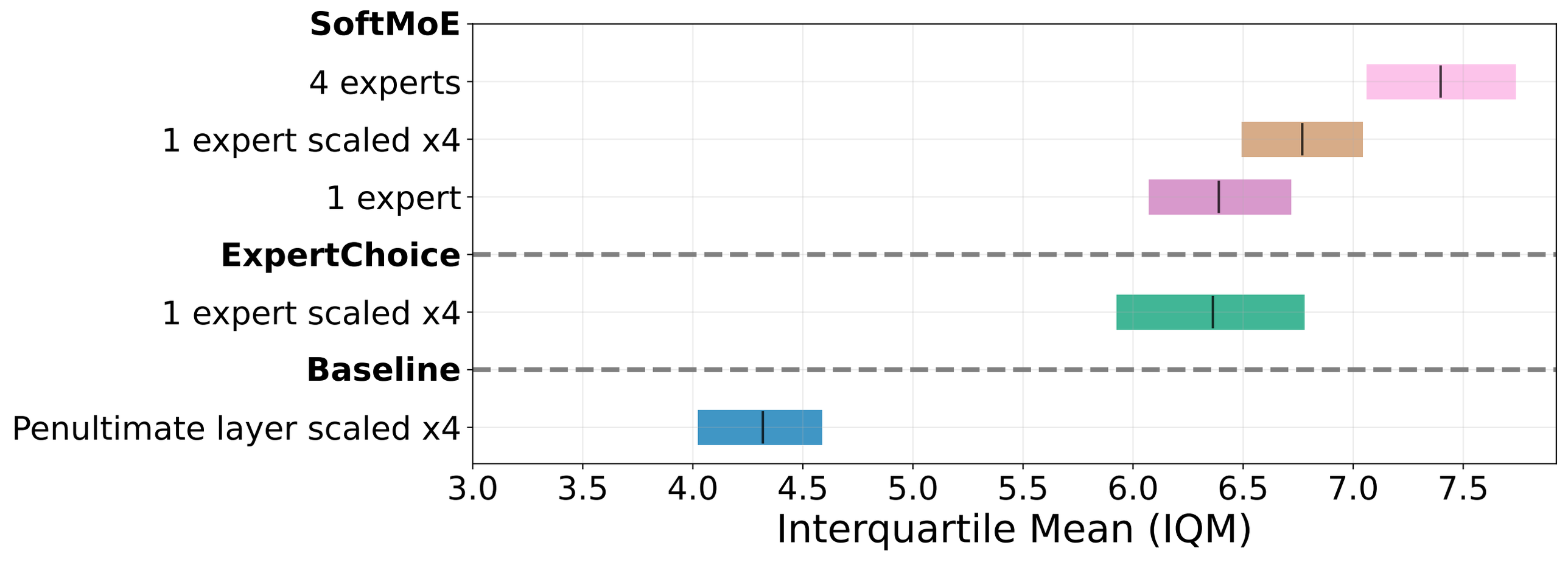
SoftMoE
4 experts
1 expert scaled x4
1 expert
ExpertChoice
1 expert scaled x4
Baseline
Penultimate layer scaled x4
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
Interquartile Mean (IQM)