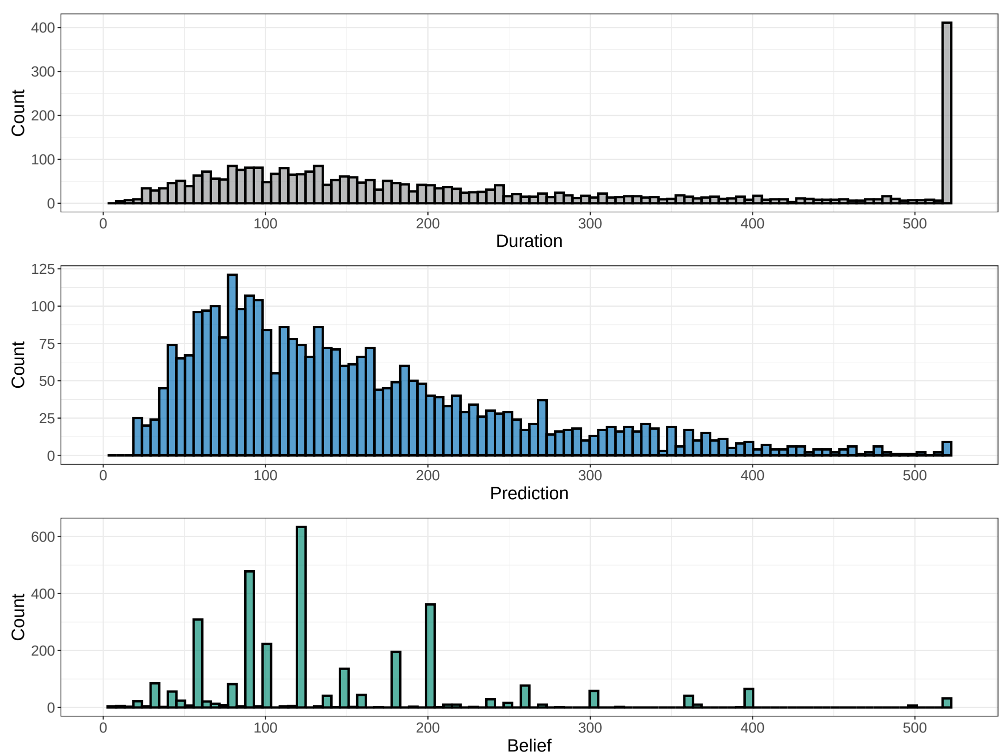

Figure A.3: **Durations, Predictions, and Prior Beliefs.** Distributions of realized unemployment duration, algorithmic predictions, and caseworker beliefs in the experimental sample. All variables are measured in days and top-coded at 520 days.

Results

	Belief Update		
Algorithm	-7.58	-5.71	-4.13
	(4.72)	(6.95)	(7.44)
Prediction < 91		6.92	12.6
		(7.72)	(11.3)
Prediction > 180		-8.30	6.50
		(10.2)	(16.8)
Algorithm, Prediction < 91		-18.0	-22.1**
		(11.9)	(10.7)
Algorithm, Prediction > 180		9.15	1.47
		(11.7)	(16.6)
Belief < 91			45.1***
			(10.8)
Belief > 180			-47.4***
			(15.5)
Algorithm, Belief < 91			-5.16
			(12.9)
Algorithm, Belief > 180			0.745
			(19.2)
Prediction < 91, Belief < 91			-8.80
			(18.9)
Prediction > 180, Belief < 91			-19.5
			(19.0)
Prediction < 91, Belief > 180			-34.1
			(25.4)
Prediction > 180, Belief > 180			-37.7
			(26.4)
Algorithm, Prediction < 91, Belief < 91			17.6
			(17.5)
Algorithm, Prediction > 180, Belief < 91			12.6
			(22.3)
Algorithm, Prediction < 91, Belief > 180			7.09
			(34.9)
Algorithm, Prediction > 180, Belief > 180			21.9
			(32.2)
N	1,174	1,174	1,174

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.4: **Belief Updating.** Estimates for belief updating between the first and second information elicitation. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Pr(ALMP)		
Algorithm	-0.002	-0.001	-0.012
	(0.011)	(0.020)	(0.030)
Prediction < 91		-0.010	0.002
		(0.024)	(0.039)
Prediction > 180		0.002	-0.014
		(0.021)	(0.031)
Algorithm, Prediction < 91		0.004	0.006
		(0.034)	(0.049)
Algorithm, Prediction > 180		-0.006	0.036
		(0.032)	(0.053)
Belief < 91			-0.052
			(0.039)
Belief > 180			-0.044
			(0.037)
Algorithm, Belief < 91			-0.002
			(0.051)
Algorithm, Belief > 180			0.055
			(0.053)
Prediction < 91, Belief < 91			-0.064
			(0.058)
Prediction > 180, Belief < 91			-0.001
			(0.049)
Prediction < 91, Belief > 180			0.059
			(0.059)
Prediction > 180, Belief > 180			0.070
			(0.048)
Algorithm, Prediction < 91, Belief < 91			0.028
			(0.075)
Algorithm, Prediction > 180, Belief < 91			-0.052
			(0.070)
Algorithm, Prediction < 91, Belief > 180			-0.069
			(0.092)
Algorithm, Prediction > 180, Belief > 180			-0.107
			(0.083)
N	3,183	3,183	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.5: **Program Assignment.** Estimates for program (ALMP) assignment. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Application Requirements		
Algorithm	-0.022	-0.030	-0.042
	(0.052)	(0.066)	(0.115)
Prediction < 91		0.029	0.0008
		(0.103)	(0.116)
Prediction > 180		-0.175*	-0.266*
		(0.096)	(0.142)
Algorithm, Prediction < 91		-0.023	-0.011
		(0.139)	(0.192)
Algorithm, Prediction > 180		0.051	0.120
		(0.128)	(0.206)
Belief < 91			-0.066
			(0.123)
Belief > 180			-0.150
			(0.206)
Algorithm, Belief < 91			0.119
			(0.176)
Algorithm, Belief > 180			-0.169
			(0.213)
Prediction < 91, Belief < 91			-0.059
			(0.168)
Prediction > 180, Belief < 91			0.056
			(0.194)
Prediction < 91, Belief > 180			0.217
			(0.337)
Prediction > 180, Belief > 180			0.322
			(0.284)
Algorithm, Prediction < 91, Belief < 91			0.043
			(0.237)
Algorithm, Prediction > 180, Belief < 91			-0.103
			(0.294)
Algorithm, Prediction < 91, Belief > 180			-0.075
			(0.390)
Algorithm, Prediction > 180, Belief > 180			-0.118
			(0.374)
N	2,572	2,572	2,572

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.6: **Application Requirements.** Estimates for application requirements. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Pr(Sanction)		
Algorithm	0.005	0.023	0.068**
	(0.011)	(0.020)	(0.032)
Prediction < 91		-0.009	0.002
		(0.019)	(0.034)
Prediction > 180		-0.0003	0.048
		(0.020)	(0.030)
Algorithm, Prediction < 91		-0.021	-0.045
		(0.029)	(0.047)
Algorithm, Prediction > 180		-0.038	-0.106**
		(0.032)	(0.047)
Belief < 91			0.032
			(0.028)
Belief > 180			0.046
			(0.038)
Algorithm, Belief < 91			-0.056
			(0.041)
Algorithm, Belief > 180			-0.104**
			(0.051)
Prediction < 91, Belief < 91			-0.033
			(0.044)
Prediction > 180, Belief < 91			-0.079*
			(0.041)
Prediction < 91, Belief > 180			0.012
			(0.050)
Prediction > 180, Belief > 180			-0.083*
			(0.043)
Algorithm, Prediction < 91, Belief < 91			-0.009
			(0.058)
Algorithm, Prediction > 180, Belief < 91			0.117*
			(0.064)
Algorithm, Prediction < 91, Belief > 180			0.111
			(0.071)
Algorithm, Prediction > 180, Belief > 180			0.106
			(0.069)
N	3,183	3,183	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.7: **Sanctions.** Estimates for sanctions. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Pr(2nd Meeting within 30 Days)		
Algorithm	0.009	-0.011	-0.070***
	(0.011)	(0.017)	(0.022)
Prediction < 91		-0.006	0.005
		(0.020)	(0.036)
Prediction > 180		0.013	-0.018
		(0.022)	(0.025)
Algorithm, Prediction < 91		0.070***	0.070
		(0.026)	(0.043)
Algorithm, Prediction > 180		-0.004	0.065*
		(0.026)	(0.036)
Belief < 91			-0.036
			(0.028)
Belief > 180			-0.043
			(0.033)
Algorithm, Belief < 91			0.125***
			(0.034)
Algorithm, Belief > 180			0.048
			(0.043)
Prediction < 91, Belief < 91			-0.011
			(0.045)
Prediction > 180, Belief < 91			0.067
			(0.044)
Prediction < 91, Belief > 180			-0.037
			(0.049)
Prediction > 180, Belief > 180			0.028
			(0.064)
Algorithm, Prediction < 91, Belief < 91			0.004
			(0.061)
Algorithm, Prediction > 180, Belief < 91			-0.156***
			(0.051)
Algorithm, Prediction < 91, Belief > 180			0.011
			(0.072)
Algorithm, Prediction > 180, Belief > 180			-0.039
			(0.076)
N	3,183	3,183	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.8: **Meeting Intensity.** Estimates for meeting intensity. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Applications		
Algorithm	-0.014	-0.086	0.109
	(0.103)	(0.144)	(0.229)
Prediction < 91		0.069	0.019
		(0.153)	(0.262)
Prediction > 180		-0.281*	-0.508**
		(0.145)	(0.208)
Algorithm, Prediction < 91		0.067	-0.171
		(0.209)	(0.357)
Algorithm, Prediction > 180		0.169	0.167
		(0.255)	(0.319)
Belief < 91			0.112
			(0.261)
Belief > 180			-0.254
			(0.289)
Algorithm, Belief < 91			-0.272
			(0.334)
Algorithm, Belief > 180			-0.495
			(0.371)
Prediction < 91, Belief < 91			-0.115
			(0.375)
Prediction > 180, Belief < 91			0.230
			(0.337)
Prediction < 91, Belief > 180			0.355
			(0.432)
Prediction > 180, Belief > 180			0.650
			(0.532)
Algorithm, Prediction < 91, Belief < 91			0.427
			(0.550)
Algorithm, Prediction > 180, Belief < 91			0.190
			(0.546)
Algorithm, Prediction < 91, Belief > 180			0.456
			(0.644)
Algorithm, Prediction > 180, Belief > 180			-0.250
			(0.608)
N	3,013	3,013	3,013

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.9: **Applications.** Estimates for applications. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Reservation Wage Adjustment		
Algorithm	-125.3	-178.8	-400.2
	(103.5)	(149.7)	(245.4)
Prediction < 91		10.8	-232.1
		(146.6)	(192.2)
Prediction > 180		-108.3	-333.3
		(162.4)	(258.6)
Algorithm, Prediction < 91		-180.7	-131.5
		(264.1)	(376.5)
Algorithm, Prediction > 180		291.5	707.8**
		(224.9)	(339.2)
Belief < 91			-60.4
			(210.7)
Belief > 180			-307.8
			(248.0)
Algorithm, Belief < 91			379.0
			(306.2)
Algorithm, Belief > 180			449.3
			(373.2)
Prediction < 91, Belief < 91			484.4*
			(255.3)
Prediction > 180, Belief < 91			-14.4
			(353.8)
Prediction < 91, Belief > 180			340.5
			(361.4)
Prediction > 180, Belief > 180			927.1***
			(337.6)
Algorithm, Prediction < 91, Belief < 91			-348.4
			(426.6)
Algorithm, Prediction > 180, Belief < 91			-324.4
			(410.5)
Algorithm, Prediction < 91, Belief > 180			313.0
			(490.3)
Algorithm, Prediction > 180, Belief > 180			-1,288.0**
			(544.6)
N	1,174	1,174	1,174

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.10: **Reservation Wage Adjustment.** Estimates for reservation wage adjustment. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Pr(UE Exit), 3 Months		
Algorithm	0.015	-0.008	-0.038
	(0.013)	(0.021)	(0.029)
Prediction < 91		-0.0009	-0.050*
		(0.021)	(0.026)
Prediction > 180		0.005	-0.022
		(0.021)	(0.030)
Algorithm, Prediction < 91		0.074**	0.090**
		(0.032)	(0.040)
Algorithm, Prediction > 180		0.001	0.019
		(0.022)	(0.035)
Belief < 91			0.077**
			(0.035)
Belief > 180			-0.096***
			(0.026)
Algorithm, Belief < 91			0.036
			(0.053)
Algorithm, Belief > 180			0.042
			(0.043)
Prediction < 91, Belief < 91			0.103**
			(0.046)
Prediction > 180, Belief < 91			0.049
			(0.062)
Prediction < 91, Belief > 180			0.001
			(0.035)
Prediction > 180, Belief > 180			0.015
			(0.046)
Algorithm, Prediction < 91, Belief < 91			0.003
			(0.076)
Algorithm, Prediction > 180, Belief < 91			-0.030
			(0.070)
Algorithm, Prediction < 91, Belief > 180			-0.018
			(0.056)
Algorithm, Prediction > 180, Belief > 180			0.017
			(0.053)
N	3,183	3,183	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.11: **UE Exit, 3 Months.** Estimates for unemployment exits to employment. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Pr(UE Exit), 12 Months		
Algorithm	0.022	0.027	0.062
	(0.016)	(0.028)	(0.043)
Prediction < 91		0.041	0.032
		(0.033)	(0.054)
Prediction > 180		0.022	-0.006
		(0.031)	(0.046)
Algorithm, Prediction < 91		0.023	-0.060
		(0.049)	(0.085)
Algorithm, Prediction > 180		-0.035	0.005
		(0.039)	(0.060)
Belief < 91			0.221***
			(0.052)
Belief > 180			-0.261***
			(0.051)
Algorithm, Belief < 91			-0.100
			(0.065)
Algorithm, Belief > 180			-0.072
			(0.069)
Prediction < 91, Belief < 91			-0.020
			(0.076)
Prediction > 180, Belief < 91			0.006
			(0.066)
Prediction < 91, Belief > 180			-0.033
			(0.077)
Prediction > 180, Belief > 180			0.050
			(0.071)
Algorithm, Prediction < 91, Belief < 91			0.170*
			(0.101)
Algorithm, Prediction > 180, Belief < 91			-0.008
			(0.081)
Algorithm, Prediction < 91, Belief > 180			0.208*
			(0.115)
Algorithm, Prediction > 180, Belief > 180			-0.031
			(0.102)
N	3,183	3,183	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.12: **UE Exit, 12 Months.** Estimates for unemployment exits to employment. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Post-UE Earnings, 24 Months		
Algorithm	75.8	97.1	125.0
	(87.7)	(137.2)	(238.6)
Prediction < 91		423.2**	314.6
		(180.8)	(244.8)
Prediction > 180		-95.3	-324.7
		(132.6)	(236.7)
Algorithm, Prediction < 91		-87.0	-215.8
		(237.7)	(354.3)
Algorithm, Prediction > 180		36.6	319.4
		(188.8)	(331.2)
Belief < 91			633.6***
			(210.3)
Belief > 180			-1,039.4***
			(271.3)
Algorithm, Belief < 91			-27.9
			(314.9)
Algorithm, Belief > 180			-372.2
			(407.1)
Prediction < 91, Belief < 91			277.1
			(370.8)
Prediction > 180, Belief < 91			240.1
			(305.9)
Prediction < 91, Belief > 180			-352.2
			(395.6)
Prediction > 180, Belief > 180			413.9
			(363.7)
Algorithm, Prediction < 91, Belief < 91			-33.9
			(515.3)
Algorithm, Prediction > 180, Belief < 91			-401.5
			(404.5)
Algorithm, Prediction < 91, Belief > 180			1,064.2*
			(591.4)
Algorithm, Prediction > 180, Belief > 180			-129.8
			(494.4)
N	3,171	3,171	3,171

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.13: **Post-UE Earnings, 24 Months.** Estimates for post-unemployment earnings. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Pr(Re-registration), 24 Months		
Algorithm	-0.016	-0.0008	0.024
	(0.019)	(0.030)	(0.049)
Prediction < 91		0.001	-0.006
		(0.033)	(0.058)
Prediction > 180		0.015	0.015
		(0.033)	(0.058)
Algorithm, Prediction < 91		-0.037	-0.090
		(0.043)	(0.076)
Algorithm, Prediction > 180		-0.013	-0.027
		(0.043)	(0.070)
Belief < 91			0.017
			(0.054)
Belief > 180			-0.002
			(0.067)
Algorithm, Belief < 91			-0.060
			(0.064)
Algorithm, Belief > 180			0.003
			(0.091)
Prediction < 91, Belief < 91			-0.030
			(0.093)
Prediction > 180, Belief < 91			0.045
			(0.082)
Prediction < 91, Belief > 180			0.104
			(0.108)
Prediction > 180, Belief > 180			-0.095
			(0.084)
Algorithm, Prediction < 91, Belief < 91			0.128
			(0.102)
Algorithm, Prediction > 180, Belief < 91			-0.010
			(0.105)
Algorithm, Prediction < 91, Belief > 180			-0.006
			(0.107)
Algorithm, Prediction > 180, Belief > 180			0.092
			(0.095)
N	2,222	2,222	2,222

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.14: **Re-registrations, 24 Months.** Estimates for reservation re-registrations. All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

Robustness

	Belief Update	Meeting	R. W. Adj.	Exit, 3 M.
Algorithm	-15.5*	0.036	-409.4**	0.048**
	(8.80)	(0.022)	(179.6)	(0.022)
Prediction	-0.012	0.000	-0.418	0.000
	(0.043)	(0.0001)	(0.527)	(0.000)
Algorithm, Prediction	0.048	-0.0002	1.74*	0.000*
	(0.049)	(0.0001)	(0.874)	(0.000)
N	1,174	3,183	1,174	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.15: **Robustness: Linear Model.** All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Belief Update	Meeting	R. W. Adj.	Exit, 3 M.
Algorithm	-37.8***	0.066**	-434.8*	0.068**
	(12.4)	(0.029)	(253.9)	(0.032)
Prediction Basis 1	-41.5*	0.020	-268.8	-0.015
	(22.7)	(0.055)	(428.6)	(0.057)
Prediction Basis 2	29.2	0.027	-118.3	-0.038
	(24.8)	(0.070)	(316.7)	(0.063)
Algorithm, Prediction Basis 1	73.0**	-0.153**	984.6	-0.151*
	(30.6)	(0.074)	(597.9)	(0.080)
Algorithm, Prediction Basis 2	-32.5	0.015	602.1	-0.025
	(27.6)	(0.081)	(445.5)	(0.071)
N	1,174	3,183	1,174	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.16: **Robustness: Splines Basis 2.** All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Belief Update	Meeting	R. W. Adj.	Exit, 3 M.
Algorithm	-42.350**	0.092**	-233.794	0.061
	(19.084)	(0.043)	(337.265)	(0.055)
Prediction Basis 1	-32.028*	0.000	-149.158	0.041
	(16.697)	(0.041)	(398.572)	(0.040)
Prediction Basis 2	-25.287	0.042	-210.656	-0.156*
	(30.066)	(0.068)	(507.548)	(0.092)
Prediction Basis 3	21.697	0.038	-134.028	-0.105
	(23.237)	(0.073)	(403.787)	(0.067)
Algorithm, Prediction Basis 1	48.134**	-0.075	667.954	-0.088
	(22.587)	(0.060)	(484.665)	(0.056)
Algorithm, Prediction Basis 2	60.598	-0.179*	288.935	-0.104
	(41.759)	(0.104)	(774.368)	(0.130)
Algorithm, Prediction Basis 3	-13.966	-0.041	387.624	-0.041
	(29.182)	(0.090)	(493.774)	(0.088)
N	1,174	3,183	1,174	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.17: **Robustness: Splines Basis 3.** All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level.

	Belief Update	Meeting	R. W. Adj.	Exit, 3 M.
Algorithm	5.08	0.057	-1,154.2***	0.101**
	(19.6)	(0.039)	(318.2)	(0.045)
...				
Algorithm, Prediction	-0.037	0.000	5.90***	-0.001*
	(0.086)	(0.000)	(1.64)	(0.000)
...				
Algorithm, Prediction, Belief	0.001	0.000	-0.029***	0.000*
	(0.001)	(0.000)	(0.009)	(0.000)
N	1,174	3,183	1,174	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.18: **Robustness (Interaction): Linear Model.** All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level. To conserve space, only interaction terms involving treatment assignment are reported.

	Belief Update	Meeting	R. W. Adj.	Exit, 3 M.
Algorithm	-99.6*	0.279***	-1,415.2**	0.187*
	(51.4)	(0.095)	(654.7)	(0.102)
...				
Algorithm, Prediction Basis 1	229.7*	-0.415*	4,496.8***	-0.484*
	(125.6)	(0.218)	(1,559.7)	(0.283)
Algorithm, Prediction Basis 2	47.8	0.547**	1,701.5	-0.345
	(81.5)	(0.236)	(1,133.7)	(0.350)
...				
Algorithm, Prediction Basis 1, Belief Basis 1	-160.2	0.792	-11,641.9***	0.896
	(249.2)	(0.513)	(3,735.2)	(0.636)
Algorithm, Prediction Basis 2, Belief Basis 1	55.1	-1.37**	-5,672.6*	0.702
	(168.9)	(0.575)	(2,971.2)	(0.773)
Algorithm, Prediction Basis 1, Belief Basis 2	723.7	0.193	-7,332.1**	0.062
	(439.7)	(0.381)	(2,998.5)	(0.368)
Algorithm, Prediction Basis 2, Belief Basis 2	566.9	-0.151	-7,980.6	-0.499
	(386.4)	(0.784)	(5,432.3)	(0.333)
N	1,174	3,183	1,174	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.19: **Robustness (Interaction): Splines Basis 2.** All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level. To conserve space, only interaction terms involving treatment assignment are reported.

	Belief Update	Meeting	R. W. Adj.	Exit, 3 M.
Algorithm	-101.4	0.072	-2,474.6**	0.242
	(98.8)	(0.275)	(1,098.6)	(0.264)
. . .				
Algorithm, Prediction Basis 1	123.2	-0.902***	-1,456.9	-0.507
	(102.2)	(0.315)	(2,906.7)	(0.494)
Algorithm, Prediction Basis 2	149.4	0.822	6,902.6**	-0.493
	(223.4)	(0.774)	(2,861.9)	(0.715)
Algorithm, Prediction Basis 3	-109.2	1.36*	3,858.8	-0.792
	(153.9)	(0.769)	(2,495.4)	(0.839)
. . .				
Algorithm, Prediction Basis 1, Belief Basis 1	-103.5	0.691**	-1,552.4	0.213
	(128.1)	(0.274)	(1,943.4)	(0.298)
Algorithm, Prediction Basis 2, Belief Basis 1	-199.1	-0.730	-8,833.0**	0.101
	(266.7)	(0.584)	(3,612.9)	(0.552)
Algorithm, Prediction Basis 3, Belief Basis 1	-201.0	-1.18**	-6,411.0**	0.234
	(170.5)	(0.493)	(2,430.2)	(0.485)
Algorithm, Prediction Basis 1, Belief Basis 2	-93.2	1.78***	1,464.7	0.916
	(270.6)	(0.657)	(5,940.7)	(0.993)
Algorithm, Prediction Basis 2, Belief Basis 2	285.0	-2.02	-9,268.5	0.919
	(612.2)	(1.60)	(6,876.6)	(1.47)
Algorithm, Prediction Basis 3, Belief Basis 2	736.4*	-2.87*	-3,330.2	1.58
	(382.9)	(1.61)	(5,948.7)	(1.64)
Algorithm, Prediction Basis 1, Belief Basis 3	168.6	0.652*	-4,955.6*	0.364
	(349.3)	(0.343)	(2,667.2)	(0.419)
Algorithm, Prediction Basis 2, Belief Basis 3	884.0	-0.565	5,007.9	0.535
	(880.3)	(0.713)	(5,493.8)	(0.638)
Algorithm, Prediction Basis 3, Belief Basis 3	1,067.5*	-0.901	5,088.2	0.380
	(604.0)	(0.877)	(4,954.7)	(0.457)
N	1,174	3,183	1,174	3,183

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.20: **Robustness (Interaction): Splines Basis 3.** All specifications include caseworker fixed effects. Standard errors are clustered at the caseworker level. To conserve space, only interaction terms involving treatment assignment are reported.