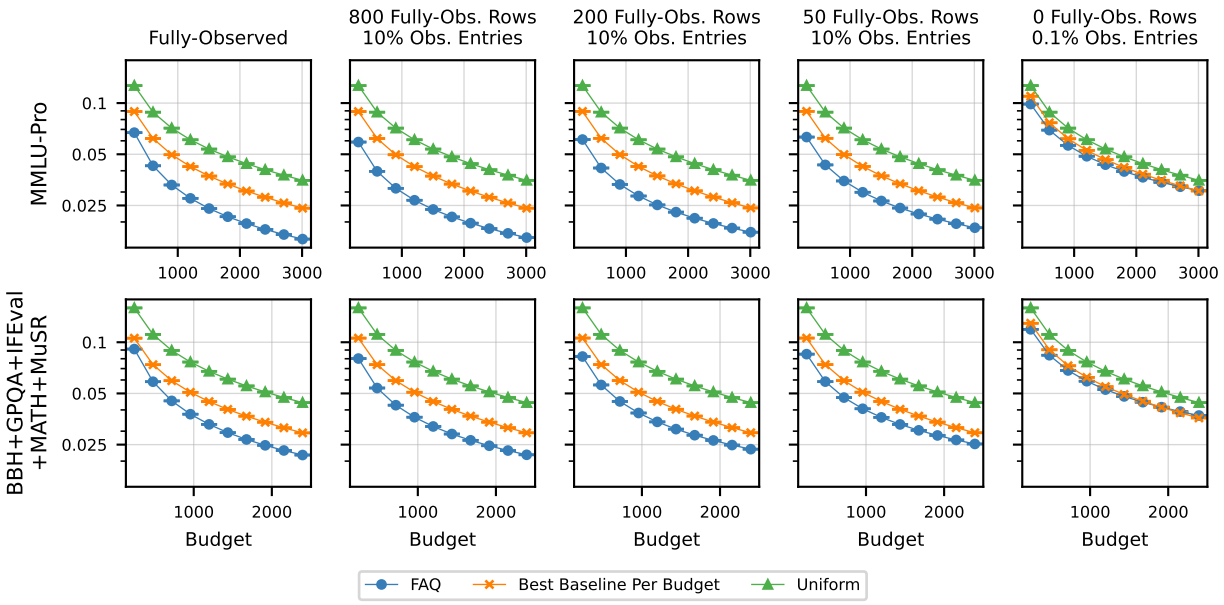
Fully-Observed
800 Fully-Obs. Rows
10% Obs. Entries
200 Fully-Obs. Rows
10% Obs. Entries
50 Fully-Obs. Rows
10% Obs. Entries
0 Fully-Obs. Rows
0.1% Obs. Entries
MMLU-Pro
BBH+GPQA+IFEval
+MATH+MuSR
0.1
0.05
0.025
1000
2000
3000
Budget
FAQ
Best Baseline Per Budget
Uniform