

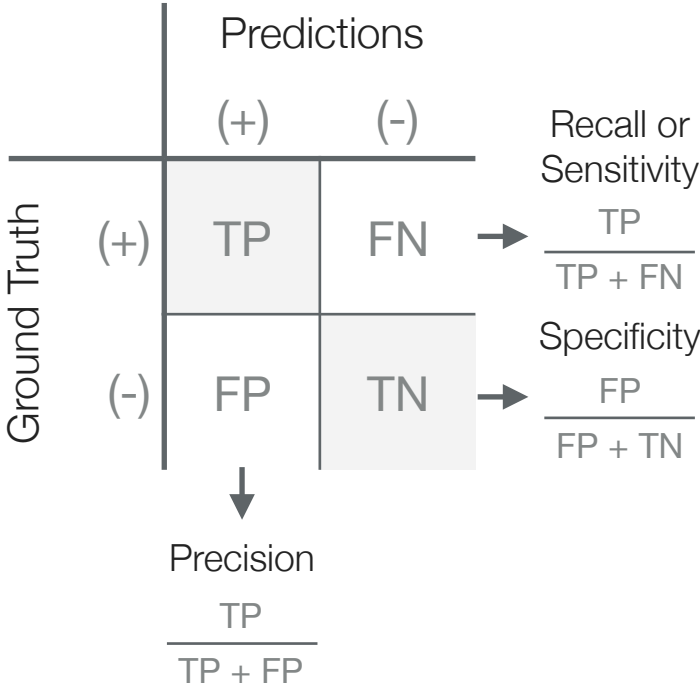
Ground Truth and
Prediction Probability

ID
1
2
3
4
5
6
7
8
9
10

Original Labels	
Ground Truth	Prediction Probability
(+)	0.99
(-)	0.70
(+)	0.38
(+)	0.33
(+)	0.26
(-)	0.16
(-)	0.15
(-)	0.14
(-)	0.12
(-)	0.07

Inverted Labels	
Ground Truth	Prediction Probability
(-)	0.01
(+)	0.30
(-)	0.62
(-)	0.67
(-)	0.74
(+)	0.84
(+)	0.85
(+)	0.86
(+)	0.88
(+)	0.93

Confusion Matrixes



Original Labels @0.50

		Predictions	
		(+)	(-)
Ground Truth	(+)	1	3
	(-)	1	5

Accuracy@0.50 = 0.600

Precision@0.50 = 0.500

Recall@0.50 = 0.250

MCC@0.50 = 0.100

PR AUC = 0.767

ROC AUC = 0.875

Inverted Labels @0.50

		Predictions	
		(+)	(-)
Ground Truth	(+)	5	1
	(-)	3	1

Accuracy@0.50 = 0.600

Precision@0.50 = **0.625**

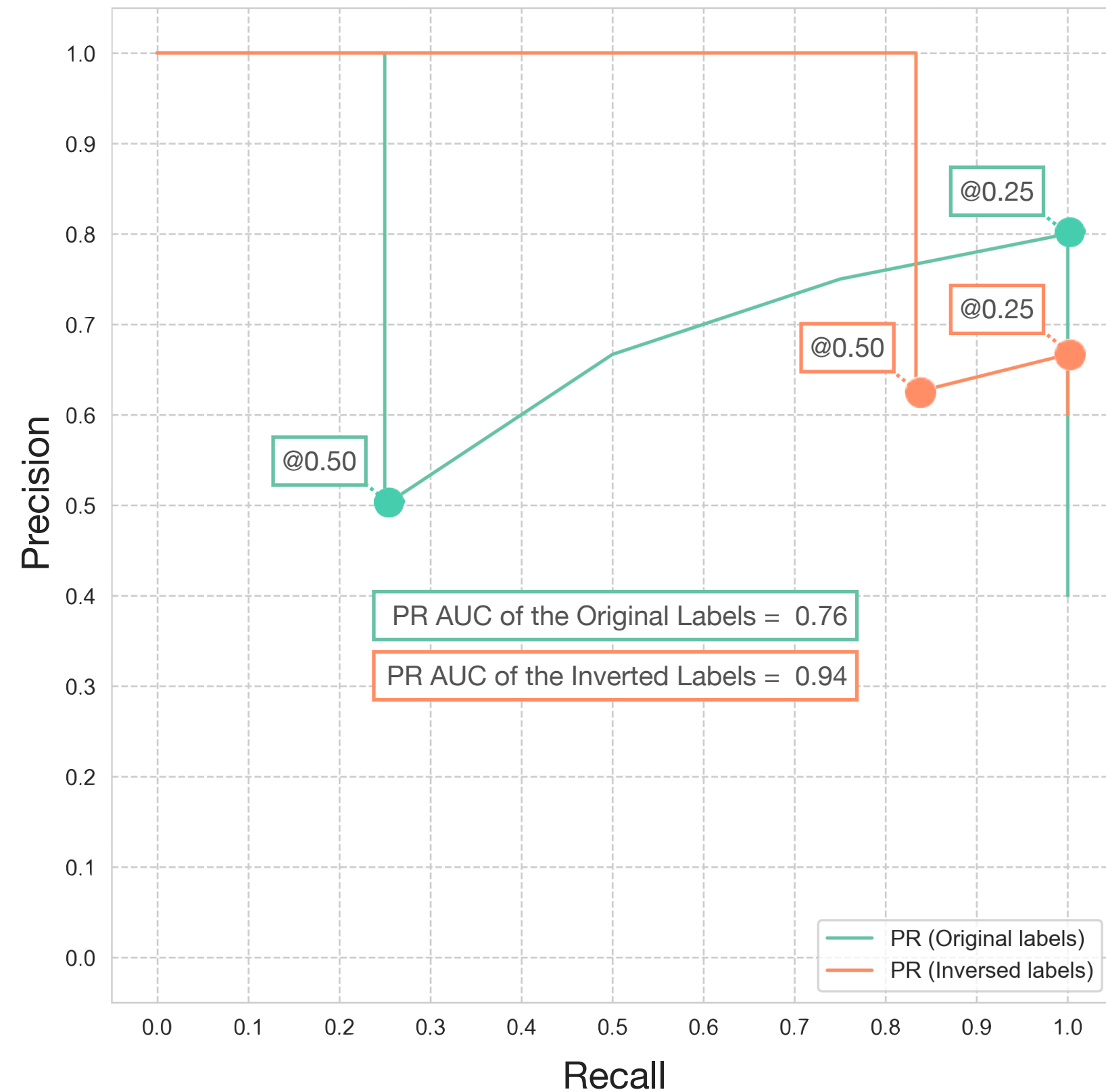
Recall@0.50 = **0.833**

MCC@0.50 = 0.100

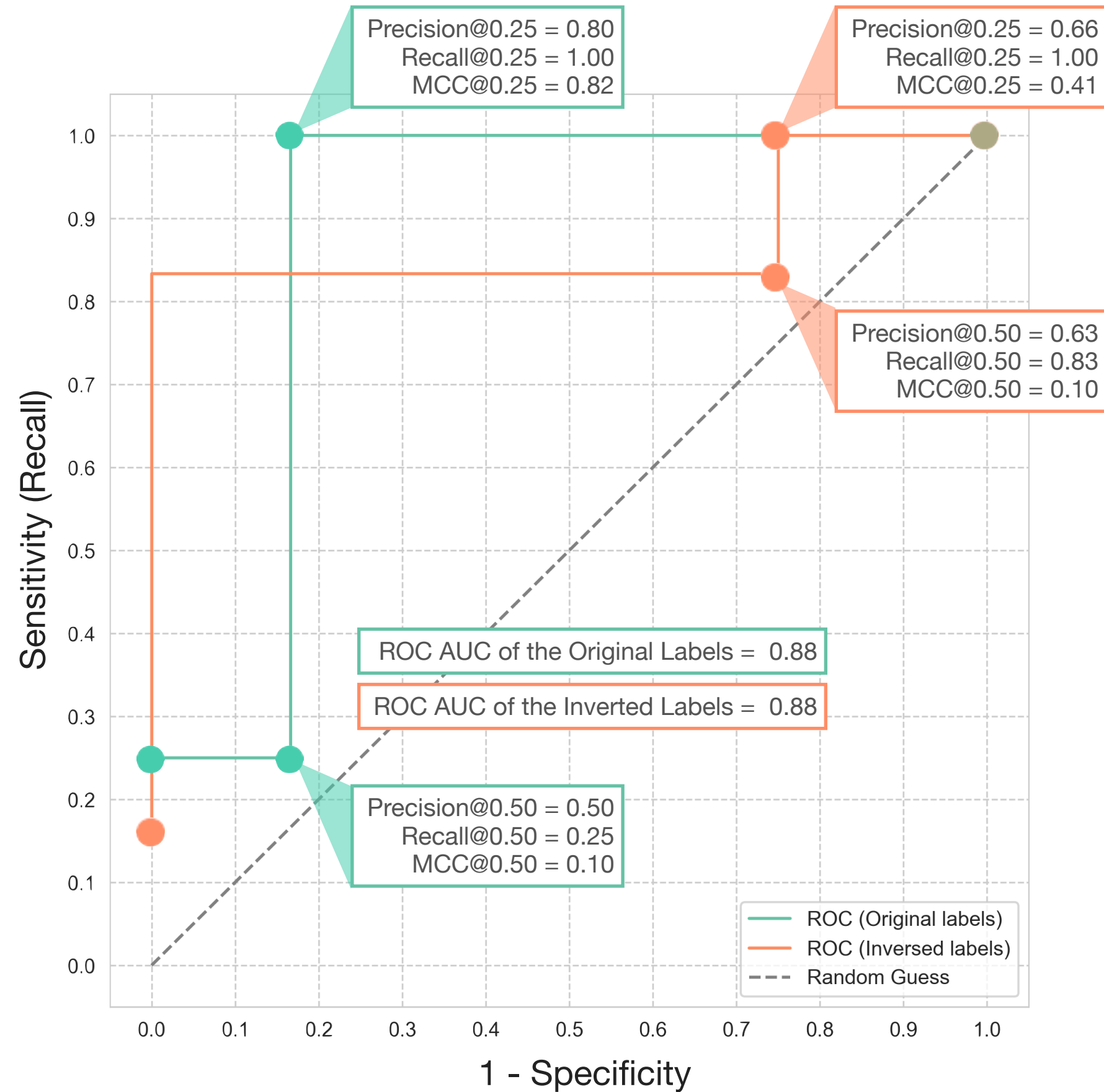
PR AUC = **0.941**

ROC AUC = 0.875

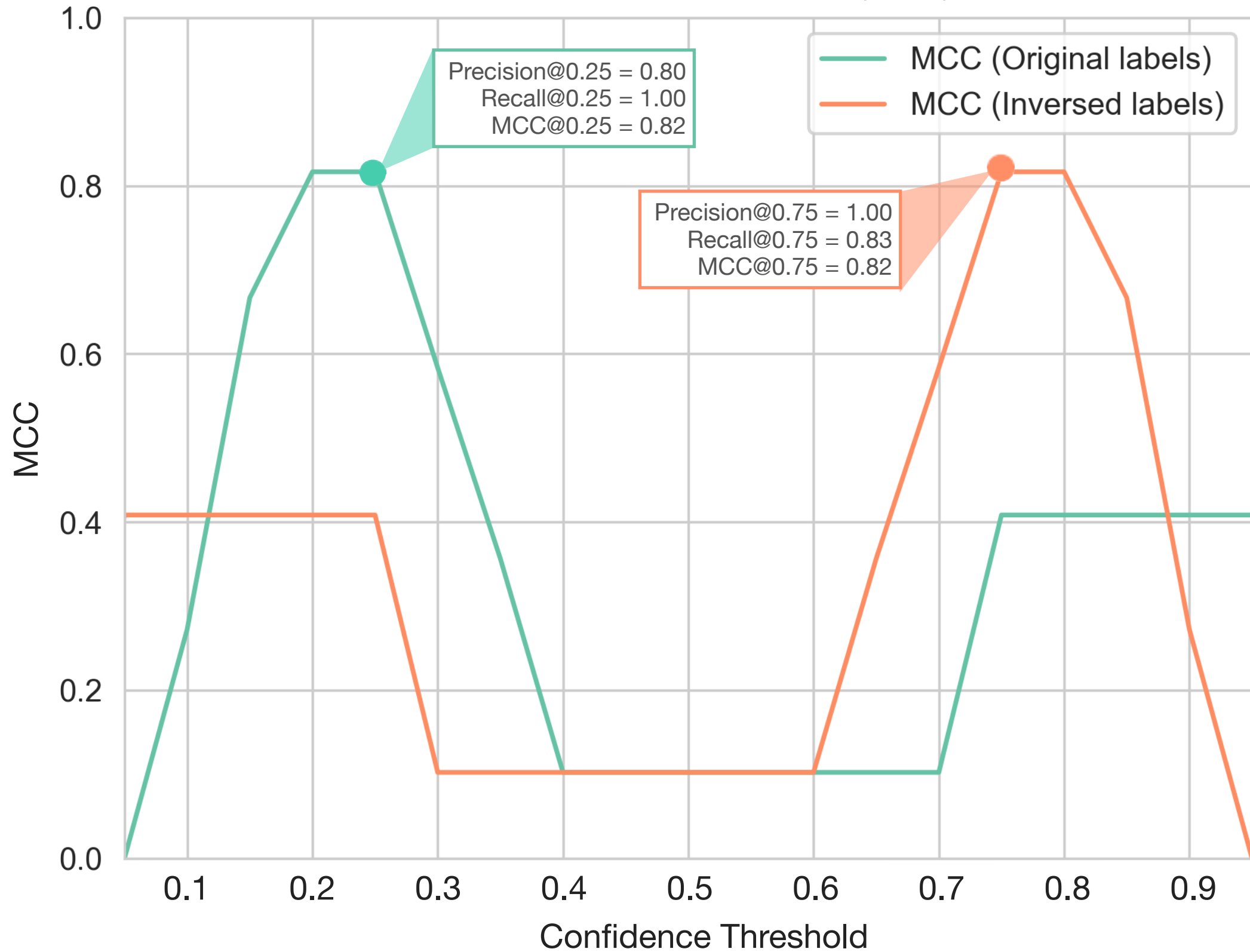
Precision-Recall (PR) Curve



Receiver Operating Characteristics (ROC) Curve



Matthews Correlation Coefficient (MCC) Curve



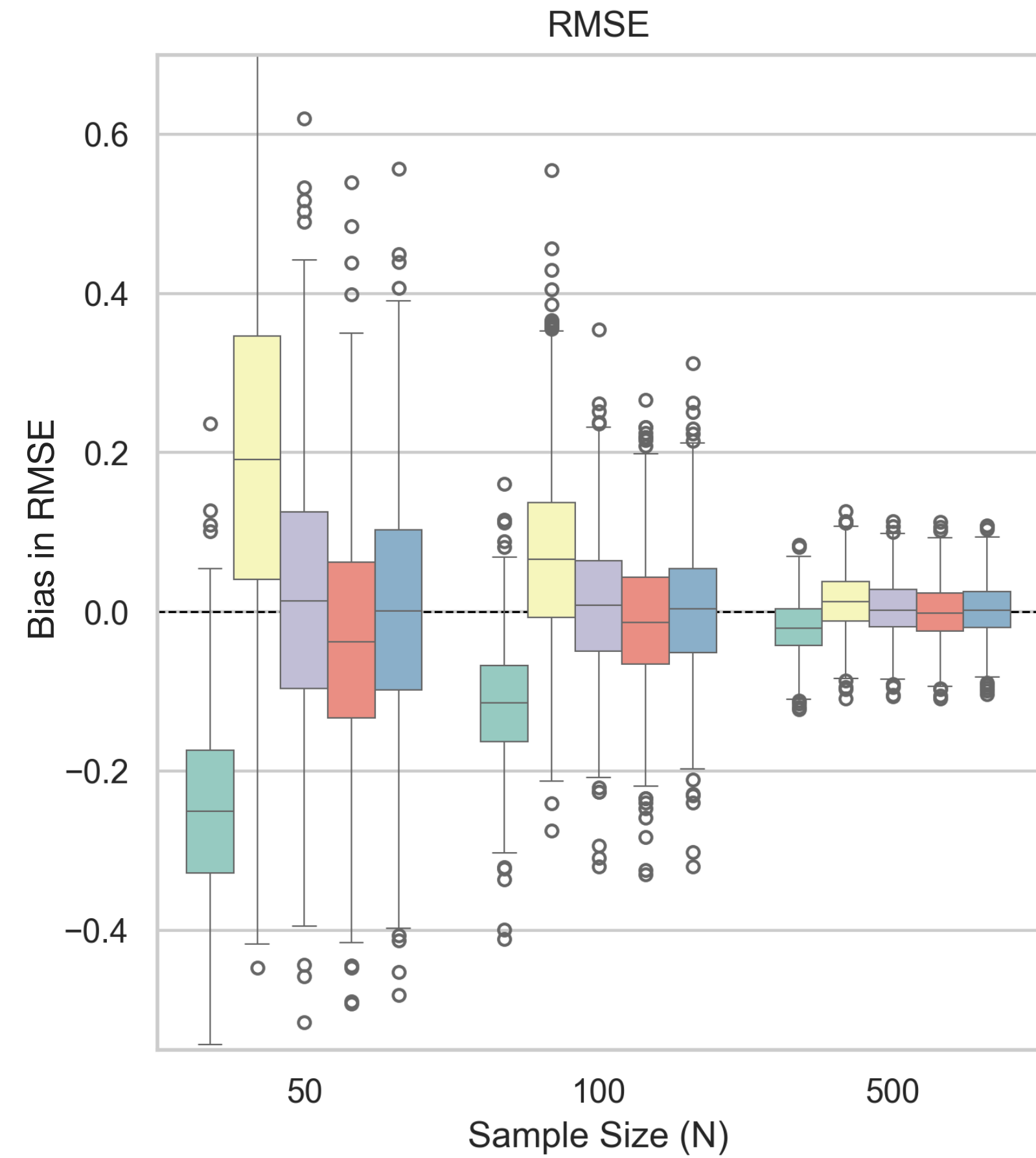
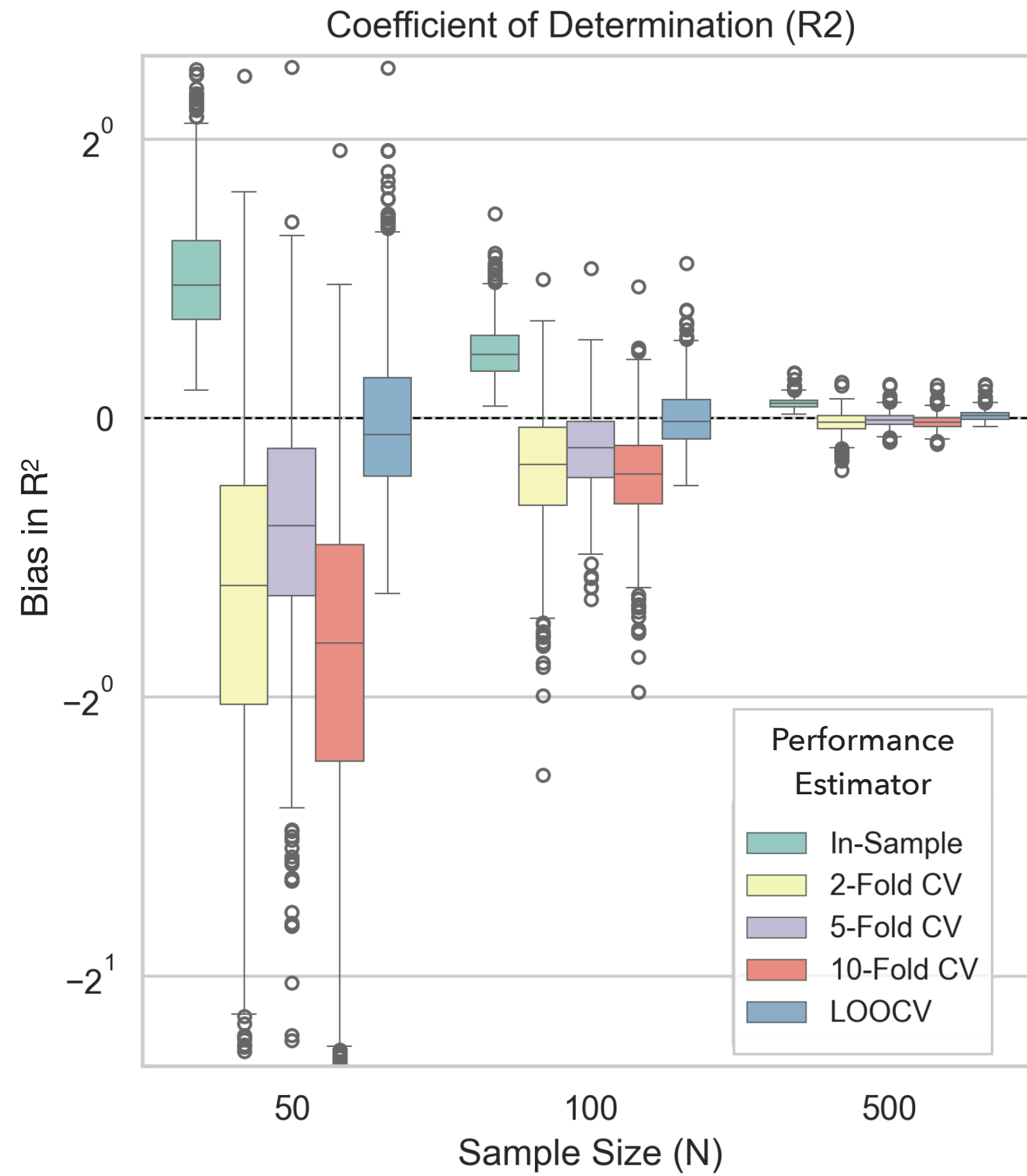
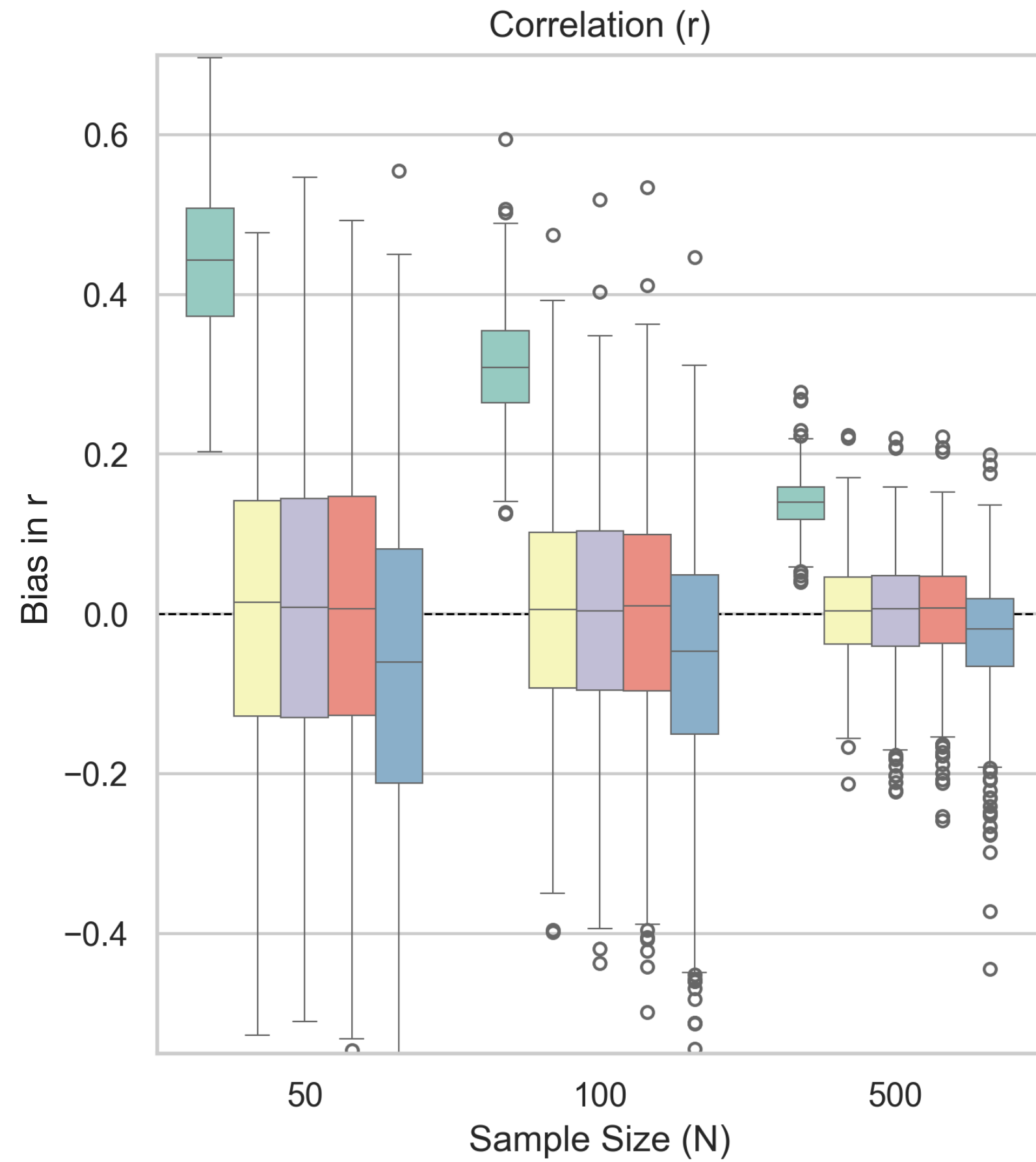
Confusion Matrix at the Optimum Threshold

Original Labels @0.25

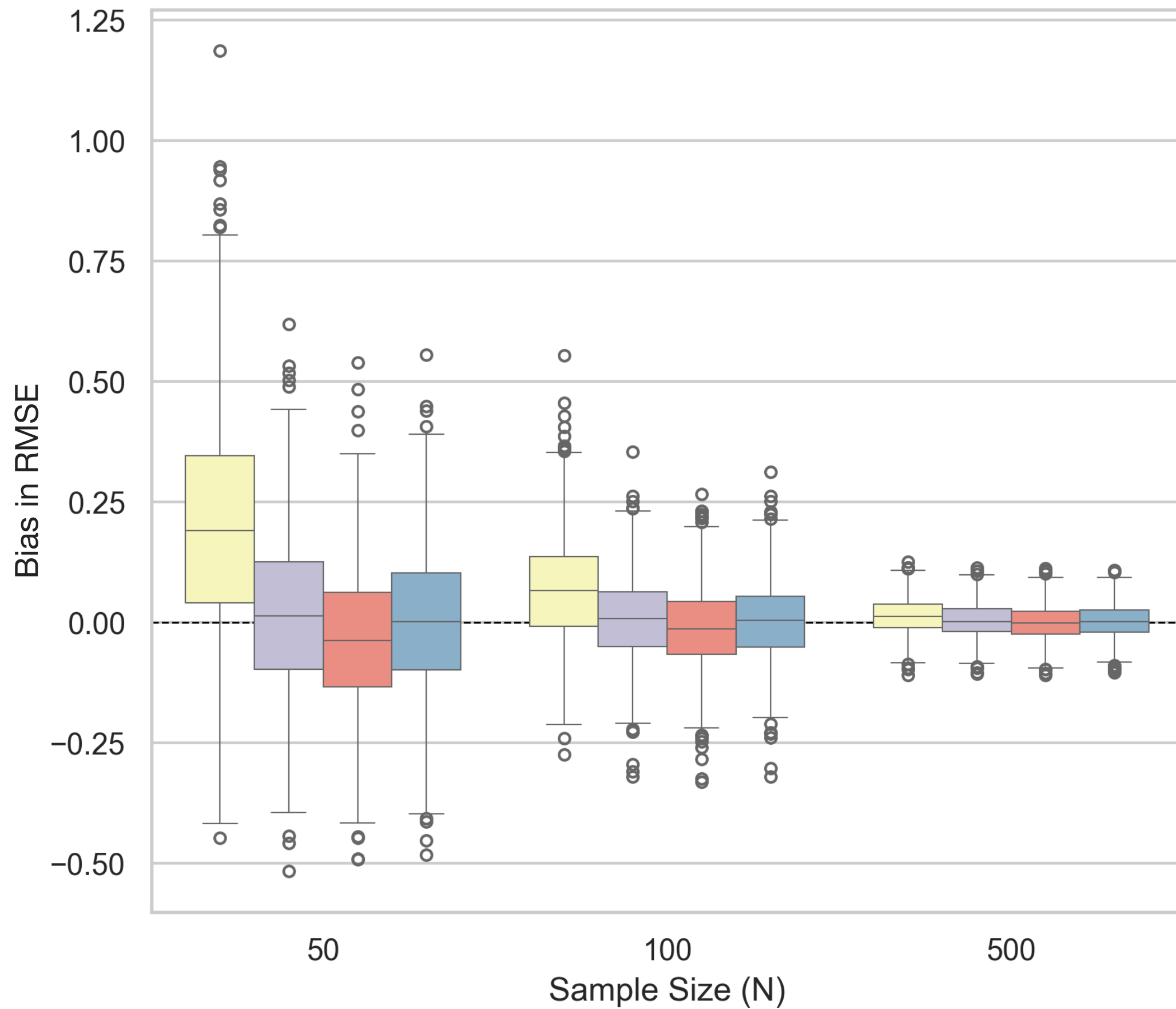
		Predictions	
		(+)	(-)
Ground Truth	(+)	4	0
	(-)	1	5

Inverted Labels @0.75

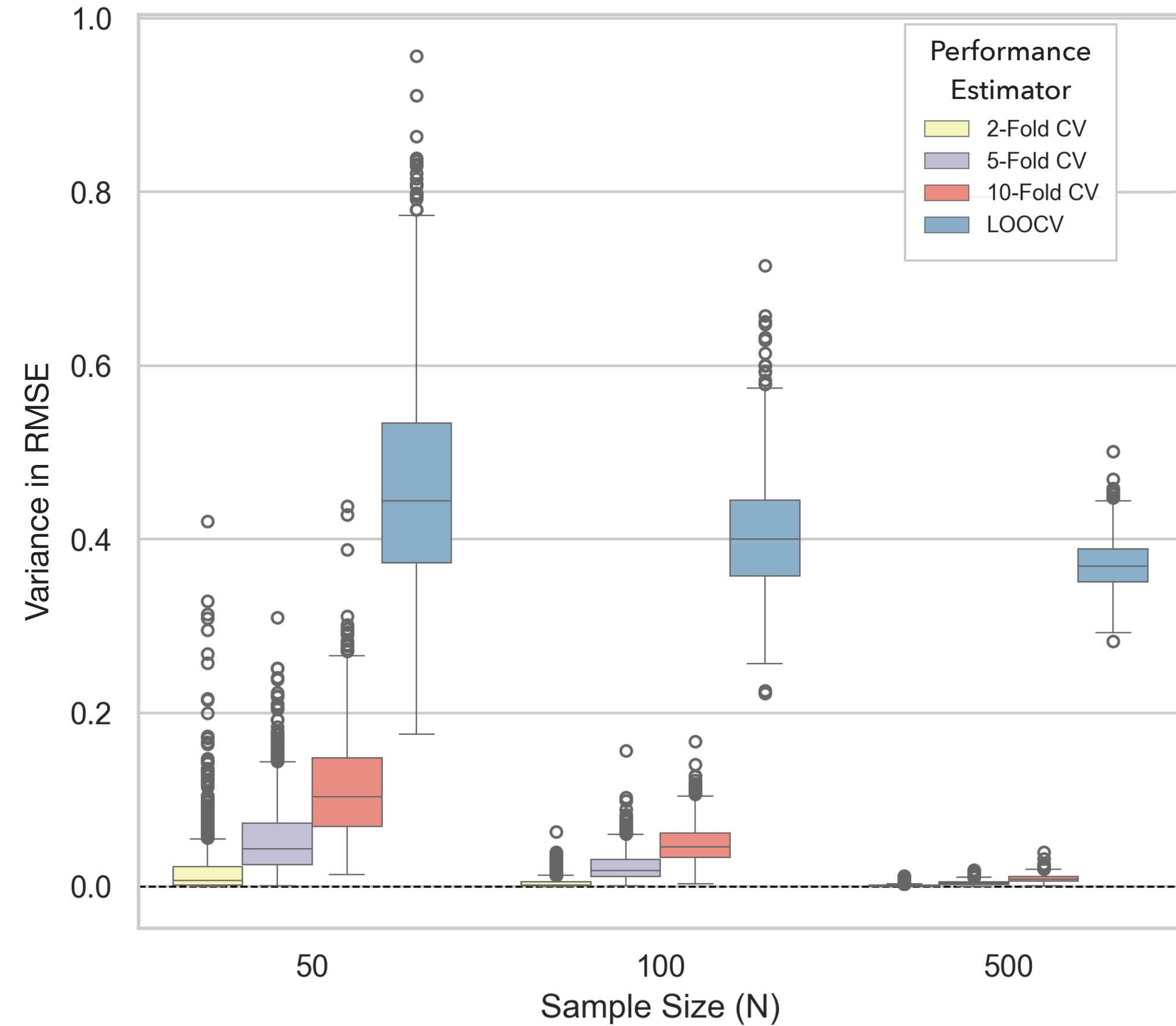
		Predictions	
		(+)	(-)
Ground Truth	(+)	5	1
	(-)	0	4



Bias



Variance



Feature Selection (FS)

Hyperparameter Tuning (HT)

Metric Report

FS=0; HT=0

Entire Dataset
(100%)

X'

Inner K-Fold CV

Training
(80%)

Test
(20%)

$\mathbb{E}[g(f_{\mathcal{D}})]$

Report the best performance
from all possible
hyperparameter combinations

FS=1; HT=0

Inner K-Fold CV

Training
(80%)

X'

Training
(80%)

Test
(20%)

$\mathbb{E}[g(f_{\mathcal{D}})]$

Report the best performance
from all possible
hyperparameter combinations

FS=0; HT=1

Entire Dataset
(100%)

X'

Inner K-Fold CV

Training
(64%)

Validation
(16%)

$f_{\mathcal{D}}$

Test
(20%)

Estimated from the
unseen dataset

FS=1; HT=1

Inner K-Fold CV

Training
(64%)

X'

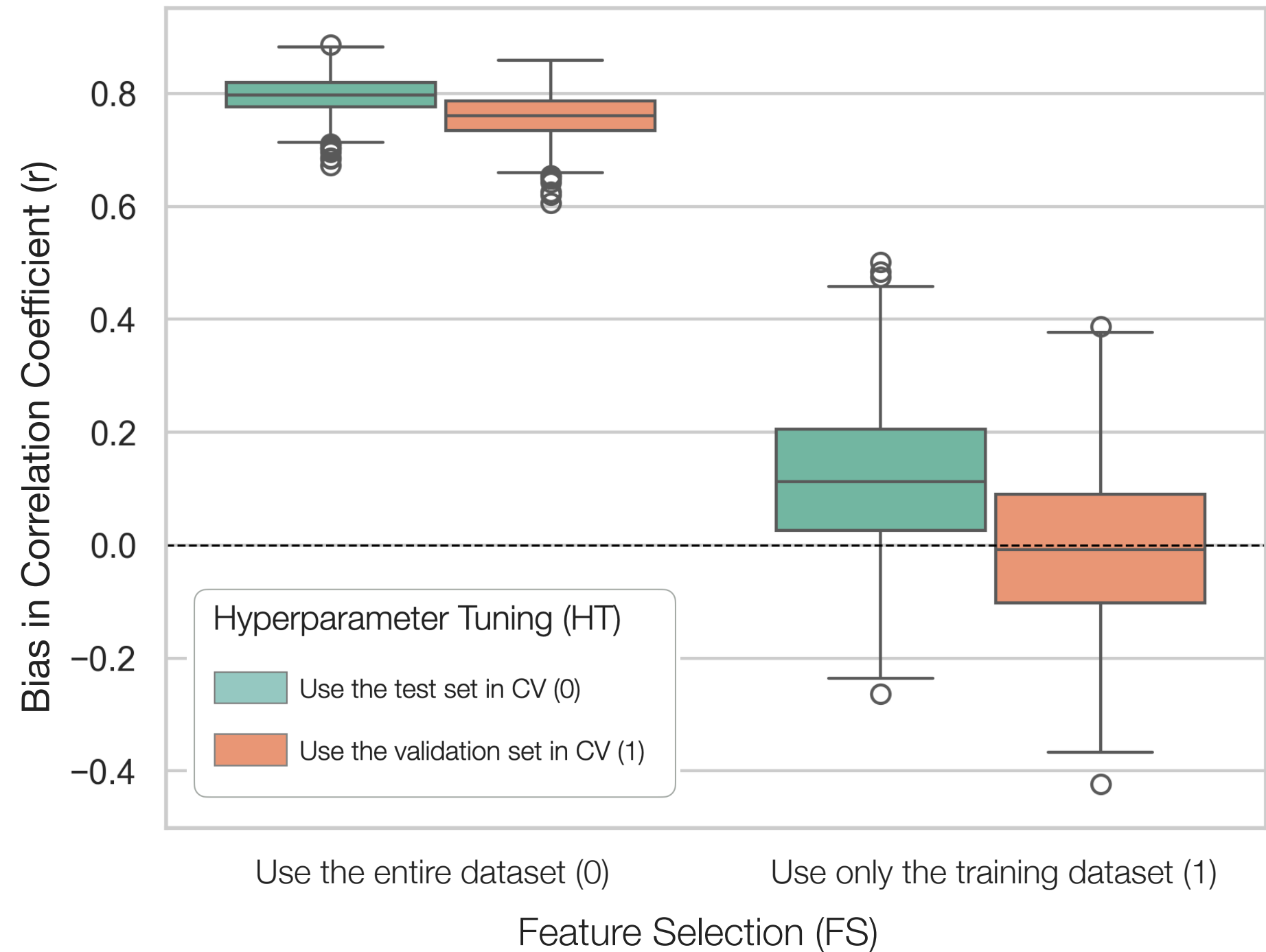
Training
(64%)

Validation
(16%)

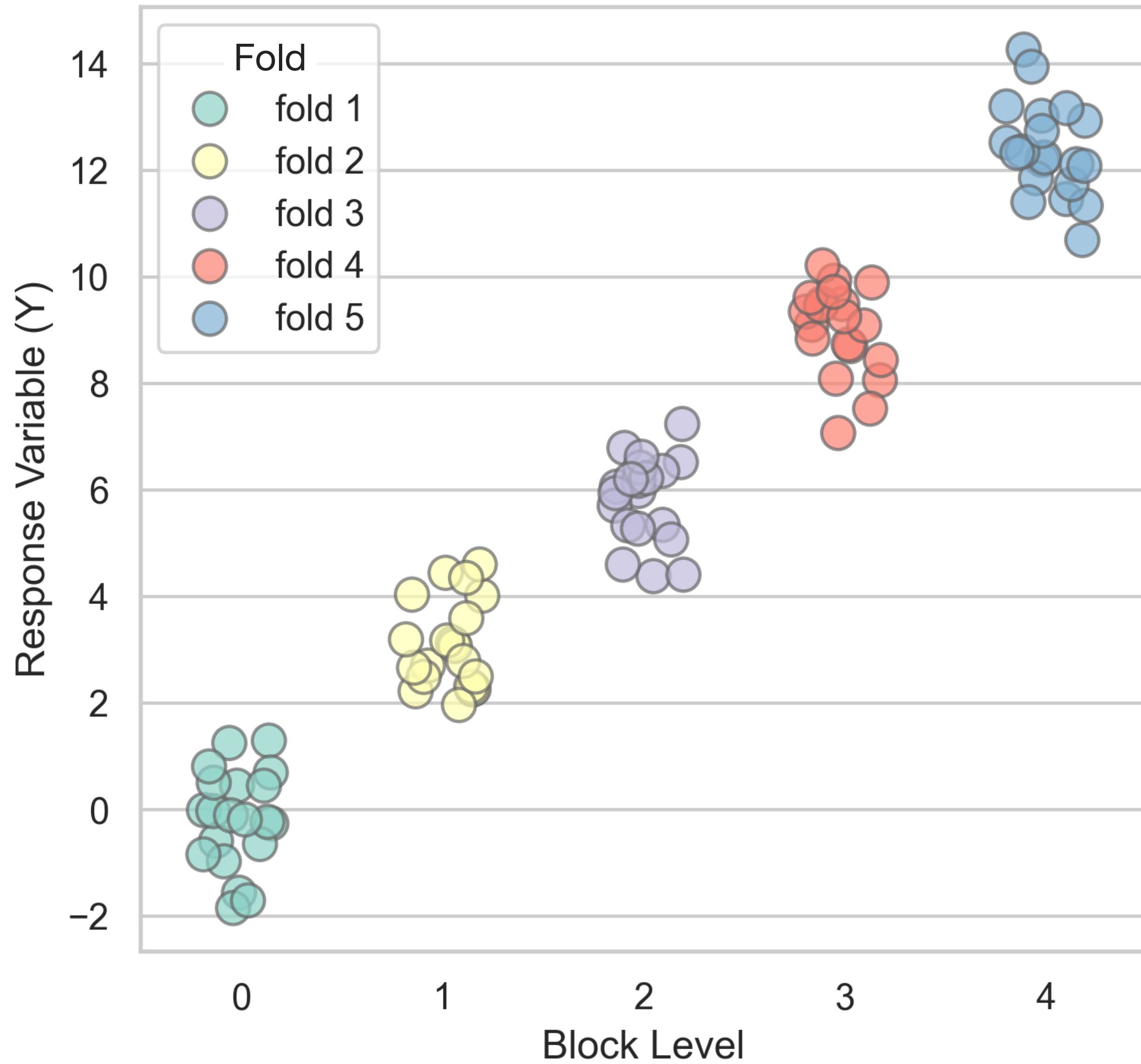
$f_{\mathcal{D}}$

Test
(20%)

Estimated from the
unseen dataset



Block Cross Validation



Random Cross Validation

