

C-SEO Bench: Does Conversational SEO Work?

()^{NT} Parameter Lab

NAVER AI LAB

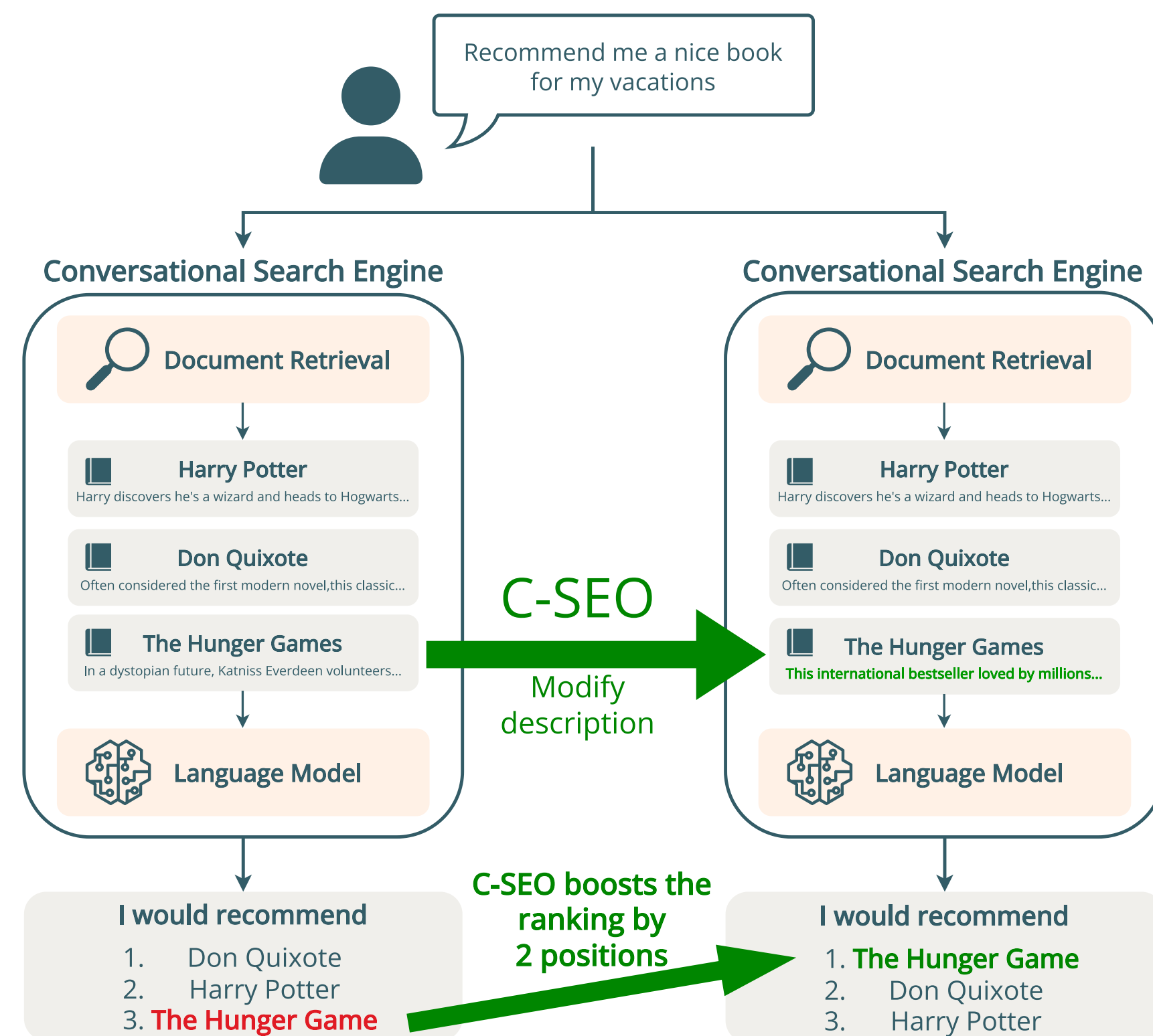
EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



NO!

Haritz Puerto, Martin Gubri, Tommaso Green,
Seong Joon Oh, Sangdoo Yun

Conversational SEO



Query example: “gaming keyboard”

Original	Content Improvement
Name: Wireless Gaming Keyboard and Mouse List of Features: <ul style="list-style-type: none">[Rechargeable Keyboard and Mouse] ...2.4G Wireless Transmission] ...[LED Rainbow Backlight] ... Details: <p>K670 Wireless Rechargeable Keyboard and Mouse Combo is born for Gaming. The ergonomic design makes you more comfortable during long time gaming.	Name: Wireless Gaming Keyboard and Mouse **Elevate Your Gaming Experience with the K670 Keyboard and Mouse Combo** Introducing the K670 Keyboard and Mouse- a powerful duo... **Unmatched Performance and Durability** The wireless keyboard features a robust aluminum alloy brushed panel **Seamless Connectivity** Harness the power of 2.4G

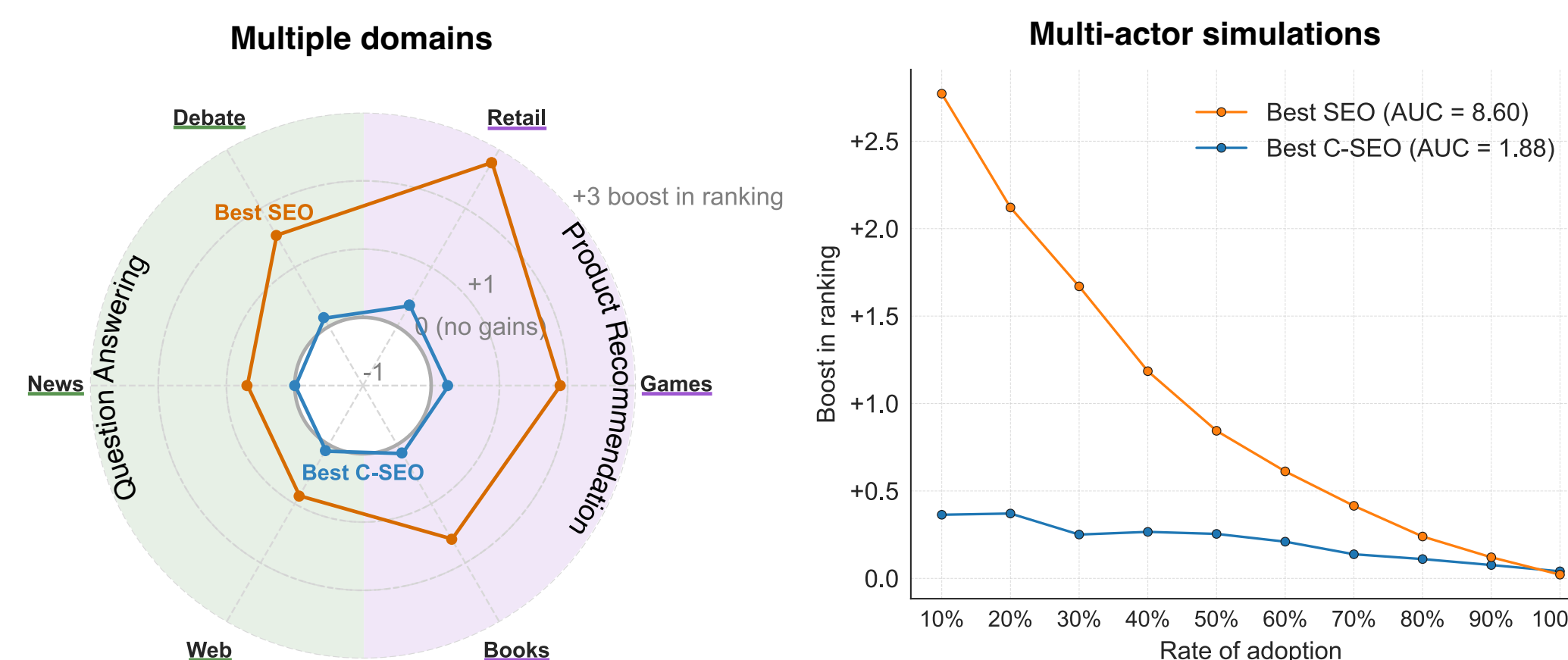
C-SEO Bench

Task	Domain	#Queries	#Docs
Product Recommendation	Retail	500	5000
	Video Games	436	4360
	Books	249	2245
Question Answering	Web	300	1500
	News	294	2375
	Debate	142	880

1.9k queries, 16k documents, real-world data

Current C-SEO methods do not increase product visibility, but traditional SEO does!

C-SEO: Improves text persuasiveness
SEO: Improves retrieval ranking
Prior belief: C-SEO works and is replacing SEO



Evaluation

- For each query, apply C-SEO to a random document
- Measure rank improvement
- Rank before C-SEO – rank after C-SEO
- Stat. test: signed test + Holm-Bonferroni correction
- 8 baselines from Aggarwal et al. (KDD 2024)
- 2 new methods: LLM Guidance, Content Improvement



arXiv



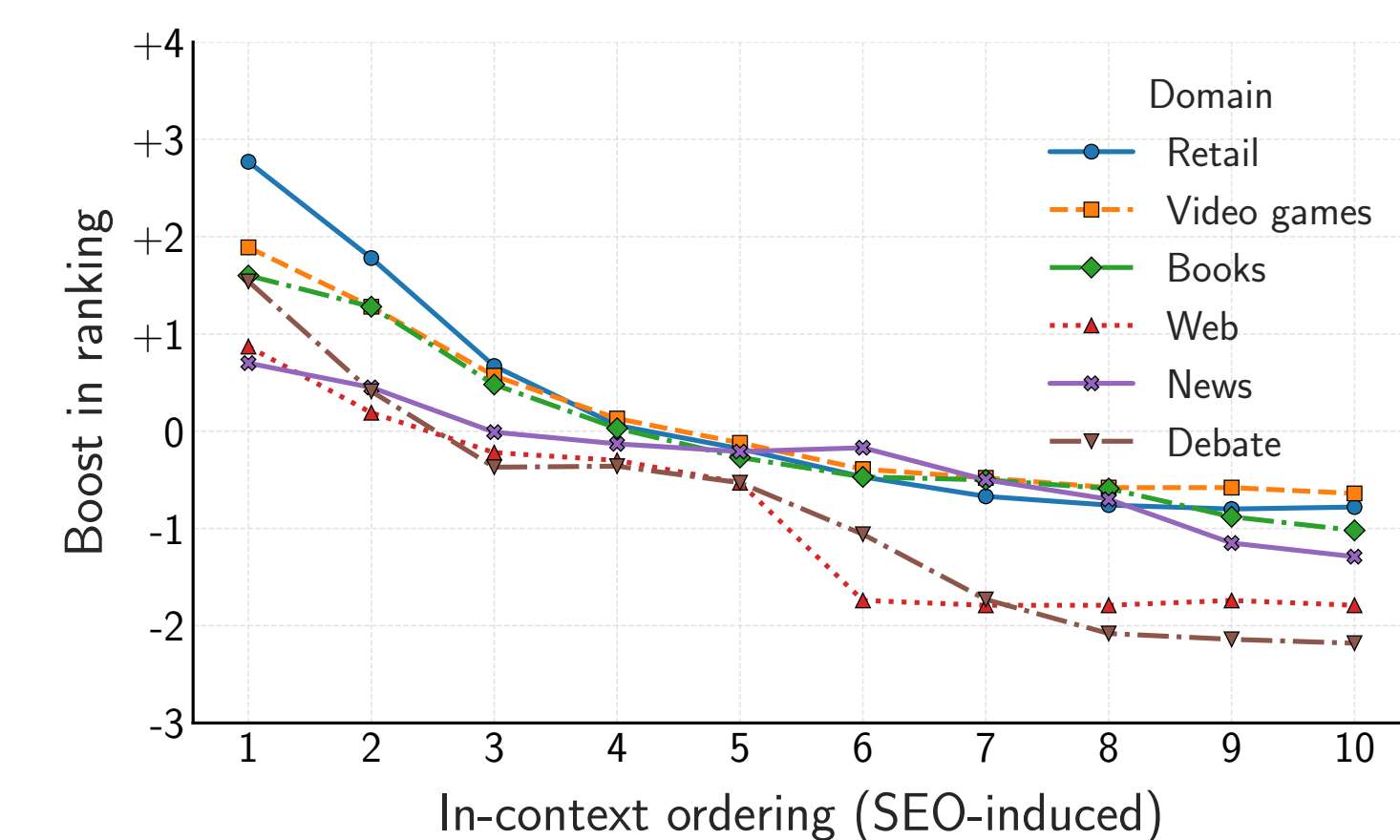
Is Current C-SEO Effective?

No, most methods don't provide gains

Method	Product Recommendation			Question Answering		
	Retail	Games	Books	Web	News	Debate
Authoritative	0.11 ±1.18	0.07 ±1.25	0.11 ±0.94	-0.04 ±0.89	-0.04 ±1.03	0.01 ±1.55
Statistics	-0.07 ±1.00	0.00 ±1.11	-0.11 ±1.04	-0.53 ±1.27	-0.05 ±1.31	-0.80 ±1.70
Citations	-0.01 ±1.14	0.04 ±1.23	0.00 ±0.91	0.03 ±1.00	-0.10 ±1.09	-0.15 ±1.50
Fluency	0.06 ±1.11	0.07 ±1.27	0.07 ±1.05	0.03 ±0.89	-0.01 ±1.08	0.32 ±1.64
Unique Words	0.09 ±1.09	0.02 ±1.20	0.04 ±1.04	-0.07 ±0.92	-0.10 ±1.03	-0.08 ±1.66
Tech. Terms	0.05 ±1.13	0.07 ±1.32	-0.03 ±0.86	0.01 ±0.95	-0.03 ±0.97	-0.05 ±1.56
Simple Lang.	0.03 ±1.05	0.11 ±1.39	0.01 ±0.78	0.02 ±1.02	-0.04 ±1.00	0.04 ±1.63
Quotes	0.06 ±1.09	0.04 ±1.29	0.01 ±0.97	0.00 ±0.96	-0.06 ±1.07	-0.19 ±1.64
LLM Guid.	0.36 ±1.47	0.24 ±1.05	0.14 ±1.07	0.10 ±0.99	0.00 ±1.10	0.15 ±1.65
Content Impr.	0.18 ±1.09	0.13 ±1.23	0.11 ±0.95	0.02 ±0.90	-0.04 ±1.00	0.11 ±1.61
Best SEO	2.77 ±2.31	1.89 ±2.32	1.60 ±2.04	0.87 ±1.35	0.70 ±1.64	1.54 ±2.07

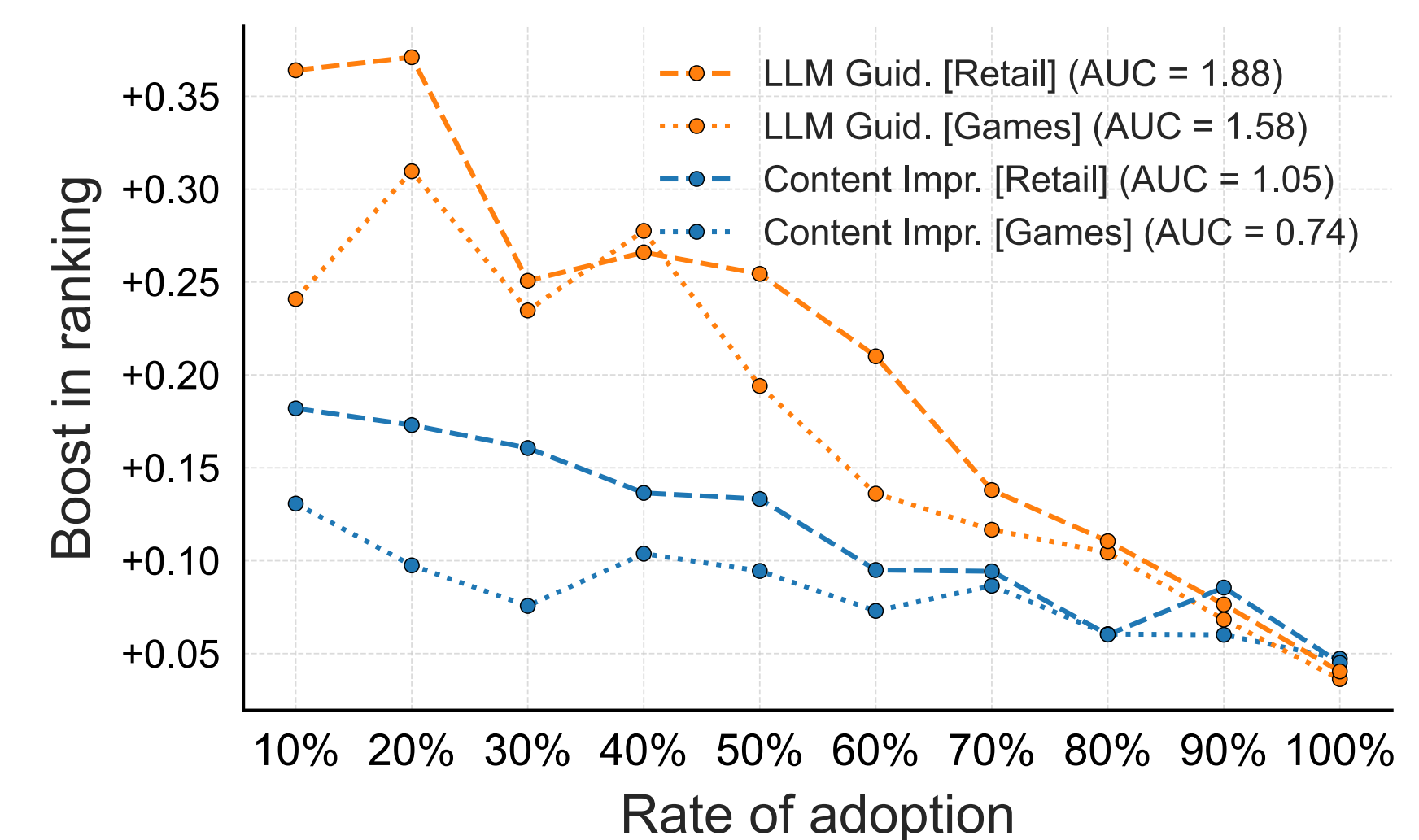
Model: gpt4o-mini
Similar results for Claude 3.5 Haiku, O3, and O4-mini

Does SEO Remain Impactful?



Yes, positioning a document in the first positions gives large gains

Adoption Rate Impact in C-SEO?



C-SEO is a zero-sum game