Risk scores for long-term unemployment and the assignment to job search counseling

Ernst, Sebastian ; Mueller, Andreas I ; Spinnewijn, Johannes

DOI: https://doi.org/10.1257/pandp.20241092

Originally published at:
Ernst, Sebastian; Mueller, Andreas I; Spinnewijn, Johannes (2024). Risk scores for long-term unemployment and the assignment to job search counseling. AEA Papers and Proceedings, 114:572-576.
DOI: https://doi.org/10.1257/pandp.20241092

AEA Papers and Proceedings 2024, 114: 572–576
https://doi.org/10.1257/pandp.20241092

Risk Scores for Long-Term Unemployment and the Assignment to Job Search Counseling†

By Sebastian Ernst, Andreas I. Mueller, and Johannes Spinnewijn*

A recent literature emphasizes the role of heterogeneity in job finding and long-term unemployment risk across job seekers (e.g., Alvarez, Borovičková, and Shimer 2023; Gregory, Menzio, and Wiczer 2021; Mueller, Spinnewijn, and Topa 2021; Ahn, Hobijn, and Şahin 2023; Mueller and Spinnewijn 2023). This literature finds that job seekers differ vastly in their probability of finding a job and thus the likelihood of becoming long-term unemployed. A recent report (Desiere, Langenbucher, and Struyvan 2019) documents the increasingly common practice of risk profiling the unemployed for targeted unemployment policies, either based on specific dimensions (e.g., age and education) as assessed by a mediating caseworker or as predicted by statistical profiling models. Yet despite the large and growing body of work on job search and unemployment policy (Card, Kluve, and Weber 2018; Kircher 2022), the use of risk profiling for targeted interventions has received little attention in the academic literature (e.g., Black et al. 2003; Mueller and Spinnewijn 2023; van den Berg et al. 2023).

In this paper, we analyze how risk profiling is used to assign job seekers to job search counseling in Flanders, Belgium. We leverage the specific context and data to shed light on a number of questions and issues. In particular, we compare algorithmic selection to self-selection and selection by mediators and caseworkers, discuss practical challenges for the implementation of risk profiling, and highlight avenues for further research.

* Ernst: London School of Economics and Political Science (email: s.y.ernst@lse.ac.uk); Mueller: University of Texas at Austin, NBER, CEPR, and IZA (email: andimueller@utexas.edu); Spinnewijn: London School of Economics and Political Science and CEPR (email: j.spinnewijn@lse.ac.uk). We thank the VDAB and Joost Bollens, Stijn Van De Velde, and Kristof Mondy in particular for their generous support. We also thank participants at the AEA meetings and Bart Cockx for valuable comments.

† Go to https://doi.org/10.1257/pandp.20241092 to visit the article page for additional materials and author disclosure statement(s).

I. Risk Profiling in Flanders

All unemployed individuals in Belgium are required to enroll with the regional Public Employment Service (PES) to receive unemployment benefits. The *Vlaamse Dienst voor Arbeidsbemiddeling en Beroepsopleiding* (VDAB) is the PES for the Flemish region, which is the northern, Dutch-speaking part of Belgium and which contains about 60 percent of the population. The VDAB operates a job search platform and organizes job search counseling and the assignment to training programs. Guidelines on unemployment policy from the European Union provide a clear recommendation for member states to prevent and reduce long-term unemployment (EUR-Lex 2018).

As soon as someone enters unemployment, the VDAB opens a new file and generates a risk score. To obtain this score, the VDAB has developed a prediction model, "Employment Prospects," which has been operational since October 2018 and predicts the probability that an unemployed person begins a period of employment lasting at least 28 consecutive days within the next six months. The model uses random forest algorithms and includes more than 400 individual features, with a strong emphasis on labor market history and the ongoing unemployment spell and also includes sociodemographic characteristics, job search preferences (e.g., occupation and geography), and behaviors (e.g., updating of profile, strategy, and CV). Risk scores are updated for job seekers who remain unemployed. The model is also retrained every month, adding the most recent spells.

After registration with the VDAB, the VDAB sends a confirmation email to the job seeker describing what they need to do to receive unemployment benefits and what will happen next. Four weeks into the unemployment spell, job seekers are sent an email instructing them to call the VDAB service line. These "inbound calls" result in an in-depth half-hour interview with a

Table 1—Predicted and Observed Job Finding Rates

	Observations	Share	Predicted JFR	Observed JFR
Full sample	84,943	1.00	0.551	0.609
Inbound call in weeks 1–3	16,656	0.20	0.511	0.556
Assigned to caseworker	7,799	0.09	0.475	0.465
Caseworker contact	6,958	0.08	0.473	0.440
Inbound call in weeks 4–5	19,799	0.23	0.573	0.608
Self-reliant	13,272	0.16	0.596	0.689
Assigned to caseworker	6,527	0.08	0.525	0.444
Caseworker contact	5,603	0.07	0.524	0.419
Considered for outbound call at 5 weeks	27,157	0.32	0.547	0.533
Prediction >65 percent	5,509	0.06	0.716	0.722
Prediction <65 percent	21,648	0.25	0.504	0.485
Calls in weeks 6–8	13,226	0.16	0.503	0.475
Self-reliant	6,125	0.07	0.536	0.590
Assigned to caseworker	7,101	0.08	0.474	0.376
Caseworker contact	6,202	0.07	0.474	0.356

Notes: This table reports mean six-month predicted and observed job finding rates after one week of unemployment. The full sample includes all new unemployment spells of fully unemployed benefit recipients aged 26–58 at the start of the spell. Spells are only included if they begin between October 3, 2020, and September 15, 2022, if they last at least one week and if a risk score is available at week 1. Individuals immediately reassigned to a caseworker from a previous spell or engaging in subsidized or temporary work by week 5 are excluded. The sample for "inbound call in weeks 4–5" is further restricted to spells lasting at least four weeks. After around five weeks of unemployment, the service line uses individuals' current risk scores to determine whom to call. Individuals who are not considered for an outbound call at this point include those who were assigned to a caseworker due to an inbound call in weeks 1–3, those who made an inbound call in weeks 4–5, those who have exited unemployment by week 5, and those who have been assigned to a caseworker by week 5 through another mechanism. Another 7 percent of the full sample is not contacted for other reasons (e.g., they no longer receive benefits and therefore cannot be made to see a caseworker). "Calls in weeks 6–8" includes both inbound and outbound calls. "Caseworker contact" refers to any contact with a caseworker at a regional office by week 16.

VDAB mediator who then decides whether the job seeker is *self-reliant*. If not, the mediator sets an appointment with a local caseworker. Five weeks into the spell, the risk scores are used to determine which job seekers, among those who have not yet contacted the VDAB, to reach out to. Any job seeker with a predicted job finding probability below 65 percent is called. These "outbound calls" proceed like the inbound calls, with the VDAB mediator determining at the end of the interview whether the job seeker is self-reliant or needs assignment to a local caseworker. Just like all job seekers deemed self-reliant, those with a risk score above 65 percent are not called until week 19 of the spell unless their risk score drops below the 65 percent threshold before that.[1]

Table 1 illustrates the different steps of the contact strategy and selection of job seekers for assignment to caseworkers. The table reports the number and share of unemployed job seekers remaining at each step and their predicted and observed job finding probability at the start of the spell. The sample includes unemployment spells starting between October 2020 and September 2022. A first observation is that only 25 percent of all unemployed job seekers are targeted by the algorithmic risk scoring at around 5 weeks of unemployment. This small share is not because of selective assignment—only 6 percent of the total sample is screened out because of their risk score. However, a large share has exited unemployment (19 percent), and many have already been assigned to a caseworker or contacted VDAB themselves (42 percent) within the first 5 weeks of unemployment. Hence, the actual use of algorithmic

[1] Job seekers can also call the service line and be assigned to a caseworker in weeks 1–3. However, the mediator cannot classify them as self-reliant before week 4. If not assigned to a caseworker, they will be instructed to call the service line again in week 4. Note also that risk score groups are in principle available to mediators but are not used in the assignment decision.

profiling by the VDAB is fairly limited. A second observation is that a substantial share of the selection is still done by the VDAB mediators themselves, either in addition to or instead of the algorithmic selection. During the calls taking place after the algorithmic selection, the VDAB mediator screens out another 46 percent of the remaining job seekers as self-reliant. For the inbound calls not preceded by algorithmic selection, this share is about two-thirds of the job seekers. A third observation is that the ultimate share of job seekers meeting with caseworkers is relatively low. In fact, the algorithmic screening has lead to an assignment to caseworkers of only 7 percent of the total sample.

II. Assessment of Algorithmic Risk Scoring

We assess the impact of the algorithmic risk scoring and compare the resulting selection with the selection by the mediators and job seekers themselves. Figure 1 compares the distribution of risk scores of job seekers selected by the algorithm (panel A), by the job seekers themselves (panel B), and by the mediator (panel C). These selections happen at different times in the unemployment spell and on different samples of unemployed job seekers. For comparability, we plot the risk scores as predicted at week 1 in all panels. Panel A naturally shows a clear distinction between job seekers with a risk score below versus above 65 percent who are targeted for outbound calls. Still, there is significant overlap as the risk scores used for selection are generated at the time of selection (at five weeks into unemployment). The predicted job finding probabilities are thus not fully persistent, as also documented in Sweden by Mueller and Spinnewijn (2023). Panel B shows the selection by job seekers themselves in deciding whether or not to make an inbound call. The distributions are relatively close to each other. Job seekers who call in during the first three weeks have lower predicted job finding probability. Job seekers who call in after being instructed to have slightly higher predicted job finding probabilities. This positive selection is the opposite of the VDAB's intended targeting strategy as applied in the algorithmic profiling. Panel C considers the selection made by the VDAB mediators during the calls and shows that the predicted job finding probabilities are substantially lower for the job seekers who are deemed not self-reliant.

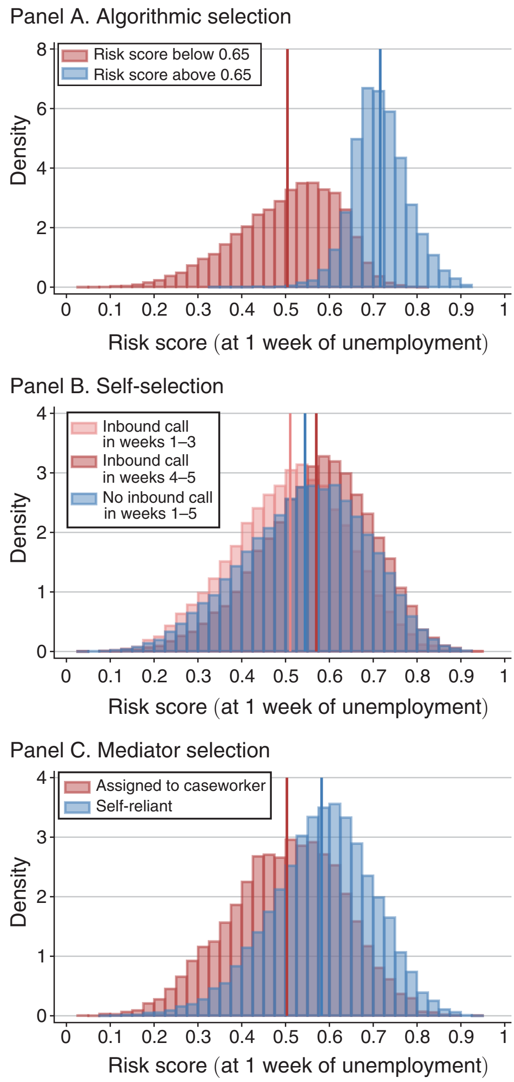

Figure 1. Selection by Risk Scores

The mediators' selection adds value with or without prior algorithmic selection. First, a comparison in Table 1 shows that the risk scores of those assigned to caseworkers *with* prior algorithmic selection (47 percent) versus *without* prior algorithmic selection (48 percent and 53 percent, respectively) are not very different. Hence, by this criterion, the mediators perform well by themselves compared to their combination with algorithmic selection. Of course, the algorithmic procedure saves on a large number of calls to be made. Second, a comparison of the predicted job finding and observed job finding of the selected groups suggests that the mediators have relevant private information. While the wedge in predicted job finding probabilities between job seekers deemed self-reliant versus not equals 7 percentage points (54 percent versus 47 percent), the wedge in observed job finding probabilities is 3 times as large at 21 percentage points (59 percent versus 38 percent). This is even a lower bound on the private information channel when caseworkers have a positive treatment effect on the observed job finding of those assigned to them.[2]

A natural but positive challenge for the contact and assignment strategy is the fact that many job seekers find employment. However, some practical challenges result in additional attrition. First, of those job seekers selected by the algorithm, a significant share are called but not reached. Second, of those job seekers selected by the mediator, a small share do not actually meet with a caseworker. The attrition muddles the intended selection for assignment to caseworkers. Figure 2, Panel A compares those who are and are not targeted for the outbound calls with those who are and are not actually reached. The densities show that the attrition at this stage is random, but there is greater similarity in predicted risk scores between those who are and are not contacted than between those who are and are not targeted for contact. Panel B of Figure 2 makes the same comparison for those assigned by mediators to caseworkers or not and those who actually meet caseworkers or not. In this case, the attrition is negatively selected, with more job seekers with low risk scores dropping out. This further undermines the targeting efforts.

[2]Results are very similar when comparing six-month predicted and observed job finding at four weeks of unemployment instead of at one week. These are generated closer to the date of mediator selection. The predicted job finding probabilities of job seekers deemed self-reliant versus not are then 52 percent versus 46 percent, and the wedge in observed job finding probabilities is still much larger at 62 percent versus 41 percent. This is in line with van den Berg et al. (2023), who show that caseworker assessments in Germany have predictive power of reemployment above and beyond algorithmic risk scores.

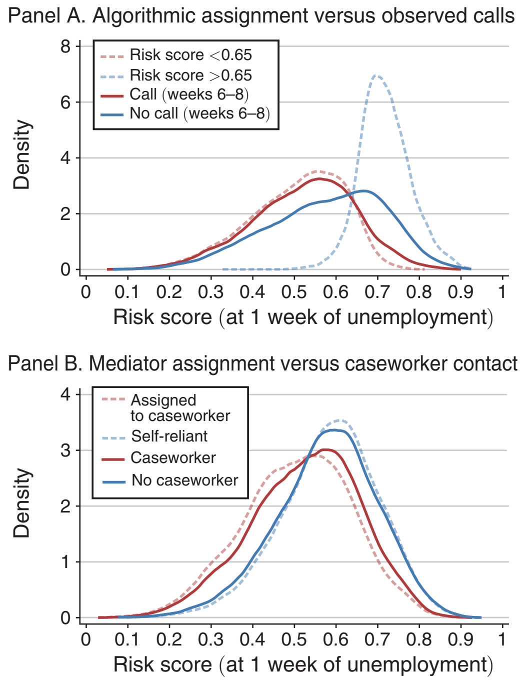

Figure 2. Attrition in Assignment

III. Discussion and Further Research

The effective use of risk scoring can improve the assignment of PES interventions in two ways: interventions can be targeted better and implemented earlier in the spell (Mueller and Spinnewijn 2023). For example, if at one week, the PES targeted the bottom 7 percent of the distribution of predicted job finding, the average risk score of this targeted group would be 28 percent. This compares to an average risk score of 47 percent for the 7 percent of our sample who are targeted by the algorithm and ultimately matched to a caseworker.[3] Of course, this presumes that individuals prone to long-term unemployment gain more from the interventions and that the gains for these individuals are not reduced when intervening earlier. While there is some evidence on the treatment effects of unemployment policies by duration of unemployment

[3]Alternatively, the PES could screen by actual duration of unemployment as the employment prospects for the sample of survivors decreases, but this type of screening comes at the cost of a delayed timing of interventions.

(e.g., Cockx, Lechner, and Bollens 2023 in the Flemish context), we are not aware of any study that evaluates the heterogeneity in treatment effects based on predicted risk scores. Given the increasing attention that PES agencies pay to algorithmic assignment of targeted policies, it is important to provide more evidence on who gains more from these targeted policies and to quantify the returns to targeting. To this purpose, in follow-up work with the Flemish data, we are looking at the treatment effects of assignment to caseworkers and how they differ by predicted long-term unemployment risk.

Another avenue for further work is to study the optimal combination of algorithmic assignment with voluntary take-up and mediator assessment. Our evidence shows that mediators appear to have private information in relation to the likelihood of job finding. It would be of great interest to understand whether mediators also have private information on who responds more to targeted policies. Similarly, individuals who self-select may also be more prone to respond to these policies.

Finally, further research is needed on how algorithmic assignment may conflict with ethical and fairness considerations (see Rambachan et al. 2020). In the Flemish context, gender and country of origin are not used in the prediction model, though it is not clear whether this avoids algorithmic bias if other features included in the model are correlated with these characteristics.

REFERENCES

Ahn, Hie Joo, Bart Hobijn, and Ayşegül Şahin. 2023. "The Dual US Labor Market Uncovered." NBER Working Paper 31241.

Alvarez, Fernando, Katarína Borovičková, and Robert Shimer. 2023. "Decomposing Duration Dependence in a Stopping Time Model." *Review of Economic Studies*. https://doi.org/10.1093/restud/rdad109.

Black, Dan A., Jeffrey A. Smith, Mark C. Berger, and Brett J. Noel. 2003. "Is the Threat of Reemployment Services More Effective Than the Services Themselves? Evidence from Random Assignment in the UI System." *American Economic Review* 93 (4): 1313–27.

Card, David, Jochen Kluve, and Andrea Weber. 2018. "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations." *Journal of the European Economic Association* 16 (3): 894–931.

Cockx, Bart, Michael Lechner, and Joost Bollens. 2023. "Priority to Unemployed Immigrants? A Causal Machine Learning Evaluation of Training in Belgium." *Labour Economics* 80: 102306.

Desiere, Sam, Kristine Langenbucher, and Ludo Struyvan. 2019. "Statistical Profiling in Public Employment Services: An International Comparison." OECD Working Paper 224.

EUR-Lex. 2018. "Council Decision (EU) 2018/1215 of 16 July 2018 on Guidelines for the Employment Policies of the Member States." https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018D1215.

Gregory, Victoria, Guido Menzio, and David G. Wiczer. 2021. "The Alpha Beta Gamma of the Labor Market." NBER Working Paper 28663.

Kircher, Philipp. 2022. "Schumpeter Lecture 2022: Job Search in the 21st Century." *Journal of the European Economic Association* 20 (6): 2317–52.

Mueller, Andreas I., and Johannes Spinnewijn. 2023. "The Nature of Long-Term Unemployment: Predictability, Heterogeneity and Selection." NBER Working Paper 30979.

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

Rambachan, Ashesh, Jon Kleinberg, Jens Ludwig, and Sendhil Mullainathan. 2020. "An Economic Perspective on Algorithmic Fairness." *AEA Papers and Proceedings* 110: 91–95.

van den Berg, Gerard J., Max Kunaschk, Julia Lang, Gesine Stephan, and Arne Uhlendorf. 2023. "Predicting Re-Employment: Machine Learning versus Assessments by Unemployed Workers and by Their Caseworkers." IZA Discussion Paper 16426.