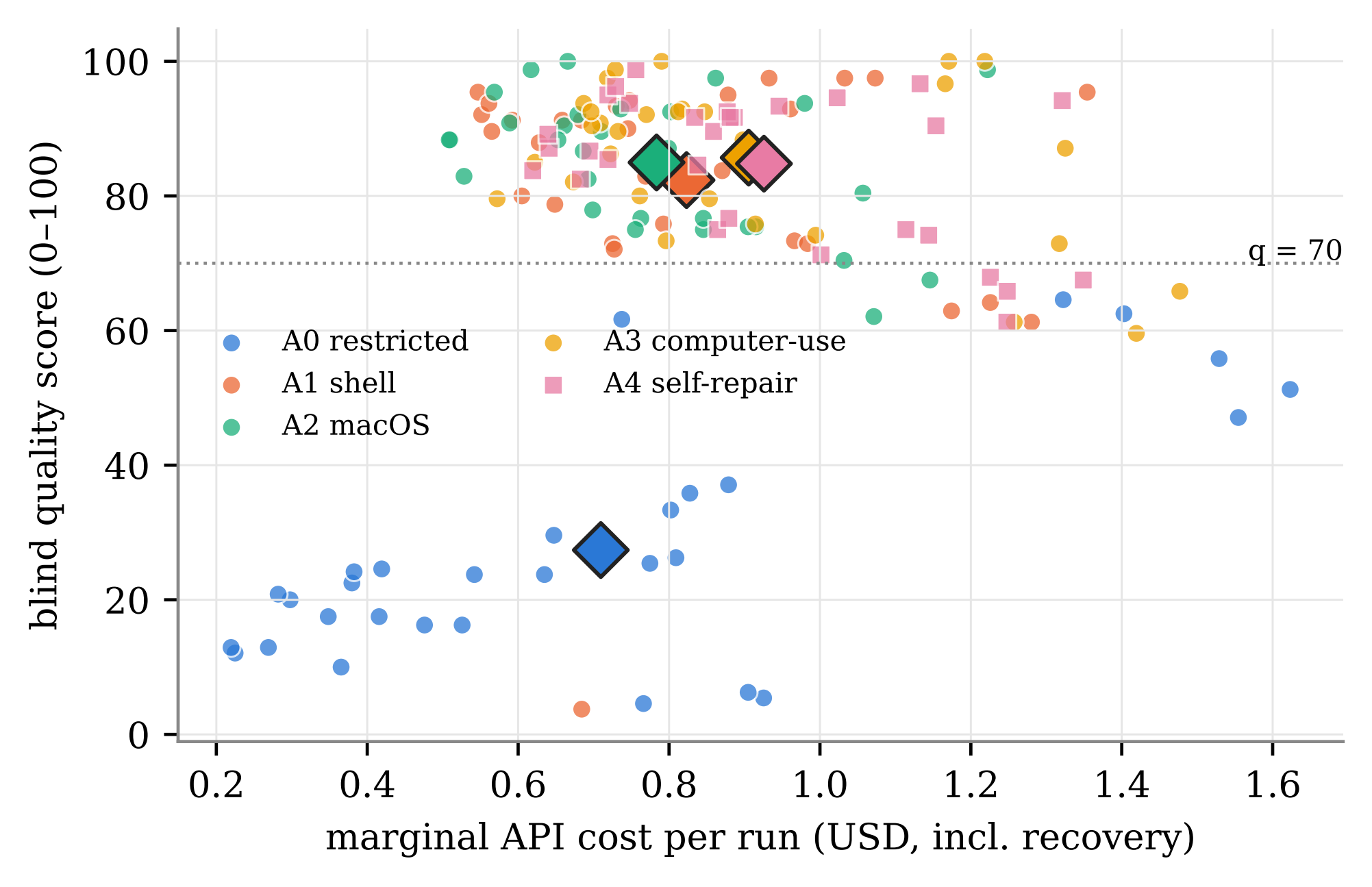

blind quality score (0–100)
marginal API cost per run (USD, incl. recovery)
q = 70
A0 restricted
A1 shell
A2 macOS
A3 computer-use
A4 self-repair
0
20
40
60
80
100
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6