Comparison of ROC Gap and AUC across models, $q_y = 0.5, k = 3, q_t = 1$

↓ Bias mitigation

↑ Discriminative performance

Legend:
- y-model (oracle)
- y-obs model
- Group 0 only
- Group 1 only
- Tested-only
- SELF
- DivideMixBasedModel
- Generalized JS
- Truncated LQ
- DragonNet
- Peer loss
- Group peer loss
- SAREM
- DCEM (ours)