Digitalisation in Public Employment and Guidance Services

Maynooth University Social Services and Department of Sociology

18 September 2020

Session 1 – Chair dr. Michael Mc Gann

The profiling system of the PES in Flanders using artificial intelligence: the accuracy-equity trade-off

Ludo.Struyven@kuleuven.be

@LudoStruyven

AI innovations in PES

- AI innovations are transforming the welfare benefits services
- Decision-making based on machines, replacing the judgment of human caseworkers
- Used to suspend welfare payments without notice
- But AI and algorithms may also enable a positive activation of jobseekers
- AI is considered a promising avenue to improve the (cost-)efficiency and customization of delivering public services

Profiling and targeting

- Classification and segmentation of jobseekers
  - Aim is to devote limited resources to ‘vulnerable’ jobseekers
- Prevention and early identification of high-risk jobseekers
  - Focus for policy intervention has been shifted to the initial period up to at most the first year of unemployment (the longer a person is out of work, the harder it is to find a job)
  - A timely support for all jobseekers becomes more urgent as benefits become more conditional or limited in time (double task of PES)

Three types of profiling

- Rule-based profiling
  - Uses administrative eligibility criteria to classify jobseekers into client groups
  - Require normative choices and are often path-dependent
- Caseworker-based profiling
  - Relies on caseworkers' judgement to profile jobseekers
  - Caseworkers' discretion leads to different outcomes for similar jobseekers
- Statistical profiling
  - Uses a statistical model that predict the likelihood of work resumption
  - Is often considered an objective approach

AI-based profiling

- Next step in the development of statistical profiling
  - Predict a jobseeker’s likelihood of resuming work within a certain period, using machine learning techniques, often including many more variables (big data)
  - To give input into targeting and tailoring to jobseekers’ need: to allocate or adapt employment services/programs to jobseekers
- Main advantages
  - AI models are more flexible, can be updated continuously
  - Will in general more accurately predict, but can they also better avoid discrimination ?

Central question

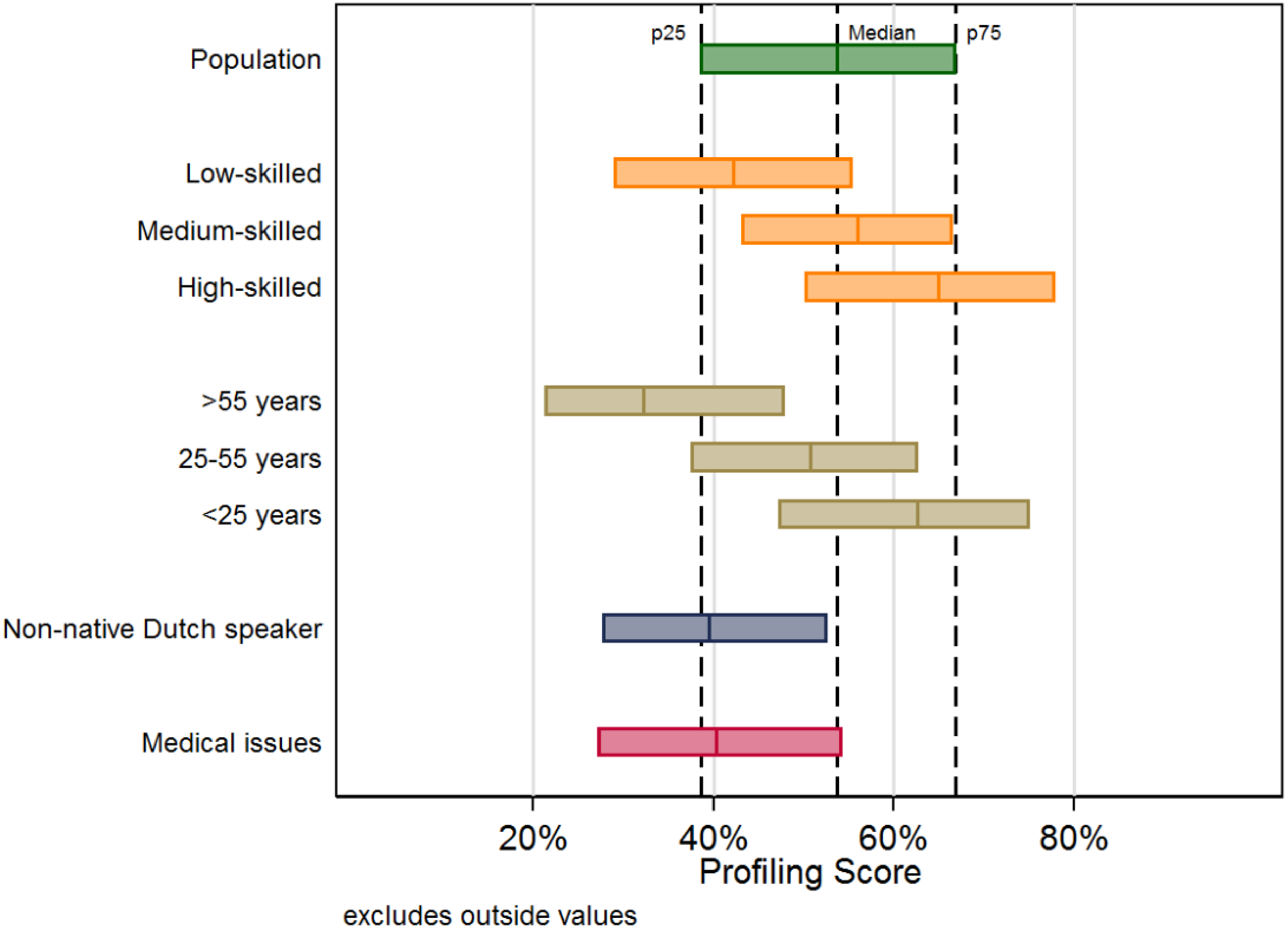
- Does AI-based profiling improve early identification of jobseekers at-risk of becoming long-term unemployed ?
  - Compared to more classical ways of classifying jobseekers
  - While not increasing discrimination
- We explore accuracy and fairness as a trade-off
  - Improving accuracy comes at the cost of discrimination
  - This is inherent to any form of profiling
- We obtained access to the output of an innovative AI-model of VDAB
  - Gradually replacing the existing rule-based model
  - Version from January 2018

Data

- Current and previous unemployment spells
- Work experience
- Record data: languages, preferred jobs and regions, studies
- Client information (age)
- Activity in personal platform (MijnLoopbaan)
  - updating cv, updating preferred jobs, competencies

- Note that sensitive information is embedded in big data
- In our research: gender, origin, nationality
- Work within 6 months, of at least 28 days
- All new jobseekers registered at VDAB in the course of 2016 N= 288 765

Belonging to a certain target group does not tell the whole story



Methodology

- We compare three profiling approaches which each classify 33.8% of the jobseekers as high-risk jobseekers
  1) **randomly** classifying jobseekers as high-risk
  2) classifying all **low-skilled** jobseekers as high-risk
  3) classifying jobseekers with a **profiling score** lower than 45% as high-risk
- We compare accuracy and fairness of AI-based profiling versus rule-based profiling
- We play with the threshold used in the AI-based profiling model to distinguish low from high-risk jobseekers
- By combining accuracy and fairness in a single graph, we derive the accuracy-equity trade-off

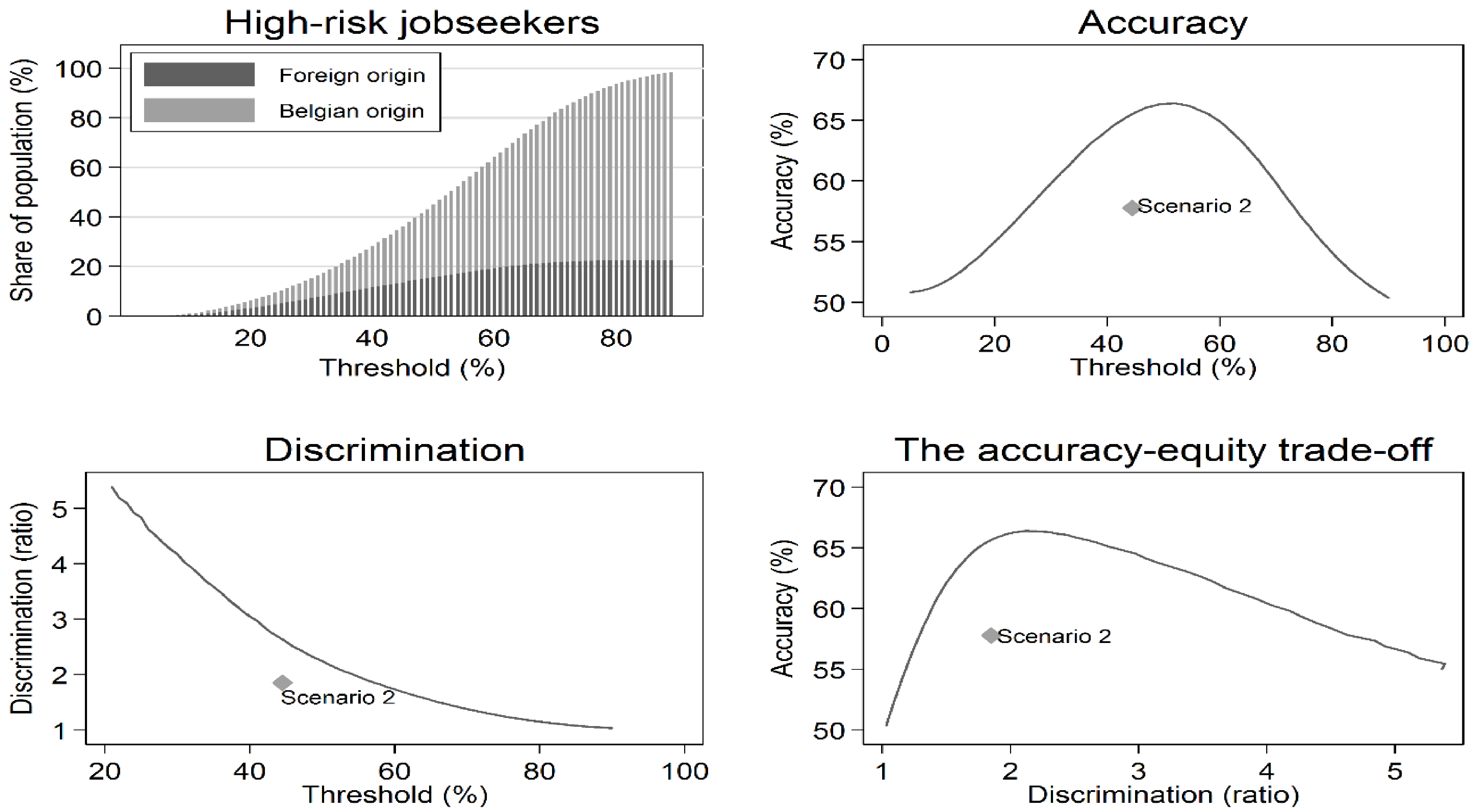
Accuracy and equity

- **Accuracy** is defined as the share of jobseekers that are correctly identified as high or low-risk jobseekers
  - We obtained data on the labour market trajectory of each jobseeker so that we could compare the predicted outcome to the real outcome
- **Fairness** is defined as follows: if jobseekers of disadvantaged groups that find a job ex-post are more likely to be misclassified as high-risk jobseekers ex ante, relative to this proportion among the dominant group (discrimination as a ratio)
  - A model is fair if the false positive rate is equal across groups = predictive equality
  - Independent from existing inequalities in the historical data

	Selection-rule 1 (randomly labelling jobseekers as high-risk)	Selection-rule 2 (labelling all low-skilled jobseekers as high-risk)	AI-based profiling (labelling all jobseekers with a profiling score lower that 45% as high-risk)
Share of jobseekers labelled as high-risk jobseeker	33.8%	33.8%	33.8%
Accuracy (share of jobseekers correctly identified as low or high-risk)			
All jobseekers	50.2%	58,0%	66,0%
Belgian origin	51.5%	59.4%	65.4%
Foreign origin	45.8%	51.5%	66.0%
Discrimination (found a job ex-post, misclassified as high-risk ex-ante)			
Belgian origin	34.0%	23.2%	14.8%
Foreign origin	34.1%	42.9%	38.9%
ratio foreign /Belgian origin	1.00	1.85	2.63

The trade-offs of an AI profiling model

- Parameters are set so that exactly the same proportion of jobseekers is labelled as high-risk
- Depending on its resources and objectives, the PES could also set another threshold to distinguish between low and high-risk jobseekers
- The threshold determines
    1) the share of jobseekers labelled as high-risk
    2) the accuracy of the profiling model
    3) its fairness
    4) the accuracy-equity trade-off
- We developed a graphical tool to visualize this trade-off


High-risk jobseekers
Foreign origin
Belgian origin
Share of population (%)
Threshold (%)
Accuracy
Accuracy (%)
Scenario 2
Threshold (%)
Discrimination
Discrimination (ratio)
Scenario 2
Threshold (%)
The accuracy-equity trade-off
Accuracy (%)
Scenario 2
Discrimination (ratio)

Conclusion

- The maximum level of accuracy is 66%
- At this level, jobseekers of foreign origin are 2.6 times more likely to be misclassified than jobseekers of Belgian origin (39% versus 15%)
- Switching to AI-based profiling could increase accuracy without increasing discrimination (by reducing the threshold), or could reduce discrimination while keeping the same level of accuracy (by increasing the threshold)
- For low values of the threshold, few jobseekers are considered high-risk, but those that are considered high-risk are predominantly of foreign origin. This explains the low accuracy and high level of discrimination for low values of the threshold

Discussion

- Discrimination matters
  - Depends on how the model is used: to automate decision-making or support caseworkers
  - Depends on the value of services offered: supporting versus monitoring jobseekers
- Role of caseworkers
  - AI models will only support caseworkers if they trust the models
  - Crucial to the operation of the PES is how to integrate these models into decision-making processes

THANK YOU

Further Reading

- Desiere, S., & Struyven, L. (2020). Using Artificial Intelligence to classify Jobseekers: The Accuracy-Equity Trade-off. *Journal of Social Policy,* 1-19. doi:10.1017/S0047279420000203
- Desiere, S., Langenbucher, K. & Struyven, L. (2019). "Statistical profiling in public employment services: An international comparison", *OECD Social, Employment and Migration Working Papers*, No. 224, OECD Publishing, Paris. http://dx.doi.org/10.1787/b5e5f16e-en
- Van Landeghem, B., Desiere, S. & Struyven, L. (forthcoming). Statistical Profiling of Unemployed Jobseekers. IZA World of Labor