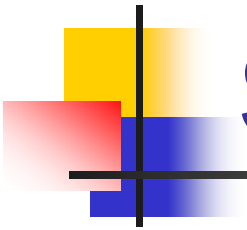
The Economics of Internet Search

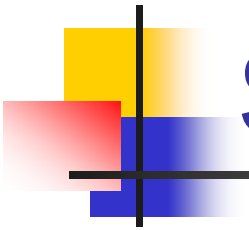
Hal R. Varian

January 20, 2007



Search engine use

- Search engines are very popular
  - 84% of Internet users have used a search engine
  - 56% of Internet users use search engines on a given day
- They are also highly profitable
  - Revenue comes from selling ads related to queries



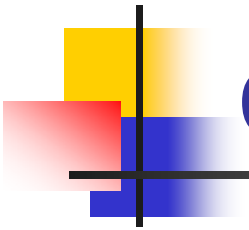
Search engine ads

- Ads are highly effective due to high relevance
  - But even so, advertising still requires scale
    - 3% of ads might get clicks, 3% of those might convert
    - So only 1 out a thousand who see an ad actually buy
    - Hence, price per impression or click will not be large
    - But this performance is good compared to conventional advertising! CPM (cost per thousand impressions.)
    - TV ~ $10 CPM = 1 cent per person per impression
- On the supply side:
  - High fixed costs for infrastructure, low marginal costs for serving

Summary of industry economies

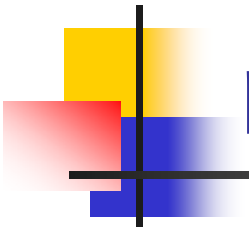
- Entry costs (at a profitable scale) are large due to fixed costs
- User switching costs are low
  - 56% of search engine users use more than one
- Advertisers follow the eyeballs
  - Place ads wherever there are sufficient users, no exclusivity
- Hence market is structure is likely to be
  - A few large search engines in each language/country group
  - Highly contestable market for users
  - No demand-side network effects that drive towards a single supplier so multiple players can co-exist

What services do search engines provide?



- Google as yenta (matchmaker)
  - Matches up those seeking info to those having info
  - Matches up buyers with sellers
  - Two sided market: advertisers want to meet buyers, Google shows content to attract potentiall buyers, charges advertisers for introduction (like other advertiser-supported media)
- How do you make valuable matches?
  - Information science: information retrieval
  - Economics: assignment problem

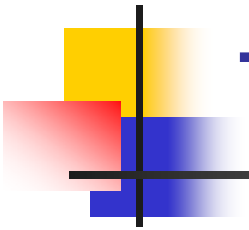
Brief history of information retrieval



- Started in 1970s, basically matching terms in query to those in document
- Was pretty mature by 1990s
- DARPA started Text Retrieval Conference at urging of FBI, CIA, NSA, etc.
  - Offered training set of query-document pairs with indicator of relevance
  - Offered challenge set of queries and documents
  - Roughly 30 research teams participated

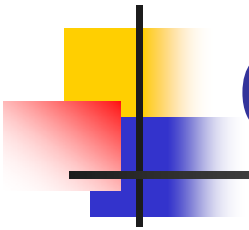
Example of IR algorithm

- Prob(document relevant) = some function of characteristics of document and query
  - E.g., logistic regression $\log(p_i/(1-p_i)) = X_i \beta$
- Explanatory variables
  - Terms in common
  - Query length
  - Collection size
  - Frequency of occurrence of term in document
  - Frequency of occurrence of term in collection
  - Rarity of term in collection



The advent of the web

- By mid-1990s algorithms were very mature
- Then the Web came along
  - IR researchers were slow to react
  - CS researchers were quick to react; NSF DL project
- Link structure of Web became new explanatory variable
  - PageRank = measure of how many important sites link to a given site
  - Also “anchor text” is very helpful
  - Improved relevance of search results significantly

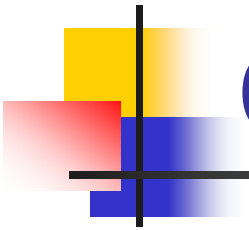


Google development

- Brin and Page tried to sell their algorithm to Yahoo for $1 million (they wouldn't buy) – assumed search was a commodity
- Formed Google with no real idea of how they would make money
- Put a lot of effort into improving algorithm
- Everybody else was convinced search had been commodified ...and there was no way to make money on it

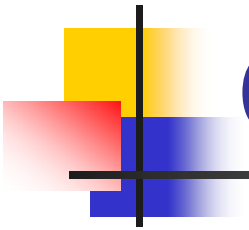
Business model

- GoTo Ad Auction – Pasadena
  - GoTo's model was to auction search result placement
  - Changed name to Overture, auctioned ads
  - Google liked the idea of an ad auction and set out to improve on Overture's model in Fall 2001
- Original Overture model
  - Rank ads by bids
  - Highest bidders get more prominent (higher up) slots
  - High bidder pays what he bid (1st price auction)



Google auction

- Rank ads by bid x expected clicks
  - Want ad with highest expected revenue in best position (price per click x number of clicks)
- Each bidder pays price determined by bidder below him
  - Price = minimum price necessary to retain position
  - Motivated by engineering, not economics
- Overture (now owned by Yahoo)
  - Adopted 2nd price model about same time
  - Currently moving to using expected revenue for ranking

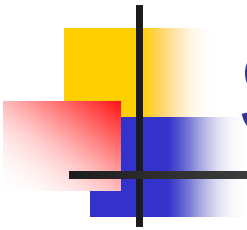


Google and game theory

- It is fairly straightforward to calculate Nash equilibrium of Google auction
  - Basic principle: in equilibrium each bidder prefers the position he is in to any other position
  - Gives set of inequalities that characterize equilibrium
  - Inequalities can be inverted to give values as a function of bids...
    - If you are in psn 3, were willing to pay price to get there, not willing to pay price to get to psn 2

Implications of analysis

- Basic result: *incremental cost per click has to be increasing in the click through rate.*
- Why? If incremental cost per click ever decreased, then someone bought expensive clicks and passed up cheap ones.
- Necessary and “sufficient” condition
- Similar to classic competitive pricing rule
  - Price = marginal cost
  - Marginal cost has to be increasing



Simple example

- Suppose all advertisers have same value for click v
  - Case 1: Undersold auctions. There are more slots on page than bidders.
  - Case 2: Oversold auctions. There are more bidders than slots on page.
- Minimum price paid
  - Case 1: The reserve price is r = 5 cents.
  - Case 2: Last bidder pays price determined by 1st excluded bidder.

Undersold pages

- Bidder in each slot must be indifferent to being in last slot

$$(v - p_s)x_s = (v - r)x_m$$

- Or

$$p_s x_s = r x_m + v(x_s - x_m)$$

- Payment for slot s = payment for last position + value of incremental clicks

Example of undersold case

- Two slots
  - x_1 = 100 clicks
  - x_2 = 80 clicks
  - v=.50
  - r=.05
- Solve equation
  - p_1 100 = .05 x 80 +.50 x 20
  - p_1 = 14 cents, p_2=r=5 cents
  - Revenue = .14 x 100 + .05 x 80 = $18

Oversold pages

- Each bidder has to be indifferent between being in his slot and not being shown:
- So $(v - p_s)x_s = 0$

$$p_s = v$$

- For previous 2-slot example, with 3 bidders, p_s=50 cents and revenue = .50 x 180 = $90
- Revenue takes big jump when advertisers have to compete for slots!
- Similar to Klemperer's example of Dutch mobile phone license auction

Why online business are different...

- Online businesses (Amazon, eBay, Google...) can continually experiment
  - *kaizen =* "continuous improvement"
- Hard to do *controlled experiments* with *product design* for traditional industries
  - Manufacturing – iPhone took 2 ½ years
- Very easy to do online
  - Leads to very rapid (and subtle) improvement
  - Learning-by-doing leads to significant competitive advantage for incumbents if they take advantage of it

Conclusion

- Marketing as the new finance
  - Data + computers + models
- Real time data allows for continuous improvement
- Market prices reflect value of ads (incremental price per click)
- Quantitative methods really work
- We are just at the beginning...