Approximate out-of-sample model performance
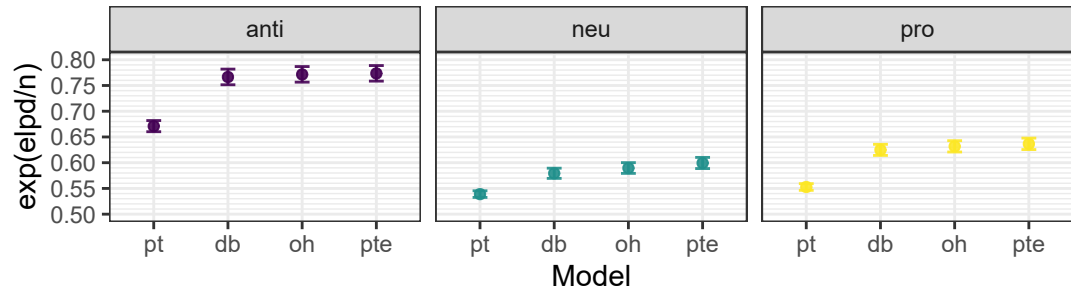
Model comparison