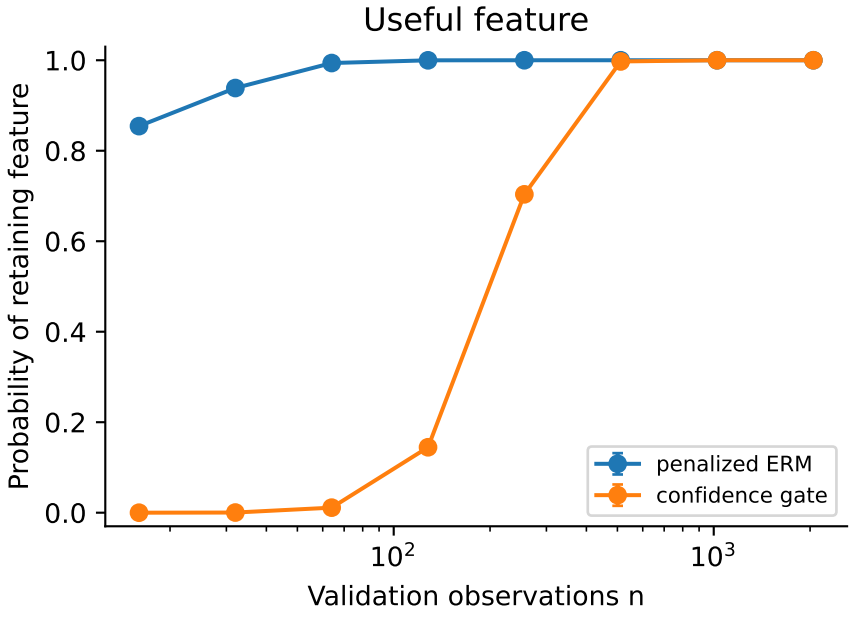
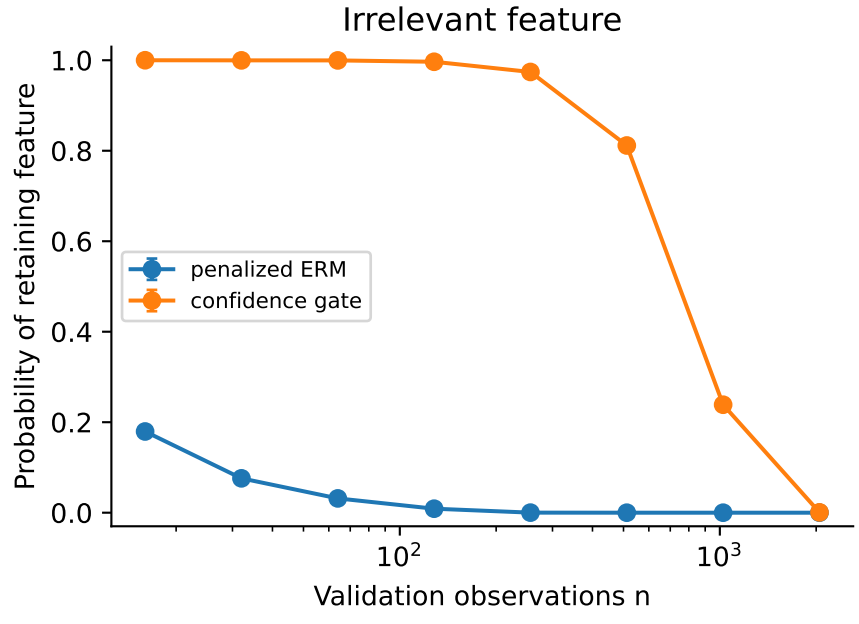

Useful feature
Probability of retaining feature
1.0
0.8
0.6
0.4
0.2
0.0
10^2
10^3
Validation observations n
penalized ERM
confidence gate


Irrelevant feature
Probability of retaining feature
1.0
0.8
0.6
0.4
0.2
0.0
10^2
10^3
Validation observations n
penalized ERM
confidence gate