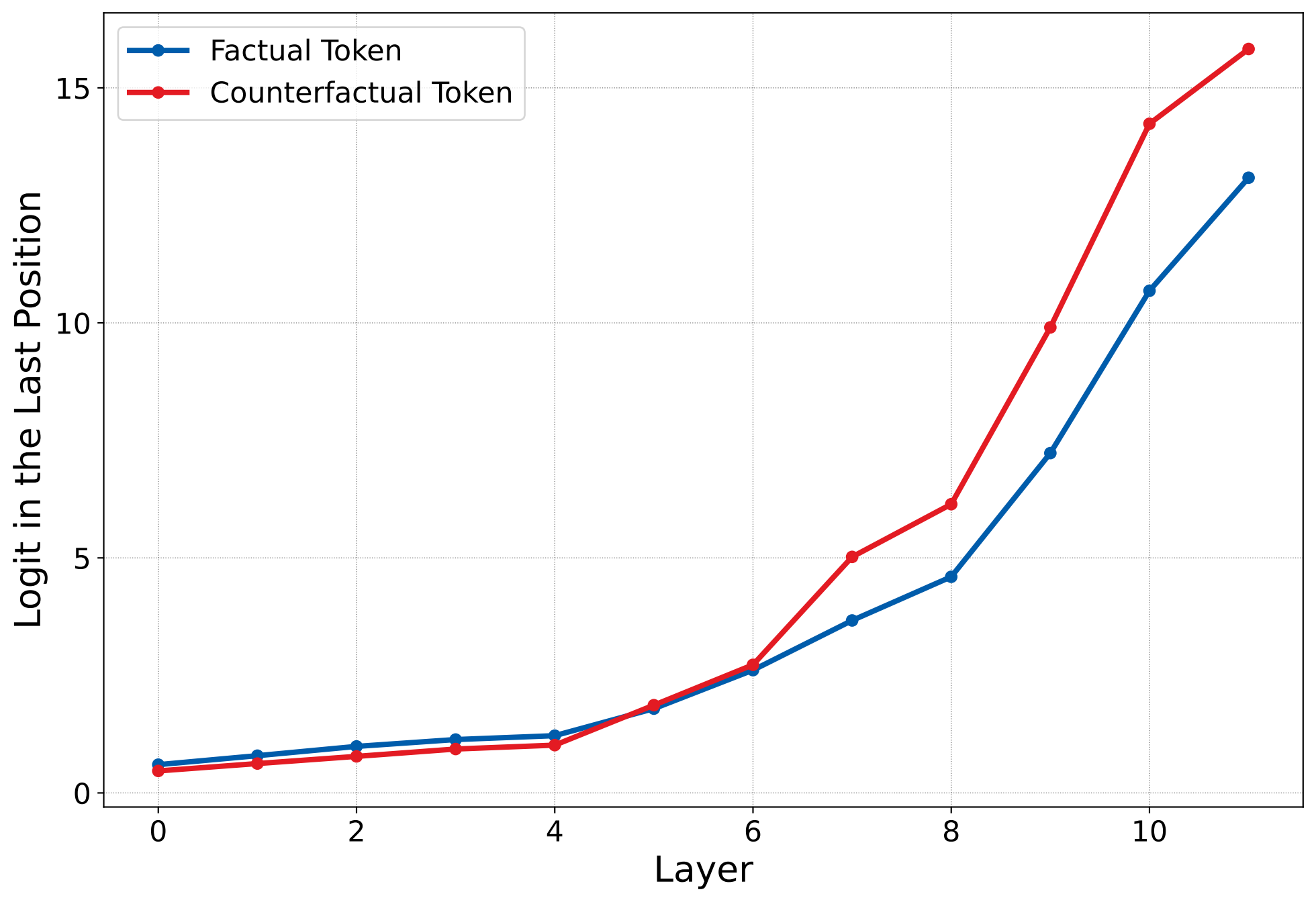
Factual Token
Counterfactual Token
Logit in the Last Position
Layer
0
5
10
15
0
2
4
6
8
10