Beyond Traditional Risk Scores

Tackling LS/CMI Offender Misclassifications with Machine Learning

William Arbour, Sébastien Brouillette-Alarie, Guy Giguère, Guy Lacroix, and Steeve Marchand

Steeve Marchand

Melbourne Institute: Applied Economic and Social Research

University of Melbourne

Applied Research in Crime and Justice Conference
Brisbane, July 1, 2024

Opinions expressed in this presentation are those of the authors only and do not necessarily reflect those of any institution.

Risk Assessment Tools

Various types of risk assessment tools used to evaluate offenders'

- Risk of committing new crimes
- Criminogenic needs
- Challenges related to specific factors (e.g. criminal history, alcohol/drug)

The evaluation may impact:

- Treatment referrals, interventions
- Early release, parole decisions
- Pre-sentencing reports

Risk Assessment Tools

One of the most common tools: **LSI-R**, **LS/RNR**, or **LS/CMI**

- Administered by a trained professional
- A lengthy evaluation

Outcomes of the evaluation:

- Referrals to treatments and interventions
- **Simplified score** that determines a **risk category** (e.g. low risk, high risk)

Impacts later outcomes (e.g. parole)

LS/CMI Risk Score and Category

43 items, leading to a score between 0 and 43

Thresholds that determine one of the five risk categories

- Very low (0-4)
- Low (5-10)
- Medium (11-19)
- High (20-29)
- Very high (30-43)

Very predictive of recidivism… but this oversimplification could misclassify some individuals

Simplifications Leading to LS/CMI Category

1) Simple aggregation of the 43 items (unweighted)

2) Some sub-items are ignored, e.g.:

   Item 7 : Disciplinary sanctions (yes/no) → One point if yes
   Subitem 7.1 : Number of sanctions → Ignored

3) Many items are rated on an ordinal scale (e.g.: 0, 1, 2, 3) then binarized (0, 1)

4) Demographic characteristics are available but ignored

This Paper - Research Questions

1) How much does the **simplification** of evaluations into risk categories lead to **misclassification** of offenders' risk

2) **Who** tends to be most misclassified by this procedure?

3) Can simple adjustments to the procedure avoid misclassification

Data

1) All **LS/CMI assessments** conducted in Québec (Canada) provincial prisons between 2008 and 2015
   - Sentences between 6 months and 2 years
   - N = 45,535
   - 93% of men; average age of 35 years
   - Observe all answers on the form

2) All **sentences** (provincial and federal) in Quebec between 2008 and 2020
   - Compute recidivism based on new sentences (reconviction)

Linked at the individual level

Methodology - Overview

1. Use **probabilistic classifier machine learning (ML) algorithm**
   Uses all information available in the evaluations to obtain individual probabilities of recidivism

2. Comparison with **simple predictions using LS/CMI risk categories**
   Allows to measure misclassification resulting from simplified procedure

3. Explore how **simpler adjustments** to the score can reduce misclassification
   - Precise score
   - Weighted score
   - Weighted score + age

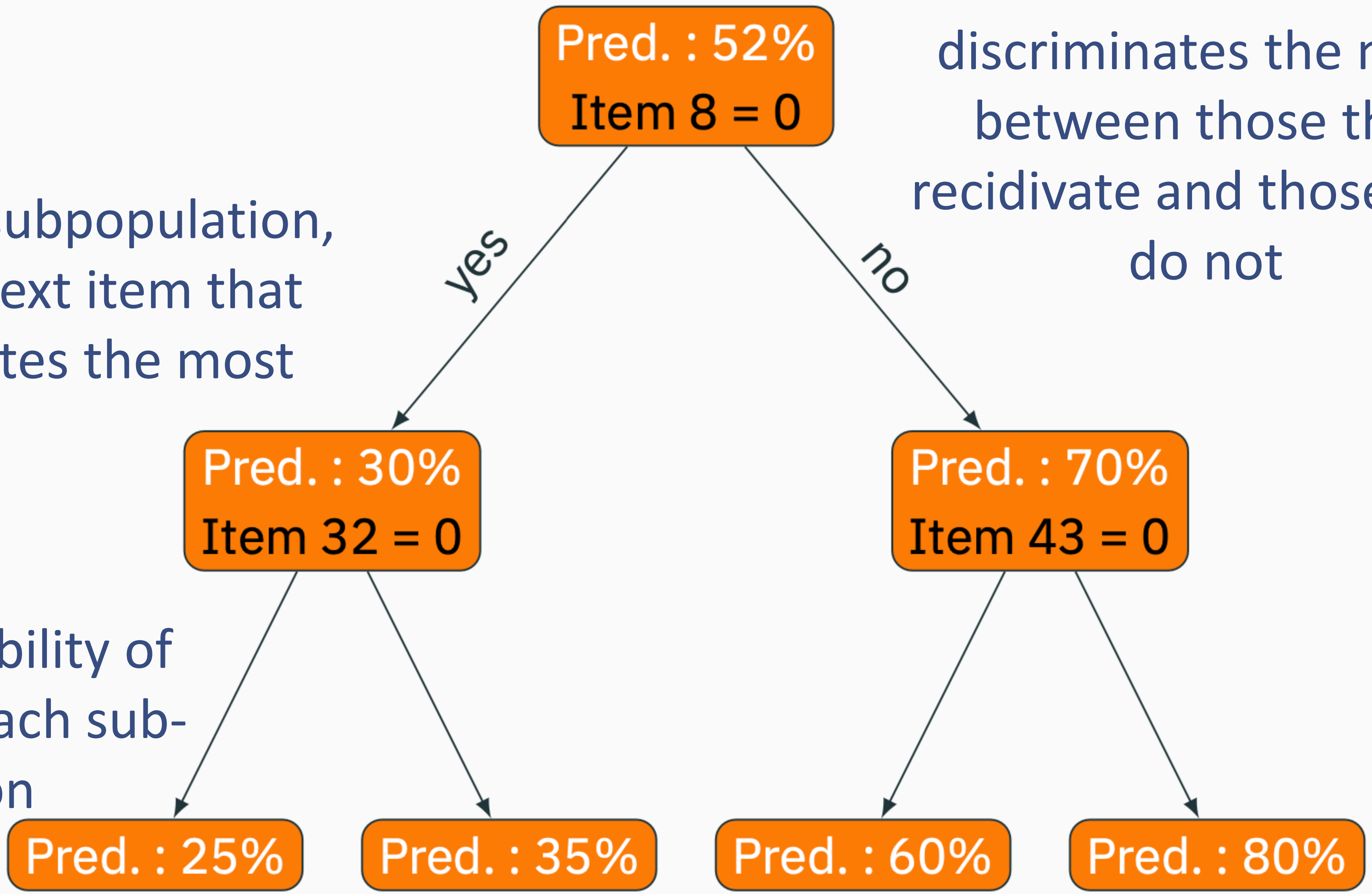
4. Construct **new risk categories** that avoid misclassification using our ML predictions

Probabilistic Classifier

Finds the item that discriminates the most between those that recidivate and those that do not

Within the subpopulation, finds the next item that discriminates the most

Obtain a probability of recidivism for each sub-population



Advantages of the Approach

Flexible and non-parametric procedure

- Interaction between items
- Implicit weighting of items

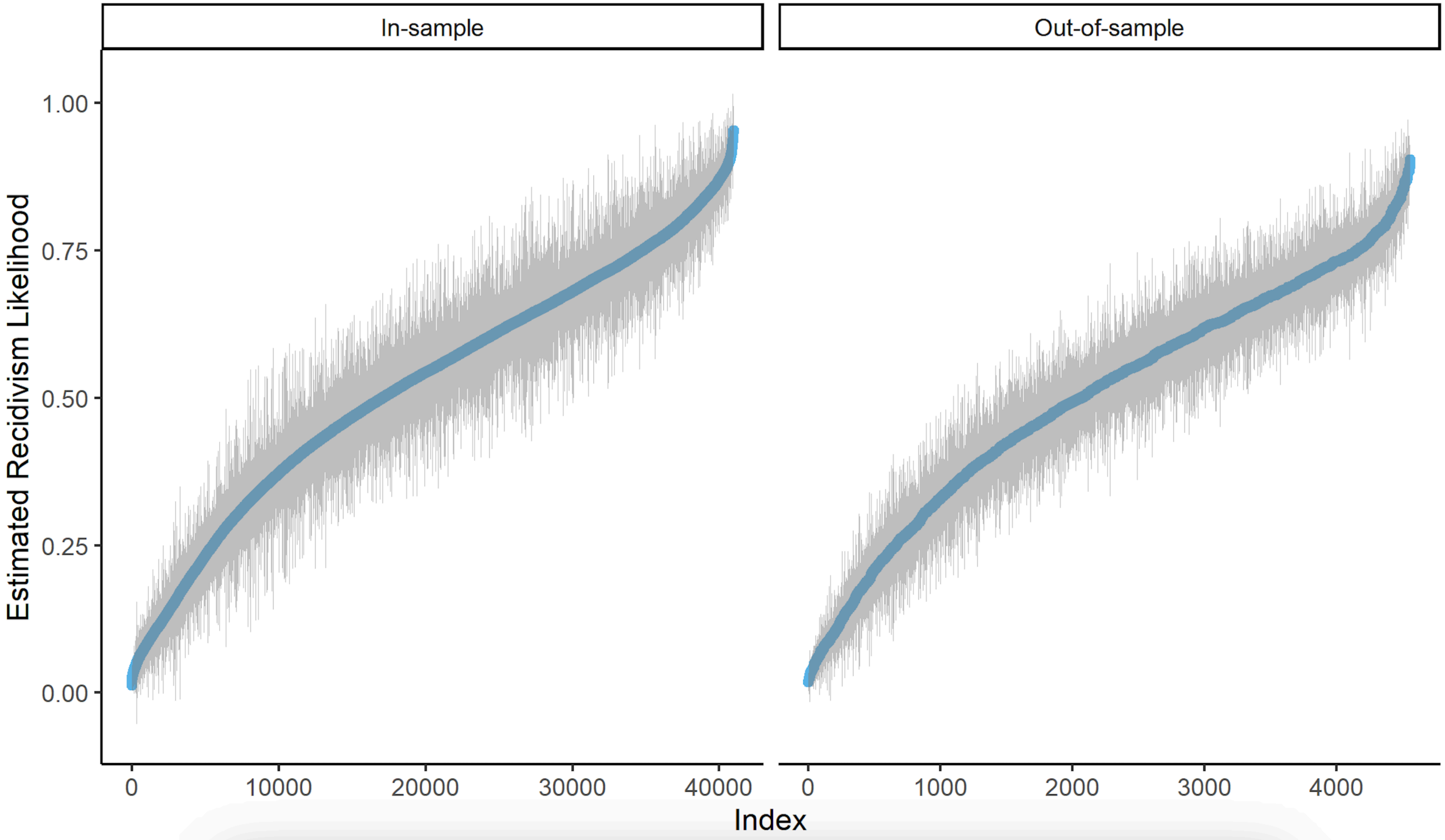
Avoids overfitting

- Estimated on training sample
- Tested on testing sample

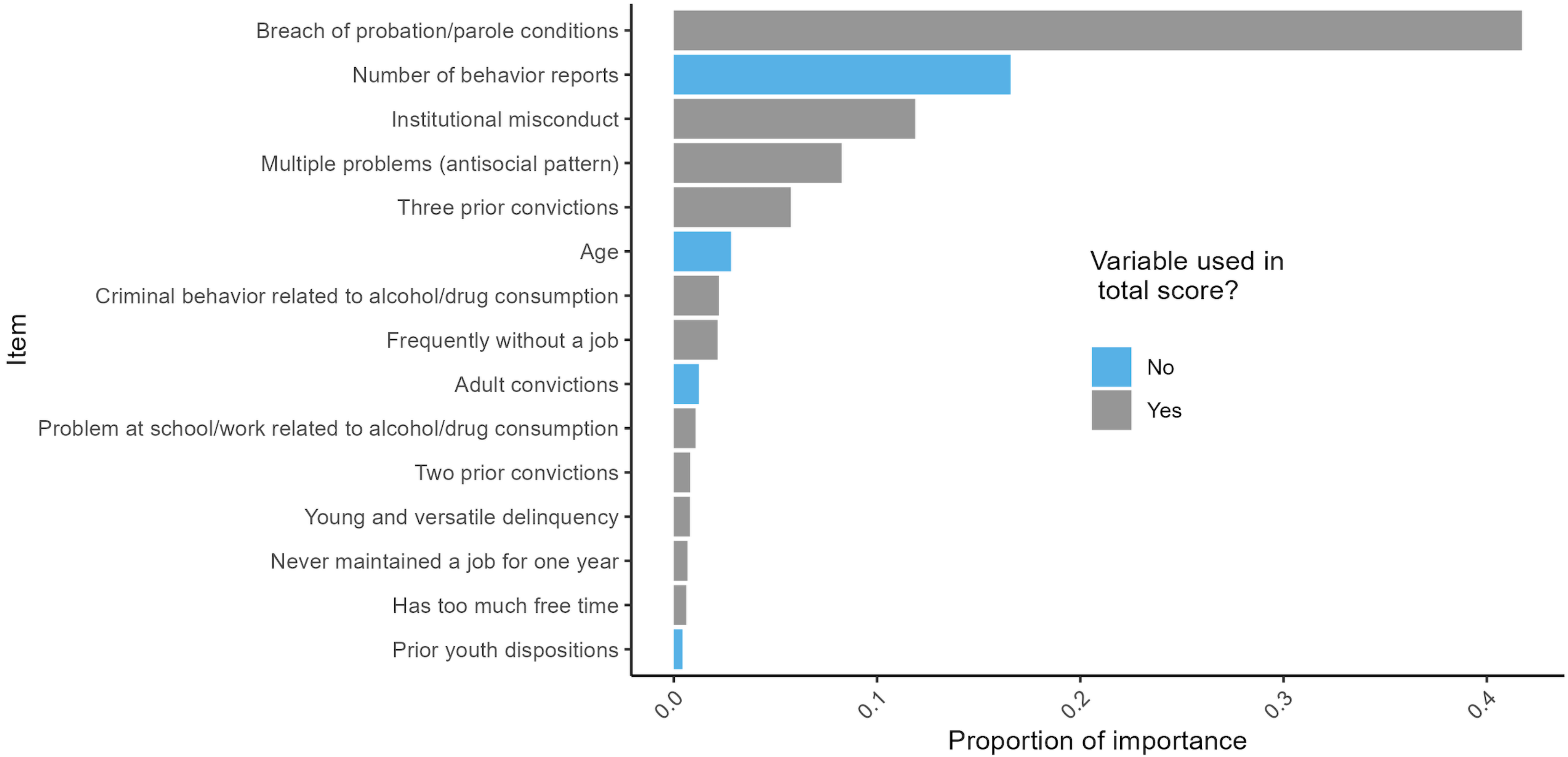
Provides individual recidivism probabilities

- More precise than simple classifiers (i.e. yes/no)
- Useful for categorising risk levels

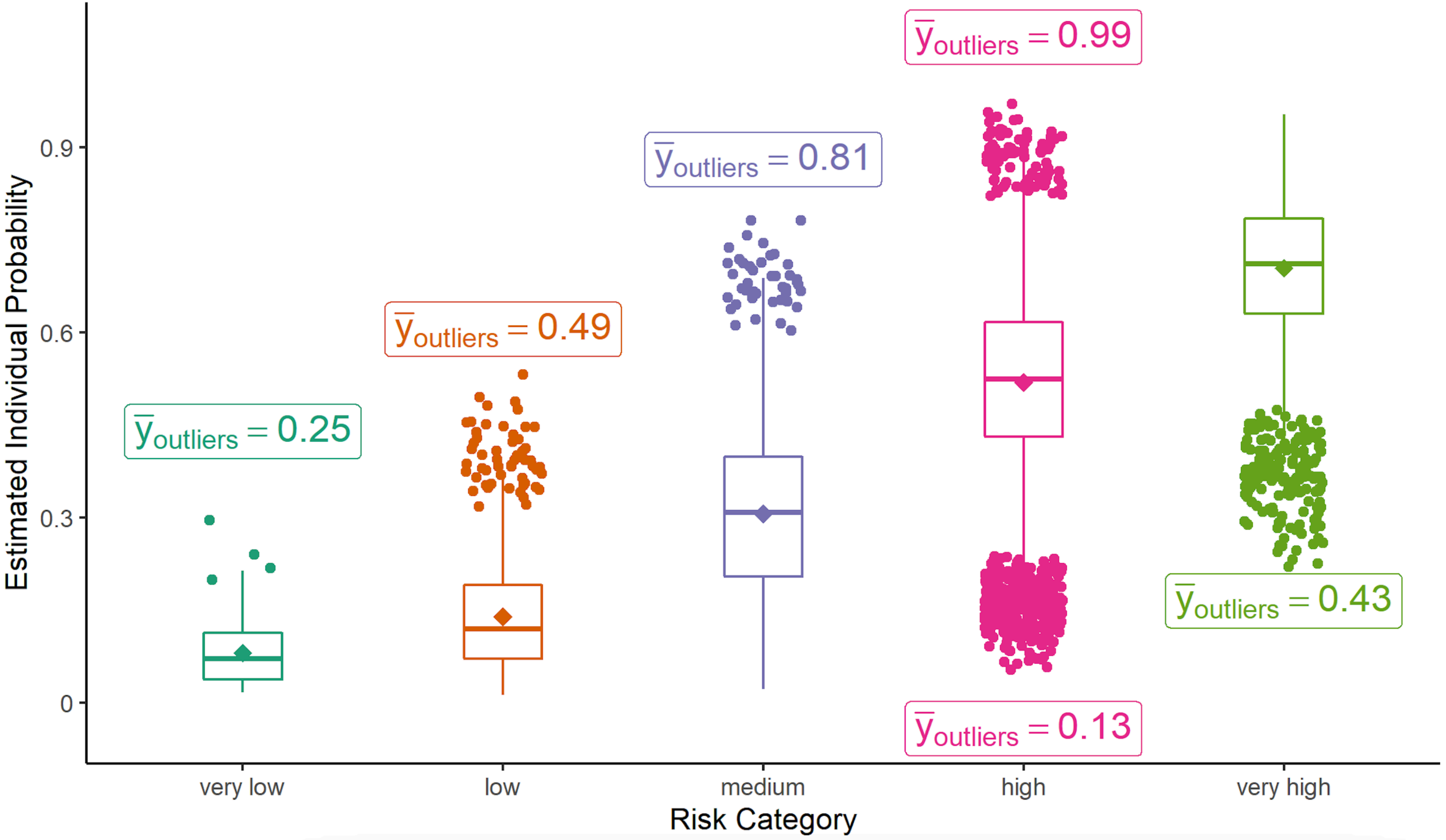
Recidivism Probabilities



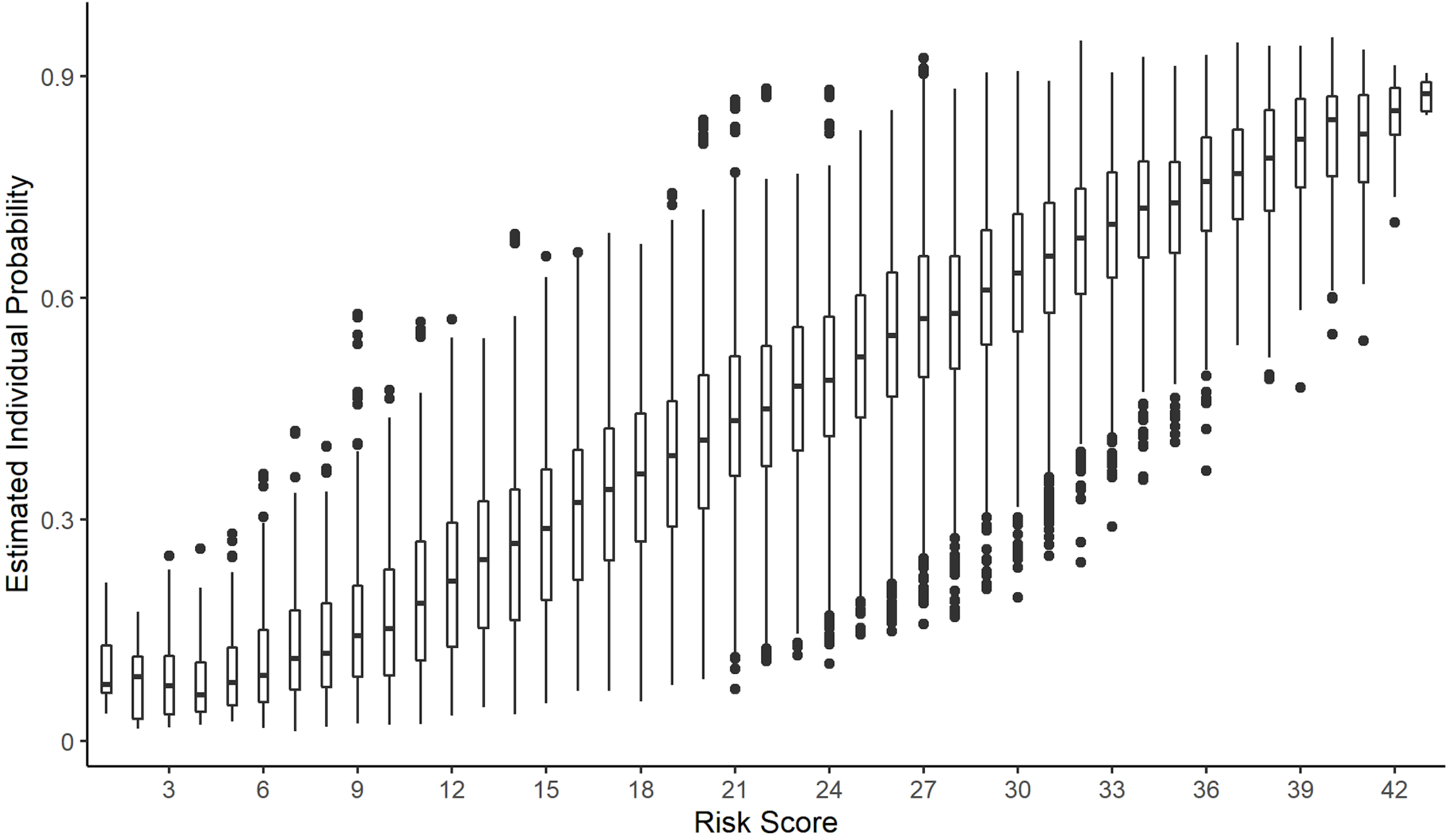
Variable Importance



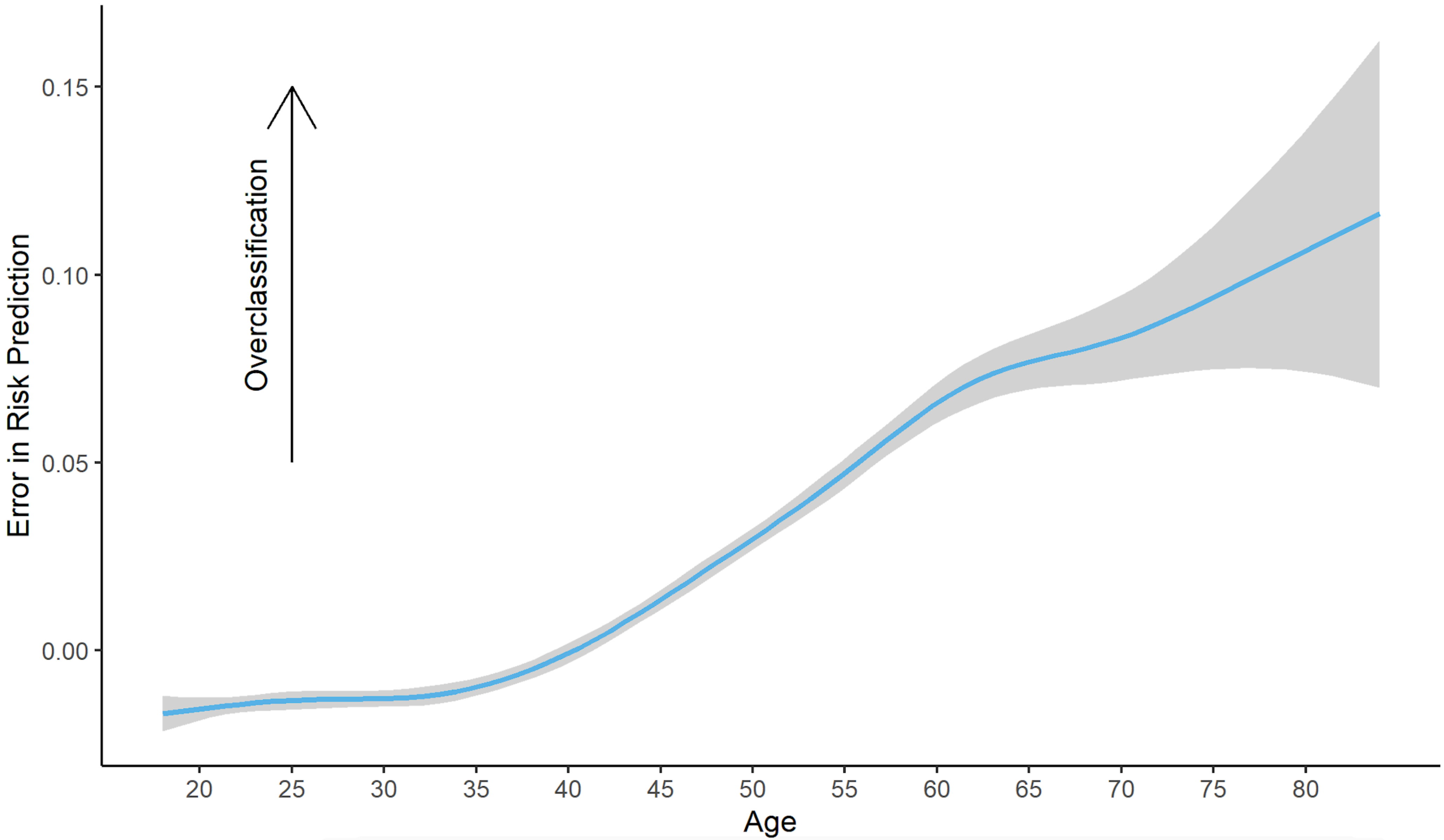
Misclassification in LS/CMI Categories



Predicted Probability by Risk Score



Classification Errors by Age



Can Simple Adjustments Avoid Misclassification?

We predict probabilities based on

1. The precise risk score (from 0 to 43)

2. A score weighted with variable importance measures

$$WeightedScore = \lfloor 43 \times \sum_{j=1}^{43} w_j \mathbb{1}\{item_j = 1\} \rfloor$$

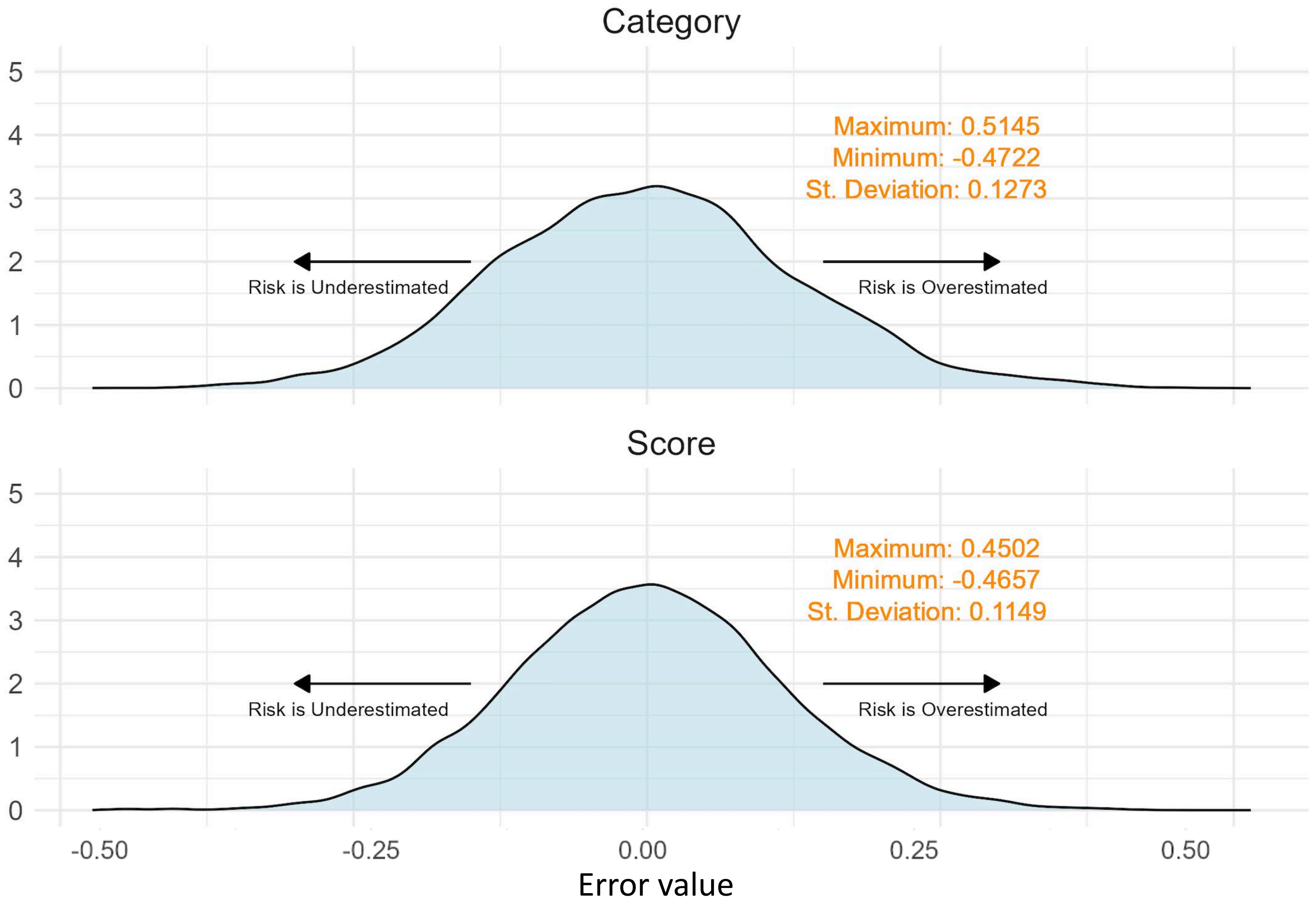
3. A logit model including the weighted score and age

$$recidivism_i = \beta_0 + \sum_{j=1}^{43} \beta_j \mathbb{1}\{WeightedScore_i = j\} + \lambda age_i + \epsilon_i$$

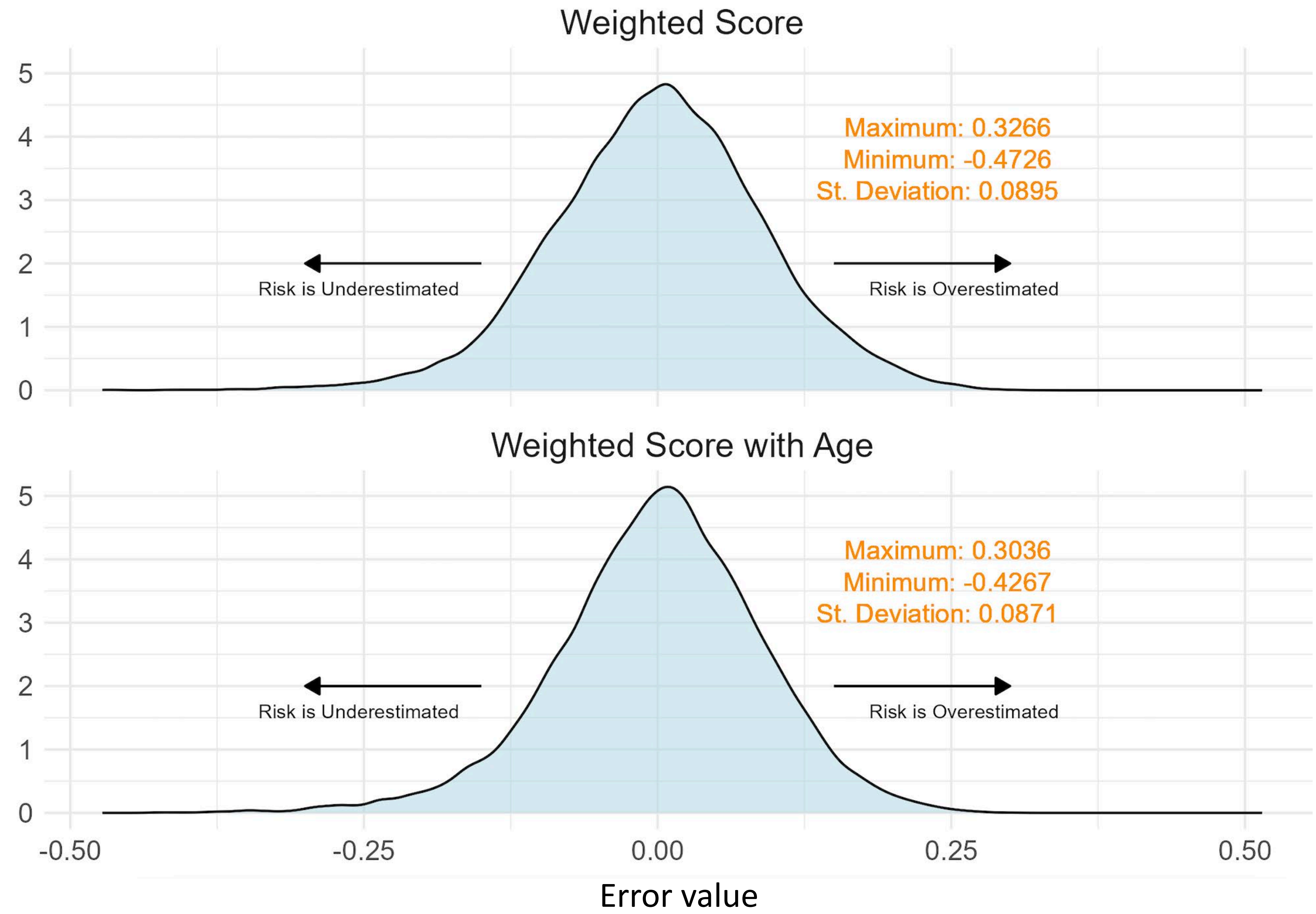
Can Simple Adjustments Avoid Misclassification?

	Correlation		AUC	
	in-sample (OOB)	*out-of-sample*	*in-sample (OOB)*	*out-of-sample*
ML	.4685	.4546	.7684 (± .0045)	.7611 (± .0138)
Risk Category	.3734	.3883	.7006 (± .0047)	.7115 (± .0140)
Risk Score	.3952	.4043	.7245 (± .0049)	.7318 (± .0144)
Weighted Score	.4105	.4078	.7318 (± .0048)	.7319 (± .0145)
Weighted Score + Age	.4200	.4099	.7380 (± .0048)	.7324 (± .0145)

Adjustments and Classification Errors



Adjustments and Classification Errors



Defining New Risk Categories

ML leads to better predictions, but individual probabilities may not be convenient in practice

We define new risk categories based on our ML probabilities

- k-mean clustering algorithm
- 5 categories (but flexible)

	Correlation		AUC	
	in-sample (OOB)	*out-of-sample*	*in-sample (OOB)*	*out-of-sample*
ML	.4685	.4546	.7684 (± .0045)	.7611 (± .0138)
Risk Category	.3734	.3883	.7006 (± .0047)	.7115 (± .0140)
Risk Score	.3952	.4043	.7245 (± .0049)	.7318 (± .0144)
Weighted Score	.4105	.4078	.7318 (± .0048)	.7319 (± .0145)
Weighted Score + Age	.4200	.4099	.7380 (± .0048)	.7324 (± .0145)
Cluster	.4557	.4437	.7555 (± .0045)	.7482 (± .0137)

Discussion (1/2)

Simplifying the evaluation to a risk category leads to misclassification

- Broad risk category
- Unequal importance of items
- Useful information ignored
- No interaction between items

Older individuals are most at risk of being misclassified

Simple adjustments reduce only partially misclassification errors

Possible to use ML to construct categories that reduce misclassification

Discussion (2/2)

Our findings do not suggest these risk assessments are wrong

They contain rich information

But the information could be better exploited

Thank you

steeve.marchand@unimelb.edu.au