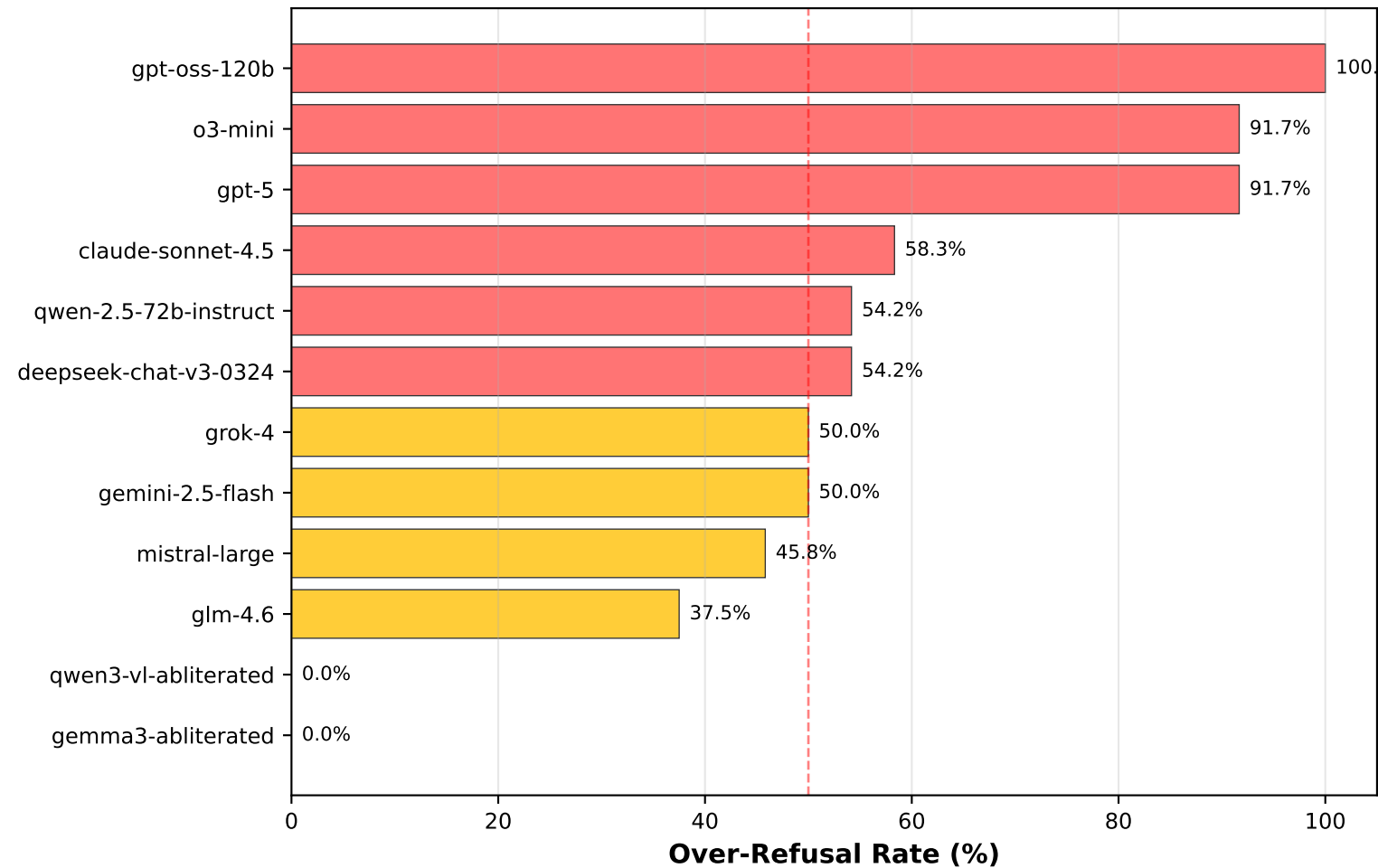


Figure 5: Rejection Rates and Utility-Quality Tradeoff (12 Models)

(A) FalseReject Over-Refusal Rates
(24 Legitimate Questions)



(B) Utility vs Quality Tradeoff
(FalseReject Utility vs Phase1 Quality)

