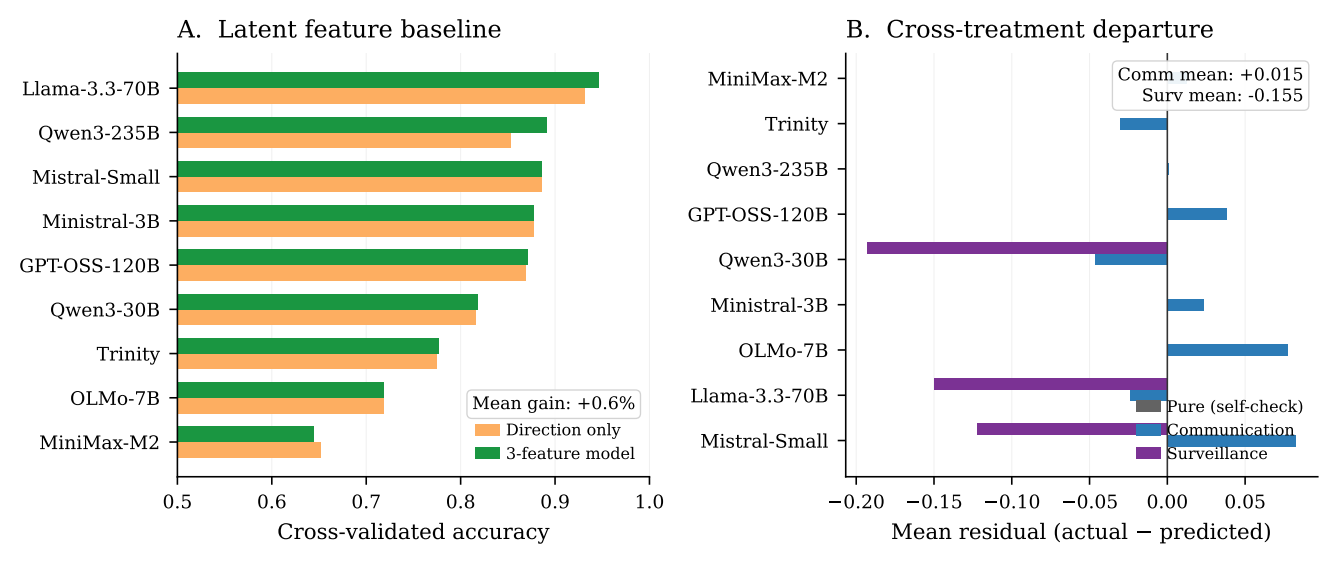
A. Latent feature baseline
Llama-3.3-70B
Qwen3-235B
Mistral-Small
Ministral-3B
GPT-OSS-120B
Qwen3-30B
Trinity
OLMo-7B
MiniMax-M2
Mean gain: +0.6%
Direction only
3-feature model
0.5
0.6
0.7
0.8
0.9
1.0
Cross-validated accuracy
B. Cross-treatment departure
MiniMax-M2
Trinity
Qwen3-235B
GPT-OSS-120B
Qwen3-30B
Ministral-3B
OLMo-7B
Llama-3.3-70B
Mistral-Small
Comm mean: +0.015
Surv mean: -0.155
Pure (self-check)
Communication
Surveillance
−0.20
−0.15
−0.10
−0.05
0.00
0.05
Mean residual (actual − predicted)