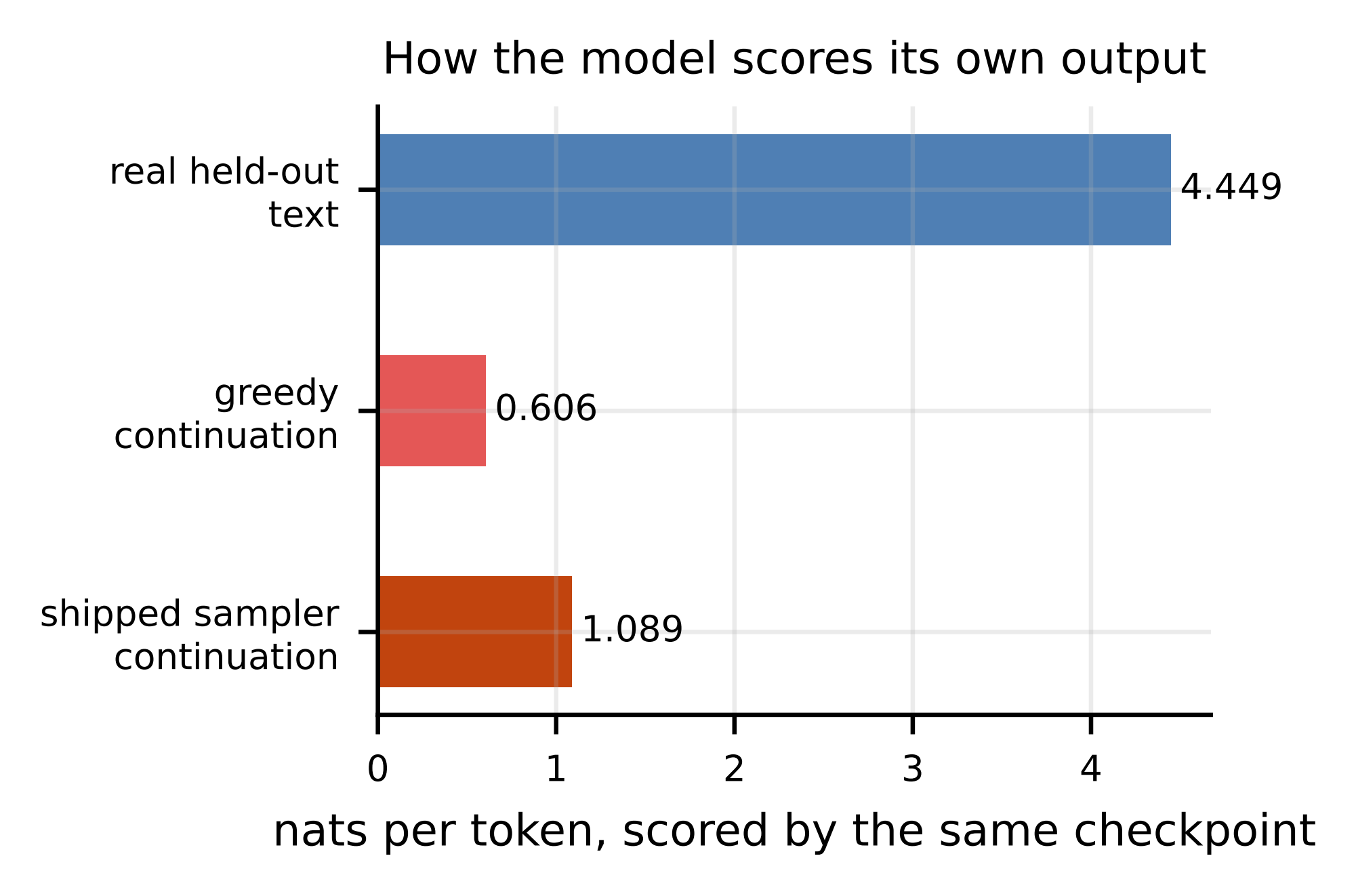

How the model scores its own output
real held-out text
4.449
greedy continuation
0.606
shipped sampler continuation
1.089
0
1
2
3
4
nats per token, scored by the same checkpoint