I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16626

Measuring Employment Readiness for Hard-to-Place Individuals

Simon Tranberg Bodilsen
Søren Albeck Nielsen
Michael Rosholm

NOVEMBER 2023

I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16626

Measuring Employment Readiness for Hard-to-Place Individuals

Simon Tranberg Bodilsen
Aarhus University

Søren Albeck Nielsen
Aarhus University

Michael Rosholm
Aarhus University and IZA

NOVEMBER 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Measuring Employment Readiness for Hard-to-Place Individuals

In an era characterized by population aging and economic challenges in welfare states across the world, sustaining these welfare systems requires a large workforce. Many individuals outside the labor market aspire to work but encounter a labyrinth of obstacles. While Public Employment Services employ Active Labor Market Policies, their effectiveness for this group remains uncertain. This study introduces the Employment Readiness Indicator Questionnaire (ERIQ), transcending traditional employment categories by assessing individuals' progress toward employment and measuring employment readiness for those labeled "hard-to-place". Integrating socially vulnerable clients into the labor market remains an unsolved challenge. ERIQ demonstrates impressive predictive abilities and points towards actionable recommendations by identifying malleable traits, such as social skills, coping strategies, goal orientation, and self-efficacy, that significantly impact employment readiness. ERIQ emerges as a valuable resource for policymakers and practitioners, advancing the goal of promoting labor market participation for socially vulnerable individuals.

JEL Classification: J60, J64, J68

Keywords: employment readiness, social assistance, machine learning, predictive algorithms

Corresponding author:
Michael Rosholm
Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

E-mail: rom@econ.au.dk

1 Introduction

In an era characterized by the constricting pressures of population aging, Baumol's cost disease and Wagner effects on welfare states, there exists an urgent need to improve the long-term sustainability of the welfare state, e.g. through larger supply of workers available for work in order to both increase the number of tax payers and reduce the number of individuals dependent on income transfers from the government (Andersen, 2015). Moreover, a significant proportion of individuals outside the traditional labor market share a common aspiration—namely, the desire to work. Hence, a transition into employment would also presumably improve the well-being and possibly even the physical and mental health of those obtaining employment. However, their journey towards employment is marred by a labyrinth of—personal as well as systemic—obstacles and challenges, making it difficult for them to find employment. While the Public Employment Services (PES) employ a comprehensive toolbox of e.g. Active Labor Market Policies (ALMPs), aiming to facilitate reintegration into the labor market, the effectiveness of these tools for this particularly vulnerable group of individuals remains uncertain, in part due to very low transition rates into ordinary employment and the associated difficulty of measuring causal effects. It is therefor imperative to develop a dynamic tool that transcends the traditional categorical views of employment, unemployment and non-participation. This tool should have the unique capacity to measure the progression of individuals on the fringes of the labor market towards sustainable employment—effectively capturing the employment readiness of those who are often labeled as "hard-to-place" but nevertheless willing to work.

Even within the context of Denmark's comprehensive Scandinavian-type welfare state structure, characterized by very extensive and relatively generous social assistance and employment services, the task of effectively integrating socially vulnerable clients receiving social assistance benefits into the labor market remains a significant and hitherto unsolved challenge.

This study validates a novel survey tool; the Employment Readiness Indicator

Questionnaire (henceforth labeled ERIQ), by offering an assessment of its ability to predict active job search and employment and compare it against a rich set of administrative data variables. The unique strength of this tool lies in its impressive forecasting ability as well as its ability to identify and direct caseworkers towards possible actions to enhance employment readiness. The latter is potentially useful to both individuals and caseworkers. These features originate in the survey's capability to pinpoint the factors that mostly influence socially vulnerable individuals in transitioning from social assistance to employment (or education).

We present compelling evidence for the validity of the ERIQ as a tool for predicting progression towards labor market reintegration. The ERIQ has an AUC-ROC (Area Under the Receiver Operating Characteristic Curve) of 83% when predicting employment, compared to an AUC-ROC of 64% using data from the administrative registers. The AUC-PR (Area Under the Precision-Recall Curve) is 28% compared to 10% for a similar model based on register data (baseline of 5.5%). In addition, once the ERIQ questionnaire is included in the predictive model, almost no predictive power is added by extending the model with data from the administrative registers.

To the best of our knowledge, we are the first to develop a freely available survey based tool that both (i) has very good predictive properties and (ii) almost directly recommends possible actions for the group of disadvantaged individuals on the edge of or outside the labor market. The latter is achieved by pointing to malleable traits that affected employment readiness, such as social skills, everyday coping strategies, goal orientation and self-efficacy.

By bridging the gap between research and practice, we contribute significantly to the broader effort of promoting inclusive and effective labor market participation for job-ready socially vulnerable clients.

Figure 1: Employment Readiness Domains

2 Literature on employment readiness and how to measure it

We investigate the relationship between distinct indicators of employment readiness included in the ERIQ and the likelihood of socially vulnerable clients commencing job search and obtaining employment using machine learning tools to predict these intermediate and final outcomes of the path towards employment readiness. The individual components of the survey were identified based on a literature survey of factors associated with employment and employment readiness (Væksthuset and NewInsight, 2012). This led to a large number of possible questions grouped within 11 overall domains, as shown in Figure 1. These questions were subsequently reduced to 11 questions for the caseworker and 11 questions for the client, which in combination covers all 11 domains.

For a group with such complex and differential problems, it is clear that for some, the path to employment is long, while for others it may be shorter.

McQuaid and Lindsay (2005) discuss different definitions of employability and argues that a proper definition should focus on demand-side factors (e.g. the needs of employers) as well as supply-side factors (e.g. network, personal and work-related competencies, motivation etc.) and, in particular, their interactions, as important for employability or employment readiness.

Relatedly, Pearson et al. (2023) argue that employment readiness should be viewed through a capabilities approach, focusing on what people *can* do rather than what they actually do.

As EIC (2020) notes, *"Each disadvantaged jobseeker faces a unique set of personal and work-related barriers; Personal circumstances, such as financial hardship, disability, caring responsibilities and substance dependence, can create barriers to employment by limiting access to opportunities and resources that improve jobseekers' employment prospects and enable them to find and retain work."* It is, however, important to understand how each of these measures associate with employment readiness, so appropriate measures to improve employment readiness can be identified.

We have only been able to identify few available tools that try to measure employment readiness in a disadvantaged population such as those on the edge of the labor market. There are commercial tools, such as the Canadian tool called the 'Employment Readiness Scale' (ERS), which contains 75 questions in total and is claimed to predict correctly in 80% of cases[1].

Wittevrongel et al. (2022) investigate two tools and their validity specifically for youth with autism spectrum disorders; the Work Readiness Inventory (WRI) and the Ansell-Casey life skills Assessment. The WRI is particularly relevant to our case. They argue that factors such as responsibility, flexibility, competencies, communication, self-efficacy and hopefulness are factors valued by potential employers.

Ding et al. (2023) conducts a scoping reivew of tolls that assess employability of cancer survivors, which is a very specific population. It focuses mostly on health related

[1]See `https://ersscale.com/` and (EIC, 2020).

indicators and thus does not offer a lot for a broader targeted tool.

Dencker-Larsen (2017) investigates whether the danish Well-being survey combined with data from administrative register can be used to predict a measure of employment readiness. She uses information on health, well-being, self-efficacy, alcohol use, and drug use and conclude that the relation between the proposed measures and subsequent employment is rather unstable. The analytical sample is, however, rather small (N<1,000). She does find some evidence that a self-efficacy factor (the belief in ones ability to find employment and to work) is significantly related to subsequent employment.

Hence, the literature on employment readiness measurement for a more general population of individuals on the margin of the labour market is, to the best of our knowledge, very limited. There is a related literature that attempts to predict employment in unemployed (employment ready) populations, such as the German TrEffeR (which also attempts to point to the potentially most effective intervention) and the danish Job Barometer (e.g. Stephan et al. (2006) and Rosholm et al. (2006), respectively). However, these models rely entirely on information available in administrative registers, which implies that their predictions are not that useful to caseworkers and can be discouraging to the clients (say, if the model predicts low employment chances due to age, gender, ethnicity, and educational background, this prediction is not very constructive in terms of how to intervene to improve the likelihood of employment).

3 Data and Methods

This study utilizes a unique data set comprising self-reported progression surveys collected every three months from social assistance clients assessed to be not ready for work, in 10 municipalities across Denmark, along with their approximately 300 attached caseworkers. The project and data collection was conceived and organized by Væksthuset (the Greenhouse) and Væksthusets Research Centre.[2] The data span a four-year period, from 2013 to 2016, with almost all clients entering the project in 2013. An essential aspect of this study is the ability to merge these survey responses with comprehensive data from administrative registers, which include detailed geographic and demographic information as well as very detailed weekly information on labor market status (employment, unemployment and other income transfers etc.), detailed educational information, and historical health and criminal records, for each recipient.

The integration of self-reported surveys with administrative data facilitates a comprehensive understanding of the recipients' progression and the intricate factors that impact their transition into employment. By combining these two data sources, this study gains valuable insights into the dynamics shaping recipients' trajectories and elucidates the correlates of successful employment outcomes. This comprehensive approach enables a more holistic exploration of the multifaceted factors that contribute to recipients' transition from social assistance to employment, providing a robust foundation for evidence-based policy recommendations and interventions.

3.1 Sample Description

The predictive model was analyzed using data collected from social assistance recipients assessed to be not ready for work in 10 municipalities across Denmark. The initial data set encompassed 15,818 unique responses from 5,512 clients. Each response ideally consisted of both 11 questions posed to the client (the client questionnaire) and 11 questions posed to the caseworker (the caseworker questionnaire). These questionnaires were answered in connection with compulsory meetings held between caseworkers and clients at the PES. To ensure the reliability and accuracy of the data, we carefully filtered out responses where either the client or the caseworker had not answered the questionnaire. Additionally, we excluded observations with a gap of more than 6 months between the completion of the two questionnaires. These data cleaning steps resulted

[2] Væksthuset is a service provider owned by a foundation, also called Væksthuset. The surplus generated by being service provider is used to finance Væksthusets Research Centre

in a refined data set, comprising 11,268 unique responses from 3,697 clients.

To focus specifically on the study of progression, we further eliminated 1,105 responses from clients who had only completed the survey once. As a result, the final population included in the statistical analysis consisted of 10,163 observations from 2,599 clients. It is important to note that whenever we analyze progression towards employment, the client's initial answers as well as the change in the answers from the first survey to the current one is included in our assessments. Thus, for the final analyses, the data set comprised 7,564 unique responses from 2,599 clients.

Appendix Table A.1 presents information on the 2,599 clients included in the analysis and their characteristics anchored at the time of the first meeting between the client and the caseworker. The clients were on average 39 years old, there is a small over-representation of women, and only 20 percent were married or stably living together with their partner. In general, the clients hold a low educational level with 71 percent having high school or less as their highest educational degree. Their employment history over the previous 5 years was very unfavorable, i.e., they were employed on average 2% of the past two years and 10% of the past five years. They had a substantial use of social assistance, which they also received at the anchor point in order to be included in this study. Interestingly, and in line with what we expected, we observe that the clients also had high usage of prescription medication (especially pain killers, life style medication, and antidepressants) and generally many contacts with the health care sector in terms of somatic and mental health diagnoses.

As common in the literature, we further divide the complete set of ERIQ responses into two separate samples for modeling purposes: a training sample representing 75% of the data, utilized for model development, and a test sample encompassing the remaining 25%, used to evaluate model performance. To avoid overfitting issues, we randomize individuals based on their (anonymized) personal ID numbers, guaranteeing that no individuals appear in both the training and test samples. This method yields a training sample comprising 5675 responses from 1930 distinct participants, while the test sample contains 1889 responses from 672 different participants.

3.2 Outcomes and Information Sets

3.2.1 Outcomes

The primary objective of this project is to examine the transition out of social assistance and into employment. Another primary outcome is the initiation of active job search. As a secondary outcome, we also consider transitions into educational programs in the ordinary educational system (although ERIQ was not developed with this transition in mind), since for social assistance recipients below 30 without qualifying education, it is a major aim for the PES to help them into the educational system.

Table 1 illustrates that only 9% of the sample successfully made a transition into either employment or education, with 6% entering employment and the remainder entering education. These figures underscore the main challenge faced when investigating the progression from social assistance to employment, as a significant portion of the recipients are very distanced from the labor market.

Job search is captured by constructing a dummy variable to indicate whether they are actively applying for jobs or not, taken from ERIQ. Table 1 demonstrates that job application prevalence is substantially higher, with 27% of the sample actively searching for jobs. We approach this measure from two perspectives. First, we examine whether applying for jobs serves as a viable intermediate goal toward the long-term objective of leaving unemployment altogether by including this dummy in the model for the transition into employment (and education). Secondly, we explore whether ERIQ can predict the likelihood of applying for jobs and, consequently, enhance the probability of successful reemployment in the long run. By investigating both of these angles, we hope to uncover valuable insights to support individuals in their progression from social assistance towards employment.

Table 1: Descriptive statistics for outcomes

	Average
Employment (1/0)	0.058
Education (1/0)	0.033
Employment or education (1/0)	0.089
Applying for jobs (1/0)	0.267
Number of observations	7.564

3.2.2 Feature Sets

We construct two distinct feature sets for our analysis. The first feature set (referred to as the "Admin" feature set) comprises a comprehensive set of characteristics of the social assistance recipients, extracted from the administrative registers. This includes information on gender, age, ethnicity, cohabitation status, municipality of residence, educational level, as well as fairly detailed employment, health, medical, and crime histories. An exhaustive list of variables included in the Admin feature set is found in Panel B of Table A.1 in the Online Appendix. The Admin feature set provides information on the participants receiving social assistance, representing characteristics that are often challenging, if not impossible, to change.

The second feature set is the employment readiness indicators (referred to as the "ERIQ" feature set). It contains all the information obtained from the two ERIQ questionnaires (one to clients and one to caseworkers). Social assistance recipients participating in ERIQ are queried approximately every three months during compulsory meetings at the PES, where they respond to a set of questions about their personal experiences. These questions cover various aspects, including social networks, coping strategies in daily life, health management, and knowledge about opportunities in the labor market as well as job search strategies. Additionally, the caseworkers are asked to evaluate the same social assistance recipients at the same meeting, using a set of indicators, some of which overlap with the participant's responses, while others explore additional dimensions, such as concentration ability and the caseworker's belief in the participant's potential for employment. The selection of questions for both the participants and caseworkers was based on a comprehensive literature review

Table 2: ERIQ indicators

Questions	**Client**	**Caseworker**
Q1	Work Ideation	Realism
Q2	Taking Initiative	Goal Orientation
Q3	Collaboration	Initiative
Q4	Networking	Self-Presentation
Q5	Everyday Coping	Collabouration
Q6	Health	Instruction
Q7	Competence	Concentration
Q8	Job Performance	Networking
Q9	Knowledge of Opportunities	Everyday Coping
Q10	Job Search	Health
Q11		Job Prospect

Notes: With the exception of client question 10, all indicators are measured using a single question employing a 5-point Likert scale (1–5). The complete formulation of each question can be found in Table A.2 in the Online Appendix. Clients were also asked about their reservation wage, Q11, however; due to insufficient response rates and measurement error issues, this aspect was not incorporated into the analysis.

Væksthuset and NewInsight (2012), aiming to identify employment readiness indicators that are malleable. The selected indicators are summarized in Table 2, while the full set of questions are available in the Online Appendix Table A.2. For descriptive statistics, please refer to Panel A in Table A.1 in the Online Appendix.

Finally, we combine the two feature sets into a third feature set ("Admin + ERIQ") to investigate whether the information contained in both sets complement each other, resulting in improved predictions. Alternatively, if no significant improvement is observed, it may suggest that one of the sets is more influential in the prediction process.

3.3 Methods

3.3.1 Prediction Models

We employ four different machine learning methods of increasing degrees of complexity to predict the primary and secondary outcomes. Importantly, all four models are implemented using the exact same sample splits and data, ensuring that the model predictions are directly comparable.

Linear Probability Model First, we consider a linear probability model (LPM) estimated using ordinary least squares. This model offers the advantage of being straightforward and interpretable, allowing us to determine the influence of each variable by examining the regression coefficients. However, the clear downside of the LPM lies in its simplicity, as it only captures linear relationships in the data and assigns non-zero weight to all variables in the feature set, which increases the risk of overfitting.

Logistic Regression Model with LASSO The second model combines the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 1996) with a logistic regression framework. This hybrid approach is well-suited for handling binary outcome variables and provides both variable selection and regularization, enhancing the precision of the predictions. To determine the optimal size of the regularization parameter λ, we employ five-fold cross-validation. Specifically, we select the value of λ that maximizes the cross-validated AUC-ROC (see below).

To implement this model, we utilize the glmnet R package, and in accordance with the authors' recommendations, we standardize all variables to have a mean of zero and a standard deviation of one. This standardization helps ensure comparability and stability in the model's performance across different variables.

Random Forest Model The third model is a random forest model, initially introduced by Breiman (2001), which employs bagging as an ensemble learning technique. Bagging involves training different individual decision trees on various random subsets of the data in parallel. Additionally, random forest models perform a random selection of explanatory variables for each decision tree, significantly reducing the risk of overfitting the model.

For the implementation of the random forest algorithm, we utilize the ranger R package (Wright and Ziegler, 2017). To optimize the model's predictive performance, two critical hyperparameters, namely the number of variables considered at each node (mtry) and the minimal node size (min.node.size), were thoughtfully selected. We employed a Bayesian optimization approach to identify the optimal hyperparameter configurations, maximizing the AUC-ROC through five-fold cross-validation. We implement the random forest algorithm using 1,000 independent trees.

Extreme Gradient Boosting Model The final and most complex predictive model is the extreme gradient boosting (XGBoost) model(Chen and Guestrin, 2016). This method uses boosting as an ensemble learning technique. Boosting combines weak models iteratively, focusing on correcting errors made by previous models, to create a strong predictive model. The XGBoost algorithm effectively handles nonlinear relationships in the data and mitigates overfitting through regularization and pruning.

To estimate the XGBoost model, we utilize the xgboost R package and fine-tune its performance by optimizing seven hyperparameters through Bayesian optimization. In accordance with the xgboost package's terminology, we explore the following hyperparameters: max.depth, eta, gamma, subsample, colsample_bytree, colsample_bynode, and min_child_weight. Specifically, we search for the hyperparameter configurations that yield the highest AUC-ROC in the training sample.

3.3.2 Model Performance

The predictive models we consider yield the probability of the transition into employment (or either of the other outcomes). To assess their performance using two different feature sets, we employ AUC-ROC and AUC-PR as performance metrics.

The ROC curve plots the true positive rate of the predictive model against its false positive rate for each decision thresholds from 0 to 1. A higher AUC-ROC indicates that the model is more likely to assign a higher predicted probability of transition out of unemployment to a randomly chosen true positive (i.e., an individual actually finding employment) than to a randomly chosen true negative (i.e., an individual not finding employment). It is essential to note that random predictions would yield an AUC-ROC of 50%.

In binary classification, the precision of a classifier is the ratio of true positives to the total number of predicted positives (true positives plus false positives), while recall corresponds to the true positive rate (true positive divided by the sum of true positives and false negatives). By adjusting the threshold between zero and one for a given prediction model, we can plot the precision-recall curve, and the area under this curve (AUC-PR) can be calculated. An optimal model would have an AUC-PR value of one, indicating perfect precision and recall, while random guessing yields a score equal to the proportion of positives in the data (in our case, 5.8% for employment). Higher AUC-PR values indicate better model performance for a specific data set, but it is crucial to compare them to the prevalence of the outcome in the data. Therefore, direct comparison of AUC-PRs between different data sets or outcomes should be avoided, as their interpretation is specific to the characteristics of each data set. However, it is valid for comparison between different features sets and model specifications.

The AUC-PR has a particular advantage in the context of highly imbalanced data, as the present case where the fraction of negatives is significantly larger than the fraction of positives (Saito and Rehmsmeier, 2015). In the ROC approach, equal importance is given to predicting both negative and positive instances correctly, which might lead to a high AUC-ROC score even when the model exhibits a significant number of false positive predictions. This is more likely to happen in severely imbalanced data sets, where true negatives outweigh false negatives. However, because the AUC-PR focuses on how well the model predicts the positives (i.e., movement into employment) the fraction of correctly predicted negatives becomes irrelevant.

In the context of transitions from social assistance to employment, it is crucial to study how well a predictive model can identify positive outcomes. Therefore, focusing on precision and recall allows us to address this aspect effectively, ensuring that the model's performance is assessed based on its ability to predict positive outcomes accurately.

3.3.3 Explaining Predictions

To elucidate the influence of different variables, including interactions between them, on the outcomes of interest, we employ Shapley additive explanation (SHAP) values (Lundberg and Lee, 2017; Lundberg et al., 2020). SHAP values offer a model agnostic approach to unravel the underlying factors shaping the predicted probabilities of the transition out of unemployment.

By utilizing SHAP values, we can gain insights into how predictive models make specific predictions for each individual in the data set. These values provide a measure of the contribution of each variable in each feature set to the final prediction. A SHAP value for a variable expresses how much its information alters the model's opinion in relation to the prediction. In other words, SHAP values illustrate how the values of individual variables influence the prediction away from the average prediction of the outcome, while accounting for correlations between variables. For comparison, the SHAP values equal the regression coefficients of a linear regression model in situations where variables are uncorrelated and there are no interactions.

The adoption of SHAP values enhances the interpretability and transparency of predictive models, enabling a deeper understanding of the factors influencing the outcome of interest. The insights gleaned from SHAP values facilitate tailored interventions and evidence-based policy decisions aimed at adaptable variables, thus potentially contributing to higher employment rates in the long run and increasing well-being of social assistance recipients not ready for work.

4 Results

In this section, we present the predictive models' performance in forecasting the primary and secondary outcomes. We examine the models' ability to predict the likelihood of obtaining employment, enrolling in education within a year after responding to the questionnaire, the combination of the two (i.e., finding employment or commencing education,) and we examine the predictive performance in regards to the inclination to apply for jobs.

4.1 Model Performance

Table 3 presents the test performance of the four predictive models on each of the four outcomes. The models are evaluated using three distinct feature sets: Admin, ERIQ, and a combination of both (Admin + ERIQ).

The ERIQ feature set demonstrates the most promising predictive performance across the primary outcomes (employment and job search) across all four predictive models. Notably, for the transition into employment and for starting job search, the XGBoost model achieves AUC-ROC scores of 83.48% and 84.24%, respectively, and AUC-PR scores of 27.51% and 62.76%, respectively, when using the ERIQ feature set. These results suggest that the ERIQ feature set, which encompasses information related to personal experiences, social networks, coping strategies, health management, and knowledge about job market opportunities obtained from the ERIQ questionnaire, plays a vital role in accurately predicting successful transitions into employment as well as active job search.

On the other hand, the Admin feature set, which includes baseline characteristics and historical data, shows comparatively lower predictive performance when predicting employment. For employment within a year and active job search, the XGBoost model achieves AUC-ROC scores of 63.63% and 68.53%, respectively, and AUC-PR scores of 9.87% and 42.46%, respectively. These results highlight that the Admin feature set alone may not fully capture the essential factors that influence successful transitions into employment or the start of active job search.

When looking at the secondary outcome, transition into education within a year, the Admin feature set outperforms the ERIQ feature set in terms of both performance measures.

The combined feature set (Admin + ERIQ) leverages the strengths of both ERIQ

Table 3: Test performance of the predictive models

	Admin				ERIQ				Admin+ERIQ			
	AUC-ROC	95% CI		AUC-PR	AUC-ROC	95% CI		AUC-PR	AUC-ROC	95% CI		AUC-PR
(a) Employed within a year												
Linear probability model	62.96	57.33	68.60	8.98	80.69	76.73	84.64	20.02	77.65	73.26	82.04	17.33
Logistic LASSO	63.63	58.24	69.02	9.31	82.51	78.43	86.59	25.59	82.23	78.27	86.19	21.35
Random forest	60.60	54.94	66.25	9.01	83.31	79.58	87.05	26.70	84.31	80.85	87.76	24.03
XGBoost	63.63	57.75	69.50	9.87	83.48	79.72	87.25	27.51	83.73	79.99	87.47	24.89
Outcome rate: 5.51%												
(b) Applying for a job												
Linear probability model	67.18	64.38	69.98	41.21	83.00	80.88	85.11	61.04	84.39	82.37	86.41	62.12
Logistic LASSO	67.67	64.89	70.46	41.61	83.43	81.37	85.49	60.93	85.54	83.62	87.45	63.66
Random forest	69.06	66.31	71.81	43.16	83.90	81.92	85.88	61.16	85.32	83.42	87.21	63.00
XGBoost	68.53	65.76	71.31	42.46	84.24	82.27	86.20	62.76	86.32	84.50	88.13	66.05
Outcome rate: 24.46%												
(c) In education within a year												
Linear probability model	80.27	74.30	86.25	20.74	74.99	68.61	81.38	10.16	82.72	76.71	88.73	26.13
Logistic LASSO	80.74	74.89	86.59	20.06	76.74	70.79	82.70	8.35	86.71	82.20	91.22	28.88
Random forest	82.67	76.82	88.52	15.74	72.79	65.61	79.98	8.18	85.50	80.38	90.62	19.67
XGBoost	80.67	74.45	86.88	16.23	75.62	69.24	82.00	9.42	86.63	81.63	91.63	24.87
Outcome rate: 3.07%												
(d) Employed or in education within a year												
Linear probability model	68.79	64.15	73.43	19.37	81.05	77.65	84.46	28.22	81.21	77.56	84.86	31.37
Logistic LASSO	68.62	64.05	73.19	19.17	81.50	78.10	84.89	32.08	82.94	79.47	86.42	36.68
Random forest	69.17	64.74	73.61	19.96	80.87	77.29	84.44	32.35	85.19	82.26	88.13	38.05
XGBoost	70.17	65.78	74.57	20.72	81.50	78.04	84.96	33.43	84.06	80.74	87.39	41.26
Outcome rate: 8.42%												

Notes: The table provides AUC scores for the four models on the two primary and two secondary outcomes.

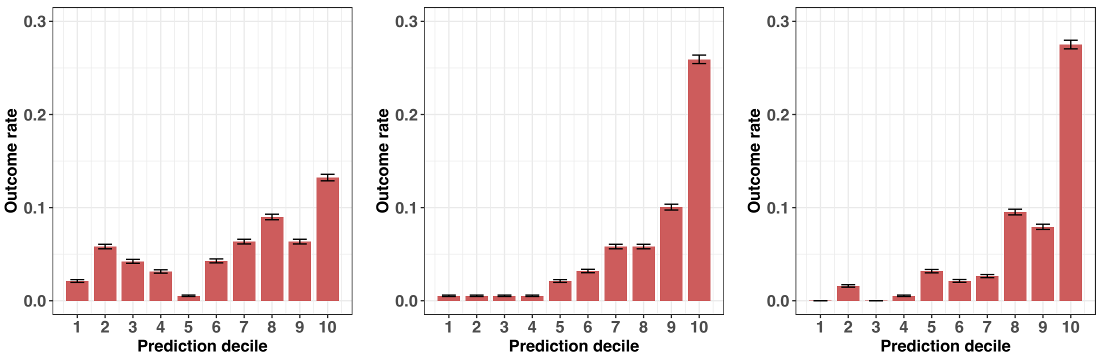
and Admin feature sets. However, even with this combination, the ERIQ feature set remains crucial for improved predictive performance. For employment within a year and active job search, the XGBoost model achieves AUC-ROC scores of 83.73% and 86.32%, respectively, and AUC-PR scores of 24.89% and 66.05%, respectively. These results indicate that when predicting future employment and active job search, very little is gained from adding Admin to the ERIQ feature set.

We divided the data into two sub-groups based on the institutional setting, as caseworkers at the PES focus primarily on helping young individuals below 30 into education and on finding employment opportunities for those aged 30 or above. Consequently, it is interesting to assess whether the ERIQ indicators serve as the best feature set for predicting the transition into employment and job search for both age groups. Online Appendix Table A.3 reveals that the ERIQ indicators have a significant impact on the model's performance across both age splits. Moreover, the noteworthy predictive performance of the full models when predicting enrollment into the secondary outcome education using the Admin feature set is largely influenced by age. Thus, the ERIQ feature set proves vital as its adaptable features provide crucial information on employment and education for both age groups. The results reaffirm the importance of the ERIQ feature set in accurately predicting successful transitions out of social assistance and job search for individuals in different age groups, reinforcing its value for targeted interventions and policy decisions aimed at enhancing reemployment and well-being for disadvantaged social assistance recipients.

Another way to illustrate the ability of the different feature sets to predict the primary outcomes is to plot the true positive rate within prediction deciles, with decile 1 being the 10% of individuals in the sample with the lowest predicted probability of a given outcome, and decile 10 is, similarly, the 10% with the highest predicted probability of a given outcome.

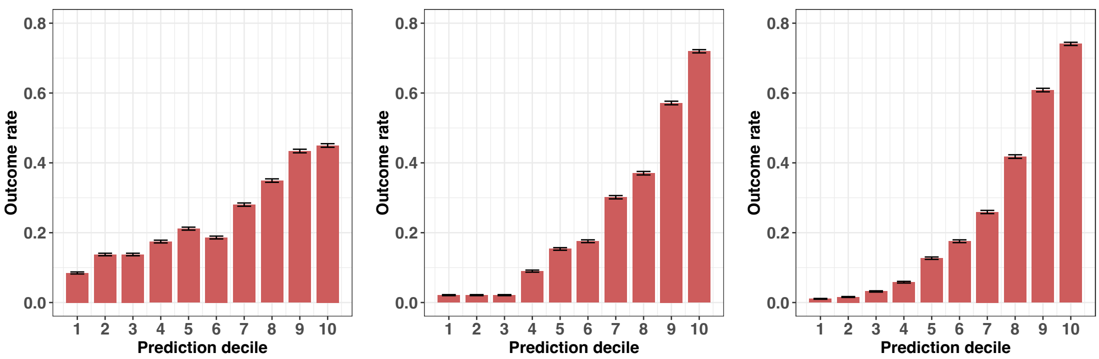
We perform this analysis for the two primary outcomes, employment and active job search, in Figures 2 and 3, respectively. The figures only show the result from

Figure 2: Fraction of true positives by prediction decile: Employed within a year



(a) Admin **(b)** ERIQ **(c)** Admin + ERIQ

Figure 3: Fraction of true positives by prediction decile: Applying for a job



(a) Admin **(b)** ERIQ **(c)** Admin + ERIQ

our preferred XGBoost machine learning model. Figure 2 shows that, for the Admin feature set in Panel (a), the true positive rate tends to increase by prediction decile, but the increase is not very steep, nor is it monotonous. In the lowest prediction decile, the true positive rate is around 2%, while it is 13% in the highest prediction decile. For the ERIQ feature set, in Panel (b), the relation is, on the other hand, monotonous and much steeper. Namely, the true positive rate is below 1% and 26% in decile 1 and 10, respectively, and the relation is convex with a large increase especially from decile 9 to decile 10. Combining the two feature sets in Panel (c), we gain a tiny bit of precision in deciles 1 and 10, but at the cost of the monotonous relationship across deciles.

For job search, shown in Figure 3, the overall picture is much the same, with gains in precision as well as monotonicity when going from admin to ERIQ feature sets. In the admin feature set, the true positive rate is 8% and 44% in deciles 1 and 10, respectively, while the same numbers for the ERIQ features set is 2% and 72%.

In sum, the ERIQ feature set emerges as the most critical factor contributing to the enhanced predictive performance of the model. Its ability to capture various aspects of an individual's life, including personal experiences, social networks, coping strategies, health management, and job market awareness, proves essential in accurately predicting the transition into employment and the initiation of job search. This understanding facilitates the development of targeted interventions and evidence-based policy decisions aimed at enhancing reemployment of socially vulnerable clients. The superior performance of XGBoost, particularly when using the ERIQ feature set, makes it the preferred model for our analysis, allowing us to gain in-depth insights into the factors influencing successful transitions.

4.2 Client vs. caseworker questionnaire

We now investigate to what extent the client and caseworker questionnaire separately contribute to predicting the outcomes of interest. Table A.4 in the Online Appendix sheds light on this by presenting AUC-ROC and AUC-PR scores for the complete ERIQ feature set, and split into client and caseworker indicators. Evidently, the confidence intervals overlap, with a marginal discrepancy in the explanatory power of the two feature subsets. Notably, client indicators slightly outperform caseworker indicators, when it comes to predicting transition out of unemployment and when predicting job search. This distinction emphasizes that in scenarios characterized by limited resources to caseworkers for filling in the questionnaire, prioritizing the collection of client indicators would be a possibility.

4.3 Subset of ERIQ indicators implemented in practice

Another possibility would be to use a subset of both questionnaires. Gathering 22 indicators after each meeting is a potentially resource-intensive endeavor, which could pose challenges for practical implementation. Drawing on insights from Rosholm et al. (2017), a preliminary exploration of the correlation between ERIQ indicators and employment, a subset of these indicators has been adopted by some PES offices in Sweden.[3] We therefore conduct an exploratory analysis to compare the performance of the full set of indicators with this restricted subset, bearing in mind that answering only a subset of the full questionnaire may slightly alter the answers to the questions. Table A.5 presents the predictive model's performance using only the aforementioned subset of ERIQ indicators. The AUC-ROC and AUC-PR scores are marginally lower compared to utilizing the complete ERIQ set, yet the confidence intervals overlap. This suggests that, within resource-constrained environments, opting for these eight indicators is relevant and viable.

5 Which variables predict employment readiness?

In the following section, we investigate which of the specific questions in the ERIQ questionnaire that contribute most to the predictive model for our employment readiness outcomes by employing SHAP values. These values unveil the most influential factors that shape the predicted probabilities concerning clients' prospects of securing a job as well as their job search activity.

In the figures below, we combined the global variable importance and local variable importance information into one main plot. The plot displays the mean of the absolute SHAP values for the ten most important variables, giving an overview of their overall impact on the model predictions. Additionally, we show the distribution of the SHAP values for the same variables using color coding to indicate the values of each variable.

5.1 Finding employment

The ERIQ indicators offer valuable insights by highlighting the specific dimensions of employment readiness that directly relate to employment attainment and the initiation of job search. The unique combination of available data enables us to track clients

[3] Specifically, client indicators 3, 5, 6, 7, 8, and 10, as well as caseworker indicators 2 and 11 in Table A.2 have often been used.

considered not immediately ready for employment over an extended duration, observing who successfully secured employment within that time frame—and who did not.

We observed in Table 1 that 5.5% of the monitored sample obtained employment within a 12-month period after answering a given questionnaire. Employing SHAP value analysis on our preferred XGBoost model provides valuable insights into the indicators that most effectively predict clients' likelihood of obtaining employment.

The SHAP analysis identifies the most important indicators for predicting employment. The ten most important indicators are shown in Figure 4 and are presented both for the model using the ERIQ feature set only (Panel (a)) and the model using the full feature set that combines ERIQ and administrative data (Panel (b)). The analysis in panel a shows that the caseworker's belief in the client's ability to find a job (job prospect) has clearly the largest impact on the likelihood of acquiring a job, as the SHAP value is more than twice as large as the second-most important variable, which is the indicator for applying for a job. It is quite impressive that the caseworkers' subjective assessments of the clients' abilities to obtain employment trumps actual job search behavior. Note, however, that also the applied job search channels appear in the figure and contribute to explaining employment. The analysis also shows that clients who are goal oriented (goal orientation), who believe themselves that they can handle a job (job performance), and who improve their ability to cope with any health challenges (health) increase the likelihood of acquiring a job.

When looking at Panel (b) in Figure 4, which combines the two feature sets, we find that still the caseworker's belief is the single most important predictor followed by some long term employment history and sex. Five of the ten most important indicators are from the ERIQ, and, more importantly, these indicators are malleable, at least to some extent, in contrast to sex, age, and employment history.

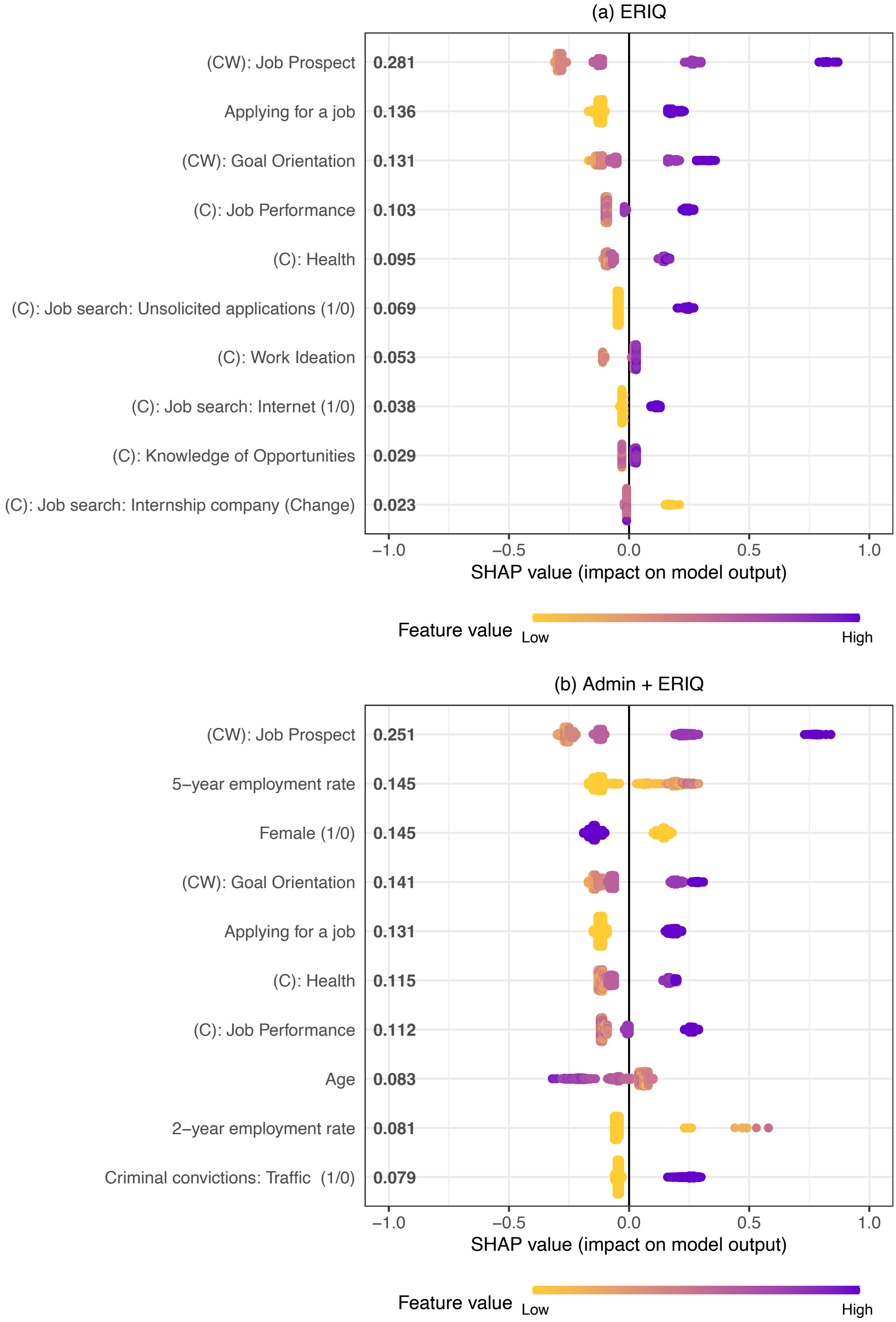
5.2 Applying for a job

Recognizing the crucial importance of active job search as a necessary step towards obtaining employment, this section delves into clients' job search activities, identifying key indicators that notably influence the probability of initiating job search.

Panel (a) Figure 5 underscores the strong connection between clients' job search activities and their response across the measured ERIQ indicators. The belief among clients in their capacity to handle a job emerges as the most predictive factor, while the beliefs of the caseworker is also on the list of most important factors. Moreover, the caseworker's assessment of the degree of goal orientation of the client is important for predicting job search. Clients who improve their ability to cope with any health challenges are also more likely to start looking for a job. Furthermore, clients who are more aware of the opportunities available in the labor market in relation to their personal resources and challenges, will increase the likelihood that they start looking for a job. Finally, factors such as everyday coping skills, work ideation, initiative, and ability to concentrate are good predictors of job search activity.

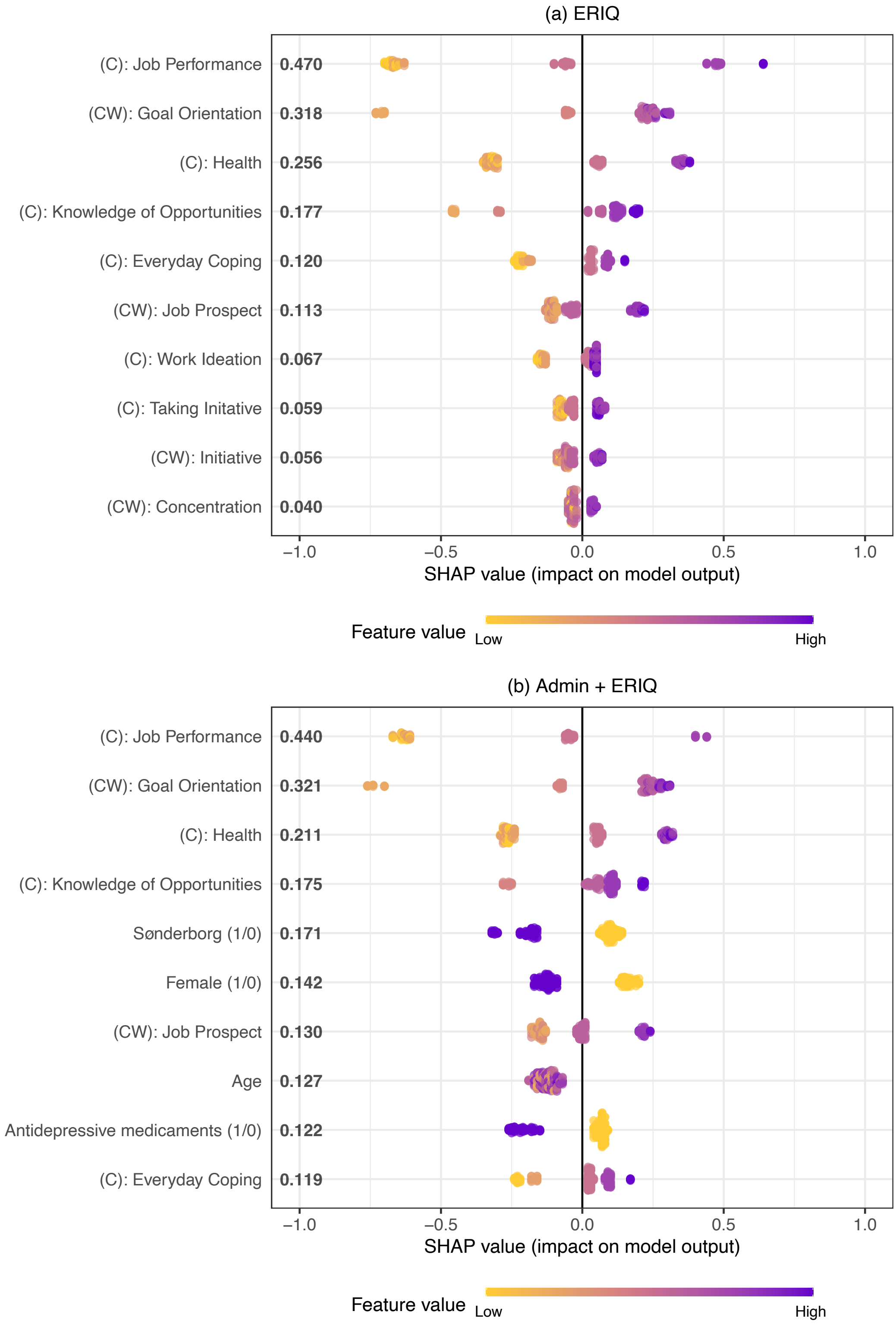
In Panel (b) of Figure 5 it is observed that six of the ERIQ questions are among the ten most important predictors when combining ERIQ and administrative data in the predictive model. The important predictors from register data are a specific geographic location, being female, age, and taking (or not) antidepressant medication.

Figure 4: SHAP values for predicting the transition into employment



Notes: This figure illustrates the SHAP values for the 10 most influential features. Panel (a) displays results from a model exclusively utilizing ERIQ features, while Panel (b) encompasses a model incorporating both ERIQ and administrative data. The y-axis is labeled with feature names, ordered by descending importance, and accompanied by their respective average impact on the model output. Meanwhile, the x-axis represents SHAP values, quantifying the magnitude of change in log odds associated with each feature. In this graphical representation, feature values are color-coded, with dark purple indicating high values and red signifying low values. (C) denotes client questions, and (CW) denotes caseworker questions.

Figure 5: SHAP values for predicting job search



Notes: This figure illustrates the SHAP values for the 10 most influential features. Panel (a) displays results from a model exclusively utilizing ERIQ features, while Panel (b) encompasses a model incorporating both ERIQ and administrative data. The y-axis is labeled with feature names, ordered by descending importance, and accompanied by their respective average impact on the model output. Meanwhile, the x-axis represents SHAP values, quantifying the magnitude of change in log odds associated with each feature. In this graphical representation, feature values are color-coded, with dark purple indicating high values and yellow signifying low values. (C) denotes client questions, and (CW) denotes caseworker questions.

6 Discussion and Conclusion

This study offers a new tool, the Employment Readiness Assessment Questionnaire, ERIQ, which has very strong predictive properties when it comes to predicting crucial measures of employment readiness—such as job search and obtaining employment—for social assistance recipients who are not assessed to be immediately ready for work.

These insights have potentially strong implications for the ability to assist clients further away from employment in their journey towards employment. As such, it offers an intermediate target outcome that can be used for measuring progress towards employment and it points to specific challenges experienced by the client or assessed by the caseworker, which—in contrast to information from administrative registers, such as age, gender, ethnicity, and labor market history—are to some extent malleable either by the client, the caseworker, or through appropriately tailored interventions. The insights gained from using ERIQ thus enables the tailoring of interventions to address the specific needs and challenges of individuals outside the labor market, ultimately increasing the efficacy of such programs.

This tool also offers an intermediate target outcome on which to measure the impact/effectiveness of active interventions (labor market and other types of interventions) targeted at overcoming specific challenges. The knowledge gained from this study can thus serve as a compass for quality assurance and evaluation efforts, guiding the selection of indicators that are most important for job search initiation and successful job acquisition for a given client with a certain combination of challenges or disadvantages.

Moreover, the results underscores the instrumental role of caseworkers in guiding social vulnerable clients' progress toward the labor market. Caseworkers serve as catalysts for clients' successful transition out of long-term unemployment, as evidenced by the high importance of both client and caseworker indicators on predicted probabilities. This confirms the importance of promoting collaborations between caseworkers and clients to maximize the impact of reintegration efforts into employment.

In conclusion, ERIQ enables a deeper dive into the dynamics governing job search activities and employment outcomes among socially vulnerable clients and is directly applicable in PES offices.

The applicability of ERIQ's predictive capabilities beyond the Danish context is yet to be established. Nevertheless, its active implementation in several PES offices in Sweden is an encouraging sign of potential cross-cultural utility. Considering the extensive volume of data within Denmark's administrative registers, the noteworthy superiority of a tool utilizing ERIQ over one relying solely on administrative registers implies that this enhanced predictive performance may extend to other nations as well. This suggests a potential superiority of ERIQ over administrative data based predictive tools in diverse international contexts.

Acknowledgements

We are grateful to Væksthuset and Væksthusets Research Centre for conducting the literature review and data collection, as well as the actual construction of the survey developed in this project. We are also grateful for getting access to the data from the project. Further, we gratefully acknowledge useful comments and suggestions from Leif Kongsgaard, Væksthuset, and Charlotte Liebak Hansen, Væksthusets Research Centre.

References

Andersen, T. M. (2015). Mauno koivisto lecture 2014: The nordic welfare model and welfare services – can wemaintain acceptable standards? *Research on Finnish Society 8*, 85–96.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Chen, T. and C. Guestrin (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Dencker-Larsen, S. (2017, Sep.). Measuring employability for disadvantaged unemployed people? evidence from survey and register data. *Nordic Journal of Working Life Studies 7*(3).

Ding, M., E. Gane, H. Wiffen, and V. Johnston (2023). Tools to assess employment readiness for colorectal cancer survivors: A scoping review. *Cancer Medicine 12*(17), 18327–18353.

EIC (2020). Inquiry into sustainable employment for disadvantaged jobseekers. Technical report.

Lundberg, S. M., G. Erion, H. Chen, and et al. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence 2*(1), 56–67.

Lundberg, S. M. and S.-I. Lee (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems.*

McQuaid, R. W. and C. Lindsay (2005). The concept of employability. *Urban Studies 42*(2), 197–219.

Pearson, S., C. Lindsay, E. Batty, A. M. Cullen, and W. Eadson (2023). Relational approaches to employability. *Journal of Education and Work 36*(4), 270–283.

Rosholm, M., J. Staghøj, M. Svarer, and B. Hammer (2006). A danish profiling system. *Nationaløkonomisk Tidsskrift 144*, 209–229.

Rosholm, M., K. L. Sørensen, and L. Skipper (2017). Bip indikatorer og jobsandsynlighed. Technical report, Væksthusets Forskningscenter.

Saito, T. and M. Rehmsmeier (2015). The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PloS one 10*(3), e0118432.

Stephan, G., S. Rässler, and T. Schewe (2006). Das TrEffeR-Projekt der Bundesagentur für Arbeit : die Wirkung von Maßnahmen aktiver Arbeitsmarktpolitik (The TrEffeR project of the Federal Employment Agency : the effect of active labour market poli. *Zeitschrift für ArbeitsmarktForschung - Journal for Labour Market Research 39*(3/4), 447–465.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological) 58*(1), 267–288.

Væksthuset and NewInsight (2012). Litteraturreview ifm. beskæftigelsesindikatorprojektet. Technical report, Væksthuset and New insight.

Wittevrongel, K., M. W., M. L. Hébert, and J. D. Nicholas, D. B.and Zwicker (2022). Acceptability of employment readiness measures in youth and young adults on the autism spectrum: a mixed-methods study. *Disability and Rehabilitation 44*, 4319–4331.

Wright, M. N. and A. Ziegler (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software 77*(1), 1–17.

Supplementary Material for

Measuring Employment Readiness for Hard-to-Place Individuals

(not intended for publication)

Table A.1: Sample description and feature sets

	Mean	Std. Dev.
Panel A: ERIQ indicators		
Client indicators		
- *Work Ideation*	3.207	1.427
- *Taking Initiative*	3.027	1.390
- *Collabouration*	3.805	1.093
- *Networking*	3.755	1.365
- *Everyday Coping*	2.820	1.275
- *Health*	2.533	1.162
- *Competence*	3.623	1.244
- *Job Performance*	2.837	1.355
- *Knowledge of Opportunities*	3.012	1.356
- *Applying for job*	0.287	0.452
Caseworker indicators		
- *Realism*	3.262	1.389
- *Goal Orientation*	2.409	1.237
- *Initiative*	2.898	1.297
- *Self-Presentation*	2.888	1.417
- *Collabouration*	3.110	1.532
- *Instruction*	3.333	1.533
- *Concentration*	2.686	1.594
- *Networking*	2.756	1.524
- *Everyday Coping*	2.650	1.237
- *Health*	2.929	1.138
- *Job Prospect*	2.548	1.239
Panel B: Admin data		
Age (years)	38.938	9.821
Age-squared/100 (years)	16.126	7.876
Female (1/0)	0.561	0.496
Married (1/0)	0.205	0.404
Non-danish ethnicity (1/0)	0.204	0.403

Continued on next page

Table A.1: Sample description (continued)

	Mean	Std. Dev.
Employment and benefit history		
- *2-year employment rate*	0.030	0.102
- *5-year employment rate*	0.101	0.170
- *5-year social assistance rate*	0.578	0.338
- *5-year self-sufficient rate*	0.080	0.165
- *5-year sickness benefit rate*	0.086	0.144
- *5-year education benefit rate*	0.040	0.124
- *5-year health program rate*	0.000	0.001
- *3-year employment rate*	0.051	0.129
- *3-year social assistance rate*	0.690	0.337
- *3-year self-sufficient rate*	0.059	0.146
- *3-year sickness benefit rate*	0.084	0.174
- *3-year education benefit rate*	0.033	0.125
Education		
- *High-school or below (1/0)*	0.705	0.456
- *Vocational training (1/0)*	0.242	0.428
- *Higher education (1/0)*	0.053	0.224
Criminal charges		
- *Penal code (1/0)*	0.133	0.340
- *Traffic (1/0)*	0.157	0.364
- *Other (1/0)*	0.088	0.284
Criminal convictions		
- *Penal code (1/0)*	0.135	0.342
- *Traffic (1/0)*	0.154	0.361
- *Other (1/0)*	0.088	0.284
Prescription medicine		
- *Life style medicine (1/0)*	0.230	0.421
- *Pain killers (1/0)*	0.303	0.460
- *Anti psychotic medicine (1/0)*	0.112	0.315
- *Anxiety medicine (1/0)*	0.069	0.254
- *Anti depressive medicine (1/0)*	0.289	0.454

Continued on next page

Table A.1: Sample description (continued)

	Mean	Std. Dev.
- *Abuse related medicine (1/0)*	0.079	0.270
Somatic diagnoses		
- *Cancer (1/0)*	0.008	0.087
- *Diabetes (1/0)*	0.009	0.096
- *Neurological (1/0)*	0.064	0.245
- *Cardiovascular (1/0)*	0.047	0.212
- *Respiratory (1/0)*	0.043	0.202
- *Musculoskeletal (1/0)*	0.220	0.415
- *Maternal issues (1/0)*	0.094	0.292
- *Digestive and urogenital (1/0)*	0.188	0.391
- *Eye and periocular (1/0)*	0.032	0.176
- *Simple healthcare utilization (1/0)*	0.574	0.495
- *Other (1/0)*	0.304	0.460
Psychiatric diagnoses		
- *Organic (1/0)*	0.004	0.065
- *Substance-induced (1/0)*	0.056	0.230
- *Various Psychotic (1/0)*	0.028	0.166
- *Affective (1/0)*	0.137	0.344
- *Anxiety and stress-related (1/0)*	0.122	0.328
- *Behavioral changes (1/0)*	0.007	0.085
- *Personality disturbances (1/0)*	0.085	0.278
- *Intellectual disability (1/0)*	0.002	0.044
- *Developmental disorders (1/0)*	0.007	0.085
- *Childhood and adolescent disorders (1/0)*	0.050	0.217
- *Other (1/0)*	0.027	0.162
Municipality		
- *Faxe (1/0)*	0.047	0.211
- *Horsens (1/0)*	0.045	0.207
- *Rebild (1/0)*	0.044	0.205
- *Silkeborg (1/0)*	0.109	0.312
- *Sønderborg (1/0)*	0.260	0.439

Continued on next page

Table A.1: Sample description (continued)

	Mean	Std. Dev.
- *Thisted (1/0)*	0.072	0.259
- *Vejen (1/0)*	0.121	0.326
- *Viborg (1/0)*	0.057	0.232
- *Vordingborg (1/0)*	0.088	0.284
- *Aarhus (1/0)*	0.157	0.364

Table A.2: Questions to client and Caseworkers

Questions to client	
No.	Questions
1	Are you aware of what type of work you would like to perform?
2	How do you feel about initiating contact with people whom you do not know?
3	How good are you at collaborating with others?
4	Do you have the support of family and friends when you need help?
5	Do you have the personal energy in your everyday life to focus on getting a job?
6	In general, how would you rate your (physical and mental) health in terms of being able to hold a job?
7	Do you think your skills can be used in a workplace?
8	Do you think you are able to carry out work at a workplace?
9	Do you know what to do in order to improve your job opportunities?
10	How do you search for a job?
Questions to Caseworkers	
No.	Questions
1	Does the client have a realistic understanding of where in the labour market, his/her competencies can be applied?
2	To what degree does the client act with determination in terms of obtaining a job?
3	How do you assess the client's ability to seek and initiate dialogue with others?
4	How good is the client in discussing about him/herself and his/her relevant competencies?
5	How do you assess the client's ability to cooperate with others?
6	How do you assess the client's ability to receive and understand instructions about a task?
7	How do you assess the client's ability to concentrate on a task without being distracted?
8	To what extent does the client have a social network that provides support for entering the labour market?
9	To what extent is the client able to master his/her own life at the same time as focusing on obtaining a job?
10	To what extent is the client able to master any (physical and mental) health problems?
11	Do you believe that the client will get a job within the next year?

Table A.3: Test performance of the XGBoost model for different age groups

	Admin				ERIQ				Admin+ERIQ			
	AUC-ROC	95% CI		AUC-PR	AUC-ROC	95% CI		AUC-PR	AUC-ROC	95% CI		AUC-PR
(a) Employed within a year												
$<$ 30 years	81.28	70.96	91.61	22.48	84.21	76.22	92.20	25.05	85.73	76.53	94.94	31.97
$\geq$ 30 years	61.45	55.21	67.69	8.07	83.43	79.28	87.58	28.08	83.38	79.32	87.45	23.74
Outcome rates: 7.03% and 5.34%												
(b) Applying for a job												
$<$ 30 years	69.82	59.27	80.36	39.86	81.23	73.35	89.11	51.83	83.10	75.43	90.76	47.71
$\geq$ 30 years	68.02	65.12	70.91	42.76	84.73	82.70	86.75	64.08	86.56	84.69	88.43	67.18
Outcome rates: 16.76% and 25.29%												
(c) In education within a year												
$<$ 30 years	58.89	46.74	71.05	23.17	73.17	63.22	83.11	43.87	69.60	59.26	79.94	35.58
$\geq$ 30 years	66.88	56.04	77.72	5.35	77.41	67.56	87.26	5.38	80.96	71.31	90.60	13.36
Outcome rates: 17.84% and 1.47%												
(d) Employed or in education within a year												
$<$ 30 years	69.67	60.00	79.35	39.45	79.46	71.72	87.21	61.15	82.48	75.77	89.19	63.68
$\geq$ 30 years	65.32	60.11	70.53	10.98	82.70	78.68	86.72	32.77	82.16	78.02	86.31	30.37
Outcome rates: 23.78% and 6.75%												

Notes: The table provides AUC scores for the XGBoost model for the four different outcomes for individuals younger than 30 years and older than 30 years.

Table A.4: Comparing Client and Caseworker indicators

	ERIQ				Client				Caseworker			
	AUC-ROC	95% CI		AUC-PR	AUC-ROC	95% CI		AUC-PR	AUC-ROC	95% CI		AUC-PR
(a) Employed within a year												
Linear probability model	80.69	76.73	84.64	20.02	80.23	76.33	84.14	18.71	78.73	74.18	83.27	18.66
Logistic LASSO	82.51	78.43	86.59	25.59	80.63	76.29	84.98	21.47	78.57	73.92	83.23	21.49
Random forest	83.31	79.58	87.05	26.70	82.51	78.78	86.23	22.69	78.13	73.41	82.85	22.68
XGBoost	83.48	79.72	87.25	27.51	82.89	79.12	86.67	24.90	77.93	73.10	82.76	21.71
Outcome rate: 5.51%												
(b) Applying for a job												
Linear probability model	83.00	80.88	85.11	61.04	81.81	79.63	83.99	57.26	79.58	77.33	81.83	55.78
Logistic LASSO	83.43	81.37	85.49	60.93	82.07	79.92	84.23	57.57	79.36	77.07	81.65	55.37
Random forest	83.90	81.92	85.88	61.16	82.39	80.29	84.49	58.49	80.12	77.92	82.32	56.26
XGBoost	84.24	82.27	86.20	62.76	82.71	80.62	84.81	58.53	80.18	77.96	82.40	56.47
Outcome rate: 24.46%												
(c) In education within a year												
Linear probability model	74.99	68.61	81.38	10.16	75.51	69.03	81.98	11.50	72.42	66.60	78.24	6.89
Logistic LASSO	76.74	70.79	82.70	8.35	74.00	67.49	80.52	8.06	74.87	69.04	80.69	7.15
Random forest	72.79	65.61	79.98	8.18	69.78	62.49	77.08	9.19	70.31	63.26	77.35	6.41
XGBoost	75.62	69.24	82.00	9.42	73.82	67.06	80.58	9.98	73.11	66.69	79.54	7.23
Outcome rate: 3.07%												
(d) Employed or in education within a year												
Linear probability model	81.05	77.65	84.46	28.22	79.33	75.86	82.81	26.44	78.95	75.37	82.53	25.01
Logistic LASSO	81.50	78.10	84.89	32.08	79.17	75.35	82.99	28.35	79.18	75.68	82.69	26.58
Random forest	80.87	77.29	84.44	32.35	78.79	75.07	82.50	27.34	77.84	74.02	81.66	28.44
XGBoost	81.50	78.04	84.96	33.43	79.53	75.76	83.30	29.37	78.54	74.87	82.20	29.23
Outcome rate: 8.42%												

Notes: The table provides AUC scores for the four models on the four different outcomes.

Table A.5: Test performance of the predictive models using only a subset of the ERIQ covariates

	Subset of ERIQ			
	AUC-ROC	95% CI		AUC-PR
(a) Employed within a year				
Linear probability model	81.05	76.98	85.12	19.77
Logistic LASSO	81.34	76.90	85.79	24.73
Random forest	83.20	79.52	86.87	23.79
XGBoost	82.78	78.67	86.89	25.85
Outcome rate: 5.51%				
(b) Applying for a job				
Linear probability model	81.65	79.46	83.83	57.65
Logistic LASSO	81.76	79.58	83.93	57.54
Random forest	81.69	79.58	83.80	55.09
XGBoost	81.52	79.37	83.67	55.01
Outcome rate: 24.46%				
(c) In education within a year				
Linear probability model	73.17	67.16	79.19	8.23
Logistic LASSO	74.02	67.75	80.28	7.13
Random forest	69.03	61.98	76.07	6.74
XGBoost	73.65	67.25	80.04	6.95
Outcome rate: 3.07%				
(d) Employed or in education within a year				
Linear probability model	78.37	74.53	82.22	28.86
Logistic LASSO	79.55	75.70	83.41	33.09
Random forest	79.19	75.47	82.91	29.96
XGBoost	79.53	75.63	83.44	32.75
Outcome rate: 8.42%				

Notes: The table provides AUC scores for the four models on three different outcomes.