Prediction versus Discretion: Human-AI Collaboration in Assignment of Unemployed Jobseekers *

Lovisa Rambjer †

August 23, 2025

Abstract

Can caseworkers improve upon algorithmic assignment of unemployed jobseekers to active labor market programs (ALMPs) at the Public Employment Service (PES)? This paper studies the impact of caseworker discretion over an algorithm that recommends assignment based on predicted reemployment probability. I set up a framework where ALMP assignment involves a trade-off between fairness (assigning ALMP to jobseekers with the lowest reemployment probabilities) and efficiency (maximizing treatment effects). This trade-off introduces tension between the algorithm's design and broader policy objectives, creating potential scope for caseworkers to improve outcomes by incorporating private information or better aligning with policymaker goals. Leveraging as-if random assignment of jobseekers to caseworkers, I reconstruct the algorithmic counterfactual and evaluate outcomes along both fairness and efficiency dimensions. The results show that caseworkers reduce fairness, despite achieving above-random prediction accuracy. At the same time, their deviations decrease employment rates. This suggests that while caseworkers may attempt to trade off fairness for efficiency, they are unsuccessful in targeting the most effective ALMPs — resulting in a reduction in both fairness and efficiency.

JEL: D8, J64, J65, J68

Keywords: algorithm, unemployment, prediction, discretion.

*I would like to thank Georg Graetz and Peter Fredriksson for their invaluable guidance and support. I am also grateful for helpful suggestions from Will Dobbie, Simon Ek, Marie-Pascale Grimon, Peter Hull, Danielle Li, Daniele Paserman, Ashesh Rambachan, Roland Rathelot, Lindsey Raymond, Olle Törnquist, Arne Uhlendorff and Johan Vikström, as well as comments from seminar and conference participants at the CREST-IFAU Workshop - Evaluating Labor Market Policies 2025 (Paris), CESifo Junior Workshop in Labor Economics 2025 (Munich), CREST PhD seminar 2025 (Paris), Doctoral Workshop on the Economics of Digitization 2025 (Tolouse), Boston University Micro workshop 2025, UCLS Annual Meeting 2024 (Uppsala), the Swedish Public Employment Service, and at the Nordic Public Policy Symposium 2024 (Copenhagen).

†Uppsala University and UCLS. Email: Lovisa.Rambjer@nek.uu.se

1 Introduction

Predictive algorithms are increasingly used to assist human decision making, with the promise of improving accuracy, efficiency, and, potentially, fairness. By processing vast amounts of data, algorithms excel in solving complex prediction tasks where humans might be limited by prejudice or inability. And as opposed to human decision makers, algorithms operate under clearly defined and controllable objective functions. Yet, while much speaks in favor of algorithms outperforming humans, the consequences of human discretion over algorithmic recommendations remain less explored.

Existing research on algorithmic decision-making focuses on settings with a well-defined prediction task and clearly aligned objectives — such as judges predicting pretrial misconduct or lenders predicting loan default. However, in many real-world policy contexts, prediction is only one component of a broader decision problem. Policymakers may pursue objectives that go beyond prediction, introducing tensions between algorithmic design and policy implementation. This trade-off complicates the analysis of caseworker discretion over algorithmic recommendations, but addressing it is central to understanding algorithmic governance in high-stakes public services

This paper studies such a setting: the assignment of active labor market programs (ALMPs) at the Swedish Public Employment Service (PES). Like in many OECD countries, the Swedish PES uses a statistical profiling tool to identify jobseekers at risk of long-term unemployment, which provides caseworkers with a recommendation on how to allocate support. But while the algorithm predicts the probability of reemployment, assignment decisions reflect two potentially competing objectives: *fairness*, defined as assigning ALMPs based on need (i.e., predicted reemployment probability), and *efficiency*, defined as maximizing employment outcomes. I analyze the impact of caseworker deviations from the algorithmic recommendation along both dimensions.

The empirical setting is the Swedish PES's assignment to either Independent Job Search or some ALMP, abstracting from the specific type of program. Following a rise in long-term unemployment during the 2010s the Swedish government refocused the PES's priorities toward one central goal: reducing long-term unemployment. In response, the PES introduced a profiling tool in 2020 that explicitly linked ALMP eligibility to predicted reemployment probability. This paper examines the implications of caseworker deviations from the assignment recommendations. The key questions is whether such deviations reflect valuable private information about jobseekers, such as effort or motivation, or simply reintroduces human bias and noise. And in deviating, whether caseworkers trade off fairness (prediction accuracy) for efficiency (maximizing total employment).

I first present a conceptual framework of the policymaker's objective. Assuming that the policymaker values both fairness and efficiency, I show that assignment based on predicted reemployment probability only maximizes policymaker utility if the causal effect of ALMP is monotonically decreasing in baseline reemployment probability. While such a relationship may hold on average, it is unlikely to do so across all jobseekers. Next I set up the caseworkers' assignment problem and map that to the empirical strategy. The empirical approach is divided into two parts. First, I study the impact of fairness and efficiency separately, second, I study the interaction between the two. I also conduct a complementary group analysis to study algorithmic and caseworker bias in terms of assigning ALMP based on reemployment probability rather than group membership.

I begin by analyzing the impact of caseworker discretion over the algorithmic recommendation in terms of fairness — predictive performance. A first empirical challenge arises. Jobseeker outcomes are endogenous to prior assignment decisions, and we only observe employment in the absence of ALMP for a selected sample. This selective labels problem (Kleinberg et al., 2018) biases comparisons of prediction accuracy. I address this challenge using the extrapolation method of Angelova, Dobbie, and Yang (2023), which leverages as if random assignment of jobseekers to caseworkers and variation in caseworker assignment propensities. This approach enables the evaluation of predictive performance by focusing exclusively on the employment outcomes of jobseekers assigned to Independent Job Search. Crucially, this method does not rely on assumptions about the efficacy of ALMPs. By estimating the employment rate of individuals assigned to Independent Job Search while holding the assignment rate constant, I can directly compare the predictive performance of caseworkers and the algorithm. Evaluating performance at a fixed assignment rate isolates differences in accuracy while holding both budgetary constraints and caseworker preferences constant.

I find that caseworker discretion decreases the conditional employment rate by 1.3 percentage point relative to the algorithmic counterfactual; showing that caseworkers marginally reduce fairness (by 2.5 % relative to the employment rate obtained by a random assignment). And only 9 % of caseworkers obtain a higher conditional employment rate than the algorithm. These findings suggest that caseworkers are either (i) making systematic prediction errors,[1] or (ii) pursuing objectives other than fairness when making assignment decisions.

Next I examine to what extent the reduction in fairness can be attributed to (i) poor prediction accuracy, and (ii) the fairness-efficiency trade-off. In order to get estimates of prediction accuracy I take advantage of a unique feature of the data - caseworkers must

[1] If caseworkers had achieved the same conditional employment rate as the algorithmic counterfactual, this would be consistent with random assignment.

enter a reason for deviating from a specific ALMP recommendation. This allows me to exclusively analyze cases where the caseworkers have deviated to Independent Search due to 'high reemployment probability', for which employment outcomes in the absence of ALMP are observed. The results show that, on average, caseworkers achieve conditional employment rates above those of random assignment. If the entire reduction in fairness was due to poor prediction, we would expect the average caseworker to perform worse than random. These findings suggest that caseworker objectives play an important role in explaining deviations.

Third, I analyze the impact of caseworker discretion on efficiency by exploiting variation in caseworkers' propensity to deviate from the algorithm. Specifically, I examine the relationship between deviation rates and employment outcomes. The approach is a typical 'judge design' (Chyn et al., 2025), and similar to Hoffman et al. (2018) who study the correlation between manager deviations from an algorithmic hiring recommendation and worker outcomes. The results show that caseworker deviation is reducing employment rates within 6 and 12 months.

Fourth, the results from the complementary group analysis show a large algorithmic bias against natives in the assignment to ALMPs. Where immigrants are disproportionately assigned to ALMPs, given their group level reemployment probability. Caseworker mitigates this bias by increasing the assignment to ALMPs among natives. However, caseworker deviations leads to a negative effect on natives total employment rates, while there is no effect on employment among immigrants.

Lastly, I examine how caseworker performance in fairness relates to efficiency. Splitting the sample into caseworkers who improve versus reduce fairness relative to the algorithm, I find a positive correlation across performance dimensions: caseworkers who outperform the algorithm in fairness also achieve higher employment rates and more accurate predictions.

Taken together, the results suggest that caseworkers, on average, reduce both fairness and efficiency. However, the full reduction in fairness cannot be explained by poor predictive performance alone, indicating that other factors drive deviations. One possibility is that caseworkers attempt to trade off fairness for efficiency, but lack the skill to target the most effective ALMPs, resulting in no gains in total employment.

Explicitly addressing this fairness-efficiency trade-off — and the resulting tension between algorithmic design and policy implementation — represents an important extension to the current literature on algorithmic decision making. The starting point of the literature on algorithmic decision-making has been well-defined prediction tasks — such as pretrial hearings, loan approvals, or medical diagnostics — where the policy objective can largely be captured by a single prediction exercise. In pretrial hearings, the causal impact of detention on misconduct is likely monotonic in misconduct risk. In health care, diagnostics is a clearly defined prediction task with no direct causal effect on the presence of disease, and is typically separated from the more complex decision of treatment assignment. This literature has addressed several first-order empirical challenges in studying algorithms using observational data—from solving the selective labels problem (e.g. Kleinberg et al. (2018) and Angelova et al. (2023)), to developing frameworks that characterizes the conditions under which prediction mistakes can be identified (Rambachan, 2024).

This paper builds on a long-standing literature documenting the limitations of human prediction in domains like medicine and criminal justice (e.g. Ohlin and Duncan, 1949; Dawes et al., 1989); (Dawes, 1979; Grove et al., 2000). The modern literature on algorithmic decision-making has carried on this work in settings including pretrial hearings (Kleinberg et al., 2018; Angelova et al., 2023; Rambachan, 2024), and medical diagnosis (Mullainathan and Obermeyer, 2022; Chan et al., 2022). Recent work in the labor market setting highlights the potential of algorithms in hiring decisions (e.g. Cowgill, 2020; Hoffman et al., 2018).

In the PES context, research has examined the effectiveness of using algorithms at the PES. Ernst et al. (2024) investigate how algorithmic risk profiling affects the selection into job search assistance in the Flemish PES, Belgium. They do not, however, deeply analyze the prediction accuracy of the algorithm and caseworkers. van den Berg et al. (2023) combine survey and administrative data from Germany to compare the stated prediction on reemployment probability of jobseekers and caseworkers with an algorithm which the authors have trained. They find that the algorithm outperforms caseworkers, but that the jobseekers' self assessment of their reemployment probabilities are informative. A key limitation of both these studies is that they use jobseeker outcomes which are endogenous to caseworker decisions, without discussing it's implication for their findings. Traditional work on PES caseworkers has focused on matching strategies (Behncke et al., 2010; Arni et al., 2022) and value-added (Cederlöf et al., 2021).

This paper makes two main contributions. First, it extends the literature on algorithmic decision-making to a setting where prediction and policy objectives are misaligned, generating a fairness-efficiency trade-off. Second, it provides the first evidence on the impact of caseworker discretion over algorithmic recommendations at the PES, quantifying the effects on both predictive performance and jobseeker outcomes.

The remainder of the paper is structured as follows: Section 2 provides an overview of the institutional context and Section 3 introduces the data. Section 4 sets up the conceptual framework, which is mapped to the empirical strategy in Section 5. Section 6 and 7 goes through identifying assumptions and descriptive statistics. Results are presented in Section 8 and, finally, Section 9 concludes.

2 Institutions - The Swedish Public Employment Service

Since 2019, the Swedish PES has undergone major reforms aimed at improving efficiency through digitization, automation, and privatization. The agency's role shifted toward labor market assessments, oversight of private providers, data analysis, and digital infrastructure, while Job Search Assistance was delegated to private actors in a competitive market. As a result, PES caseworkers now focus primarily on assessing eligibility for ALMPs. The reform aligned with growing political attention to long-term unemployment. Government directives moved from broad goals — such as overall ALMP volumes and employment growth — to a clear priority: reducing long-term unemployment. To support this shift, the PES introduced a statistical profiling tool to target programs toward those most at risk.

This section first outlines the standard process for newly registered jobseekers, followed by a detailed description of the algorithmic profiling tool.

2.1 Process for Newly Registered Jobseekers

The process for a typical newly registered jobseeker can be summarized in three main steps. After the summary follows a detailed description of each step of the process. [2]

1. **Registration and as-if random assignment of jobseekers to caseworkers at remote offices.**
   A jobseeker registers online with the Public Employment Service (PES) on the first day of their unemployment spell. They fill in information about themselves and book a first remote meeting with a caseworker at a central office.

2. **Meeting and assignment decision.**
   A first remote meeting is held within two weeks. The caseworker uses the algorithmic recommendation together with private information to make a labor market assessment and assignment decision.

3. **Follow-up and conclusion.**
   If a jobseeker is assigned to an ALMP, the responsibility for the jobseeker and follow-up is transferred to a local office. An assignment to Independent Job Search is assessed every six month at a central office. There are no personal caseworkers. The as-if

[2] Based on interviews with PES staff and internal documents: Memo Gemensam process för beslutshandläggare – KROM och STOM (2020); Handläggarstöd Placeringsprocessen (2020); Handläggarstöd Innehåll och Avrop (2020); Presentation - Så funkar det AMPB 1.0 (2020).

random assignment is in place for each meeting. At each follow-up, the algorithm is rerun and the algorithmic recommendation updated. Finally, the case is closed when the jobseeker secures employment or enrolls in education.

Figure 1 illustrates the procedure from registration to the first meeting and assignment decision.

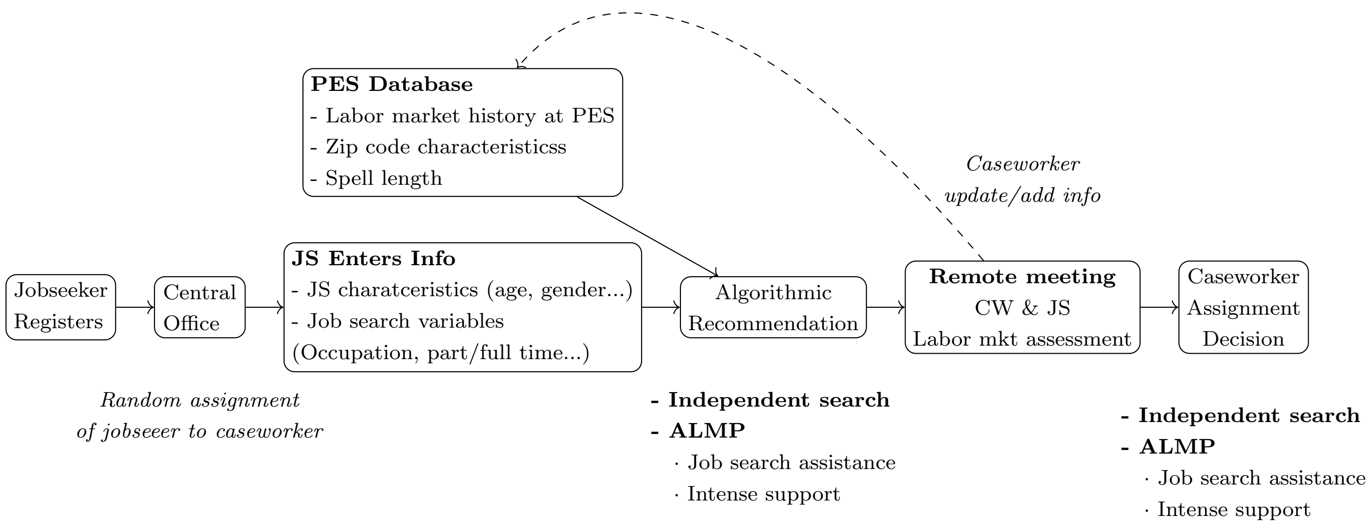


Figure 1: Process newly registered jobseekers

2.1.1 Registration and As-If Random Assignment

Standard Central-Office Pathway

An unemployed jobseeker registers online with the Public Employment Service (PES) on the first day of their unemployment spell. Registering is a requirement to apply for unemployment insurance benefits. At registration, jobseekers immediately book a first remote meeting by selecting a free slot associated with a central-office caseworker. Slots are allocated across caseworkers in a non-systematic manner, creating an as-if random assignment of jobseekers to central-office caseworkers. All central-office meetings occur during regular working hours; there are no evening or weekend slots. The central offices are located throughout Sweden and organized in 7 administrative offices. [3] Central offices are not tied to specific geographic regions but instead serve jobseekers across the entire country. In this paper, I focus exclusively on assignment decisions made at central offices.

[3] Arvidsjaur and Göteborg; Malmö and Karlskrona; Örebro, Karlskoga and Skara; Stockholm; Luleå and Lycksele; Örebro, Karlskoga and Skara, Södertälje and Sollefteå.

Exemptions & Local-Office Pathway

If a jobseeker cannot register online or attend a remote meeting—e.g., because they lack an electronic ID or require a translator—they are directed to a local PES office. Alternatively, they may visit a local office in person to register. Local offices serve a restricted geographical region and hold meetings either in person or remotely. Jobseekers assisted at local offices form a selected group with generally lower reemployment probabilities. Since the local-office slot-booking system does not guarantee as-if random assignment, I exclude assignment decisions made there from my analysis.

The Role of the Caseworker

The main task of PES caseworkers is to make labor market assessments and determine eligibility to ALMPs. The caseworker sets up an activity plan, with activities based on the assignment decision. Caseworkers also assist jobseekers with practical and administrative tasks, such as choosing a private provider. They do not provide any assistance themselves. The monitoring of jobseekers' job search activities is handled by a separate unit at the PES.

2.1.2 Statistical Profiling, Meeting and Assignment

Statistical profiling

During online registration, jobseekers provide self-reported information on personal characteristics (age, gender, country of birth, disability status, and education) and job search details (desired occupation, relevant experience, and preferred number of working hours). These self-reports are combined with administrative data from PES registries, including labor market history (previous interventions, registration periods, spell length, and ongoing programs) as well as zip code–level characteristics (local unemployment rate, median income, educational attainment, and the share of foreign citizens).

The jobseeker is informed that the statistical profiling serves as a key input for the caseworker's assignment decision and that the self reported information is to be used for the statistical profiling.

In total, the algorithm uses 26 input variables to predict each jobseeker's reemployment probability (see Table A1 for the full list). The algorithm produces a continuous profiling score that is mapped to an algorithmic recommendation using cutoffs that vary by spell length. Caseworkers see only the categorical recommendation, not the underlying score.

The recommendation has three main values (from highest to lowest predicted reemployment probability):

1. **Independent Search**
   Searching for a job mostly independently, with the help of digital self-service tools.

2. **Job Search Assistance**
   Privately provided job search assistance including coaching and guidance, CV writing, interview training, and job referrals. The exact activities included are decided on by the private providers together with the jobseeker.

3. **Intensive Support**
   For jobseekers in need of more extensive and intense support such as labor market training, labor market education, language training. Intense support also includes referral to a local office for an in depth labor market evaluation.

Section 2.2 provides a detailed description of the algorithm, the mapping to recommendations, and its predictive performance.

Eligibility Criteria & Standardized Questions

Eligibility for active labor market programs is determined primarily by the jobseeker's predicted reemployment probability. In addition, there are some ineligibility criteria such as part-time unemployment or start of employment or studies within 90 days, or other reasons that hinders the jobseeker from participating in an ALMP. To elicit whether the jobseeker is ineligible for ALMP and to facilitate the labor market assessment, the jobseeker is asked to answer 11 standardized questions. The caseworker retains final discretion, basing the decision on a holistic assessment of the jobseeker. Thus, the stated eligibility criteria serve as guidelines rather than rigid rules. Appendix Section A.2 contains the full list of criteria and standardized questions.

Jobseeker Preferences

Jobseeker preferences should be considered in the assignment decision, and jobseekers must be given the opportunity to express and justify any disagreement. Caseworkers should assess whether the stated reasons affect the labor market assessment or the evaluation of the eligibility criteria. If the caseworker concludes that the assignment to an ALMP remains appropriate and the jobseeker still refuses to participate, the case should be referred to the PES monitoring unit, which will decide on a potential sanction.

Labor Market Assessment & Assignment Decision

The first meeting is typically held within two weeks after registration. At the first meeting, the caseworker conducts a labor market assessment by combining the algorithmic recommendation with private information elicited during the meeting. If needed, the caseworker can update the input variables used in the profiling tool and rerun the algorithm. See Appendix Section A.1 for the empirical distribution of number of days between registration and first caseworker profiling, as well as change in profiling score due to caseworker updating of the algorithm.

Based on this assessment, the caseworker makes an assignment decision, which mirrors the algorithmic recommendation. In case of an assignment to Job Search Assistance, the private provider take over the responsibility to decide on specific activities. While an assignment to any Intense Support is more specific, for example labor market training, labor market education, or language training. In the paper I focus on the decision to assign either to Independent Search or any ALMP, abstracting from the choice between the different programs.

If the caseworker determines that an in-person evaluation is necessary, they may refer the jobseeker to a local office. In this paper, such referrals are categorized as assignments to an ALMP, as they indicate that the jobseeker's labor market attachment is too weak to justify Independent Job Search.

Caseworkers are generally advised to follow the algorithmic recommendation and make few deviations. Caseworker discretion was limited such that any deviation from Independent Search or Intense Support to Job Search Assistance required approval from a committee. Otherwise they had free discretion.

2.1.3 Follow-Up and Case Closure

If a jobseeker is assigned to an ALMP, the responsibility for the jobseeker and follow-up is transferred to a local office. An assignment to Independent Job Search is assessed every six months at a central office. No jobseeker is paired with a single, dedicated caseworker; instead, they are as-if randomly assigned to a caseworker each meeting by simply signing up for a remote meeting slot online. Jobseekers may also request a meeting at any time by calling or booking online.

At each follow-up, the algorithm is rerun and the algorithmic recommendation is updated. Based on the new recommendation and information gathered during the meeting, the caseworker may initiate an ALMP assignment. A case is closed once the jobseeker has found part- or full-time employment (depending on their stated preference) or enrolled in an education program eligible for student aid. This includes adult education organized by the municipality *(kommunal vuxenutbildning)*, non-formal adult education *(folkhögskola)* at the primary or secondary level, or tertiary education such as higher vocational training or university studies.

2.2 Algorithmic Profiling Tool

The algorithmic profiling tool aims to predict the probability that a jobseeker will secure employment independently within six months from the profiling date. The purpose being to allocate resources efficiently, by prioritizing jobseekers at higher risk of long-term unemployment.

2.2.1 Input Variables and Data

The algorithm is trained on the inflow of unemployed jobseekers and incorporates four main groups of input variables:

- **Self-reported characteristics:** age, gender, country of birth, disability status, and education.
- **Job search information:** desired occupation, relevant experience in that occupation, and preferred number of working hours (full-time or part-time).
- **Administrative data from PES registries:** labor market history including prior interventions, registration periods, current spell length, and ongoing programs.
- **Zip code–level characteristics:** local unemployment rate, median income, educational attainment, and share of foreign citizens.

In total, 26 input variables feed into the prediction. The full list appears in Table A1.

The data infrastructure at the PES posed a major obstacle to the design and training of the machine learning algorithm, both in terms of data quality and availability. Input variables were limited to those available in a single database, and self-reported jobseeker characteristics could not be verified. The lack of individual-level data made the algorithm heavily reliant on zip code–level characteristics, limiting its ability to differentiate between jobseekers within the same area. (Arbetsförmedlingen, 2020).

The algorithm is trained on binary outcome data (0/1) for several events related to employment or full-time studies within a specified time frame. The two primary outcomes are: (i) full-time employment and (ii) part-time employment within six months. The algorithm's output is a profiling score, calculated as the sum of the predicted probabilities for these two outcomes. This score ranges from 0 to 1. (Arbetsförmedlingen, 2020).

The outcomes data used to train the algorithm come from PES registries, which contain substantial measurement error. A manual quality check conducted by the PES revealed that 20–25% of case files included notes about positive jobseeker outcomes (employment or full-time studies) that were never recorded in the registries (Arbetsförmedlingen, 2020).

According to PES (Arbetsförmedlingen, 2023), the algorithm's accuracy (share of correctly classified outcomes) is 68 %, compared to a baseline random-classification accuracy of 57 %. Due to poor performance the PES updated the algorithm in April 2023. Appendix Section A.5 contains a summary of the differences between the two algorithmic regimes.

2.2.2 Score to Recommendation Mapping

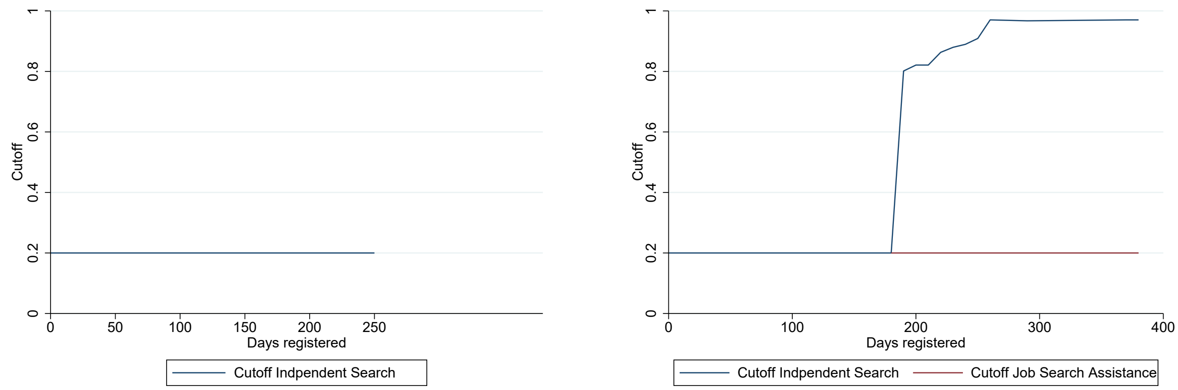
The algorithm produces a continuous profiling score for each jobseeker, which is a number ranging from 0 to 1 where a higher score implies a higher reemployment probability. As mentioned above, the score is simply calculated as the sum of the predicted probabilities for the two main outcomes. Thus, the score itself should not be interpreted as a probability.

Rather than including the current unemployment duration as an input, the system applies time-varying cutoffs to the profiling score. Specifically:

- Cutoff thresholds vary by the number of days since the start of the unemployment spell; jobseekers with shorter duration face lower thresholds than jobseekers with longer duration. Such that at a given score, the probability to be recommended an ALMP increases with the length of the spell.

- Jobseekers with profiling scores above the cutoff associated with their duration are recommended Independent Search. Those whose scores fall below the cutoff are recommended some form of ALMP.

Figure 2 shows the time-varying cutoffs used in 2022 to decide whether a jobseeker is recommended Independent Search or an ALMP track. From January 1 through September 7, the cutoff remained fixed at a profiling score of 0.2 for all jobseekers up to 250 days of unemployment. Beginning September 8, the cutoff stayed at 0.2 only until 180 days, after which the required score increased with the spell length. Consequently, jobseekers needed at least 250 days of unemployment under the January 1–September 7 rules — and 180 days under the September 8 onward rules — to become eligible for Job Search Assistance. Any jobseeker assigned to an ALMP earlier in the spell could only be placed into the Intense Support category.[4] Appendix Section A.6 contains a description of the distribution of profiling scores across categories.



(a) January 1 - September 7, 2022

(b) September 8 - December 31, 2022

Figure 2: Cutoffs based on days unemployed

Note: Cutoffs by days unemployed. From January 1 to September 7, jobseekers became eligible for Job Search Assistance only after 250 days. Up to 250 days, the cutoff between Intense Support and Independent Search was fixed at a profiling score of 0.2. Beginning September 8, the cutoff remained at 0.2 up to 180 days; after 180 days, jobseekers with scores above 0.2 became eligible for Job Search Assistance, and the upper cutoff (between Job Search Assistance and Independent Search) increased with additional days in unemployment. These post–180-day cutoffs are inferred empirically from the data. Applying these cutoffs reproduces the actual assignment recommendation in 99.99 % of non-randomized cases and 99.5 % when randomized cases are included.

3 Data

3.1 Data Sources and Coverage

I observe the universe of PES profiling events from the introduction in early 2020 through December 2022. On the jobseeker side, data include:

- Detailed PES administrative records (profiling scores, appointment dates, program starts, referrals to local offices, etc.).
- On the jobseeker side I have linked Swedish registry data (demographics, prior employment history, and long-term earnings).
- On the caseworker side, I have limited personnel data (tenure, wage and office) from the PES human-resources registry.

[4]Certain groups with generally weak labor market attachment — such as long-term unemployed youth, long-term unemployed adults, and newly arrived immigrants — face less restrictive time-varying cutoffs.

The data also contains a reason for deviating, whenever the caseworker has deviated from Job Search Assistance to either Independent Search or Intense Support.

I restrict the analysis to jobseekers starting their unemployment spell between January 1 and December 31, 2022, excluding the roll-out period 2020 and 2021.[5]

3.2 Sample Selection

The main analysis focuses on newly registered jobseekers who are subject to as-if random assignment at central offices. I apply the following inclusion and exclusion criteria:

Exclude subgroups not due to as-if random assignment:

- Jobseekers aged 16–24 or enrolled in the Youth Job Guarantee (23% of the initial pool).
- Jobseekers with prior long-term unemployment who are routed to specialized programs (8%).
- Jobseekers missing a valid search-occupation code (2.5%).

These exclusions remove 32% of profiled cases.

Exclude non-UI beneficiaries: Jobseekers without unemployment-insurance (UI) entitlement who receive benefits are excluded ($< 1\%$).

Caseworker-level restrictions: To ensure sufficient volume per evaluator, I limit the analysis to caseworkers who:

1. Handled at least 75 cases during 2022.
2. Were active (i.e., recorded at least one profiling) in at least 10 of the 12 months.

These requirements exclude 67 % of caseworkers, yielding a final set of 528 active caseworkers.

Identifying Decision Events: Many jobseekers receive multiple profilings even when no assignment decision occurs (e.g., during routine administrative updates). To isolate profiling events associated with a caseworker decision, I drop any profiling that is *identical* on three dimensions—(1) original score, (2) algorithmic recommendation, and (3) inferred assignment — if it occurs within one calendar month of the previous identical profiling. This rule removes 4.5% of profiling events; the maximum number of profilings per jobseeker falls from 12 to 6, while the median remains 1.

[5] With the intention to extending to post-April 2023 observations once the data become available.

3.3 Reason for Deviating

Whenever the caseworker deviates from Job Search Assistance to either Independent Search or Intense Support, they are required to enter a reason for deviating. Caseworkers can choose from predefined categories or enter a free-text explanation. In total, 2,025 deviations were recorded. In 87% of these cases, a free-text reason was entered; in the remaining 13%, one of eight predefined categories was selected. All free-text responses were manually reviewed and classified.

3.4 Defining Key Variables

3.4.1 Defining Caseworker Assignment

In the raw data, each profiling event includes the algorithmic recommendation (five-category variable). However, there is no single variable which records the actual assignment decision. Thus, to infer the *actual assignment* made by the caseworker, I use the following rules:

Any ALMP:

- Jobseekers whose algorithmic recommendation was Job Search Assistance and who did not have a recorded deviation to Independent Search, inferred by the reason for deviating.
- Jobseekers who start any ALMP (matching services, training, language course, etc.) within three months of the profiling date.
- Jobseekers referred to a local PES office.

Independent Search:

- Jobseekers whose algorithmic recommendation was Job Search Assistance but for whom the caseworker recorded a deviation to Independent Search, as inferred by the reason for deviating.
- Jobseekers who do not start any ALMP within three months of the profiling date.

3.4.2 Defining Deviations

I can only directly observe deviations when caseworkers deviate from Job Search Assistance. Other deviations are inferred from the recorded algorithmic recommendation (observed) and the caseworker assignment (defined above).

3.4.3 Defining Employment

The main outcome variable used in the paper is full time employment starting within 6 months from the profiling date and lasting at least 4 months. Monthly full time employment is defined using payroll tax data and defined as earning at least 75 % of the collective bargain minimum wage in retail, which equals about 17,800 SEK (1,877 USD) per month. Alternative definitions of employment are used for robustness checks, such as data from the PES registry including exiting the PES to start studying. For 50,113 cases (50 % of the sample) there is tax data to follow up 12 months after the profiling date. Unfortunately, due to data limitations it is not possible to study outcomes after 12 months.

3.5 Final Sample

After all exclusions and definitions, the 2022 sample contains:

- 99,959 decision events (first and follow-up profilings).
- 528 caseworkers at central offices.

Most jobseekers appear only once or twice in the profiling data. Figure A9 displays the distribution of profiling counts per caseworker and per jobseeker in this final sample.

4 Conceptual Framework

In this section, I formalize the policymaker's problem in a conceptual framework where the policymaker maximizes utility by allocating jobseekers to ALMPs, subject to a budget constraint. The policymaker faces two potentially conflicting goals: (i) assigning ALMPs based on reemployment probability (fairness), and (ii) maximizing overall employment (efficiency). I examine the role of both the algorithm and caseworkers in this utility maximization problem.

The main purpose of the framework is to formalize the fairness-efficiency trade-off and to guide the empirical strategy. In Section 5, I connect the analysis of caseworker discretion over algorithmic recommendations to this framework and show how it maps into the empirical design. In Section 6, I discuss and assess the identifying assumptions underpinning the strategy.

4.1 Setting up the Policymaker's Assignment Problem

The policy problem for the policymaker is how to allocate ALMPs across jobseekers. For each unemployed jobseeker i, the policymaker chooses either an ALMP or assignment to

Independent Search. In practice, a set of ALMPs is available, and their effects are likely to be heterogeneous across jobseekers. In the conceptual framework, I simplify by assuming that each jobseeker faces only one ALMP — for example, the one that maximizes employment. Let $D_i = 1$ indicate assignment to an ALMP and $D_i = 0$ indicate assignment to Independent Search. Using potential outcomes notation, the employment outcome of jobseeker i is $Y_i = Y_i^0 \cdot (1 - D_i) + Y_i^1 \cdot D_i$. Where Y^1 denote employment outcome under assignment to ALMP, and Y^0 denote employment under assignment to Independent Search. The probability of employment is denoted: $y_i = E[Y_i]$, $y_i^0 = E[Y_i^0]$ and $y_i^1 = E[Y_i^1]$. Treatment effect $\gamma_i = E[Y^1 - Y^0]$ is a function of baseline reemployment probability, $\gamma_i = k(y_i^0)$, and $k : [0, 1] \rightarrow \mathbb{R}$.

The policymaker is facing a budget constraint which requires that only a share ϕ can be assigned to ALMPs. Let N denote the total number of jobseekers.

I assume that the policymaker is *risk neutral*, caring only about maximizing expected values and disregarding uncertainty in predictions. Second, I assume that potential outcomes are independent of other jobseekers' treatment assignments, i.e. there is no interference or spillover across jobseekers, so that the SUTVA condition holds.

4.1.1 Fairness

Fairness is defined as assigning ALMPs based on need for support, meaning that jobseekers with the lowest baseline reemployment probabilities are prioritized for assignment. Fairness is achieved by solving the following maximization problem.

$$
\begin{aligned}
&\underset{D_i,..,D_N}{max} \sum_{i=1}^{N} D_i(1 - y_i^0) \\
&s.t. \sum_{i=1}^{N} D_i = N \times \phi
\end{aligned}
\tag{1}
$$

This can be solved by ranking jobseekers according to their baseline reemployment probability and applying a threshold rule that satisfies the budget constraint. Let q^{y0}_{p} denote the p'th percentile of the distribution of y^0. The threshold which ensures that the budget constraint is satisfied can be written as $s^F = q^{y0}_{\phi}$. Fairness is maximized by using the following decision rule: $D_i = 1[y_i^0 \leq s^F]$.

4.1.2 Efficiency

Efficiency is defined as assigning ALMPs in a way that maximizes total employment.

$$\max_{D_i,..,D_N} E\Big[\sum_{i=1}^{N} Y_i\Big]$$
$$s.t. \sum_{i=1}^{N} D_i = N \times \phi \tag{2}$$

Using policymaker risk neutrality and SUTVA allows me to reformulate the maximization problem in terms of maximizing the sum of reemployment probabilities y_i^0 and y_i^1. $E\Big[\sum_i^N Y_i\Big] = \sum_i^N y_i^0 + D_i(y_i^1 - y_i^0) = \sum_i^N y_i^0 + D_i\gamma_i$. Since the term $\sum_i^N y_i^0$ is constant across assignments, the maximization problem reduces to:

$$\max_{D_i,..,D_N} \sum_{i=1}^{N} D_i\gamma_i$$
$$s.t. \sum_{i=1}^{N} D_i = N \times \phi \tag{3}$$

This can be solved by ranking jobseekers according to their treatment effects γ_i and applying a threshold rule that satisfies the budget constraint. Let $s^E = q^{y0}_{1-\phi}$ denote the corresponding threshold. The maximization problem is solved under the following decision rule: $D_i = 1[\gamma_i \geq s^E]$. [6]

4.1.3 Policymaker's Maximization Problem

For each jobseeker i the policymaker derives utility from assignment to the ALMP:

$$U(D_i) = \alpha \underbrace{f(y_i^0)}_{Fairness} + (1-\alpha) \underbrace{g(\gamma_i)}_{Efficiency} \tag{4}$$

I model the utility function as a weighted sum of the utility from *fairness* and *efficiency*. The parameter $\alpha \in [0, 1]$ represents the weight that the policymaker places on fairness relative to efficiency. As $\alpha \to 1$, the policymaker values fairness alone; as $\alpha \to 0$, the policymaker values efficiency alone.

The policymaker's problem is to choose, for each jobseeker i, whether to assign them to the ALMP or Independent Search, in order to maximize total utility subject to the budget constraint.

[6] Total employment is maximized by assigning jobseekers based on their treatment effects, provided there are no sharp employment cutoffs — i.e., jobseekers do not require a minimum reemployment probability to obtain employment.

$$\begin{aligned} \underset{D_i,..,D_N}{max} \; & E\Bigg[\sum_{i=1}^{N} \alpha f(y_i^0) + (1-\alpha)g(\gamma_i)\Bigg] \\ s.t. \; & \sum_{i=1}^{N} D_i = N \times \phi \end{aligned} \tag{5}$$

A solution to the problem is to rank jobseekers based on the utility received when assigning them to the ALMP (versus Independent Search), given by $U(D_i)$, and assign them if the utility exceeds some threshold s. The ranking then determines how ALMP assignment change with s. The threshold parameter s, in turn, is set to satisfy the budget constraint. The threshold which ensures that the budget constraint is satisfied can be denoted $s = q^U_{1-\phi}$. Then, policymaker's expected utility is maximized under the following decision rule:

$$D_i = 1\Bigg[\Big(\alpha f(y_i^0) + (1-\alpha)g(\gamma_i)\Big) \geq s\Bigg] \tag{6}$$

4.2 The Algorithm

The algorithm predicts jobseekers' baseline reemployment probability, i.e. $a(X_i) \equiv Pr(Y_i^0 = 1|X_i)$, using observable characteristics $X_i \in \mathcal{X}$. It recommends ALMP for all jobseekers whose probability, or profiling score, is below a threshold s_a, which is determined by the policymaker. The algorithmic recommendation thus follows a simple decision rule, where the cutoff s_a is set to satisfy the budget constraint.

$$A_i = 1[a(X_i) \leq s_a] \tag{7}$$

4.3 Setting up the Caseworkers' Assignment Problem

In this section I set up the caseworkers' problem and show how it corresponds to the policymaker's problem and the algorithmic recommendation. Let there be a set of caseworkers j making an assignment decision analog to the policymaker's problem. For each jobseeker there is a vector of characteristics $X_i \in \mathcal{X}$ observable to both the algorithm and the caseworker and a vector of characteristics $V_{i,j} \in \mathcal{V}$ only observable to the caseworker.

Caseworkers first conducts a labor market assessment, predicting reemployment probability under assignment to Independent Search $D_i = 0$, as well as reemployment probability under the ALMP $D_i = 1$. The caseworkers take the algorithmic recommendation A_i, the public information X_i and the private information $V_{i,j}$ into account and predicts the reemployment probabilities. Note that caseworkers do not observe the profiling score, but only the assignment recommendation.

$$\hat{y}_{i,j}^{D} = h_{D,j}(A_i, X_i, V_{i,j}) \tag{8}$$

The private information $V_{i,j}$ contains (i) characteristics directly important for the jobseekers search, (ii) jobseeker preferences, (iii) jobseekers motivation, effort och search ability, and (iv) information important to determine eligibility to ALMP. I will allow $V_{i,j}$ to vary by caseworker to reflect differential caseworker ability to exert private information in the meeting. Each caseworker has individual evaluation functions $h_{D,j}(.)$, reflecting differences in the ability to predict reemployment probabilities, given the information at hand.

Next, the caseworker must choose whether to assign the jobseeker to Independent Search or the optimal ALMP. The caseworker assigns the jobseeker to the ALMP if the payoff exceeds some assignment threshold s_j.

$$D_{i,j} = 1\left[\left(\alpha_j f(\hat{y}_{i,j}^{0}) + (1-\alpha_j) g(\hat{\gamma}_{i,j})\right) \geq s_j\right] \tag{9}$$

I assume that the caseworker utility function take the same form as the policymaker's, but allow for the parameters α_j and s_j to vary by caseworker.

Similarly to the policy maker, the utility of assigning jobseeker i to the ALMP is decreasing in predicted reemployment probability and increasing in the treatment effect. The evaluation will generate a ranking of jobseekers in terms of "utility score", which maps into the assignment decision with the help of a cutoff s_j.

The threshold s_j is associated to a caseworker's *assignment rate*, for a given draw of jobseeker. The assignment rate is defined as the propensity to assign jobseekers to Independent Job Search:

$$S_j = 1 - E[D_{i,j}] \tag{10}$$

The caseworkers will vary both regarding their weighting parameter α_j, and in terms of their chosen cutoff s_j. The variation in α_j simply illustrates that caseworkers are allowed to put different weights on fairness versus efficiency. The threshold s_j can be interpreted as the caseworker's overall belief about the effectiveness of assignment to ALMPs. A caseworker who favors assignment to ALMP will pick a lower threshold s_j.

4.3.1 Aligning Caseworker Decisions with Policymaker Goals

The policymaker ensures that caseworker decisions align with their objectives through both informal and formal channels. Informally, the fairness-efficiency trade-off is communicated via government-set goals and strategic steering documents. In the model, this is captured by assuming that caseworkers face the same utility function as the policymaker.

Formally, the policymaker influences ALMP assignment volumes through the algorithmic recommendation A_i, by setting the threshold s_a. I assume that the probability of assignment to an ALMP is monotonically increasing in the algorithmic recommendation, i.e. there are no defiers.

4.4 Optimal Assignment Mechanism

Under what assumptions is the algorithm's decision rule utility maximizing for the policymaker? And when does there instead exists a fairness-efficiency trade-off? As in the policymaker's problem, the profiling score $a(X_i)$ generates a ranking of jobseekers — in this case, based on predicted reemployment probabilities. The algorithm can only be utility maximizing, for any given threshold s_a, if it induces the same ranking of jobseekers as the policymaker's utility score. The ranking is identical under the following two conditions:

Condition 1. The treatment effect of the ALMP is monotonically decreasing in baseline reemployment probability. That is, the treatment effect of an ALMP is always larger for jobseekers with lower reemployment probabilities. A monotonic relationship between treatment effects and baseline predicted outcomes is commonly, albeit implicitly, assumed in standard algorithmic decision-making settings with well-defined interventions, such as pretrial detention. Detention reliably incapacitates defendants, implying that its marginal benefit is highest for those at greatest risk of misconduct and lowest for those at low risk. Accordingly, the causal effect of detention is increasing in baseline misconduct risk.

However, in the context of ALMPs, the assumption on monotonic treatment effects is less straight forward. First, it is not obvious that the *average* relationship between baseline reemployment probability and treatment effect is monotonically decreasing. One can for example plausibly imagine an inverse U-shaped relationship, where the causal impact of ALMP is highest for individuals with intermediate baseline reemployment probabilities. [7] Second, for the condition to hold, it is not sufficient that the relationship is decreasing on average; it must hold across all jobseekers.

Condition 2. The policymaker exclusively values fairness. The second condition under which assigning ALMPs based on predicted reemployment probability

[7] Moreover, the shape of this relationship is likely to depend on the time horizon considered. For instance, jobseekers with low reemployment probabilities may require more intensive or longer-term interventions — such as vocational training — with delayed payoffs due to lock-in effects. By contrast, jobseekers with intermediate reemployment probabilities might benefit more immediately from short-term support such as job search assistance.

maximizes the policymaker's utility is if they care exclusively about fairness and place no weight on efficiency. In this case, the objective reduces to assigning ALMPs purely based on baseline reemployment probability. The goal of the PES, as stated by the government is to reduce long-term unemployment. Which would speak against them putting zero weight on efficiency.

These conditions are formally established and further discussed in Appendix G.

If the policymaker indeed values both fairness and efficiency, it is unlikely that an algorithmic recommendation based solely on baseline reemployment probability is utility-maximizing. Even if the relationship between y_i^0 and γ_i is on average monotonically decreasing, there likely exists marginal cases for which it is optimal to trade off fairness for efficiency (see Figure G1). This fairness-efficiency trade-off creates scope for human decision-makers to improve upon the algorithm by incorporating broader policy objectives. The setting thus offers a natural case where human–algorithm collaboration could lead to improved outcomes.

4.5 Model Implications

The algorithm assigns ALMPs based on reemployment probability and applies the specific cutoff s_a. Following the conceptual framework, a deviation from the algorithmic recommendation comes from a difference along one of the following three dimensions:

1. Prediction of baseline reemployment probability $\hat{y}^0_{i,j}$.
2. The chosen cutoff s_j, measured by assignment rate — propensity to assign jobseekers to an ALMP versus independent search.
3. Parameter α_j— relative weight on efficiency versus fairness in assignment. And caseworker prediction of treatment effects $\hat{\gamma}_i$.

I will examine jobseeker outcomes in terms of fairness and efficiency in order to (a) analyze whether jobseeker discretion improves or reduce policymaker utility compared to the algorithmic counterfactual, and (b) examine the presence of a fairness-efficiency trade-off and the optimal assignment mechanism design.

5 Empirical approach: Caseworkers versus Algorithm

In this section I will build on the conceptual framework to set up the empirical strategy to evaluate the impact of caseworker discretion in terms of (1) fairness and (2) efficiency.

5.1 Fairness

To analyze the impact of caseworker discretion over algorithmic recommendations on the fairness of the assignment, I will compare the observed caseworker assignment with an estimated algorithmic counterfactual assignment. The empirical approach follows Angelova, Dobbie, and Yang (2023). A fair assignment mechanism will assign jobseekers to an ALMP based on baseline reemployment probability y^0. In order for an assignment mechanism to be fair at any cutoff s, it must correctly rank jobseekers in terms of their reemployment probability. For each caseworker j, we cannot observe their ranking of jobseekers, instead we observe the assignment decision for each jobseeker together with their assignment rate. However, for the algorithm we observe the profiling score, which maps into the complete ranking of jobseekers in terms of the predicted reemployment probability. An important insight is that for the algorithm, each potential cutoff s_a is associated to some assignment rate. Meaning that for each caseworker, I can back out what cutoff would result in their assignment rate S_j, and reconstruct the counterfactual assignment of jobseekers under the algorithm. Any difference will be due to caseworker deviations from the recommendation.

To facilitate the comparison, I define the caseworker conditional employment rate as the employment rate among jobseekers assigned to Independent Search, for whom I can observe the employment outcomes in the absence of any ALMP. Caseworkers whose ranking of jobseekers are closer to a ranking in terms of reemployment probability will achieve a higher conditional employment rate, given their cutoff s_j and pool of jobseekers.

$$M_j = E[Y_i | D_{i,j} = 0] \tag{11}$$

Next, in order to compare the assignment of a caseworker versus the algorithm, I need to define a measure of algorithmic counterfactual conditional employment rate that is comparable to the caseworker's conditional employment rate.

The algorithm's conditional employment rate at the assignment rate of caseworker j is:

$$M_{a(j)} = E[Y_i | D_{a(j),i} = 0] \tag{12}$$

Note that, due to caseworker deviations, the pool of jobseekers assigned to Independent Search under the algorithmic counterfactual can be different from the actual pool of jobseekers assigned to independent search by the caseworker. Leading to a different conditional employment rate at the same cutoff, $s_j = s_a$.

Under random assignment of jobseekers to caseworkers, the average $M_{a(j)}$ will differ from M_j if the caseworker disagrees with the algorithmic prediction or ranking of jobseekers.

The impact of caseworker discretion on the fairness of the assignment, at each caseworker's assignment rate, will be measured by the difference in conditional employment rate:

$$\Delta M_{j,a(j)} = M_j - M_{a(j)} \tag{13}$$

Where $\Delta M_{j,a(j)} > 0$ implies that a caseworker assignment achieve higher fairness than the algorithm. The intuition behind this metric is that at a given assignment rate, a higher conditional employment rate implies better identification of jobseekers likely to find employment independently. Therefore, caseworker discretion improves fairness if their conditional employment rate exceeds that of the algorithm at the same assignment rate.

Note that the difference in fairness results from (i) caseworker weight on fairness α_j is smaller than 1, or (ii) that caseworkers are worse at predicting reemployment probability. Importantly, this metric does not allow me to separate the two channels. Under the assumption $\alpha_j = 1$, the whole difference could be attributed to a difference in prediction accuracy.

5.1.1 Estimating the Algorithmic Counterfactual: the Selective Labels Problem

There is one main complicating factor in estimating the algorithmic counterfactual employment rate among jobseekers assigned to Independent Search. A missing data problem, commonly referred to as the selective labels problem (Kleinberg et al., 2018). Specifically, a jobseeker's employment outcome under Independent Search Y_i^0 is not observable if the jobseekers was assigned to an ALMP.

To address the selective labels problem, I will follow the extrapolation approach suggested by Angelova, Dobbie, and Yang (2023), and which builds on work by Arnold et al. (2022) and Hull (2020). Angelova et al. (2023) study human discretion over an algorithm in the context of pre-trial hearings in the US. The identification strategy relies on random assignment of jobseekers to caseworkers in order to ensure unbiased estimates, and differences in caseworkers assignment rates to get exogenous variation in the assignment of jobseekers.

There are two main alternatives to estimate the algorithmic counterfactual employment rate under Independent Search. First, one could compute the employment rate using only jobseekers for which we observe the outcomes under Independent Search. To be unbiased, this approach requires assuming that caseworker assignment decisions are uncorrelated with true employment potential — that is, effectively random. Second, one could compute the employment rate using observed outcomes for all jobseekers. But such comparison requires assumptions about the causal effects of ALMPs.

In contrast, estimating the algorithmic counterfactual using the Angelova et al. (2023) extrapolation approach requires no assumptions about caseworkers' skill and use of private information or ALMP effectiveness.

5.1.2 Estimating the Caseworkers Parameters

The caseworker-specific assignment rates and employment rates can be estimated using the following specification.

$$1 - D_{i,j} = S_j + u_{i,j} \tag{14}$$

$$Y_{i,j}^{D=0} = M_j + v_{i,j} \tag{15}$$

Where $(1 - D_{i,j})$ and $Y_{i,j}^{D=0}$ are indicators for assignment to Independent Search and employment for jobseeker i conditional on assignment to Independent Search, respectively. S_j is the estimated assignment rate and M_j is the estimated conditional employment rate of caseworker j.

The caseworker conditional employment rate will be estimated with sampling error, due to a finite number of observations per caseworker. The sampling error will inflate the variance in the caseworker employment rate, which could bias the comparison to the algorithmic counterfactual. To correct for the sampling error I use linear empirical Bayes shrinkage, as defined by Walters (2024), shrinking the parameters towards the mean. See Appendix B for details. I use the shrinked conditional employment rates when calculating the share of caseworkers with higher/lower conditional employment rate than the algorithm as well as the difference in conditional employment rate between the caseworkers and algorithm.

When applying shrinkage to variables that are subsequently used in regression, one has to be mindful. While shrinking an independent variable reduces attenuation bias in the slope parameter, shrinkage of a dependent variable will instead induce bias (Walters, 2024). Thus, the empirical Bayes shrinkage is not applied when studying the relationship between caseworker conditional employment rate and assignment rate. The regression is instead weighted by the inverse variance of the estimated conditional employment rate, to put more weight on more precise estimates.

5.1.3 Estimating the Algorithmic Counterfactual Using Extrapolation: Intuition

In order to provide intuition for the estimation of the algorithmic counterfactual employment rate, consider an idealized setting where there is random assignment of jobseekers to caseworkers and where there is one caseworker, caseworker j^*, who assigns all jobseekers to Independent Search. The employment rate among the jobseekers assigned to caseworker j^*

should be close to the average employment potential in the population of jobseekers.

The same logic is applicable to any score threshold. Consider instead that caseworker j^* assigns all jobseekers at or above a given threshold s_a to Independent Search, such that the assignment rate at the threshold s_a is equal to one. Such behavior is consistent with (i) having a strong preference for assigning jobseekers to Independent Job Search, such that their assignment rate is close to one, or (ii) with the caseworker being totally compliant with the algorithm at the threshold s_a. With random assignment of jobseekers to caseworkers, the employment rate among that subsample of jobseekers should be close to the average employment rate for the subset of jobseekers with profiling scores at or above the threshold.

In the absence of a caseworker j^* the employment potential can be estimated using extrapolation to assignment rate 100 % within each subsample.

5.1.4 Estimating the Algorithmic Counterfactual Using Extrapolation: Implementation

The implementation of the empirical strategy can be summarized in the following steps:

1. **Restrict the sample to jobseekers with scores at or above the given threshold,** $a(x_i) \geq s$
   For the algorithm, we know that each assignment threshold, s, is associated with some algorithmic assignment rate, S. Thus, etimating the algorithmic counterfactual employment rate at some assignment rate, S, is equivalent to doing so for the subset of jobseekers with score at or above the associated assignment threshold.

2. **Estimate the caseworker level assignment rates and employment rates within the sub-sample.**
$$1 - D_{i,j} = S_j + u_{i,j}$$
$$Y_{i,j}^{D=0} = M_j + v_{i,j}$$

3. **Estimate the average employment potential in the sub-sample**
   Regress the estimated caseworker employment rates on assignment rates. The average employment potential within the sub-sample of jobseekers with $a(x_i) \geq s_a$ is equivalent to the employment rate of a caseworker with an assignment-rate of 100 % within the sub-sample (due to random assignment of jobseekers to caseworkers). In the absence of such a caseworker, we need to estimate the employment rate using extrapolation.

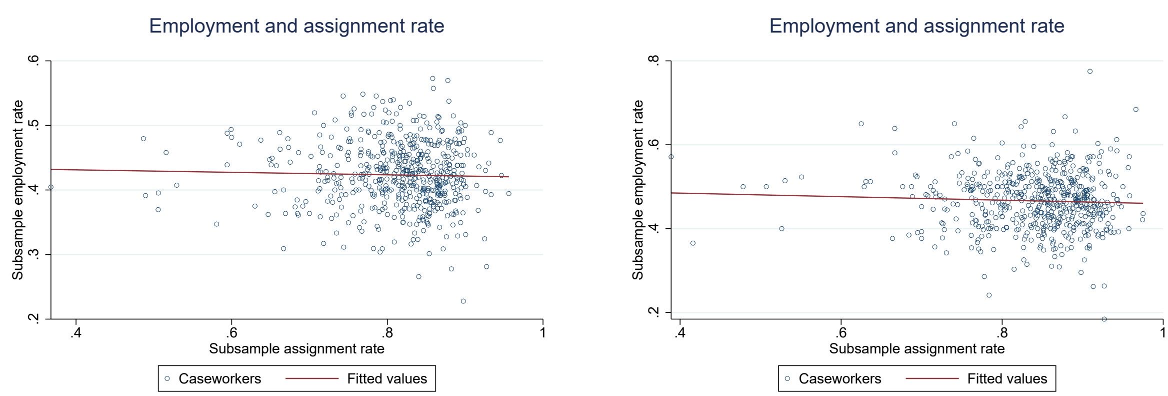
   For linear extrapolation, center the assignment rate estimates at $S_j = 1$, i.e. by defining a new variable $\hat{S}_j^C = \hat{S}_j - 1$. Then run the regression:

$$\hat{M}_j = \beta_0 + \beta_1 \hat{S}_j^C + e_i \tag{16}$$

Due to centering around $\hat{S}_j = 1$, β_0 is the estimated employment rate at an assignment rate of 100 % within the sub-sample, which provides the estimate of the threshold-specific average employment rate in the sub-sample of jobseekers with $a(x_i) \geq s$. That is, β_0 provides the estimate of $M_{a(j)}$. The regression is weighted by the inverse variance of the estimated conditional employment rate, to put more weight on more precise estimates.

Repeat the procedure for many thresholds such that the whole span of caseworker assignment rates is covered. The estimate of the population average employment potential can be obtained by using the whole sample of jobseekers and extrapolate to $S_j = 1$. For robustness, extrapolate using different functional forms.

Figures below show the extrapolation to assignment rate equal to one for two subsamples. Jobseekers in panel (b) is a subset of jobseekers in panel (a). Note, however, that the caseworker assignment rate and conditional employment rate could differ between the two subset for the same caseworker.



(a) a(X$_i$) $\geq$ 0.25

(b) a(X$_i$) $\geq$ 0.49

Figure 3: Subsample assignment rates and conditional employment rates.

Note: Extrapolation using linear regression and weighted by the inverse variance of the caseworker estimates. Includes balanced panel of caseworkers with at least 75 profilings per year.

Standard errors and confidence intervals for the algorithmic counterfactual conditional employment rate within each sub-sample are constructed using bootstrapping with random draws of caseworkers.

5.1.5 Comparing Caseworker versus Algorithmic Predictive Performance

I will quantify the comparison of caseworker versus algorithmic predictive performance using two summary measures. First, I will estimate the share of caseworkers with higher conditional employment rates than the algorithm. Second, I will estimate the difference in the case-weighted conditional employment rate between caseworkers and algorithm. Standard errors for both measured are constructed using bootstrapping with 1,000 random draws of caseworkers.

5.2 Efficiency

In this section I present the empirical strategy to study efficiency. An assignment is efficient if it assigns ALMP to jobseekers such that the total employment rate is maximized. Thus, the efficiency can be evaluated by comparing the total employment rate between caseworkers and the algorithmic counterfactual. However, as opposed to studying fairness, we cannot directly estimate the total employment under the algorithmic counterfactual assignment. Instead I will compare the total employment outcomes of caseworkers with different levels of deviation rates, the share of cases for which the caseworker deviate from the algorithmic recommendation. Where a deviation is defined in terms of the choice between Independent Search and some ALMP. The empirical approach relies on a typical 'judge design' (Chyn et al., 2025), exploiting random assignment of jobseekers to caseworkers and variation in caseworker propensity to deviate.

The employment rate of each caseworker, $E_j = E[Y_{ij}]$, can written as the weighted sum of the conditional employment rate for jobseekers assigned to Independent Search, M_j, and the conditional employment rate for jobseekers assigned to an ALMP.

$$E_j = S_j \cdot E[Y_i|D_{i,j} = 0] + (1 - S_j) \cdot E[Y_i|D_{i,j} = 1] \tag{17}$$

The assignment rates S_j will depend on the caseworkers' beliefs and predictions regarding treatment effects of ALMPs. Thus, we should let the assignment rate channel also impact the total employment rate. Therefore, I will define the algorithmic counterfactual rate as the zero deviation employment rate.

Define the algorithmic assignment rate for each caseworker as the rate of assignment they would have achieved under strict adherence to the algorithmic recommendation. S_j^A varies across caseworkers due to sampling variation.

$$S_j^A = 1 - E[A_i|cw_j = 1] \tag{18}$$

Then the algorithmic counterfactual employment rate will be a weighted sum of the employment rate conditional on Independent Search, and the employment rate under the ALMP. Note that the algorithmic recommendation does not recommend a specific treatment, but only recommends Independent Search or an ALMP.

$$E_{a(j)} = S_j^A \cdot E[Y_i|A_i = 0] + (1 - S_j^A) \cdot E[Y_i|A_i = 1] \tag{19}$$

5.2.1 Estimating the Algorithmic Counterfactual: Deviation rates

Under random assignment of jobseekers to caseworkers with different deviation rates, a positive association between employment rate and deviation rate implies that caseworker deviation on average improve efficiency. While a negative association implies that caseworkers are reducing efficiency.

This approach is similar to the setup in Hoffman et al. (2018), with two important differences. First, in this paper the institutional setup ensures random assignment of jobseekers to caseworkers, which alleviate worries that jobseekers would systematically differ on unobservable characteristics. Second, caseworkers can impact jobseeker outcomes via the choice of ALMP.

Random assignment of jobseekers to caseworkers ensures that the first term of Equation 19 is comparable across caseworkers with low and high deviation rates, at a given algorithmic assignment rate. In terms of the second term, a caseworker with zero deviation rate is only a valid estimate of the algorithmic counterfactual if caseworkers are on overage equally skilled in picking the employment maximizing ALMP, such that the expected outcomes are equal at the algorithmic ranking of jobseekers. If caseworkers with higher skill are systematically deviating more in terms of choosing between Independent Search and some ALMP, then we will overestimate the effect of caseworker deviations on total employment. Since highly skilled caseworkers would have obtained higher employment rate also under zero deviations. That would imply that the standard exclusion restriction that the decision maker only impact outcomes through the decision studied, here a deviation, is violated. Due to the institutional setting, we can boil down this violation of the exclusion restriction to a single, testable, channel - whether the caseworker skill in assigning the optimal ALMP and deviation rate is correlated. I find a small positive and statistically insignificant correlation between deviation rate and caseworker value added, defined as employment rate within 6 months among jobseekers assigned to ALMP. See Figure C2.

I will evaluate the impact of caseworker deviation on efficiency holding the algorithmic assignment rate constant. Which allows the caseworkers to impact efficiency through two channels, (i) by adjusting the assignment rate, and (ii) by changing the composition of jobseekers who are assigned to Independent Search or ALMP. I run a regression of jobseeker employment outcomes on the leave-one-out caseworker deviation rate and leave-one-out algorithmic assignment rate at the jobseeker level.

$$y_i = \beta_0 + \beta_1 \text{Deviation rate}_{-i,j} + \beta_2 \text{Alg assignment rate}_{-i,j} + \varepsilon_i \tag{20}$$

β_1 will recover the causal relationship between caseworker deviation rate and total employment. Unlike with the extrapolation approach, I will not get an estimate of the distance from the algorithmic counterfactual for each jobseeker.[8]

6 Identifying Assumptions

The identifying assumptions needed for the two empirical strategies are largely overlapping. Both rely on (i) random assignment of jobseekers to caseworkers, (ii) relevance — caseworker assignment rate (deviation rate) increase the probability of assignment to Independent Search (deviation). (iii) Exclusion restriction — caseworkers only affect outcomes via the decision, assignment or deviation, and not directly. (iv) Lastly, in the absence of caseworkers with 100 % assignment rates, the extrapolation approach requires that average employment parameters can be accurately extrapolated from the data.

Neither of the approaches require a monotonicity assumption. In this setting monotonicity would require a similar ranking of jobseekers across caseworkers, which would be largely inconsistent with the conceptual framework where we allow caseworkers to vary by both their private information and skill in predicting reemployment probabilities.

6.1 Random Assignment of Jobseekers to Caseworkers

The key identifying assumption for the paper is random assignment of jobseekers to caseworkers. The institutional design of assignment of jobseekers at the Swedish PES central offices suggests that the assumption of random assignment holds. Table 1 shows the result of balance tests of both single and joint significance of jobseekers characteristics. The table show results of single regressions of (1) caseworker leave-out assignment rate , (2) leave-out deviation rate, and (3) profiling score on predetermined jobseeker characteristics. The

[8] With a larger dataset—and under valid identifying assumptions—it would be possible to apply the extrapolation approach of Angelova et al. (2023) to estimate outcomes for caseworkers with zero deviation rates. Because caseworker deviation rates are positively correlated with assignment rates, it would be necessary to control for the latter. One non-parametric approach is to extrapolate to zero-deviation caseworkers within each level of algorithmic assignment rate. However, due to the relatively small sample, this estimation is too noisy to be informative.

results show that while jobseeker characteristics are strongly correlated with the profiling score, they are largely uncorrelated with the leave-one-out caseworker assignment rate and deviation rate. Even though some single regressions are statistically significant, the magnitudes are economically insignificant. Regarding joint significance of jobseeker characteristics, the f-test for assignment rate is small and insignificant, while the F-test for deviation rate is small but statistically significant with a p-value if 0.023.

Table 1: Test random assignment of jobseekers to caseworkers

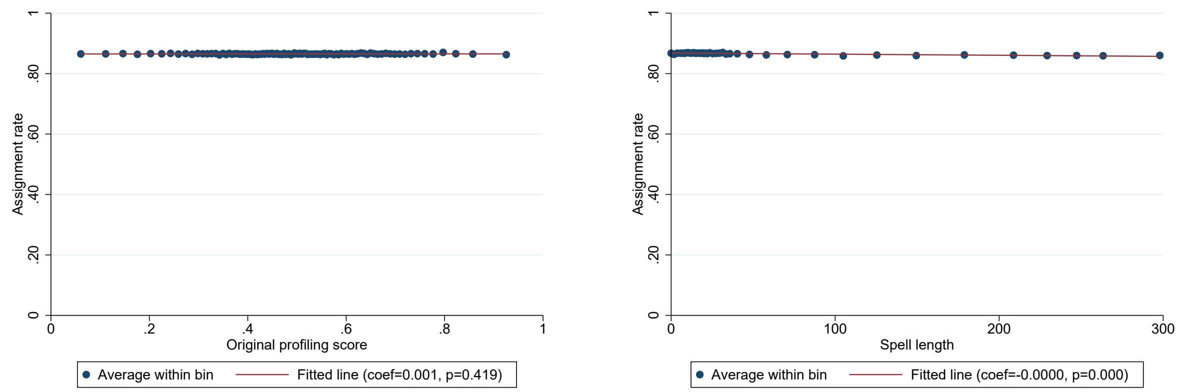
	Assignment Rate	Deviation Rate	Profiling Score
Male	-0.000	-0.001**	0.029***
	(0.000)	(0.000)	(0.001)
Historical functional disability	0.000	-0.000	-0.210***
	(0.000)	(0.000)	(0.002)
Immigrant	-0.000	0.000	-0.087***
	(0.000)	(0.000)	(0.001)
High school	-0.001	0.001	0.004
	(0.001)	(0.001)	(0.004)
Tertiary education	0.001**	-0.001***	0.025***
	(0.000)	(0.000)	(0.001)
Age 25-34	0.000	-0.000	0.051***
	(0.000)	(0.000)	(0.001)
Age 35-44	0.000	-0.000	-0.015***
	(0.000)	(0.000)	(0.001)
Age 45-54	-0.000	0.000	-0.033***
	(0.000)	(0.000)	(0.001)
Service and admin	-0.000	0.000	-0.022***
	(0.000)	(0.000)	(0.001)
Service, care and shop	-0.000	0.000	-0.057***
	(0.000)	(0.000)	(0.001)
Building and manufacturing	0.000	-0.001	0.071***
	(0.000)	(0.000)	(0.002)
Previously registered	0.000	0.000	-0.021***
	(0.000)	(0.000)	(0.002)
Previous treatment	-0.000	0.000	-0.074***
	(0.000)	(0.000)	(0.001)
F-test	1.37	1.93	2,074
P-value f-test	0.163	0.023	0.000
N	99,959	99,959	99,959
N caseworkers	528	528	528
Mean depvar	0.87	0.15	0.49

Note: Robust standard errors reported in parentheses. ***, **, * indicate significance at 1%, 5% and 10% respectively. Single regressions of the leave-one-out assignment rate and profiling score on predetermined jobseeker characteristics. The F-test tests for joint significance of jobseeker characteristics. Includes balanced panel of caseworkers with at least 75 profilings per year.

Figure 4 shows a plot of assignment rates against (a) the profiling score, and (b) the length of the unemployment spell. There is no correlation between leave-out assignment rate and profiling score. There is a small but statistically significant negative relationship between leave-out assignment rate and spell length. Note that the distribution of spell length is highly skewed. An increase in spell length from the 25th to the 75th percentile (0 to 19 days) is associated with a 0.05 standard deviation decrease in the average assignment rate.

Which can be considered a very small effect. See Figure C1 for the equivalent plots for deviation rate.

Altogether, the balance tests confirm that the institutional setting approximates random assignment of jobseekers to caseworkers.



(a) Profiling Score

(b) Spell Length

Figure 4: Plot assignment rate against profiling score

Note: Binscatter and fitted line of assignment rate against (a) profiling score and (b) length of unemployment spell. Assignment rate is the leave-out mean. Regressions are weighted by the inverse of caseworker case load. Coefficients reported from regressions on the jobseeker level. Robust standard errors used in panel (a) and bootstrapped standard errors in panel (b). Includes balanced panel of caseworkers with at least 75 profilings per year. Note the distributed of number of days unemployed at profiling is highly skewed, with 75 % of observations with less than 19 days. An increase in spell length from the 25th to the 75th percentile (0 to 19 days) is associated with a 0.05 standard deviation decrease in the average assignment rate (equivalent to 0.0019 percentage points).

6.2 Relevance

There is a positive and statistically significant relationship between caseworkers' leave-out assignment rate and the probability that a jobseeker is assigned to Independent Search. With a coefficient of 0.461 and standard error of 0.031. Likewise, there is a positive and statistically significant relationship between caseworkers' leave-out deviation rate and the probability of deviation with a coefficient of 0.424 and standard error of 0.033. See Table C1. It shows that being assigned to a caseworker with a higher assignment rate (deviation rate), significantly increases the probability that a single case gets assigned to Independent Job Search (deviation) and implies that the relevance assumption holds.

6.3 Average Employment Parameters Can Be Accurately Extrapolated

In the absence of caseworkers with 100 % assignment rates, the average employment potential within each subsample is estimated using extrapolation. Whether the extrapolation provides a consistent estimate of the average employment potential will depend on (i) approximating the correct functional form, and (ii) the distance between caseworker observations and the point we are extrapolating to, i.e. whether there are sufficiently many caseworkers with high assignment rates.

Regarding caseworker assignment rates, there is significant variation in assignment rates, from 71 % to 97 % in the full sample, which is similar to Angelova, Dobbie, and Yang (2023). Figure 5 shows the distribution of assignment rates at caseworker level for four subsamples. Each subsample consists of all jobseekers with profiling scores above the given cut-off and the assignment rates are calculated within the specific subsample. A higher profiling score indicates a higher predicted reemployment probability. The figure illustrates how the distribution of assignment rates shift to the right as the cut-off increases and the subsample gets more and more restrictive.

Regarding approximating the correct functional form, I will use linear extrapolation as the main specification and do robustness checks using quadratic and local linear extrapolation.

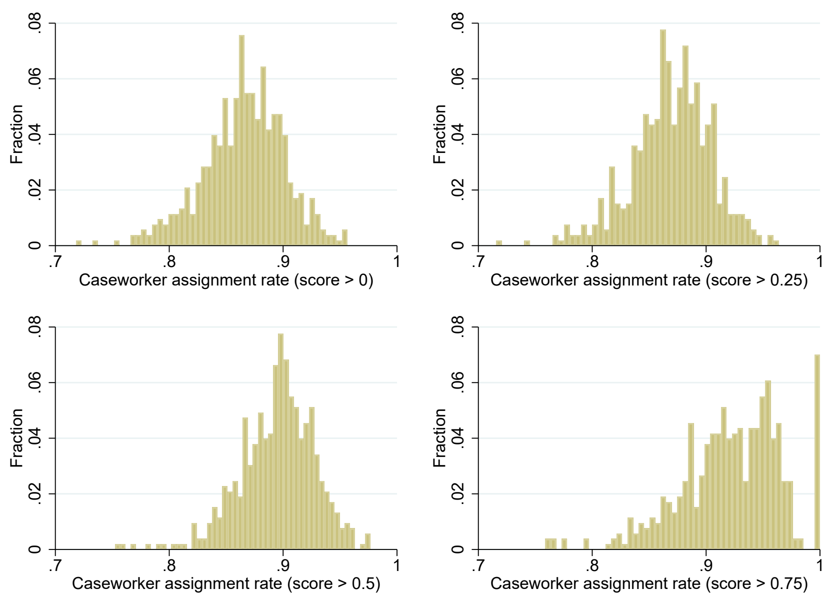


Figure 5: Caseworker assignment rates within subsamples

Note: Caseworker level. Includes balanced panel of caseworkers, who are active at least 10/12 months, with at least 75 profilings per year.

6.4 Exclusion Restriction

The institutional setting makes the exclusion restriction a credible assumption. Caseworkers main task is to make labor market assessments and aid jobseekers in administrative matters. Assistance and other active labor market programs are carried out by other state or municipal agencies or private companies. Monitoring of job search activities are handled by a separate monitoring unit at the PES. Interviews with caseworkers show that one margin they could use to adjust the assistance to jobseekers assigned to Independent Search is the number of contacts and follow-up meetings. They did not however provide any practical help with things such as CV, cover letter or job referrals. Altogether, it is therefore unlikely that caseworkers have any large direct effects on the outcomes of jobseekers beyond their assignment decision.

6.5 Caseworker Objectives

It is possible to evaluate caseworker assignment in terms of fairness and efficiency without assuming anything regarding caseworker objectives. In this paper I assume that the policymaker values fairness and efficiency similarly across different subgroups. However, caseworkers might weigh the utility of different demographic groups differently, such that they would prioritize assistance to certain groups such as jobseekers with children or jobseekers who are natives. Any such distractions from the utility maximization is considered as a result of caseworker bias.

7 Descriptive Statistics: Caseworker Deviations

In order to interpret the results as a result of caseworker deviations from the algorithmic recommendation, I need to adjust the profiling score and take the duration specific cut-offs into account. This section starts with a description of how the adjusted score is computed, followed by descriptive statistics of caseworker deviations.

7.1 Adjusted Profiling Score

Each profiling event yields an *original profiling score*, between 0 and 1, and a categorical *algorithmic recommendation* (Independent Search, Job Search Assistance, or Intensive Support). Since the cutoff between Independent Search and any ALMP varies by the number of days since the start of the unemployment spell, see Figure 2, I construct an *adjusted score* that captures the distance (in profiling-score units) to the independent-search cutoff at each jobseeker's spell length.

Formally, let $s_{i,t}$ be the original profiling score for jobseeker i at time t, and let τ_d be the cutoff in place for a jobseeker with d days of unemployment. As described above, τ_d changed over the study period, why I allow it to vary over time. Define the *adjusted score*:

$$\widetilde{s}_{i,t} \;=\; s_{i,t} \;-\; \tau_{d_{i,t}}\,.$$

By construction, $\widetilde{s}_{i,t} > 0$ implies "recommended Independent Search," while $\widetilde{s}_{i,t} \leq 0$ implies "recommended ALMP." In the subsample of jobseekers with $\leq$ 180 days of unemployment — when the cutoff is flat at 0.2 — the original and effective scores yield identical rankings. Almost 95 % of decisions happens before 180 days, see Figure A4 for the full distribution. In the full sample, the correlation between the original and adjusted score is as high as 0.92, indicating that they largely agree on the ordering of jobseekers.

I then normalize the adjusted score to lie in $[0, 1]$ for easier interpretation. In the main analysis I will use this normalized adjusted score, instead of the original profiling score. The inferred algorithmic recommendation based on $\widetilde{s}_{i,t}$ matches the recorded recommendation in 99.8% of cases.

7.2 Caseworker Deviations

Let's start by descriptively looking at caseworker deviations. Table 2 shows the matrix of the algorithmic recommendation and caseworker assignment decisions. The table confirms that caseworkers had very limited ability to assign jobseekers to Job Search Assistance if the algorithm recommendation was Independent Search or Intense Support. Instead we observe a substantial number of deviations from independent search to intense support, where caseworkers had free assignment discretion. Likewise, the deviation rate from Intense Support to Independent Search is very high at 72 %.

Table 2: Algorithmic recommendation and caseworker assignment

		Caseworker assignment		
		Independent search	Job search assistance	Intense support
Algorithmic recommendation	Independent search	81,504	0	9,412
	Job search assistance	2,025	1,362	1,318
	Intense support	3,132	0	1,206

Figure 6 shows the distribution of deviation rates across caseworkers.

Figure 7 plots the adjusted score against the probability of being recommended (panel a) or actually assigned (panel b) to Independent Search. In panel (a), the algorithmic recommendation displays a sharp discontinuity at the cutoff: as the score crosses zero, the probability of being recommended jumps from 0 to 0.95. Strikingly, panel (b) shows that

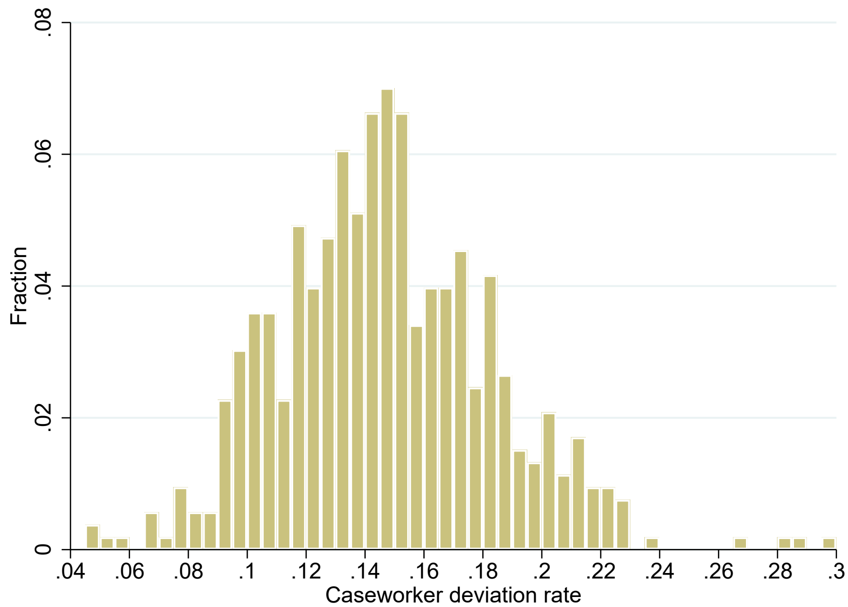


Figure 6: Caseworker deviation rate

Note: Deviation rate at caseworker level. Mean=14.6 %.

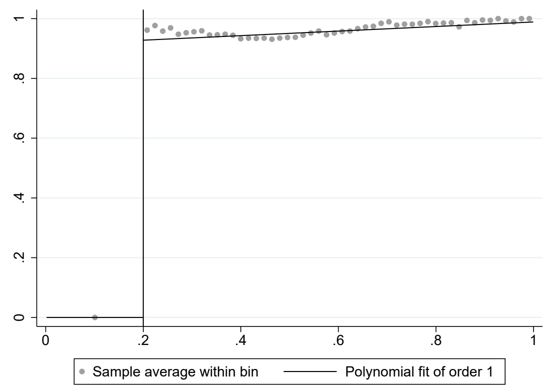
caseworker assignment is a smooth, continuous function of the score, with no discrete jump at the cutoff. The smoothness around the cutoff implies that for the marginal cases, the algorithmic recommendation is irrelevant for the caseworkers' assignment decision. Empirically, the caseworkers are deviating more close to the cutoff, which is consistent with the smoothness around the cutoff. See Appendix Figure A4 for a figure of how the deviation rate changes against adjusted score and the length of the unemployment spell.

8 Results

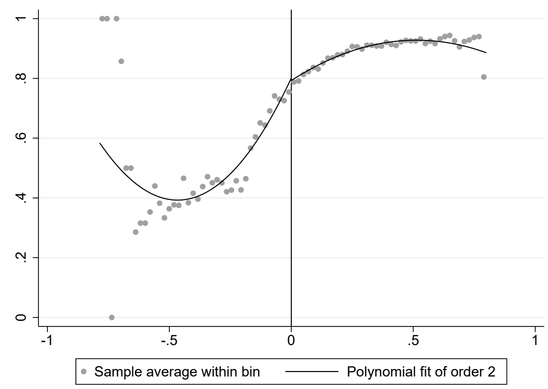
The empirical results will be presented in the following order. First the results on *fairness* and *efficiency* will be presented separately. After which I will examine the interaction and the fairness-efficiency trade-off. Second, I will examine a different margin of fairness - whether access to ALMP is similar across groups, conditional on reemployment potential. Lastly I will discuss potential mechanisms for heterogeneity in caseworker performance and some robustness checks.

8.1 Fairness: Predictive Performance

First I will examine the results on fairness, following the empirical approach presented in Section 5. Figure 8 compares caseworker and algorithmic performance by plotting the condi-



(a) Algorithmic recommendation



(b) Caseworker assignment

Figure 7: Probability of Recommendation and Assignment to Independent Search

Note: Probability of being (a) recommended and (b) assigned to independent search using the normalized effective score which accounts for varying cutoffs. RD-plot using 40 bins and a quadratic functional form fit.

tional employment rate against the assignment rate. Each point represents one caseworker's observed conditional employment rate at their own assignment rate. The orange line is a linear-regression fit through those points, showing the average caseworker conditional employment rate for any given assignment rate. The blue line is the algorithmic counterfactual estimate, indicating what the conditional employment rate would be at each assignment rate if assignments followed the algorithm's ranking instead of the caseworkers'. The algorithmic counterfactual is estimated using extrapolation from linear regression, and the points are combined using local linear interpolation.

The dashed horizontal line shows the conditional employment rate under random assignment — which is equal to the the overall population employment potential. The logic is that if caseworkers would randomly assign jobseekers to Independent Search, they should obtain the average conditional employment in the population. By construction, all three lines (the caseworker average, the algorithmic counterfactual, and the random-assignment benchmark) coincide at an assignment rate of 1, since assigning everyone to Independent Search yields the same expected employment rate.

The red points indicate the caseworker and algorithm average. Overall the algorithmic counterfactual would have obtained an assignment rate of 91 % and a conditional employment rate of 41 %. Caseworker deviations lead to an assignment rate of 86.5 and conditional employment rate of 40 %.

If a decision maker is good at ranking jobseekers by their reemployment probability, we would expect to see a negative relationship between the conditional employment rate and the assignment rate. The intuition is that a decision maker with a low assignment

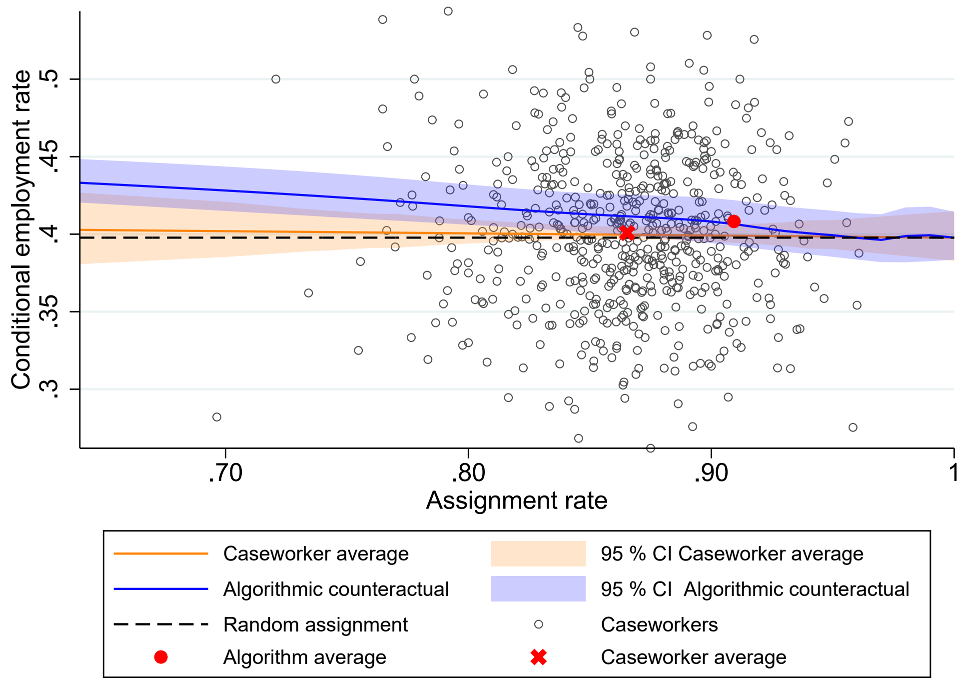


Figure 8: Full time employment within six months lasting four months

Note: Employment outcomes are based on monthly payroll tax data. The fitted lines are estimated using linear regression. Algorithmic counterfactual estimated using linear regression within each subsample. Includes balanced panel of caseworkers with at least 75 profilings per year. For caseworker average, confidence intervals are constructed using robust standard errors. For the algorithmic counterfactual confidence intervals constructed using bootstrapping.

rate, who is more selective, should assign the jobseekers with the highest reemployment probability to Independent Job Search first. So that their pool of jobseekers, for whom the employment rate is calculated, are more likely to become employed on their own. Notably, there is a negative relationship between conditional employment rate and assignment rate for the algorithm. Across caseworkers the relationship is more or less flat, indicating that caseworker discretion from the algorithmic recommendation leads to as assignment equivalent to random assignment.

Table 3 shows the comparison of algorithm and caseworkers in terms of fairness in assignment to ALMPs, using (i) the share of caseworkers achieving a higher conditional employment rate than the algorithm, and (ii) the difference in caseworker conditional employment rate after applying linear empirical Bayes shrinkage. It shows the comparison to both the original algorithm and the resulting recommendation after applying the PES's time-varying cutoffs.

Only 9.2 % of caseworkers perform better than the algorithm and 48.8 % of caseworkers perform better than a random assignment. Resulting in a 1.3 percentage point lower conditional employment rate relative to the algorithm and a similar conditional employment rate as a random assignment, conditional on assignment rate.

Compared to a random assignment, using the algorithm increases the conditional employment rate by 1.3 percentage points. However, when comparing the original algorithm (based on the raw profiling score) to the implemented version used by the PES — which incorporates time-varying cutoffs — I find that the implementation distorts the ranking of jobseekers and significantly worsens predictive performance. Specifically, the PES implementation shifts access to ALMPs toward jobseekers with longer current unemployment durations, while reducing access for those at high risk of long-term unemployment but with shorter durations.

Table 3: Caseworker vs algorithmic prediction accuracy

	(1) Caseworker vs Original alg	(2) Caseworker vs PES alg	(3) Caseworker vs Random	(4) Original alg vs Random	(5) PES alg vs Random
Share CW higher emprate	0.014***	0.092***	0.488***		
	(0.005)	(0.015	(0.025)		
Average difference emprate	-0.019***	-0.013***	-0.000	0.019***	0.013***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
N	528	528	528		

Note: ***, **, * indicate significance at 1%, 5% and 10% respectively. Caseworker vs algorithmic counterfactual, holding the assignment rate constant. Row (1) shows the share of caseworkers with a higher employment rate than (1) the original algorithm, (2) the PES algorithm taking days registered specific cutoffs into account, and (3) a random assignment of jobseekers. Row (2) shows the weighted average difference between caseworker conditional employment rate and algorithmic counterfactual and random assignment, as well as the algorithm versus a random assignment. Employment outcome is full time employment within six months lasting four months, based on monthly payroll tax data. The caseworker conditional employment rate are estimated using linear empirical Bayes shrinkage. Includes balanced panel of caseworkers with at least 75 profilings per year. Standard errors constructed using bootstrapping.

8.1.1 Prediction Accuracy

To assess whether the reduction in fairness stems from poor predictive performance, I exploit a unique feature of the data: when caseworkers deviate from assigning Job Search Assistance, they must record their reason, often in free-text form. Focusing on cases where caseworkers assign jobseekers to Independent Search due to "high reemployment probability" allows me to estimate their prediction accuracy. As long as the random rate of 40 % is an accurate estimate of the average employment potential also within this subset of jobseekers, we can use that as a benchmark to evaluate the prediction accuracy. If caseworkers were randomly deviating, they would obtain a conditional employment rate in this group of 40 %. In this group, the observed employment rate is 46.8%, significantly above the random benchmark of 40% (see Table 6). This indicates that caseworkers are relatively successful at identifying jobseekers likely to find employment on their own — suggesting they are not merely making noisy or systematically incorrect predictions.

Yet, despite this above-random predictive performance, caseworkers on average reduce fairness relative to the algorithm. This implies that deviations are not solely about predicting reemployment probability, but may reflect deliberate efforts to trade off fairness for efficiency (i.e., total employment).

8.2 Efficiency: Total Employment Outcomes

In this section I analyze the impact of caseworker deviations on efficiency, by examining the relationship between caseworker leave-one-out deviation rate and employment rate at the jobseeker level.

There is a significant negative correlation between assignment rate and deviation rate (Figure E2). In addition, locking in effects of ALMP induces a positive correlation between employment outcomes and assignment rate in the short run.

To ensure unbiased estimates I control for the leave-out algorithmic counterfactual assignment rate of each caseworker, i.e. the assignment rate of each caseworker if they had perfectly complied with the algorithmic recommendation.

$$y_i = \beta_0 + \beta_1 \text{Deviation rate}_{j,-i} + \beta_2 \text{Alg assignment rate}_{j,-i} + \varepsilon_i \tag{21}$$

Table 4 shows the regression results with and without controlling for algorithmic assignment rate. The outcome tested is full time employment within 6 and 12 months. Results are robust to controlling for leave-out algorithmic assignment rate. The coefficient in column 2 (-0.091) can be interpreted as when moving from the 10th to the 90th percentile of deviation rate, full time employment is reduced by 1pp or 2 % of the mean employment rate. Showing that the average impact of caseworker deviation on employment is small, but negative.

Column 4 shows that the negative association between deviation rate and employment holds also when considering employment 12 months post profiling.

Table 4: Deviation Rate

	Full time employment			
	6 months	6 months	12 months	12 months
CW deviation rate	-0.105**	-0.091*	-0.137**	-0.165**
	(0.045)	(0.050)	(0.065)	(0.072)
Alg. assignment rate		X		X
Number of caseworkers	99,954	99,954	50,109	50,109
Number of jobseekers	528	528	528	528

Note: ***, **, * indicate significance at 1%, 5% and 10% respectively. Robust standard errors in parentheses. Includes balanced panel of caseworkers with at least 75 profilings per year. Mean deviation rate is 0.147 and mean algorithmic assignment rate is 0.909 in the full sample and both measures are leave-one-out means.

Due to data limitations it is not possible to study longer run effects, and I can thus not rule out that caseworker deviations are more important for long run outcomes.

8.3 Mechanisms

In order to inform policy on algorithmic governance we need to understand what drives differing caseworker performance. First I will study different dimensions of caseworker performance. Second, I take a closer look at the reasons for deviating, and third, how high and low performing caseworker differ on characteristics and behaviors.

8.3.1 High and Low Performing Caseworkers

In this analysis I examine the interaction between fairness and efficiency. Specifically, I split the sample by caseworkers who are improving versus reduce fairness in assignment to ALMP relative to the algorithmic counterfactual. I study three types of outcomes, (i) overall employment outcomes, (ii) outcomes among jobseekers assigned to ALMP, and (iii) the predictive performance of caseworkers. Table 5 report the results by group of outcomes.

Results on overall jobseeker outcomes show that caseworkers who improve fairness relative to the algorithm are achieving higher employment rates and cumulative earnings both in the short and medium run. The employment rates can be compared to the estimated random rates, that is the employment potential if everyone was assigned to Independent Search which is 40 % at 6 months and 55 % at 12 months. The results indicate that the low performing caseworkers would have obtained a higher employment rate by simply assigning all jobseekers to Independent Search. These results show that there is a positive relationship between fairness and efficiency across caseworkers.

Table 5: Caseworker Predictive Performance and Value Added

	Below algorithm	Above algorithm	P-value diff
Jobseeker Outcomes			
Ft employed 6m	0.366	0.437	[0.000]
	(0.039)	(0.028)	
Ft employed 12m	0.503	0.570	[0.000]
	(0.063)	(0.053)	
Cum earnings 6m	76,151	85,293	[0.000]
	(7,872)	(6,213)	
Cum earnings 12m	177,194	201,235	[0.000]
	(21,383)	(18,642)	
Jobseeker Outcomes - ALMP			
Ft employed 6m	0.188	0.180	[0.572]
	(0.088)	(0.119)	
Ft employed 12m	0.280	0.301	[0.408]
	(0.162)	(0.230)	
Cum earnings 6m	40,614	39,187	[0.509]
	(13,748)	(16,374)	
Cum earnings 12m	89,900	97,429	[0.254]
	(40,864)	(58,740)	
Predictive Performance			
Emprate deviation high reemployment probability	0.455	0.622	[0.155]
	(0.432)	(0.456)	
Number of jobseekers	90,777	9,182	
Number of caseworkers	482	46	

Note: Robust standard errors reported in parentheses. ***, **, * indicate significance at 1%, 5% and 10%. Observation are missing for some variables. There is data 12 months post profiling for 50,113 jobseeker, across 482 low performing and high performing caseworkers. There are deviations due to high reemployment probability in 356 cases, across 178 low performing caseworkers and 15 high performing caseworkers.

The comparison of outcomes of jobseekers assigned to ALMP will tell us something about whether caseworkers differ in their skill of assigning jobseekers to their employment maximizing ALMP. Results show that there are no significant differences between the high and low performing caseworkers.

Lastly i will discuss the result on predictive performance. Predictive performance is measured by using employment rates among jobseekers for whom the caseworker has deviated from an ALMP to Independent Search due to "high reemployment probability". These results are noisy due to a very small sample and should be interpreted cautiously. As above, we can benchmark the results against the random rate within 6 months at 40 %. Where an employment rate below 40 % implies that caseworkers are making systematic prediction mistakes, while an employment rate above 40 % implies that they are successful in predicting reemployment probability. The results show that high performing caseworkers achieve 35 % higher conditional employment rate, indicating a substantial difference in predictive performance. The difference is however not statistically significant.

8.3.2 Reason for Deviating

To examine reasons for deviating I will exploit a unique feature of the data. When caseworkers deviate from Job Search Assistance they are required to report their reason for deviating, typically through free-text responses. This section focus on the reasons for deviating from Job Search Assistance to Independent Search, for which we can observe jobseeker outcomes.

Table 6 shows the conditional employment rate of jobseekers recommended job search assistance but assigned to independent search by reason for deviation. The most common reason for deviating is that the jobseeker is already working to some degree, or will start a job within the next few months.

Notably, when caseworkers deviate due to jobseekers declining to participate in the ALMP, the employment rate is very low. To what extent a caseworker takes jobseeker preferences into account when making the assignment decision is ultimately up to the caseworker. However, the caseworker has a right to make an assignment decision against the will of the jobseeker. And if the jobseeker refuses they are due to sanctions of their UI benefits.

Table 6: Conditional Employment Rate Deviation to Independent Search

	Conditional employment rate	Number of jobseekers
Job	0.501**	925
	(0.016)	
High remp prob	0.468**	356
	(0.026)	
JS declines	0.133**	147
	(0.028)	
Company	0.251**	353
	(0.023)	
Transition org	0.279**	244
	(0.029)	

Note: The conditional employment rate of jobseekers recommended job search assistance but where the caseworker deviated to independent search. By reason for deviation. 'Job' indicates that the jobseeker (i) is already working to some degree, or (ii) will start a job within the next few months. 'High reemp' indicates high reemployment probability assessed by the caseworker. 'JS declines' means that the jobseeker has declined the offer. 'Transition org' means that the jobseeker is already receiving help from a transition organization financed by the employer and union. 'company' means that the jobseeker is not eligible due to being in a board of a company. Bootstrapped standard errors in parentheses. ** Indicates that the conditional employment rate is significantly different from the random rate at a 5 % level.

8.3.3 High and Low Performing Caseworkers

An important question to public policy going forward is what we can learn from the higher performing caseworkers. One explanation could be that caseworkers who are more experienced might be performing better on average, and thus also in terms of their predictive performance. Table E2 shows the difference in means between caseworkers who obtain a higher versus a lower conditional employment rate relative to the algorithm. The results show that there are no differences in the monthly wage or experience of caseworkers. They also have a similar number of cases in total. Their behaviors are slightly different though. Where caseworkers who have higher predictive performance deviate less (12 % versus 15 %) and they assign more jobseekers to independent search.

8.4 Fairness in Assignment to ALMP: Complementary Group Analysis

This section examines fairness at the group level. The results show that caseworkers reduce fairness on average. But another important dimension of fairness is that access to ALMP should be equal across complementary groups, for example by gender or nativity, conditional on reemployment probability. As a public agency, the PES has as a responsibility to uphold this type of fairness. I examine differences in ALMP assignment across complementary subgroups, focusing on (1) gender, (2) nativity, and (3) age. Due to limited sample size, I do not analyze fairness with respect to disability status.

I measure bias by comparing the share of each subgroup assigned to an ALMP, conditional on their group-level reemployment potential. Reemployment potential is estimated using the same extrapolation method as in the main analysis. As in the main analysis, this implies that the subgroup employment potential equals the conditional employment rate under random assignment.

The share in subgroup g assigned to ALMP is simply equal to one minus the subgroup specific assignment rate, S_g. Let μ_g denote the subgroup's employment potential and $\bar{\mu}$ the average employment potential in the population. Then, the *bias* of an assignment mechanism can be defined as the difference in ALMP assignment shares across subgroups, weighted by the ratio of subgroup to average employment potential.

$$\Delta = (1 - S_{g1}) \times \frac{\mu_{g1}}{\bar{\mu}} - (1 - S_{g2}) \times \frac{\mu_{g2}}{\bar{\mu}} \tag{22}$$

This definition indicates that a mechanism is fair, if the assignment to ALMP is proportional to the difference in employment potential. By setting $\Delta = 0$ and solving for the ratio in assignment to ALMP we get that:

$$\frac{1 - S_{g1}}{1 - S_{g2}} = \frac{\mu_{g2}}{\mu_{g1}} \tag{23}$$

Such that if $\mu_{g1} = \mu_{g2} = \bar{\mu}$, the share assigned to ALMP should be equal.

Figure 9 shows the raw difference in ALMP assignment share by subgroup (solid blue bars) and the bias measure defined in Equation 22 (dashed red bars), separately for the algorithm and caseworkers. Differences are expressed in percentage points. The figure note reports the ALMP assignment shares by subgroup for both the algorithm and caseworkers.

The most striking result is the substantial bias against natives in the algorithmic assignment: immigrants are assigned to ALMPs at rates nearly 10 percentage points higher — around 140% more. After adjusting for subgroup employment potential, the bias remains large at roughly 6 percentage points. This suggests that the algorithm overestimates the reemployment gap between natives and immigrants. Caseworkers correct for this by significantly increasing ALMP assignment for natives and slightly reducing it for immigrants.

Regarding gender, caseworkers increase ALMP assignment more for men, introducing a bias against women. For age, caseworkers exacerbate the algorithmic bias against younger jobseekers.

The subgroup deviation rates are similar between male (14.3 %) and female (14.8 %), and between old (15.3 %) and young (13.9 %) jobseekers. While the difference between natives (12.9 %) and immigrant (19.3 %) is substantial.

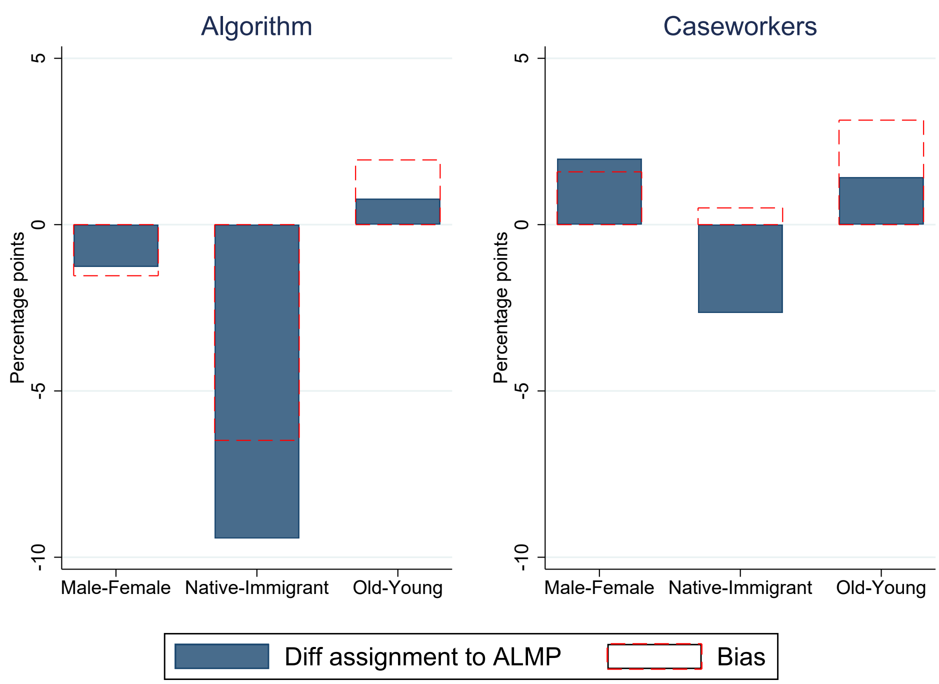


Figure 9: Caseworker versus Algorithmic Bias in Assignment to ALMP

Note: Includes balanced panel of caseworkers with at least 50 profilings per year. Assignment rates algorithm: Males 0.084; Females 0.097; Natives 0.066; Immigrants 0.161; Old 0.094; Young 0.087. Caseworkers: Males 0.143; Females 0.123; Natives 0.127; Immigrants 0.153; Old 0.140; Young 0.127.

The definition of bias in Equation 22 does not account for the *source* of bias. It captures both (i) discrimination and (ii) differences in the skill of predicting reemployment probability within each subgroup. Arnold et al. (2022), in their analysis of racial disparities in judges' pretrial decisions, construct a measure of disparate impact that adjusts for judge-level misconduct prediction accuracy — thereby isolating discrimination from skill differences.

In contrast, I adopt a broader definition of bias for two reasons. First, when comparing the algorithm and caseworkers, I am interested in the total difference in ALMP access conditional on reemployment potential — regardless of whether it stems from discrimination or predictive ability. Second, given that caseworkers appear highly noisy in their predictions, controlling for skill would imply that *any* group-level difference in ALMP assignment reflects discrimination. If caseworkers are no better than random, fairness would require equal assignment shares across subgroups.

Table 7: Caseworker Deviation Rate: Natives and Immigrants

	Full time employment			
	Natives		Immigrants	
	6 months	12 months	6 months	12 months
CW deviation rate	-0.139***	-0.231***	0.008	0.032
	(0.053)	(0.083)	(0.085)	(0.140)
Alg. assignment rate	X	X	X	X
Number of jobseekers	74,433	37,280	25,521	12,829
Number of caseworkers	528	528	528	528

Note: ***, **, * indicate significance at 1%, 5% and 10% respectively. Robust standard errors in parentheses. Includes balanced panel of caseworkers with at least 75 profilings per year. Deviation rate and algorithmic assignment rate are leave-one-out means for the whole sample.

Zooming in on the effects in the groups natives and immigrants, Table 7 shows the impact of caseworker deviation rate on employment. It is clear that while caseworker deviation has zero effect on employment of immigrants, it has a small negative effect on natives. Implying that while caseworkers are reducing fairness by increasing the access to ALMPs of natives, the cost is reduced employment.

8.5 Interpreting Results

In this subsection I will use the conceptual framework to analyze the empirical findings in terms of (a) policymaker utility, (b) the presence of a fairness-efficiency trade-off, and (c) whether assignment based on reemployment probability is utility maximizing for the policymaker.

The results show that on average:

i. Caseworker discretion is reducing fairness.

ii. Caseworker discretion is reducing total employment.

iii. Caseworkers are assigning more jobseekers to ALMPs.

iv. Caseworkers achieve an above random prediction accuracy.

v. Caseworkers mitigate algorithmic bias in assignment of ALMPs between natives and immigrants, thus improving fairness at the group level. However, caseworker deviation is simultaneously decreasing employment rates among natives.

Interpreting the results in the light of the model, I can conclude the following.

First, the reduction in fairness together with the above random prediction accuracy implies that caseworker objectives are not aligned with the algorithm, i.e. they are not simply attempting to assign jobseekers based on reemployment probability. In the light of the model, a plausible explanation is that caseworkers are valuing efficiency.

However, if caseworkers were knowingly reducing fairness in order to increase total employment outcomes, we would expect to see positive effects on total employment. Instead, we see that caseworker deviation is associated with a reduction in employment within both 6 and 12 months. Which could imply that while caseworkers are attempting to trade off fairness for efficiency, they are unsuccessful in predicting the treatment effects of ALMPs.

A limitation is that I can only follow up jobseekers 12 months post profiling, but it is possible that treatment effects take more than 12 months to be realized.

8.5.1 Policymaker Utility

Under the assumption that policymaker utility is increasing in fairness and efficiency, caseworker discretion is reducing policymaker utility.

8.5.2 Optimal Assignment Mechanism

The results of the paper highlights some important factors when designing an assignment mechanism. The treatment effect of assignment to an ALMP ultimately depends on the caseworkers skill in assigning jobseekers to the one which maximizes their employment probabilities. If such assignment decision is hard to make, such that the average treatment effect of the realized assignment is small — strictly assigning jobseekers based on reemployment probability will increase both fairness and efficiency.

Because if ALMPs have on average little effect, the total employment outcomes will be increasing in fairness. I.e. the driving factor behind the total employment will be to maximize employment among jobseekers assigned to Independent Search.

8.5.3 Alternative Objectives

Could it be that caseworkers have objectives other than fairness and efficiency? Alternative objectives could be that caseworkers are weighing jobseekers differently by characteristics, i.e. they have a preference for assigning jobseekers with children or jobseekers with other certain demographics (e.g. nativity, gender, age) to an ALMP. As a government agency, such differential access would violate basic principles, and will not be further considered here. One mechanism at play could be that caseworkers take jobseeker preferences into account, which reduces both fairness and efficiency. Another alternative is that eligibility rules are playing an important role.

8.6 Robustness Checks

See Section D.1 for robustness checks with respect to defining the outcome variable. The result that caseworkers are reducing fairness is robust to using different definitions of employment.

8.6.1 Bounding

An alternative to extrapolation for estimating the algorithmic counterfactual conditional employment rate is a bounding approach. I apply worst- and best-case bounds, using all observed outcomes of jobseekers assigned to Independent Search at each cutoff. The bounds are calculated by assuming that outcomes for unobserved jobseekers are either all negative (worst case) or all positive (best case). However, due to substantial caseworker deviations, the resulting bounds are wide and include both the extrapolated algorithmic counterfactual and the caseworker average, rendering them largely uninformative. Nevertheless, as shown in Figure D3, this exercise rules out that caseworker deviations improve fairness relative to the algorithm.

9 Conclusion

This paper extends the study of human-AI collaboration beyond well-defined prediction tasks by examining caseworker discretion over algorithmic recommendations in the assignment of unemployed jobseekers at the PES. The decision problem involves a fairness-efficiency trade-off, and I show that assigning ALMPs purely based on predicted reemployment probability is only utility-maximizing for the policymaker under strict assumptions. This creates scope for caseworkers to improve outcomes by aligning more closely with policy goals — implying that evaluating discretion solely through predictive accuracy (fairness) may be misleading.

The results show that caseworker deviations reduce fairness — assignment based on reemployment probability — as well as reducing efficiency — measured by total employment. These reductions cannot be fully explained by poor prediction accuracy but appears to reflect an unsuccessful fairness-efficiency trade-off, where caseworkers fail to predict treatment effects of ALMPs accurately. Resulting in assignment to ALMPs with on average small or zero individual treatment effects. Alternatively, the time frame of 12 months studied here is too short for treatment effects of ALMPs to materialize. And while on average caseworkers deviate in 15% of cases, these deviations have only marginal effects on jobseeker outcomes — suggesting that discretion is rather noisy.

How to design an optimal assignment mechanism for ALMPs is still an open question. I show that in theory, human decision makers have scope for improving outcomes by better aligning with policymaker objectives. And that by trading off fairness for efficiency (for marginal jobseekers) it is possible to improve total policymaker utility. However, if caseworkers are not targeting the most effective ALMPs for jobseekers, they risk reducing both fairness and efficiency.

One way of improving assignment is to learn from successful caseworkers. I find that caseworkers who are improving fairness relative to the algorithm are (i) achieving a higher prediction accuracy, and (ii) obtaining higher employment rates. They are not, however, obtaining higher employment rates among jobseekers assigned to ALMPs. They are also deviating less from the algorithmic recommendation.

These findings highlight the importance of further research to better understand when and why human discretion improves upon algorithmic recommendations — and how policymakers can foster more effective human-AI collaboration in public service delivery.

References

ANGELOVA, V., W. S. DOBBIE, AND C. YANG (2023): "Algoriothmic recommendations and human discretion," *NBER working paper*, September 2023, 1–87.

ARBETSFÖRMEDLINGEN (2020): *Så funkar AMBP 1*, Presentation, Arbetsförmedlingen.

——— (2023): *Arbetsförmedlingens nya modell för bedömning av arbetssökandes stödbehov*, Arbetsförmedlingen.

ARNI, P., G. J. VAN DEN BERG, AND R. LALIVE (2022): "Treatment versus regime effects of carrots and sticks," *Journal of Business & Economic Statistics*, 40, 111–127.

ARNOLD, D., W. DOBBIE, AND P. HULL (2022): "Measuring racial discrimination in bail decisions," Tech. rep.

BEHNCKE, S., M. FRÖLICH, AND M. LECHNER (2010): "A caseworker like me–Does the similarity between the unemployed and their caseworkers increase job placements?" *The Economic Journal*, 120, 1430–1459.

BENNMARKER, H., M. LUNDIN, T. MÖRTLUND, K. SIBBMARK, M. SÖDERSTRÖM, AND J. VIKSTRÖM (2021): *Krom-erfarenheter från en ny matchningstjänst med fristående leverantörer inom arbetsmarknadspolitiken*, IFAU-Institutet för arbetsmarknads-och utbildningspolitisk utvärdering.

CEDERLÖF, J., M. SÖDERSTRÖM, AND J. VIKSTRÖM (2021): "What makes a good caseworker?" Tech. rep., Working Paper 2021: 9, IFAU, Uppsala.

CHAN, D. C., M. GENTZKOW, AND C. YU (2022): "Selection with variation in diagnostic skill: Evidence from radiologists," *The Quarterly Journal of Economics*, 137, 729–783.

CHYN, E., B. FRANDSEN, AND E. LESLIE (2025): "Examiner and judge designs in economics: a practitioner's guide," *Journal of Economic Literature*, 63, 401–439.

COWGILL, B. (2020): "Bias and productivity in humans and algorithms: Theory and evidence from resume screening," *Columbia Business School, Columbia University*, 29.

DAWES, R. M. (1979): "The robust beauty of improper linear models in decision making." *American psychologist*, 34, 571.

DAWES, R. M., D. FAUST, AND P. E. MEEHL (1989): "Clinical versus actuarial judgment," *Science*, 243, 1668–1674.

Egebark, J., L. Laun, L. Liljeberg, M. Rödin, M. Söderström, E. Videnord, and J. Vikström (2024): "En effektutvärdering av arbetsförmedling med fristående leverantörer," Tech. rep., IFAU-Institute for Evaluation of Labour Market and Education Policy.

Ernst, S., A. I. Mueller, and J. Spinnewijn (2024): "Risk score for long-term unemployment and the assignment to job search counceling," *Working paper*.

Grove, W. M., D. H. Zald, B. S. Lebow, B. E. Snitz, and C. Nelson (2000): "Clinical versus mechanical prediction: a meta-analysis." *Psychological assessment*, 12, 19.

Hoffman, M., L. B. Kahn, and D. Li (2018): "Discretion in hiring," *The Quarterly Journal of Economics*, 133, 765–800.

Hull, P. (2020): *Estimating hospital quality with quasi-experimental data*, Unpublished Working Paper.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2018): "Human decisions and machine predictions," *The quarterly journal of economics*, 133, 237–293.

Mullainathan, S. and Z. Obermeyer (2022): "Diagnosing physician error: A machine learning approach to low-value health care," *The Quarterly Journal of Economics*, 137, 679–727.

Ohlin, L. E. and O. D. Duncan (1949): "The efficiency of prediction in criminology," *American Journal of Sociology*, 54, 441–452.

Rambachan, A. (2024): "Identifying prediction mistakes in observational data," *The Quarterly Journal of Economics*, qjae013.

Regeringen (2025): *Regleringsbrev för budgetåret 2025 avseende Arbetsförmedlingen*, Regeringen, Arbetsmarknadsdepartementet.

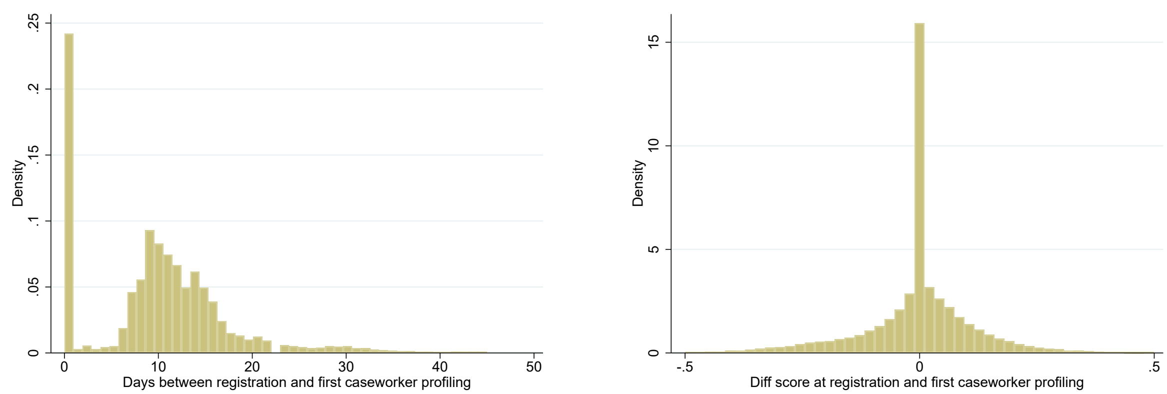
van den Berg, G. J., M. Kunaschk, J. Lang, G. Stephan, and A. Uhlendorff (2023): "Predicting re-employment:machine learning versus assessments by unemployed workers and by their caseworkers," *IFAU working paper*, 1–58.

Walters, C. (2024): "Empirical Bayes methods in labor economics," in *Handbook of Labor Economics*, Elsevier, vol. 5, 183–260.

A PES Institutions and the Profiling Tool

A.1 Timing of First Profiling and Caseworker Updating of the Profiling Score

In the central-office population, 72% of jobseekers receive their first profiling at initial registration (i.e., same day as they sign up online), which occurs between day 0 and 9 of the unemployment spell (median = 0). Among jobseekers whose first profiling occurs during the caseworker meeting, the spell length at profiling ranges from 0 to 287 days on (median = 15). During the meeting, the profiling score is updated only if new information emerges; otherwise, the registration score stands. The median time between registration profiling and the first caseworker profiling is 10 days, and 5% of these first-meeting profiling events lead to a change in the algorithmic recommendation.



(a) Number of days between

(b) Difference in profiling score

Figure A1: Profiling at registration and first caseworker profiling.

Note: Sample of jobseekers with their first profiling at registration (72% of total sample). The plot excludes the 5 % most extreme values. The number of days ranges from 0 to 375 days. And the difference in score ranges from -0.76 to 0.80.

A.2 Eligibility Criteria & Standardized Questions

Eligibility for active labor market programs is determined primarily by the jobseeker's predicted reemployment probability. In addition, a jobseeker is ineligible if any of the following conditions apply.

Eligibility criteria.

1. Already part-time employed or have secured employment starting within 90 days or enrolled in a full time studies starting within 90 days.

2. Is going on parental leave within 90 days.
3. Is moving abroad.
4. Is a citizen of another EU/EES country, an asylum seeker, or a Ukrainian refugee in Sweden under the EU Temporary Protection Directive.
5. Receiving assistance from a former employer, a transition organization[9], another government agency, or the municipality.
6. Registered on the board of a company.

Standardized questions. The following standardized questions are asked when caseworkers conduct the labor market assessment. The questions are answered either during the online registration or at the first meeting with the caseworker.

1. What did you do before becoming unemployed?
2. Do you know what type of job you would prefer?

2a. Do you know what you would like to study? (If no to question 2)

3. How would you rate your Swedish language skills?
4. Do you know how many jobs you plan to apply for this month?
5. Do you have any obstacles to relocating or commuting long distances?
6. Do you think you will get a job or start studying within three months?
7. What were your average monthly labor earnings during the past year?
8. Do you face any obstacles that prevent you from applying for jobs or getting hired?

9a. Do you have any health issues that may affect your chances of getting a job?

9b. Do you have a disability that may affect your chances of getting a job?

[9]Transition organizations — owned by employer associations and labor unions and regulated through collective bargaining agreements — offer services focused on educational and career counseling and also provide financial support.

A.3 Reform Context - Introduction of the Algorithm

The Swedish PES has two main goals related to job search stated by the government. Firstly, they shall prioritize jobseekers who already are, or are at high risk of becoming, long term unemployed. Secondly, they shall improve matching on the labor market, by increasing the number of participants in ALMP and education (Regeringen, 2025).

A.3.1 PES Mandate Shift

In 2019, a political agreement fundamentally altered the Public Employment Service's (PES) mandate: PES would focus on labor market assessments, monitoring private providers, producing statistics and analyses, and maintaining digital infrastructure, while job search assistance would be delivered primarily by private providers operating in a competitive market.[10] As a result, PES caseworkers' main task became conducting labor market assessments and determining eligibility for job search assistance and other active labor market programs. These organizational reforms coincided with substantial budget cuts, leading to a reduction in the number of local offices, further centralization of service delivery, and increased digitization and automation.

A.3.2 Rising Long-Term Unemployment & Policy Priorities

The reform of the PES also coincided with a shift in policy makers' priorities. Beginning in 2019, long-term unemployment in Sweden began to rise and reached record-high levels during the COVID-19 pandemic. In response, policymakers wanted to direct PES resources toward jobseekers already in, or at high risk of, long-term unemployment. Government steering documents shifted their focus from broad targets — such as overall ALMP volumes and general employment growth — to a specific priority: reducing long-term unemployment.

A.3.3 Matching Services Pilot

To align service delivery with this new priority, PES introduced three major changes:

1. Explicit linking of ALMP eligibility to predicted reemployment probability.
2. Implementation of an algorithmic profiling tool to ensure assignment to ALMPs was based on each jobseeker's predicted reemployment probability.

[10] The January Agreement: `https://www.socialdemokraterna.se/download/18.1f5c787116e356cdd25a4c/1573213453963/Januariavtalet.pdf`

3. Differentiation of provider compensation according to jobseekers' reemployment probabilities, thereby incentivizing private providers to prioritize higher-need cases rather than easier placements.

Reports indicate that, following these changes, the share of long-term unemployed participants in ALMPs increased significantly (Bennmarker et al., 2021).

Concurrently, PES piloted a new privately provided job search assistance program — Matching Services — from March 2020 to December 2021. To evaluate Matching Services' impact, the algorithmic recommendation was randomized for jobseekers whose profiling scores lay near the category cutoffs; see Egebark et al. (2024) for a detailed description of that evaluation design. In 2022, Matching Services became the largest ALMP at PES, enrolling over 70 000 participants (Egebark et al., 2024).

A.4 Algorithm Input Variables

Table A1: Input variables algorithm

Date of registration PES
Country of birth
Disability
Number of PES programs or interventions 10 years
Number of days in PES programs or interventions 10 years
Number of days registered at the PES
Current or the number of days registered at the latest registration three years
Number of days since the last registration three
Age
Municipality
Zip code
Search category
Target occupation
Education in target occupation
Experience in target occupation
Searching for work to the extent of (full time, part time)
Level of education
Field of education
Gender
Applying for jobs abroad
Applying for unemployment insurance
Share of unemployed within the zip code
Median income zip code
Share with secondary education within zip code
Share of households types within zip code
Share with non-European citizenship within zip code.

A.5 Two Regimes

The algorithm described so far was in place up until Marsh 2023. In order to improve the predictive performance and incorporate unemployment duration in a systematic way, the PES developed a new algorithm which was introduced in April 2023. When changing the algorithm the PES also changed the name of the matching services program, from KROM to ROM2. I will use the program names to refer to the two algorithmic regimes.

Regime 1: KROM (March 2020 – April 2023)

- Trained on the population of newly registered jobseekers only.
- Used time since the start of the unemployment spell solely to determine cutoff thresholds.
- Output: assignment recommendation (Independent Search, Job Search Assistance including sub-tracks 2a–2c, or Intensive Support).
- Caseworker discretion was limited — any deviation from Independent Search or Intensive Support to Job Search Assistance required approval from a committee.

Regime 2: ROM2 (April 2023 – present)

- Trained on the entire stock of registered jobseekers, incorporating spell length as an explicit input variable.
- Output: assignment recommendation plus a ranked list of the most important features for the prediction.
- Caseworker discretion is unrestricted — caseworkers may override the algorithmic recommendation in any category and direction without external approval.

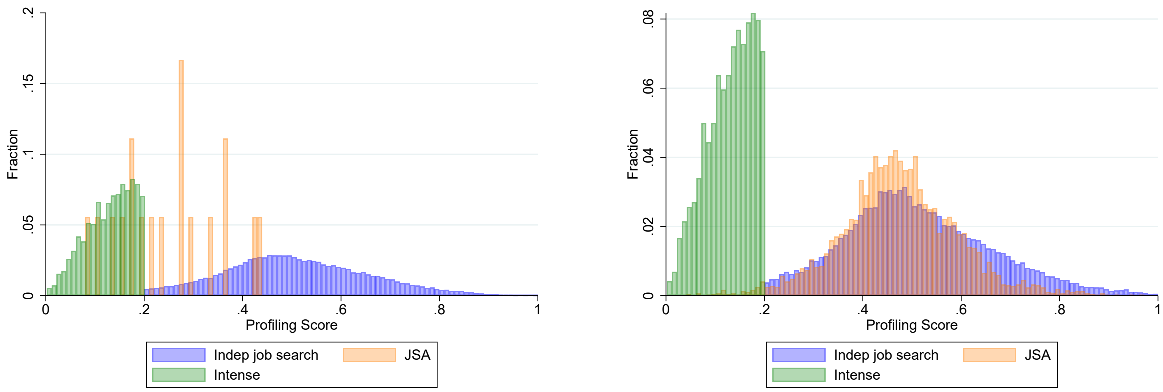
According to PES, ROM2-algorithm's accuracy (share of correctly classified outcomes) is 74.9 %, compared to KROM-algorithm's 68 % — both measured against a baseline random-classification accuracy of 57 %. These performance metrics do not adjust for PES interventions. Under the assumption that the PES assistance and interventions improve the labor market outcomes of jobseekers, the PES argues that these measures should be viewed as conservative. (Arbetsförmedlingen, 2023)

A.6 Profiling score and algorithmic recommendation

The algorithmic recommendation had three main categories; (1) Independent Search, (2) Job Search Assistance, and (3) Intense Support. Within the Job Search Assistance recommendation category, the algorithm further subdivides jobseekers into three sub-tracks—2a (highest reemployment probability), 2b (intermediate), and 2c (lowest)—based on finer-grained cut-off bands. All sub-tracks offer the same services in principle; however, private providers receive higher compensation for sub-tracks associated with lower predicted reemployment probabilities to encourage needs-based service delivery rather than prioritizing easier cases.

Jobseekers close to the cutoffs between categories were randomized to either group, to facilitate an evaluation of the Job Search Assistance program.

Figure A2 illustrates how the time varying cutoffs affect the distribution of underlying profiling scores across recommendation categories. Jobseekers with scores below 0.2 are recommended Intense Support, whereas those with higher scores may be recommended either Job Search Assistance or Independent Search. Note that the score distributions for Job Search Assistance and Independent Search overlap, reflecting that some jobseekers with similar scores receive different recommendations depending on their unemployment duration. During Jan 1 - Sep 7, jobseekers in the sample were only recommended Job Search Assistance if randomized across a threshold.



(a) January 1 - September 7, 2022

(b) September 8 - December 31, 2022

Figure A2: Profiling score and algorithmic recommendation

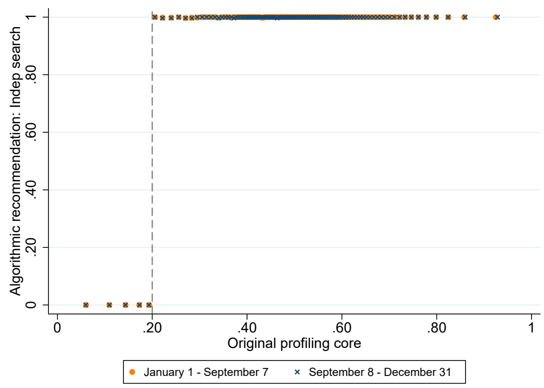
Note: Distribution of profiling scores across the three main categories of the algorithmic recommendation. The fraction within each category and period sums to 1. Includes randomized cases. Number of observations 1 Jan - 7 Sep: Independent Search 66,661; JSA 20; Intense Support 3,023. Sep 8 - Dec 31: Independent search 34,674; JSA 4,953; Intense Support 1,511.

A.7 Adjusted Profiling Score

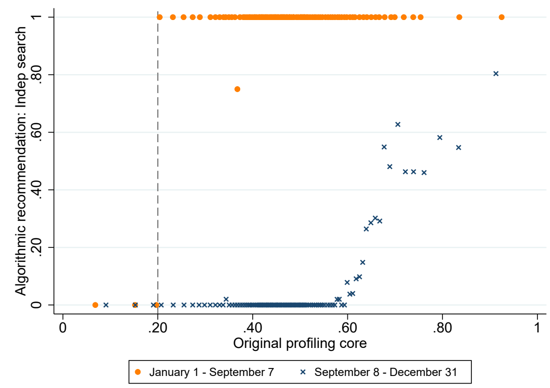
Figure A3 illustrates the relationship between the original profiling score and the algorithmic recommendation. The algorithmic recommendation is based on (i) the profiling score and (ii) cutoffs depending on the length of the spell. Jobseekers with a score less than 0.2 were always recommended Intense Support.

Panel (a) illustrates that for jobseekers with less than 180 days of unemployment, there is practically a sharp cutoff of 0.2 between any ALMP and independent search. This applies to a large majority of jobseekers. For jobseeker with unemployment spells longer than 180 days, the sharp cutoff of 0.2 applies until September 7. From September 8th and onwards the cutoff increases with the length of the spell. Panel (c) verifies that the sharp cutoff of 0.2

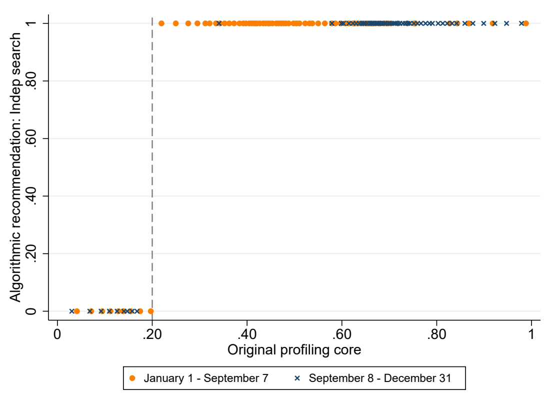
always applies to intense support. Jobseekers with score above 0.2 get recommended either job search assistance or independent search, illustrated in panel (d).



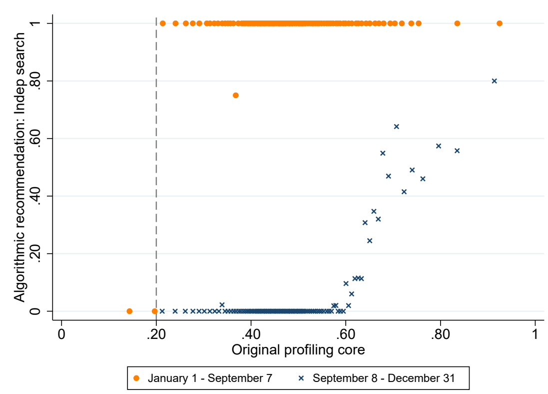
(a) 0-179 days unemployed



(b) $\geq$ 180 days unemployed



(c) $\geq$ 180 days unemployed
Intense support vs independent search



(d) $\geq$ 180 days unemployed
Job search assistance vs independent search

Figure A3: Profiling score and probability of being recommended independent search

Note: Binscatter of the algorithmic recommendation to independent search against the original profiling score. Sample split into two periods, panel (a) showing jobseekers with 0-179 days of unemployment (N=99,056) and panel (b) showing jobseekers with at least 180 days of unemployment (N=5,786). Panel (c) shows jobseekers wiht at least 180 days of unemployment and the recommendation between intense support and independent search. Panel (d) shows jobseekers with at least 180 days of unemployment an the recommendation between job search assistance and independent search. In the data the maximum number of days of unemployment is 376 days. The number of jobseekers with more than 180 days of unemployment from January 1 to September 7 is 418. The number of jobseekers with more than 180 days of unemployment from September 8 to December 31 is 5,368.

Figure A4 shows the distribution of number of days unemployed at the profiling date.

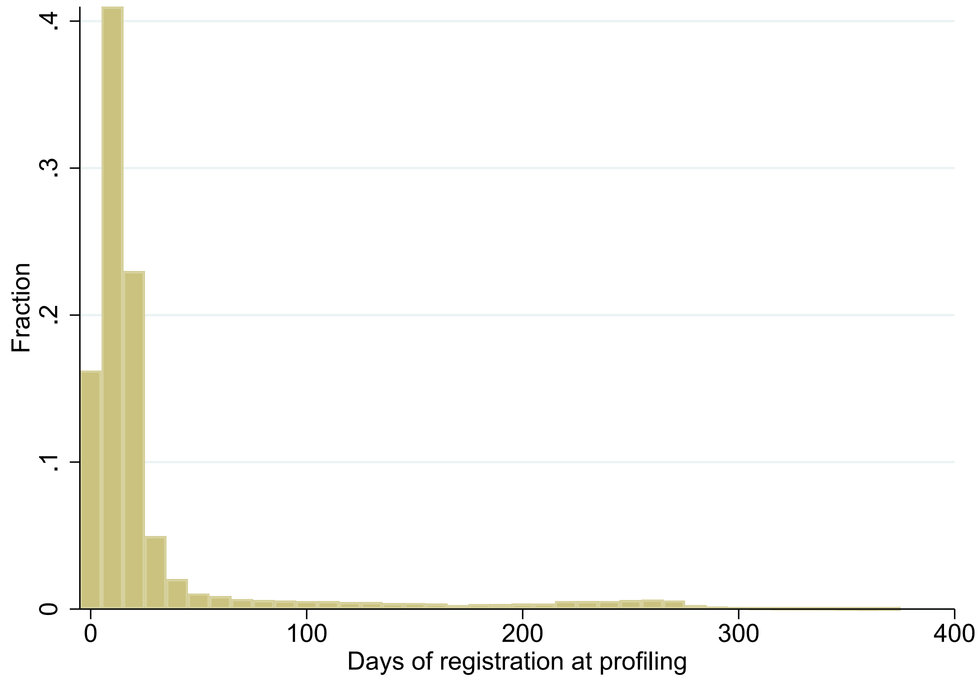
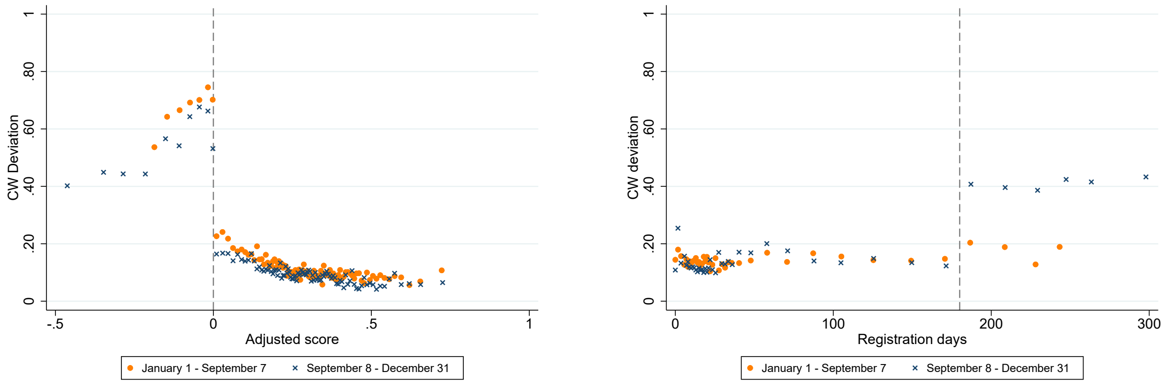


Figure A4: Distribution number of days unemployed

Note: Length of current unemployment spell at the profiling date. The median length at profiling is 13 days. 10 % of profilings are done at day 80 or later and 5 % of profilings are done at day 196 or later.

A.8 Caseworker Deviations

Figure A5 shows the probability of caseworker deviations at different levels of the adjusted score and days in unemployment. Panel (a) shows that the probability of deviation jumps to around 70 % just left of the cutoff. The results are similar if we would instead look at the probability of deviation for scores below 0.2 (at the cutoff to intense support). Panel (b) plots the probability of deviations against number of days in unemployment. The sharp increase at 180 days for the period 7 September - 31 December is driven by deviations from Job Search Assistance to Independent Job Search. The fact that deviations are concentrated right around the cutoff explains the smooth distribution over the cutoff in Figure 7 panel (b).



(a) Deviation and adjusted score

(b) Deviation and days registered

Figure A5: Caseworker Deviations

Note: Caseworker deviations against the adjusted profiling score and days of registration. See Figure A6 for a plot of caseworker deviations and the original profiling score.

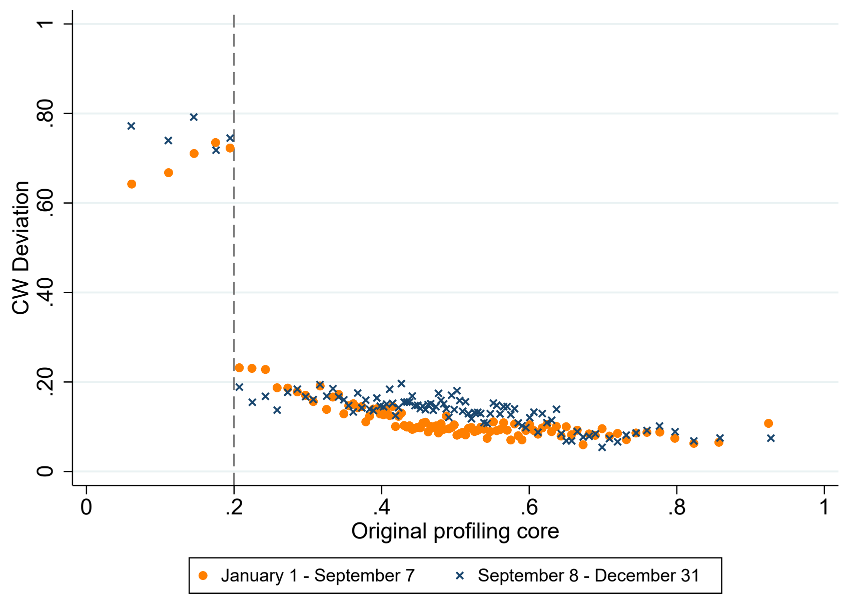


Figure A6: Caseworker deviation and original profiling score

Note:

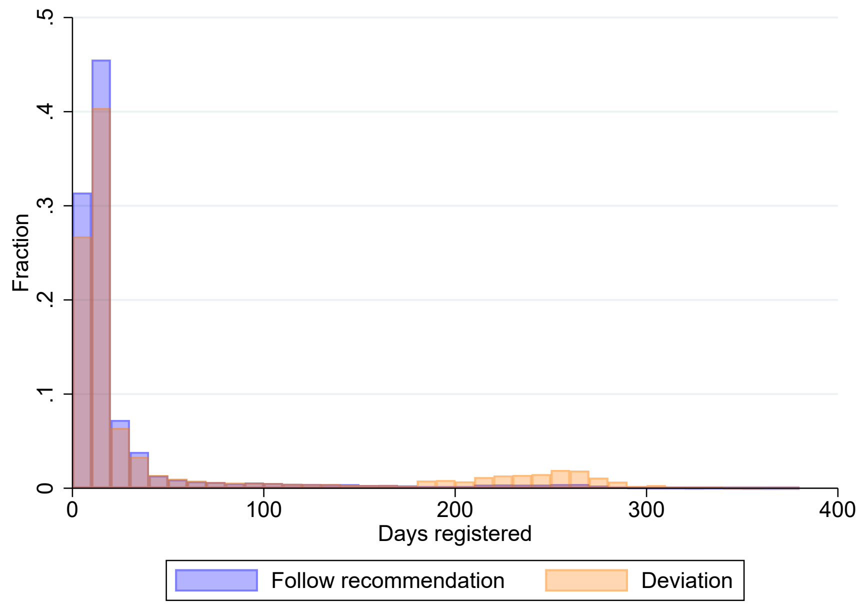
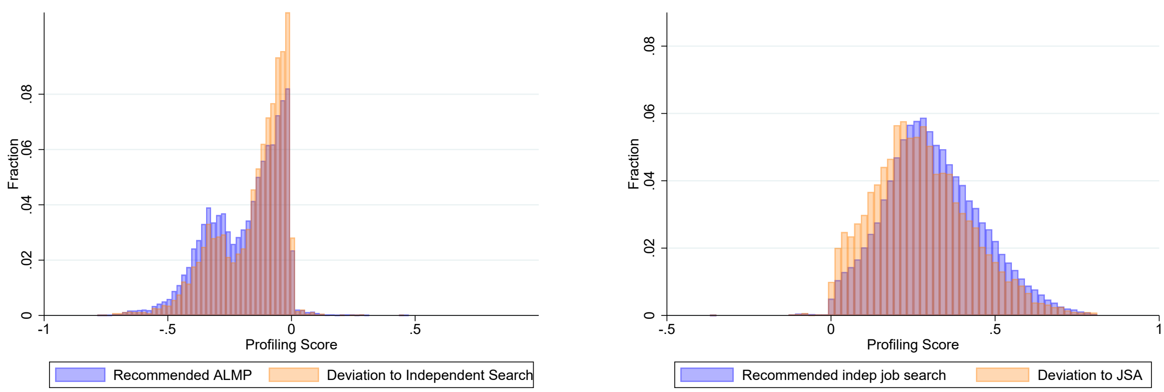


Figure A7: Distribution of Days Registered by Deviation

Note:



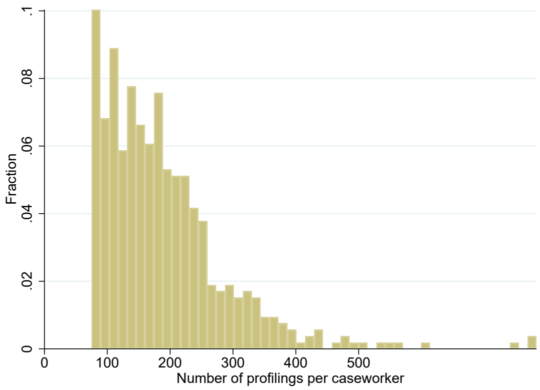
(a) Deviation to Independent Search

(b) Deviation to ALMP

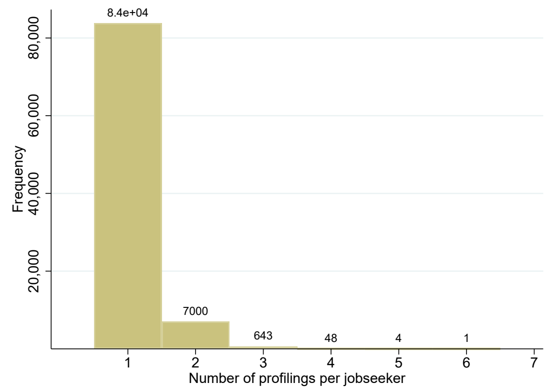
Figure A8: Distribution of Scores by Deviation

Note: In each panel, deviations are a subset of cases in each cateogry - i.e. recommendation to Independent Search, panel (a), and recommendation to ALMP, panel (b). Score is the new score.

A.9 Case Load - Final Data



Number of decision events per caseworker



Number of decision events per jobseeker

Figure A9: Distribution of Decision Events Across Caseworkers and Jobseekers

Note: Number of decision events in the final sample.

B Linear Empirical Bayes Shrinkage

I use linear empirical Bayes shrinkage, following the three steps lined out by Walters (2024).

Step 1. Estimation

Estimate the caseworker conditional employment rate, M_j.

$$Y_i = M_j + \varepsilon_i \tag{24}$$

Treating that $\hat{M}_j$ is normally distributed with variance s_j^2

Step 2. Deconvolution

Estimate the two parameters μ_M and σ_M.

$$\hat{\mu}_M = \frac{1}{J}\sum_{j=1}^{J} \hat{M}_j \tag{25}$$

$$\hat{\sigma}_M^2 = \frac{1}{J}\sum_{j=1}^{J} \left[(\hat{M}_j - \hat{\mu}_M)^2 - s_j^2\right] \tag{26}$$

Where subtracting s_j^2 corrects for excess noise in the $th\hat{e}ta_j$ estimates.

Step 3. Shrinkage

To form the posterior mean for each caseworker estimate $\hat{M}_j$, we take a weighted average of the noisy estimates $\hat{M}_j$ and the prior mean μ_M. How much the caseworker estimates are shrunk towards the prior mean depends on the precision of the estimates relative to the precision of the prior mean.

$$M_j^* = \hat{M}_j \left(\frac{\sigma_M^2}{\sigma_M^2 + s_j^2}\right) + \hat{\mu}_M \left(\frac{s_j^2}{\sigma_M^2 + s_j^2}\right) \tag{27}$$

C Empirical Strategy

C.1 Identifying assumptions

C.1.1 Relevance

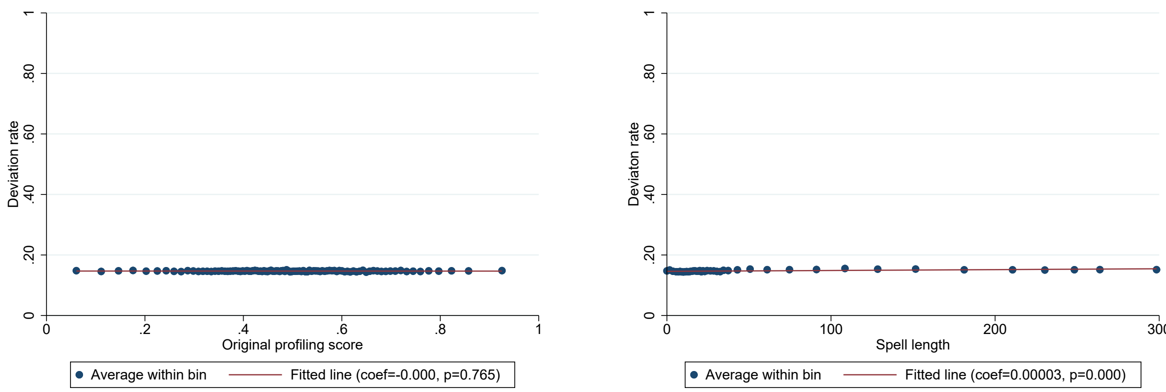
Table C1 shows the first stage regression of assignment to independent search on leave-out assignment rate. It tests the *relevance assumptions*. The positive and significant coefficients indicates that the relevance assumption holds.

Table C1: First stage

	Independent jobsearch	Deviation
CW assignment rate	0.461***	
	(0.032)	
CW deviation rate		0.424***
		(0.033)
N	99,959	99,959
Mean assignment rate	0.87	0.15
Number of caseworkers	528	528

Note: Robust standard errors reported in parentheses. ***, **, * indicate significance at 1%, 5% and 10% respectively. Dependent variable is the leave-out caseworker assignment rate. Includes balanced panel of caseworkers with at least 75 profilings per year.

C.1.2 Random Assignment



(a) Profiling score

(b) Original profiling score

Figure C1: Plot assignment rate against profiling score

Note: Binscatter and fitted line of leave out deviation rate against (a) profiling score and (b) length of unemployment spell. Coefficients reported from regressions on the jobseeker level. Robust standard errors used in panel (a) and bootstrapped standard errors in panel (b). Includes balanced panel of caseworkers with at least 75 profilings per year.

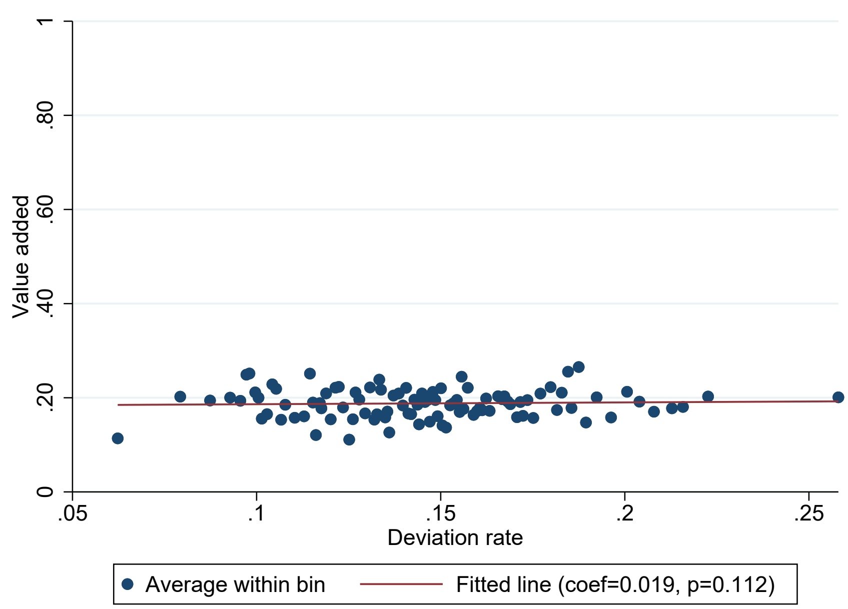


Figure C2: Deviation Rate and Caseworker Value Added

Note: Includes caseworkers with at least 75 profilings per year.

C.2 Illustration of the empirical strategy

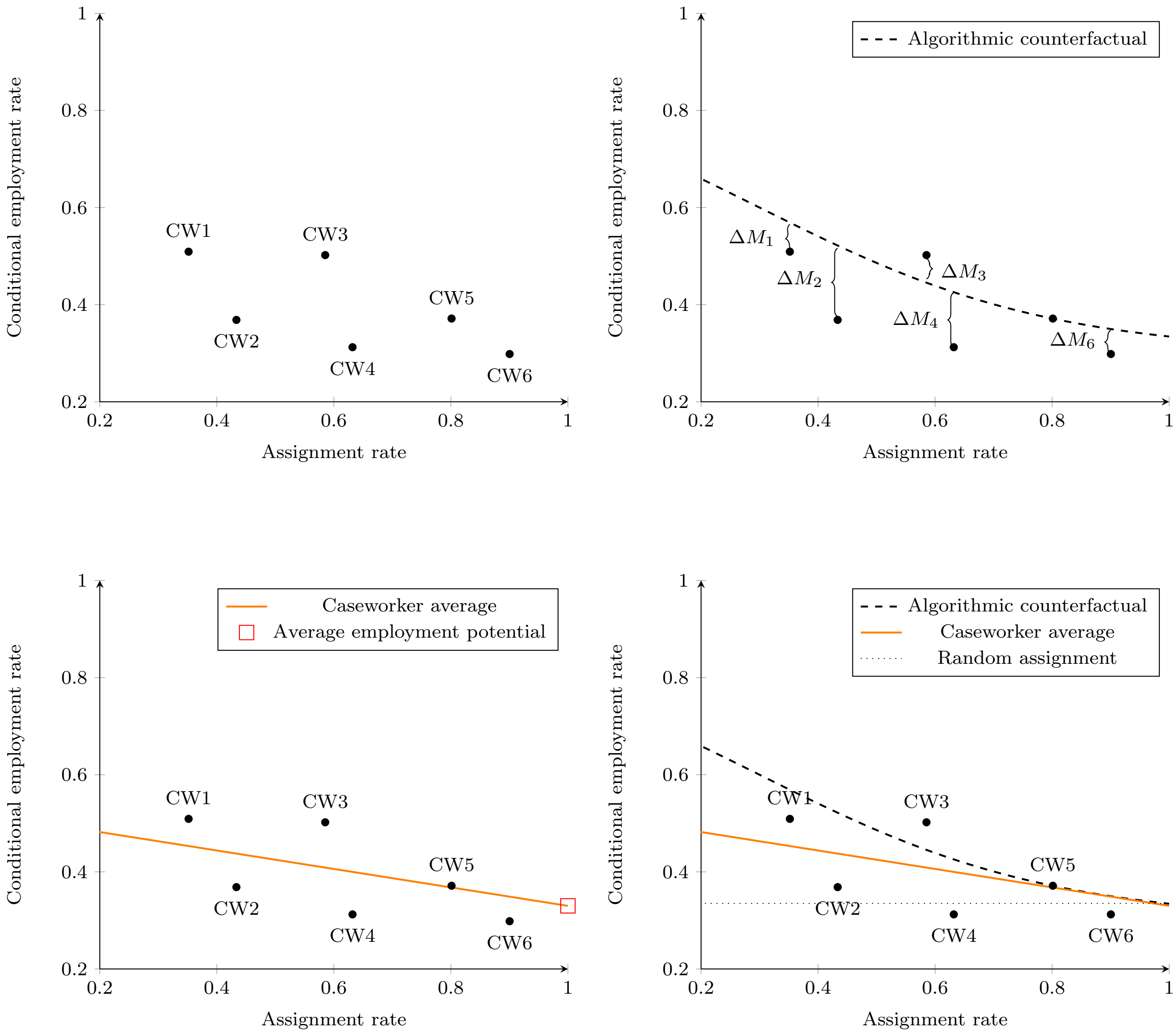
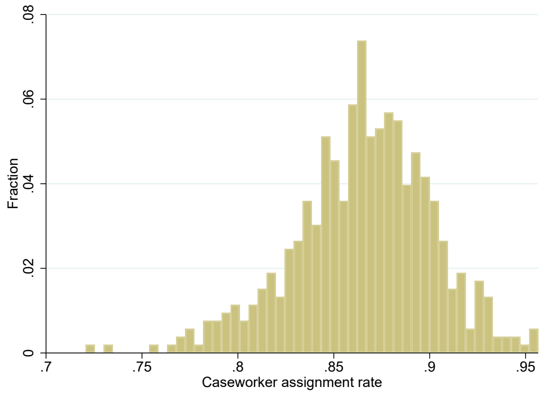
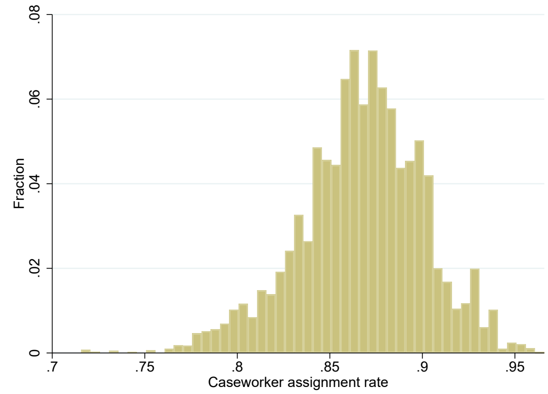


Figure C3: Empirical Strategy step by step

Note: Conditional employment rate is defined as $M_j = E[Y_i = 1|D_{ij} = 1$. Assignment rate is defined as $S_j = E[D_{ij} = 1]$.

C.3 Assignment rate





Caseworker level

Case level

Figure C4: Caseworker assignment rate

Note: Includes caseworkers with at least 75 profilings per year.

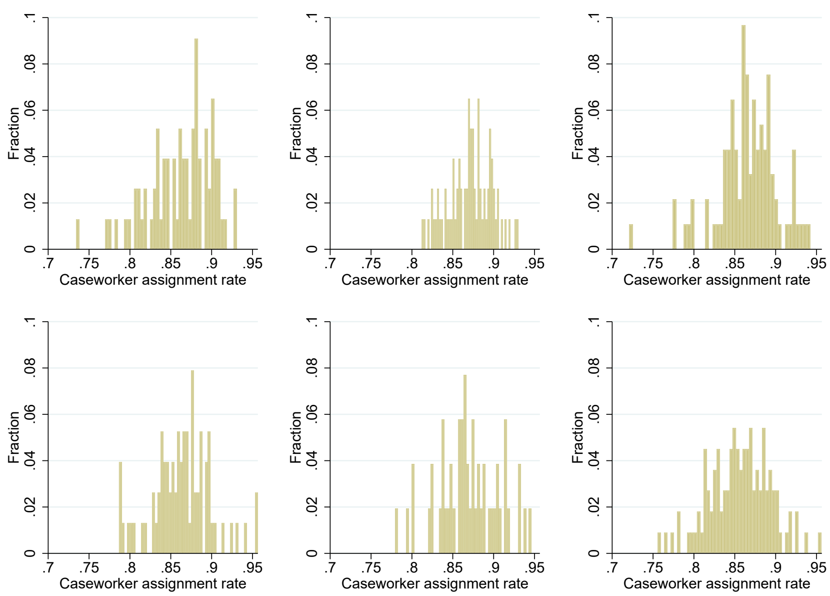
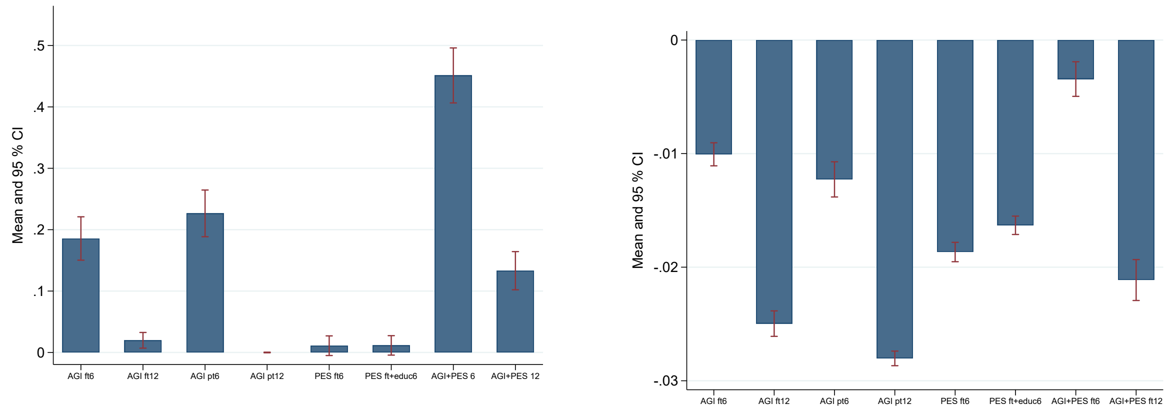


Figure C5: Caseworker assignment rate by office

Note: Includes caseworkers with at least 75 profilings per year..

D Robustness Checks

D.1 Robustness: Outcome Defintion



(a) Share of CWs higher Employment Rate than the Algorithm

(b) Diff Employment Rate

Figure D1: Caseworkers versus Algorithmic Counterfactual

Note:

D.2 Robustness: Extrapolation

D.2.1 Functional Form

Table D1: Extrapolation - Function Form

	(1)	(2)	(3)	(4)	(5)	(6)
	Caseworkers vs Algorithm			Caseworkers vs Random		
	Linear	Local Linear	Quadratic	Linear	Local Linear	Quadratic
Share CW higher emprate	0.092***	0.000	0.000	0.488***	0.060***	0.000
	(0.014)	(0.000	(0.000)	(0.022)	(0.012	(0.000)
Average difference emprate	-0.013***	-0.034***	-0.054***	-0.000	-0.014***	-0.048***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
N	528	528	528	528	528	528

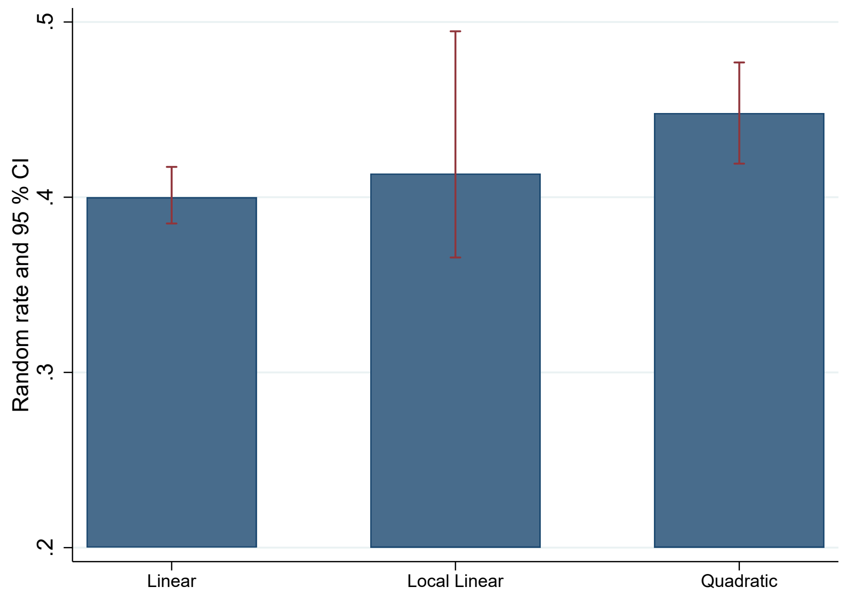


Figure D2: Extrapolation Functional Form - Random Rate

D.2.2 Bounding

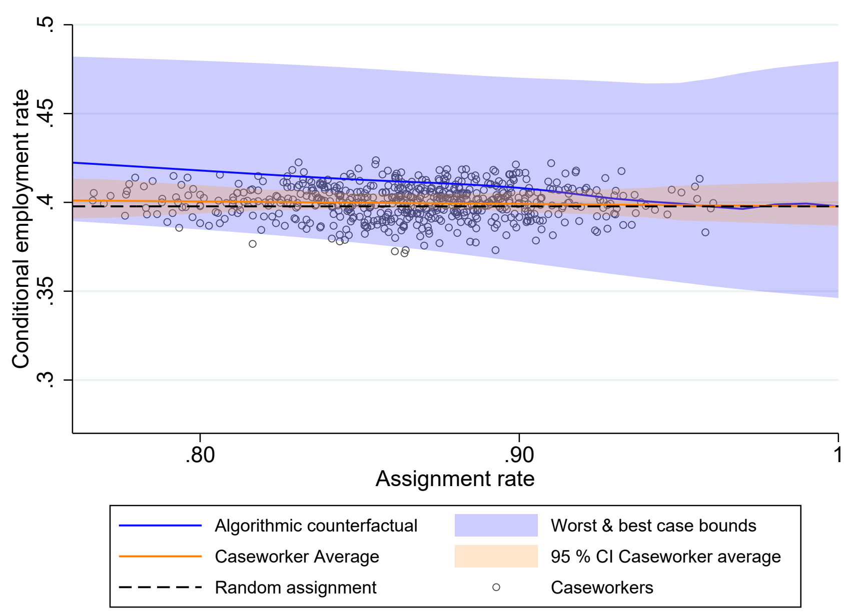


Figure D3: Worst- and Best case Bounding

E Results

E.1 Fairness - Subgroup

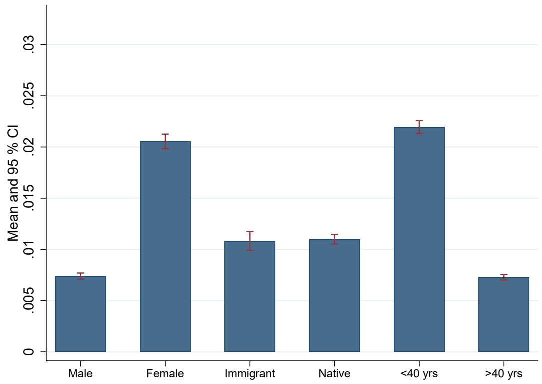
This section examines how caseworkers and the algorithm perform across different subgroups of jobseekers. The algorithmic counterfactual is re-estimated within each subgroup, allowing comparisons between the algorithm and random assignment, and between caseworkers and the subgroup-specific algorithmic counterfactual.

Figure E1 Panel (a) shows the difference in conditional employment rate between the algorithm and a random assignment. The difference is calculated as a case-weighted average of the conditional employment rate gains at each caseworker's assignment rate. It answers the question: *If each caseworker were replaced by the algorithm, holding their assignment rate constant, how much would the conditional employment rate improve relative to random assignment?* The algorithm's predictive performance varies notably across subgroups. It performs better for females than males, and for younger rather than older jobseekers.

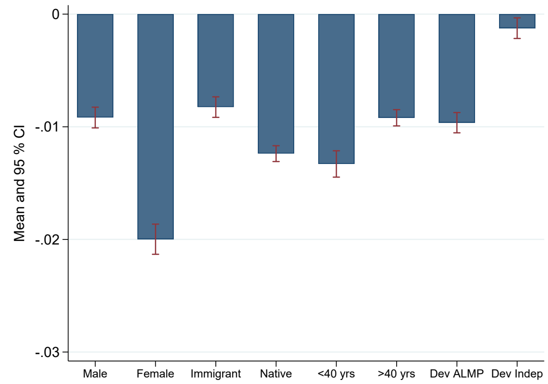
Figure E1 panel (b) compares caseworker and algorithm performance across subgroups. It displays the case-weighted average difference in conditional employment rate at each caseworker's assignment rate.

Regarding the demographic groups, caseworkers are decreasing reducing the

It is also worth noting that caseworker prediction accuracy is significantly higher in deviations to Independent Search than in deviations to ALMP.



(a) Algorithm versus Random Assignment



(b) Caseworkers versus Algorithm

Figure E1: Subgroup Predictive Performance - Conditional Employment Rate

Note: 95% confidence intervals computed using bootstrapping. The algorithmic counterfactual and random rate is estimated within each subgroup. Panel (a) show the difference in conditional employment rate between the algorithm and the random rate. The difference is calculated as a case-weighted average of the conditional employment rate gains at each caseworker's assignment rate. It answers the question: *If each caseworker were replaced by the algorithm, holding their assignment rate constant, how much would the conditional employment rate improve relative to random assignment?* Panel (b) displays the case-weighted average difference in conditional employment rate between caseworkers and algorithm at each caseworker's assignment rate.

Table E1

	Share Above Algorithm		Diff Cond. Employment Rate	
	Above	Below	Above	Below
Male	0.587	0.102	0.002	-0.010
Female	0.478	0.033	0.001	-0.022
Immigrant	0.390	0.226	-0.004	-0.009
Native	0.478	0.015	-0.000	-0.014
Young	0.587	0.104	0.003	-0.015
Old	0.391	0.050	-0.001	-0.010
N	46	481	46	481

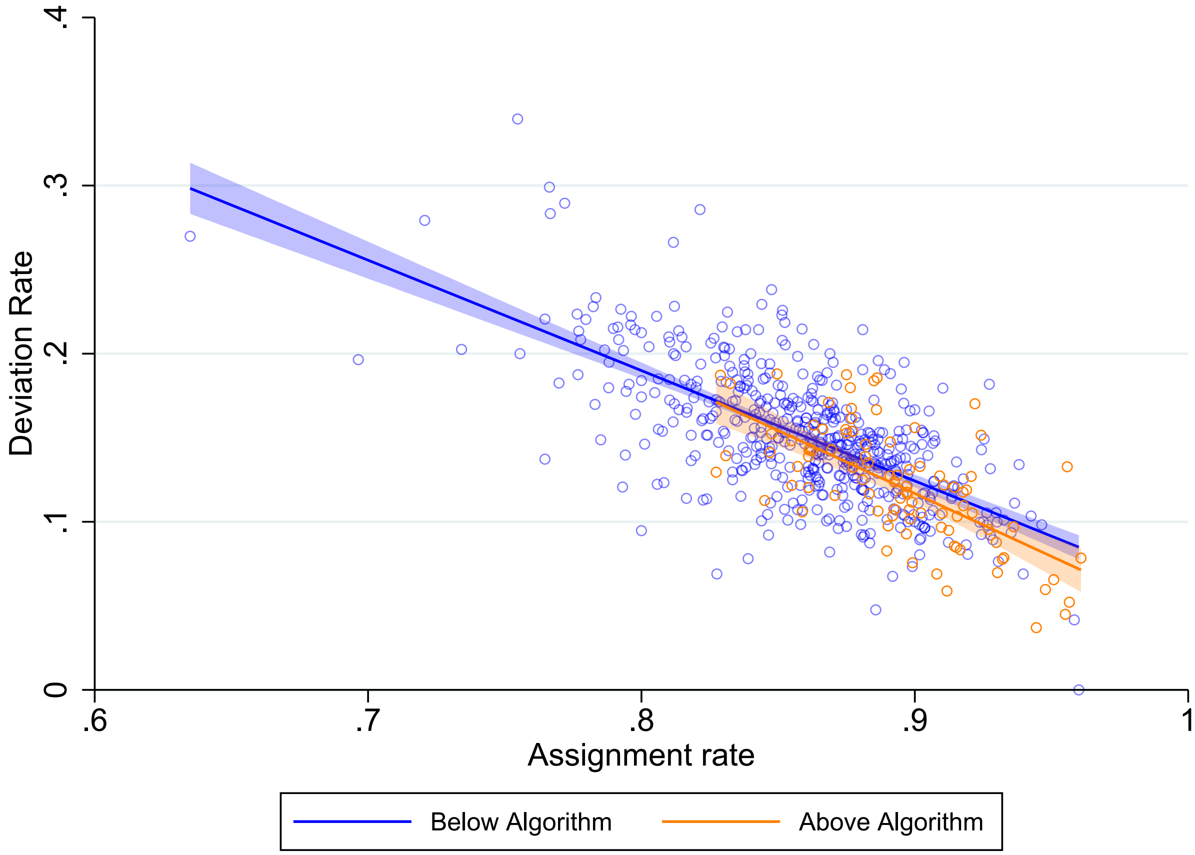


Figure E2: Deviation Rate and Assignment Rate

Note:

E.2 Low and High Performing Caseworkers

Table E2: Caseworker Characteristics

	Below algorithm	Above algorithm	P-value diff
Caseworker characteristics			
Experience years	8.162	8.279	[0.923]
	(7.480)	(7.805)	
Monthly wage (SEK)	30,301	30,201	[0.789]
	(2,353)	(2,016)	
Assignment rate	0.862	0.899	[0.000]
	(0.035)	(0.028)	
Number of cases	188	200	[0.463]
	(98)	(116)	
Deviation rate	0.150	0.117	[0.000]
	(0.036)	(0.035)	
Deviation rate indep search	0.052	0.047	[0.141]
	(0.022)	(0.025)	
Deviation rate treatment	0.098	0.070	[0.000]
	(0.029)	(0.021)	
Deviation rate ranking	0.178	0.125	[0.000]
	(0.051)	(0.048)	
$\Delta Assignmentrate$	0.050	0.025	[0.000]
	(0.041)	(0.032)	
Number of jobseekers	90,770	9,184	
Number of caseworkers	482	46	

Note: Robust standard errors reported in parentheses. ***, **, * indicate significance at 1%, 5% and 10%. Observation are missing for some variables. Caseworker experience and wage comes from the personel registries, in which there is data on 394 low performing and 43 high perfoming caseworkers. There is data 12 months post profiling for 50,109 jobseeker, across low performing and high performing caseworkers. There are deviations due to high reemployment probability in 356 cases, across low performing caseworkers and high performing caseworkers.

Table E3: Jobseeker Characteristics

	Below algorithm	Above algorithm	P-value diff
Male	0.497	0.495	[0.782]
	(0.500)	(0.500)	
Historical functional disability	0.051	0.050	[0.637]
	(0.220)	(0.217)	
Immigrant	0.256	0.249	[0.172]
	(0.436)	(0.433)	
High school	0.024	0.023	[0.385]
	(0.154)	(0.150)	
Tertiary education	0.168	0.179	[0.008]
	(0.374)	(0.383)	
Score	0.663	0.668	[0.000]
	(0.130)	(0.124)	
Service and admin	0.124	0.119	[0.219]
	(0.329)	(0.324)	
Service, care and shop	0.280	0.276	[0.347]
	(0.449)	(0.447)	
Building and manufacturing	0.083	0.085	[0.658]
	(0.276)	(0.278)	
Age 25-34	0.379	0.371	[0.115]
	(0.485)	(0.483)	
Age 35-44	0.257	0.256	[0.973]
	(0.437)	(0.437)	
Age 45-54	0.195	0.200	[0.316]
	(0.397)	(0.400)	
Previously registered	0.051	0.050	[0.677]
	(0.220)	(0.217)	
Days of registration	31.548	28.085	[0.000]
	(58.481)	(52.737)	
Number of jobseekers	90,777	9,182	
Number of caseworkers	482	46	

E.3 Efficiency

Caseworker deviations can be composed into two sources. Caseworkers might want to change the assignment rate, but agree on the ranking of jobseekers. Or they might also want to change the ranking of jobseekers. Table E4 show the results of regressing employment within 6 and 12 months on the "deviation rate ranking", which compares caseworker assignment to the algorithmic counterfactual assignment at the caseworker's deviation rate and measures to what extent they disagree. I also regress employment outcomes on the percent difference in assignment rate relative to caseworker's algorithmic counterfactual. The results imply that the ranking difference is driving results on employment rates.

Table E4: Deviation Rate Decomposition

	Full time employment					
	6 months	6 months	6 months	12 months	12 months	12 months
Deviation Rate Ranking	-0.100***		-0.105***	-0.120***		-0.127***
	(0.031)		(0.034)	(0.045)		(0.049)
Δ Assignment rate		-0.039	0.016		-0.047	0.022
		(0.041)	(0.045)		(0.059)	(0.065
Number of caseworkers	99,954	99,954	99,954	50,109	50,109	50,109
Number of jobseekers	528	528	528	528	528	528

Note: ***, **, * indicate significance at 1%, 5% and 10% respectively. Robust standard errors in parentheses. Includes balanced panel of caseworkers with at least 75 profilings per year. Mean Deviation Rate Ranking is 0.173 and mean Δ asssignment rate is 0.048 in the full sample and both measures are leave-one-out means.

F Robustness: First Profiling

This section runs the analysis using only the first profiling per jobseeker.

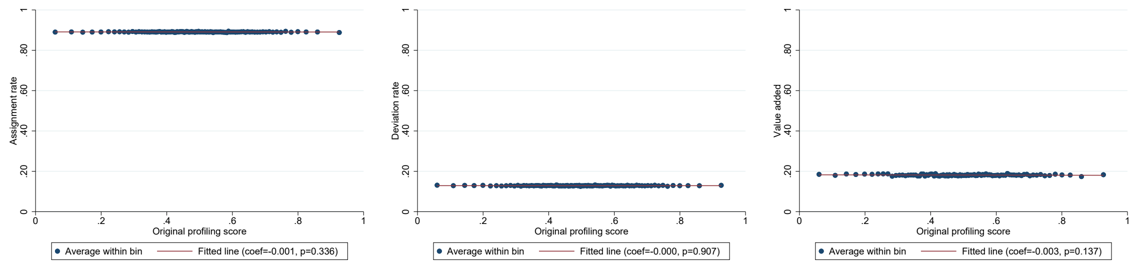
F.1 Identifying Assumptions

F.1.1 Balance

Table F1: First Profiling: Balance

	Assignment Rate	Deviation Rate	Profiling Score
Male	-0.000	-0.000	0.028***
	(0.000)	(0.000)	(0.001)
Historical functional disability	0.000	0.000	-0.217***
	(0.000)	(0.000)	(0.003)
Immigrant	0.000	0.000	-0.090***
	(0.000)	(0.000)	(0.001)
High school	-0.001	0.001	0.004
	(0.001)	(0.001)	(0.004)
Tertiary education	0.001**	-0.001**	0.026***
	(0.000)	(0.000)	(0.001)
Age 25-34	-0.000	0.000	0.051***
	(0.000)	(0.000)	(0.001)
Age 35-44	-0.000	0.000	-0.016***
	(0.000)	(0.000)	(0.001)
Age 45-54	0.000	-0.000	-0.034***
	(0.000)	(0.000)	(0.001)
Service and admin	-0.000	0.000	-0.024***
	(0.000)	(0.000)	(0.002)
Service, care and shop	-0.000	0.000	-0.059***
	(0.000)	(0.000)	(0.001)
Building and manufacturing	-0.000	-0.000	0.071***
	(0.000)	(0.000)	(0.002)
Previously registered	-0.000	0.000	-0.019***
	(0.000)	(0.000)	(0.003)
Previous treatment	-0.000	0.000	-0.076***
	(0.000)	(0.000)	(0.001)
F-test	0.77	1.08	1,795
P-value f-test	0.695	0.375	0.000
N	83,447	83,447	83,447
N caseworkers	482	482	482
Mean depvar	0.89	0.13	0.50

Note: Using only first profiling per jobseeker.



(a) Assignment Rate (b) Deviation Rate (c) Value Added

Figure F1: Balance: JS Profiling Score and CW Rates

Note:

F.1.2 First Stage

Table F2

	Independent jobsearch	Deviation
CW assignment rate	0.330***	
	(0.038)	
CW deviation rate		0.305***
		(0.039)
N	83,447	83,447
Mean assignment rate	0.89	0.13
Number of caseworkers	482	482
Number of jobseekers		

F.1.3 Assignment Rate

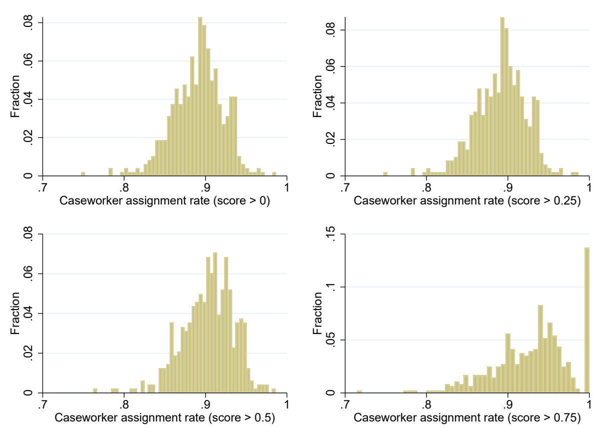


Figure F2: First Profiling: Assignment Rate

Note:

F.1.4 Deviations

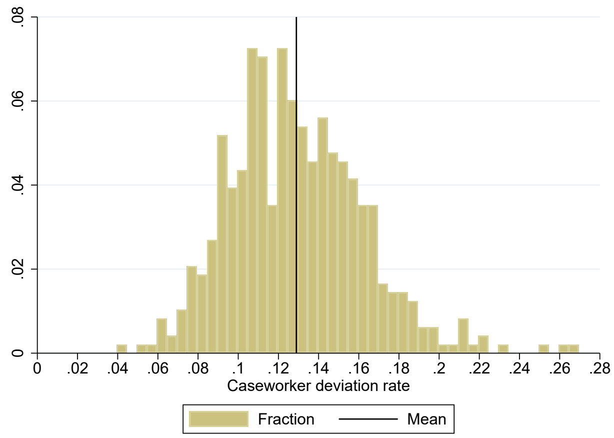


Figure F3: First Profiling: Deviation Rate

Note:

F.1.5 Exclusion Restriction

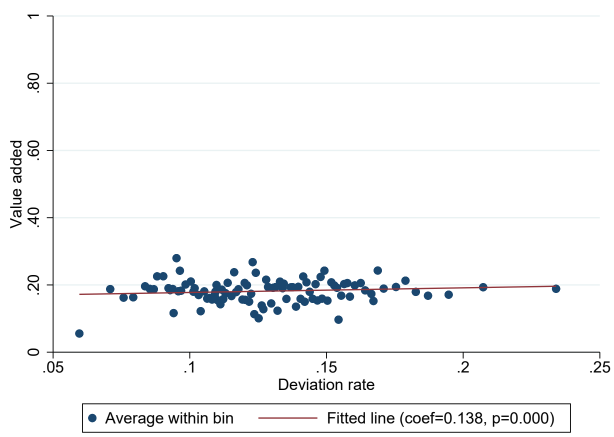


Figure F4: First Profiling: Deviation Rate and Value Added

Note:

F.2 Results

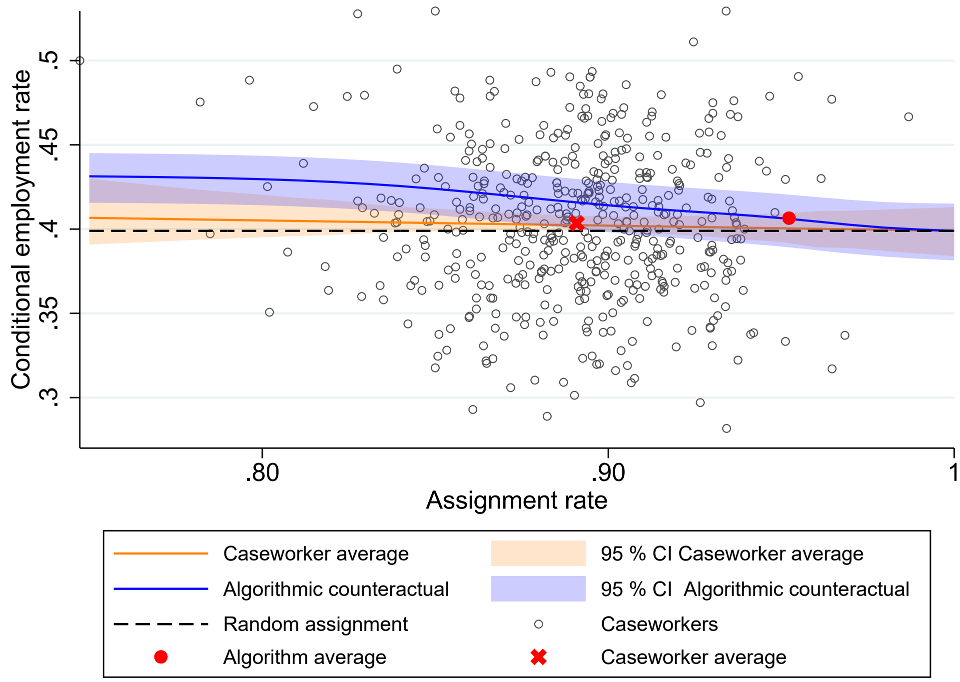


Figure F5: First Profiling: Fairness Results

Note:

Table F3: First Profiling: Efficiency Results

	Full time employment			
	6 months	6 months	12 months	12 months
CW deviation rate	-0.065	0.004	-0.120	-0.012
	(0.055)	(0.060)	(0.076)	(0.083)
Alg. assignment rate		X		X
Mean deviation rate	0.129	0.129	0.134	0.134
Number of caseworkers	83,447	83,447	44,420	44,420
Number of jobseekers	482	482	482	482

G Theoretical Appendix

This section outline model implications and proofs for propositions used in the conceptual framework.

G.1 Preliminaries

Let q_p^Z denote the p'th percentile of the empirical distribution of a given variable Z, in order to define the threshold parameter s_a in a unit-free manner. For a given share to be assigned to an ALMP, fairness is maximized if all jobseekers with $y_i^0 \leq q_p^{y^0}$ is assigned. Efficiency is maximized if all jobseekers with $\gamma_i \geq q_{(1-p)}^{\gamma}$ is assigned.

G.2 Proposition

Proposition 1 The predictive algorithm is utility maximizing, for any threshold s_a, if and only if (i) the treatment effect of ALMP is monotonically decreasing in the baseline reemployment probability y_i^0, *or* the policymaker exclusively values fairness. [11]

Proof

1. By definition, a monotonically decreasing function is order-reversing: if $y_i^0 \leq y_j^0$, then $k(y_i^0) \geq k(y_j^0)$ for all i, j. Consequently, for any p, the subset of jobseekers with $y_i^0 \leq q_p^{y^0}$ coincides with the subset of jobseekers with $\gamma_i \geq q_{(1-p)}^{\gamma}$. When these subsets fully overlap, both components of the policymaker's utility function are maximized. In this case, no fairness–efficiency trade-off arises.

2. If treatment effect γ_i is not monotonically decreasing in reemployment probability, then there exists jobseekers i, j for which $y_i^0 \leq y_j^0$, but $\gamma_i \leq \gamma_j$. In this case, the subset of jobseekers $y_i^0 \leq q_p^{y^0}$ does not coincide with the subset with $\gamma_i \geq q_{(1-p)}^{\gamma}$. Hence, if the policymaker values both fairness and efficiency, an assignment mechanisms assigning solely on y_i^0 cannot be utility maximizing for the policymaker at all cutoffs s_a.

3. If the policymaker solely values fairness, the problem reduces to assigning jobseekers based on their reemployment probability y_i^0. Making the use of a predictive algorithm utility maximizing.

[11] Provided that the algorithm achieves a sufficient level of prediction accuracy.

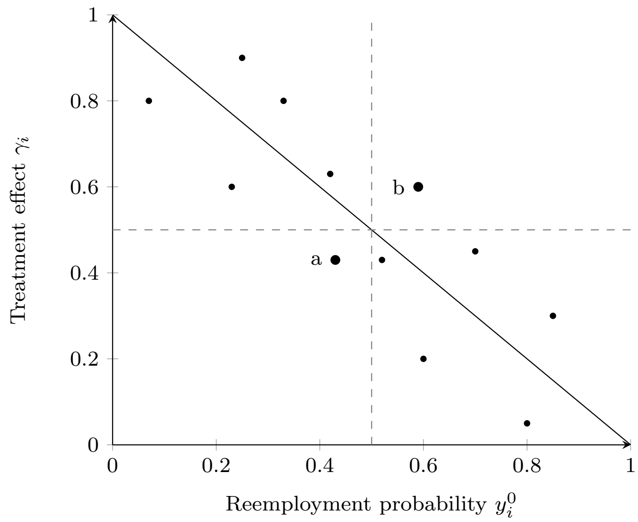


Figure G1: Illustration Proposition 1

Figure G1 illustrates a key insight from proposition 1: unless treatment effect is monotonically decreasing in baseline reemployment probability across all jobseekers, an assignment purely based on reemployment probability is not going to be utility maximizing for the policymaker. Each point in the figure represents a jobseeker i, showing their individual treatment effect of ALMP and baseline reemployment probability. In this example, the treatment effect is on average decreasing in reemployment probability, yet there remain individuals for whom the relationship is reversed.

The dashed gray lines indicate cutoffs s, corresponding to assignment rules based either on baseline reemployment probability or on treatment effect. Under a purely fairness-oriented objective, all jobseekers with $y_i^0 \leq 0.5$ would be assigned to an ALMP; under a purely efficiency-oriented objective, assignment would instead target those with $\gamma \geq 0.5$. When both objectives matter, assigning jobseekers in the upper-left quadrant yields strictly higher utility than assigning those in the lower-right quadrant.

Even with perfect predictive accuracy, assignment based solely on reemployment probability is not utility-maximizing whenever treatment effects and baseline reemployment probabilities are positively correlated for some individuals. This is illustrated by points a and b in Figure G1, which are marginal cases between which there is instead a positive relationship between reemployment probability and treatment effects. In such cases, the optimal assignment depends on the policymaker's relative weight on fairness versus efficiency.

While the treatment effect may be *on average* decreasing in baseline reemployment probability, it is a strong assumption that this relationship holds for all jobseekers i.