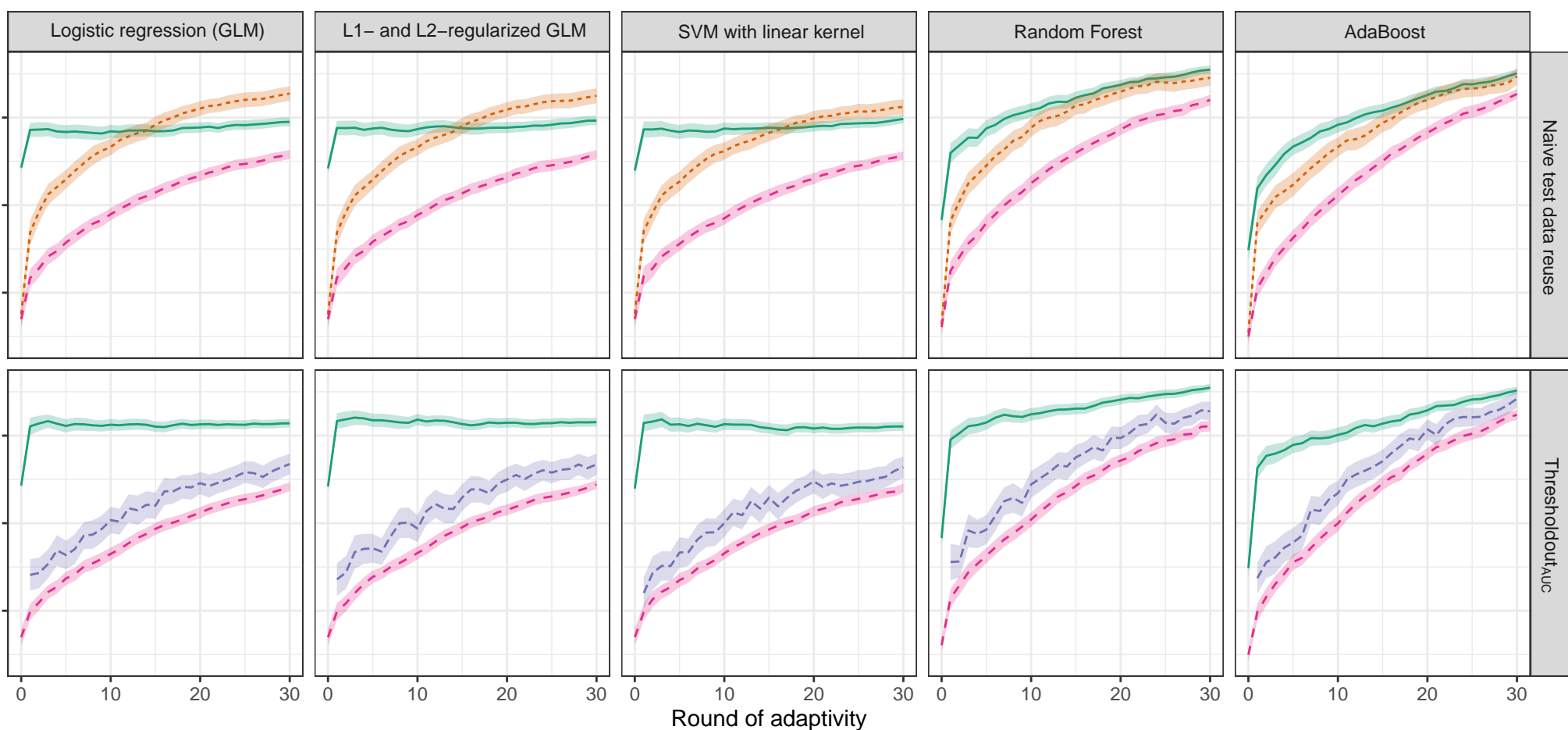


Mean AUC  $\pm$  2SE



— Training performance (cross-validation)    - - - Test performance (naive test data reuse)    - - - Thresholdout<sub>AUC</sub> output    - - - True performance (large independent dataset)