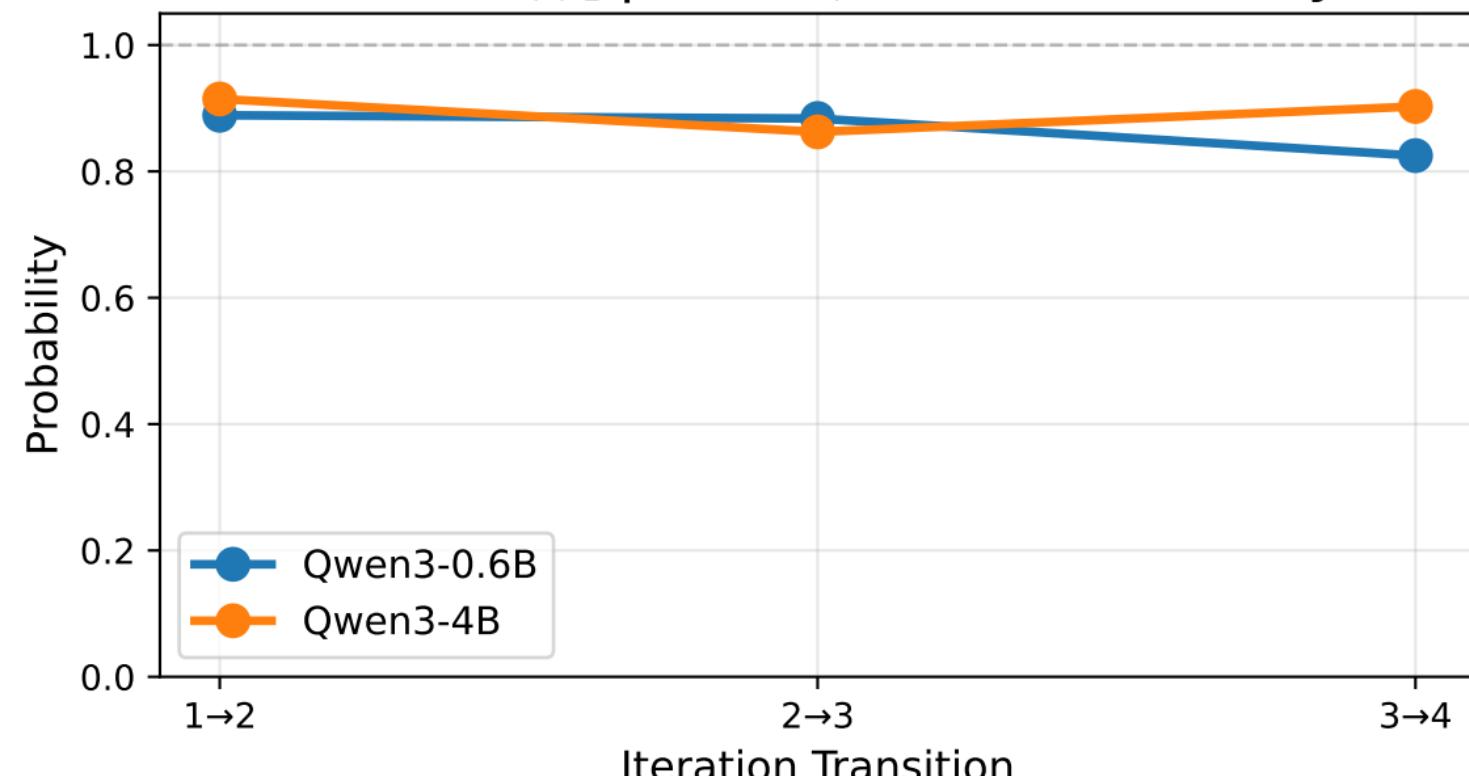


# MMLU Task: Conditional Probabilities Comparison

$P(\text{correct}_{i+1} \mid \text{correct}_i)$  - Answer Stability



$P(\text{correct}_{i+1} \mid \text{incorrect}_i)$  - Refinement Improvement

