The Role of Search Engine Optimization in Search Rankings

Zsolt Katona (with Ron Berman)

University of California

Berkeley

Haas School of Business

TELECOM ParisTech

September 15, 2010

Introduction

- Search Advertising
    - Web sites bid for sponsored links on the "results" page of a search engine.
    - Sponsored links are assigned to sites based on their (pay-per-click) bids for specific keywords.
    - Over 11% of advertising money is spent online, over 40% of that on search advertising

- Search Engine Optimization
    - Webites can also affect their position on the organic list.
    - "Black hat" vs. "White hat"
    - SEO is over a billion $ business

SEO example - What Robots See

3er BMW
3er BMW - Fahrspaß pur!
Autohaus BMW
Sie suchen ein BMW Autohaus?
Behörde Fahrzeuge Anschaffung
Fahrzeugverkauf an Behörden
Behörde Fahrzeuge Beschaffung
Fahrzeugbeschaffung für Behörden
bei BMW
Beschaffung Fahrzeuge Behörde
Beschaffung von
Behördenfahrzeugen bei BMW
Blaulichtfahrzeug BMW
Für Polizei, Notarzt und Feuerwehr -
Blaulichtfahrzeuge von BMW
Blaulichtfahrzeug Notarzt
Schnell am Einsatzort - mit einem
Notarztwagen von BMW
BMW 316
Informationen zum BMW 316
BMW 318
Informationen zum BMW 318
BMW 318i
BMW 318i - Design und Fahrkultur
BMW 320
Probefahrt mit einem BMW 320?
BMW 330
Faszination BMW 330
BMW 3er
BMW 3er Editionen
BMW 5er
BMW 5er Serie

BMW Neuwagen

Sie suchen einen BMW Neuwagen? Unsere Suche nach BMW Händlern in Ihrer Nähe bietet schnellen Zugriff auf BMW-Autohäuser in Ihrer Nähe, wo Sie sich die BMW Neuwagen in aller Ruhe und Ausführlichkeit ansehen können. BMW Neuwagen - Sie erhalten von uns Adresse, Telefon und Website der BMW-Händler in Ihrer Nähe. Suchen Sie über Postleitzahl, Stadt oder Name des BMW-Partners. BMW Neuwagen - In jeder Abteilung unserer BMW Niederlassung arbeiten Fachleute für Sie. Unsere Mitarbeiter werden durch intensive Schulungen der BMW AG immer auf dem aktuellen Stand des Wissens gehalten. BMW Neuwagen - Mit diesem Know-how erarbeiten sie garantiert immer die Lösung, die sich am besten an Ihre Bedürfnisse anpasst. In einem unserer vielen Autohäuser in ganz Deutschland können Sie sich rasch und unproblematisch für eine Probefahrt in Ihrem Lieblings-BMW Neuwagen anmelden. Egal, worum es geht: um die Absprache eines Service-Termins, eine Reparatur oder die Finanzierung Ihres Neuen oder Ihres neuen Gebrauchten. BMW Neuwagen - Sympathisch und kompetent. Unser Team ist immer im Einsatz für Sie. Bei uns ist immer was los! Langeweile kommt in unserer Niederlassung nicht auf.
Hier finden Sie Informationen zum Thema: BMW Neuwagen gesucht?.

Ein BMW Neuwagen gesucht?

BMW Neuwagen - Regelmäßig bieten wir Ihnen neue Angebote, bringen Ihnen Aktuelles über die neuesten BMW Modelle nahe und organisieren Veranstaltungen aller Art. Bei uns werden Sie gut informiert und gut unterhalten. BMW Neuwagen - Egal, welche Frage Sie an unsere Profis haben: in unseren Filialen sind Sie als unser Kunde oder als neugieriger Interessent immer herzlich willkommen. BMW Neuwagen - Wenn Ihnen das BMW-Portal im Internet Appetit gemacht hat, besuchen Sie doch einmal eine Niederlassung ganz in Ihrer Nähe - hier können Sie unseren Mitarbeitern Löcher in den Bauch fragen. BMW Neuwagen - Haben Sie Interesse an einer Probefahrt in Ihrem Wunsch-BMW? Kein Problem! Bei unseren Niederlassungen können Sie sich jederzeit für eine Probefahrt in einem unserer Automobile anmelden. In unseren Niederlassungen bekommen Sie

Google Bombs

Web Images Groups News Froogle Local more »

Google

miserable failure   [Search]   Advanced Search
                                   Preferences

Web   Results **1 - 10** of about 969,000 for **miserable failure**. (0.06 seconds)

[Biography of President George W. Bush](#)
Biography of the president from the official White House web site.
www.whitehouse.gov/president/gwbbio.html - 29k - Cached - Similar pages
        Past Presidents - Kids Only - Current News - President
        More results from www.whitehouse.gov »

[Welcome to MichaelMoore.com!](#)
Official site of the gadfly of corporations, creator of the film Roger and Me
and the television show The Awful Truth. Includes mailing list, message board, ...
www.michaelmoore.com/ - 35k - Sep 1, 2005 - Cached - Similar pages

[BBC NEWS | Americas | '**Miserable failure**' links to Bush](#)
Web users manipulate a popular search engine so an unflattering description leads
to the president's page.
news.bbc.co.uk/2/hi/americas/3298443.stm - 31k - Cached - Similar pages

[Google's (and Inktomi's) **Miserable Failure**](#)
A search for **miserable failure** on Google brings up the official George W.
Bush biography from the US White House web site. Dismissed by Google as not a ...
searchenginewatch.com/sereport/article.php/3296101 - 45k - Sep 1, 2005 - Cached - Similar pages

Google Bombs

Common Black Hat SEO techniques

- Content spam
  - Cloaking
  - Keyword stuffing
  - Hidden or invisible unrelated text
  - Doorway pages
  - Scraper sites
- Link spam
  - Link farms
  - Hidden links
  - Spam blogs
- UGC spam
  - Comment spam
  - Wiki spam

Research Questions

- What are sites' incentives to invest in SEO?

- How does SEO affect the results?

- Do (which) sites profit from it?

- What is the relationship with sponsored links and search engine revenues?

- How does searcher behavior matter?

- Conventional wisdom: SEO is good for low quality sites, bad for SE and consumers.

Related Literature

- Search Advertising
  - Edelman, Ostrovsky and Schwarz (2007), Varian (2007),Athey and Nekipelov(2010)
  - Chen and He (2006)
  - Jeziorski and Segal (2009)
  - Goldfarb and Tucker (2008), Rutz and Bucklin (2007)
  - Yao and Mela (2009), Ghose and Yang (2009)
- Role of organic links
  - Katona and Sarvary (2010), White (2009), Xu, Chen and Whiston (2009),
  - Yang and Ghose (2010)
  - SEO: Xing and Ling (2006), Sen (2005)
- Non-desirable behaviors
  - Wilbur and Zhu (2009a,2009b)
- All-pay auctions
  - Hillman and Riley (1987)
  - Barut and Kovenock (2008), Baye et al. (2006), Clark and Riis (2008)

Model

- Players: Search Engine (SE), n Web sites
- Site i has relevance (quality) q_i with $q_1 \geq q_2 \geq \ldots \geq q_n$.
- Relevance: consumers are satisfied with probability q_i if clicking on link i.
- SE displays k organic links and l sponsored links.
- SE places the best links on the organic side and maximizes revenue on the sponsored side. Total traffic (T) is an incerasing function of expected consumer utility $f(U)$.
- Consumers:
  - with probability ψ, click on the best link (conscientious clicking)
  - otherwise, click on the ith organic link (i) with prob $\gamma\beta_i$ and on the jth sponsored link with prob $(1-\gamma)\beta_j$ (random clicking)
- Site i gets utility $R_i(t)$ from t visitors, where $r_i() = R_i'()$ is decreasing.

Model cont'd

- SE cannot measure q_i perfectly,

  error term: ε_i (i.i.d., centered around 0)

- First, SE assigns (public) scores $s_i^S = q_i + \sigma\varepsilon_i$.

- After observing these, sites can invest in SEO: b_i.

- Final scores: $s_i^F = s_i^S + \alpha b_i$.

- Effectiveness of SEO: α (cost: $\frac{1}{\alpha}$).

- SE orders sites according to s_i^F.

Simple case: $n = 2, k = 1, l = 0$

- Two sites: $q_1 \geq q_2, \ v_i = R_i(f(q_i))$, one link, $\varepsilon_i = \pm 1$.
- When there is no SEO ($\alpha = 0$), $a_i^S = a_i^F = q_i \pm \sigma$.
- If $\sigma > (q_1 - q_2)/2$ then 1 wins with prob $3/4$.
- Probability of desired outcome $P(\alpha) = Pr(1 \text{ wins}|\alpha)$, that is, $P(0) = 3/4$.
- Expected traffic $ET(\alpha) = P(\alpha)(f(q_1) - f(q_2)) + f(q_2)$.
- What is the optimal α?
- How does efficiency change with q_1, q_2, v_1, v_2?

Relation to all-pay auctions (contests)

- Players compete for an item, highest bid wins, full information. Their investment is sunk.

- Two players $v_1 > v_2$: No pure strategy eq.

- Mixed strategies: both players mix between 0 and v_2.

- 1 wins with probability $1 - \frac{v_2}{2v_1}$.

- Expected payoffs $v_1 - v_2$ and 0.
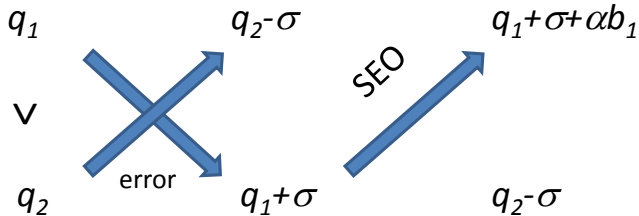
With headstart h:

- Player wins if bid+headstart is higher than others'.

- If $v_1 + h > v_2$, player 1 has a higher chance to win.

- Expected payoffs $min(v_1 - v_2 + h, v_1)$ and 0.

Results

- There always exists an $\hat{\alpha} > 0$ such that $P(\hat{\alpha}) \geq P(0) = 3/4$.

- If $v_1 > (3/2)v_2$ then for any $\sigma > (q_1 - q_2)/2$ there is an $\hat{\alpha} > 0$ such that $P(\hat{\alpha}) > P(0) = 3/4$.

- If the most relevant site values traffic highly, some positive level of SEO always improves the efficiency

$$R_1(f(q_1)) > \frac{3}{2}R_2(f(q_2))$$

- In any other case, low levels of SEO do not hurt.

q_1

$\vee$

q_2

$q_2\text{-}\sigma$

error

$q_1\text{+}\sigma$

SEO

$q_1\text{+}\sigma\text{+}\alpha b_1$

$q_2\text{-}\sigma$

Comparative statics

- $P(\alpha, \sigma, v_1, v_2, q_1, q_2)$ is increasing in v_1, q_1 decreasing in v_2, q_2.

- The more relevance is aligned with site valuation, the more efficient the ranking.

- Intuition: the auction is a mechanism that favors players with high valuation. If these coincide with the most relevant sites, then the ranking is improved.

- If the non-relevant site has a high valuation $P(\alpha)$ can be reduced to 0.

- Spammer sites can take advantage (low relevance, high valuation).

- In reality, some search engines accept paid links in the organic list.

Error and efficiency

- How does the optimal SEO effectiveness level change with the variance of the measurement error?

- Optimal SEO for searche engine: $\hat{A}$.

- Result: If $v_1 > (3/2)v_2$ then $\hat{A} = \hat{A}(\sigma)$ is increasing in σ.

- Intuition: a more effective correction mechanism may be useful if there is more error.

- Implication: Investment in search algorithms and against SEO are **complements**

$P(\alpha)$

$P(\alpha, \sigma)$

So far...
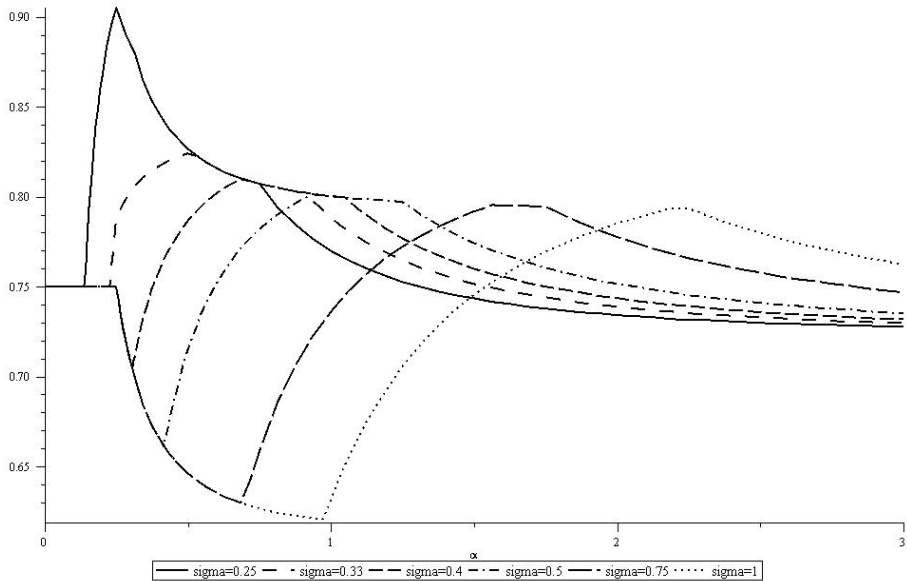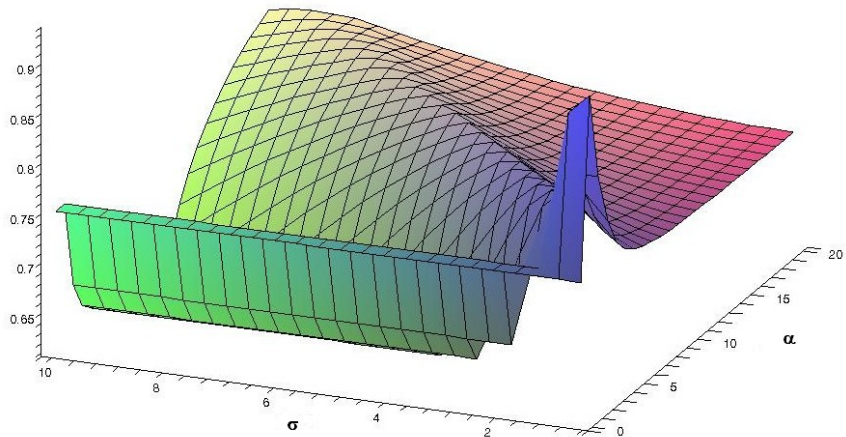
- We have shown that SEO might be good for search engine and consumers, if

- high quality sites value consumer high enough.

What about sites? Do they profit from SEO?

- Site 2 (lower quality) always loses with more SEO if $v_1 > v_2$.

- Site 1 is always strictly better off if $v_1 > 4v_2$ or $\sigma < \frac{v_1}{v_2}\frac{q_1-q_2}{2}$.

- If valuations are close, SEO makes sites fight hard for the top spot.

SEO and Sponsored Links Setup ($k = 1, l = 1, \psi = 0$)

- Timing: After the SEO process, second-price auction for the sponsored link.

- Sites bid given their organic traffic. If marginal valuation for traffic is constant, they are independent.

- If a site loses on the organic side, it compensate on the sponsored side.

    If $R_1(f(q_1)) > (3/2)R_2(f(q_2)$ and $\gamma < 1$ is high enough, there is $\hat{\alpha}$, s.t. $ET(\hat{\alpha}) > ET(0)$.

What about SE profits?

- Controversy: would search engines want to provide low quality sponsored links to gain sponsored revenue?

- $E\pi(\hat{\alpha}) > E\pi(0)$ iff
  $R_1(f(q_1)) - R_1(\gamma f(q_1)) \geq R_2(f(q_2)) - R_2(\gamma f(q_2))$.

- Higher traffic results in higher profits if
  - $|r_1'|$ is small: Revenues don't decline rapidly with traffic,
  - q_1/q_2 is close to 1: The difference between sites is not big, or
  - $f'()$ is not too high: consumers are not too sensitive to expected quality .

$$r_1(t) = 1 - t, \quad r_2(t) = r(1 - t).$$

- Consumers can tell which is the best search result without trying.
- ET is increased by SEO when r is low and γ is high or when r is high and γ is low (increasingly the later as ψ increases)
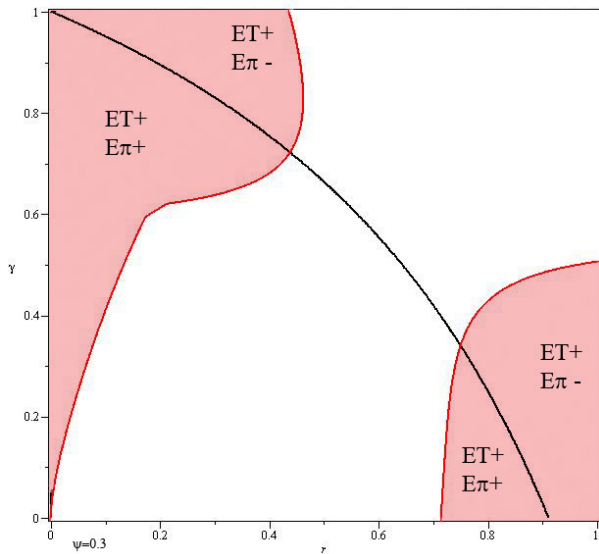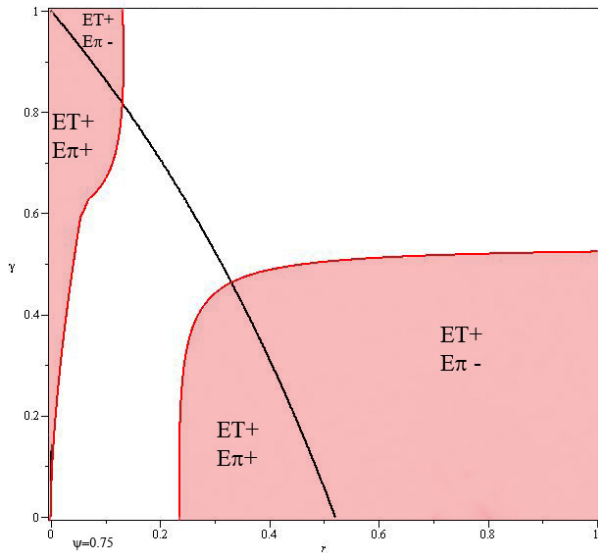- $E\pi$ is increases with ET when $\gamma < \hat{\gamma}$, where $\hat{\gamma}$ is decreasing in r and ψ.
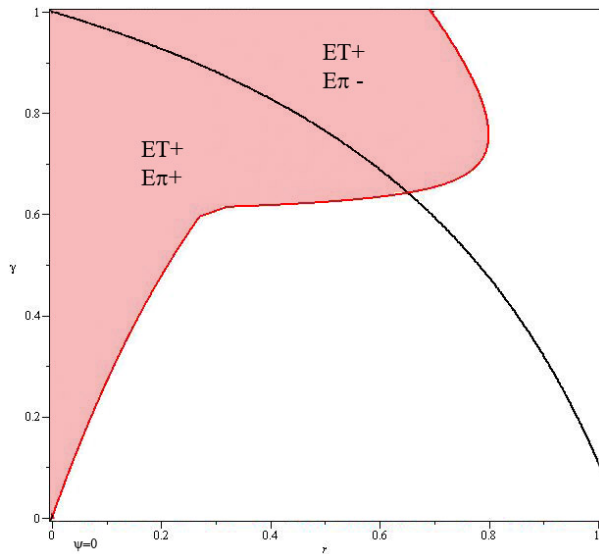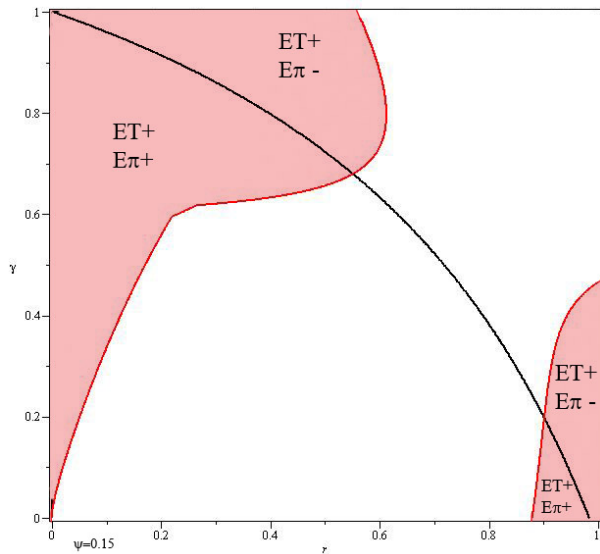
SEO benefits ($\psi = 0.30$)

SEO benefits ($\psi = 0.45$)

General model
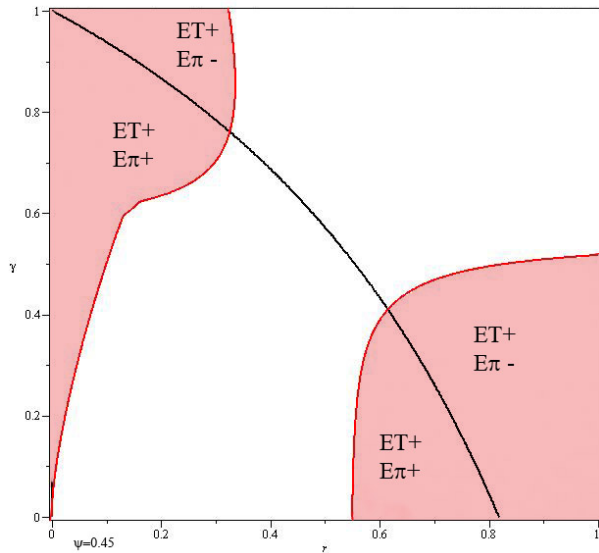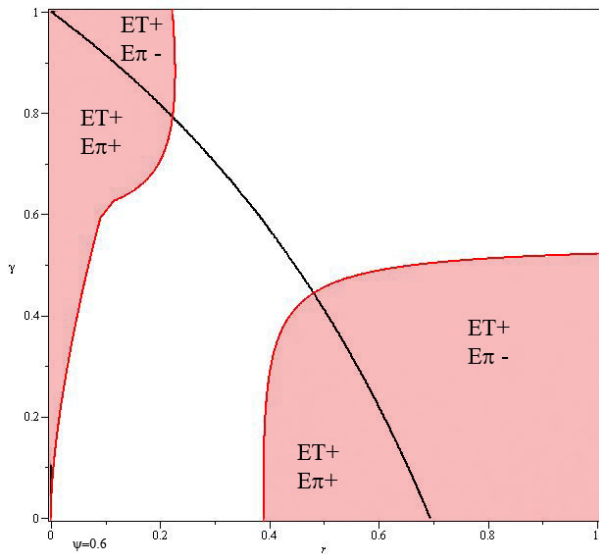
- Multiple organic links

- Multiple players

- Error has a finite support

- $\overline{\varepsilon} < \frac{q_1 - q_k}{k}$.

- Results are similar: some positive level of SEO can improve the organic ranking.

Conclusion

- We model the economic incentives behind SEO.

- We show that some positive level of SEO can be beneficial to
  - the search engine by increasing traffic and even profits
  - consumers
  - but generally not to sites, who have to defend their position

- SEO and sponsored links
  - Sites can go for either or both.
  - Higher quality organic links results in higher sponsored revenues when consumers trust the SE, valuations are different, there is enough clikc on sponsored links.

Future directions

- Different assumptions on timing and information

- Content investment / white hat SEO

- More active search engine

- Dynamics

- Heterogeneos consumers