An Empirical Analysis of Sponsored Search Performance in Search Engine Advertising

Anindya Ghose
Sha Yang

Stern School of Business
New York University
• Background
• Research Question and Summary of Results
• Theory and Econometric Model
• Data
• Results
• Takeaways
• Future and Ongoing Work
Search engines act as intermediaries between advertisers and users.

- Refer consumers to advertisers based on user-generated queries and keyword advertisements.

- Consumer behavior from search to purchase:
  - Search -> Impressions -> Clicks -> Conversions
• Pay per click (PPC) is where advertisers only pay when a user actually clicks on its ad listing to visit its website.

• Keyword: “Used cars San Diego”
Characteristics of Keywords

Classification of user queries in search engines (Broder 2002)

- Navigational
- Transactional
- Informational

Prior theory to motivate study using keyword attributes

- Presence of **Retailer** information (Retailer name)
  - “K-Mart bedding”
- Presence of **Brand** information (Manufacturer/Product specific brand)
  - “Nautica bedsheets”
- Specific search or Broad search (Length of keyword in words)
  - “Cotton bedsheets” vs. “300 count Egyptian cotton bedsheets”.
Prior theory to motivate study using keyword attributes

- Presence of Retailer information
  - Presence of Brand information
  - Specific search or Broad search

Loyal/Aware Consumers/White Pages

Competitive/Searchers/Yellow Pages
How does sponsored search advertising affect consumer behavior on the Internet?

- What attributes of a sponsored advertisement influence users’ click-through and conversion rates?
- How do the “keyword attributes” influence the advertiser’s cost-per-click, and the search engine’s ranking decision?
- Policy simulations to impute optimal CPC for the advertiser
Summary of Findings and Contributions

Hierarchical Bayesian model to empirically estimate the impact of various keyword attributes (Wordographics).

- Retailer information increases CTR.
- Brand information increases conversion rates.
- Increases in keyword length decreases CTR.
- Increase in Rank decreases both CTR and conversion rates.

Also analyze the impact of these covariates on firm level decisions – `CPC’ and `Rank’.

- Policy simulations suggest that the advertiser can make improvements in its expected profits from optimizing its CPC.
- Search engines take into account both the bid price as well as prior CTR before setting the final rank of an advertisement.
Empirical Methodology

Framework

- Hierarchical Bayesian model
  - Rossi and Allenby (2003)

- Markov Chain Monte Carlo methods
  - Metropolis-Hastings algorithm with a random walk chain to generate draws (Chib and Greenberg 1995)

Models of Decision Making

- Consumer level decision: Click-through
- Consumer level decision: Conversion
- Advertiser decision: Cost-per-click
- Search Engine decision: Keyword Rank
• First, a user clicked and made a purchase. The probability of such an event is $p_{ij} q_{ij}$.

• Second, a user clicked but did not make a purchase. The probability of such an event is $p_{ij} (1-q_{ij})$.

• Third, an impression did not lead to a click-through. The probability of such an event is $1-p_{ij}$.

• Then, the probability of observing $(n_{ij}, m_{ij})$ is given by:

$$f(n_{ij}, m_{ij}, p_{ij}, q_{ij}) = \frac{N_{ij}!}{m_{ij}!(n_{ij} - m_{ij})!(N_{ij} - n_{ij})!} \{p_{ij} q_{ij}\}^{m_{ij}} \{p_{ij} (1-q_{ij})\}^{n_{ij} - m_{ij}} \{1 - p_{ij}\}^{N_{ij} - n_{ij}}$$

N = number of impressions
n = number of clicks
m = number of conversions
p = probability of click-through
q = probability of conversion conditional on click-through
Empirical Models

\[ p_{ij} = \frac{\exp(\beta_{i0} + \beta_{i1} \text{Rank}_{ij} + \alpha_1 \text{Retailer}_i + \alpha_2 \text{Brand}_i + \alpha_3 \text{Length}_i + \varepsilon_{ij})}{1 + \exp(\beta_{i0} + \beta_{i1} \text{Rank}_{ij} + \alpha_1 \text{Retailer}_i + \alpha_2 \text{Brand}_i + \alpha_3 \text{Length}_i + \varepsilon_{ij})} \]

\[ q_{ij} = \frac{\exp(\theta_{i0} + \theta_{i1} \text{Rank}_{ij} + \overline{\delta}_1 \text{CTR}_{ij} + \delta_1 \text{Retailer}_i + \delta_2 \text{Brand}_i + \delta_3 \text{Length}_i + \eta_{ij})}{1 + \exp(\theta_{i0} + \theta_{i1} \text{Rank}_{ij} + \overline{\delta}_1 \text{CTR}_{ij} + \delta_1 \text{Retailer}_i + \delta_2 \text{Brand}_i + \delta_3 \text{Length}_i + \eta_{ij})} \]

\[ \ln(CPC_{ij}) = \omega_{i0} + \omega_1 \text{Rank}_{ij-1} + \omega_2 \text{Profit}_{ij-1} + \lambda_1 \text{Retailer}_i + \lambda_2 \text{Brand}_i + \lambda_3 \text{Length}_i + \mu_{ij} \]

\[ \ln(\text{Rank}_{ij}) = \phi_{i0} + \phi_1 \text{Bid Price}_{ij} + \overline{\phi}_2 \text{CTR}_{ij-1} + \tau_1 \text{Retailer}_i + \tau_2 \text{Brand}_i + \tau_3 \text{Length}_i + \nu_{ij} \]
• Large nationwide retailer (Fortune-500 firm) with 520 stores in the US and Canada.

• 3 months dataset from January 07 to March 07 on Google Adwords advertisements (Also data on Yahoo and MSN).

• 1800 unique keyword advertisements on a variety of products.

  - **Keyword level (Paid Search):** Number of impressions, clicks, Cost per click (CPC), Rank of the keyword, Number of conversions, Revenues from a conversion, quantity and price in each order.

  - **Product Level:** Quantity, Category, Price, Popularity.

• These are clustered into six product categories
  - Bath, bedding, electrical appliances, home décor, kitchen and dining.
Results

- Retailer-specific information increases CTR by 26.16%

- Brand-specific information increases conversion rates by 23.76%

- Increase in rank decreases both CTR and conversion rates
### Table 4: Coefficient Estimates on Bid Prices

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Retailer</th>
<th>Brand</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega_0 )</td>
<td>(-1.285)</td>
<td>(-1.036)</td>
<td>(-0.171)</td>
<td>(0.095)</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.089)</td>
<td>(0.043)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>(-0.027)</td>
<td>(0.110)</td>
<td>(0.013)</td>
<td>(-0.003)</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.039)</td>
<td>(0.013)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>( \omega_2 )</td>
<td>(-0.020)</td>
<td>(-0.049)</td>
<td>(-0.005)</td>
<td>(0.003)</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.033)</td>
<td>(0.022)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>

### Table 5: Coefficient Estimates on Keyword Ranks

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Retailer</th>
<th>Brand</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi_0 )</td>
<td>(2.119)</td>
<td>(-0.636)</td>
<td>(-0.434)</td>
<td>(-0.109)</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.152)</td>
<td>(0.076)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>( \phi_k )</td>
<td>(-3.025)</td>
<td>(1.787)</td>
<td>(0.307)</td>
<td>(0.455)</td>
</tr>
<tr>
<td></td>
<td>(0.353)</td>
<td>(0.390)</td>
<td>(0.179)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>( \phi_2 )</td>
<td>(-1.328)</td>
<td>(0.088)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Coefficient Estimates on Bid Prices

Table 5: Coefficient Estimates on Keyword Ranks
Policy Simulations

Overview

- Determine optimal bid price
- Impute profits with optimal bid and actual CPC

Findings

- Differences between optimal bid and actual CPC
  - Average deviation is 24 cents per bid
  - Generally CPC higher than optimal bid price (94%)

- Differences in ‘Expected Profits’ and ‘Actual Profits’ per keyword
  - Regressions with optimal prices show that firm should *increase* bid price with **Retailer** or **Brand** information, and decrease with **Length**.
Some Limitations

- No data on Competition.
- No explicit data on landing page quality score.
  - Content analysis based on metrics on Google Adwords (but noisy?)
- No data on text of the ad copy
• **Empirically estimate the impact of various keyword attributes on consumers’ search and purchase propensities.**
  
  – Retailer-specific information increases CTR and brand-specific information increases conversion rates.
  
  – Increase in Rank decreases both CTR and conversion rates.
  
  – What are the most “attractive” keywords from an advertiser’s perspective?
  
  – Implications for products of interest to “loyal consumers” versus “shoppers/searchers”.
• **Analyze the impact of these covariates on advertiser and search engine decisions such as CPC and Rank.**

  - Evidence that while the advertiser is exhibiting some naïve learning behavior they are not bidding optimally.

  - How should it bid in search engine advertising campaigns to maximize profits?