Panel: Inference from Internet Auction Data

Galit Shmuéli
Decision & Information Technologies Dept
Robert H Smith School of Business
University of Maryland, College Park

The Economics of Internet Auctions, FTC, 27 Oct 2005

Three (statistically significant) points

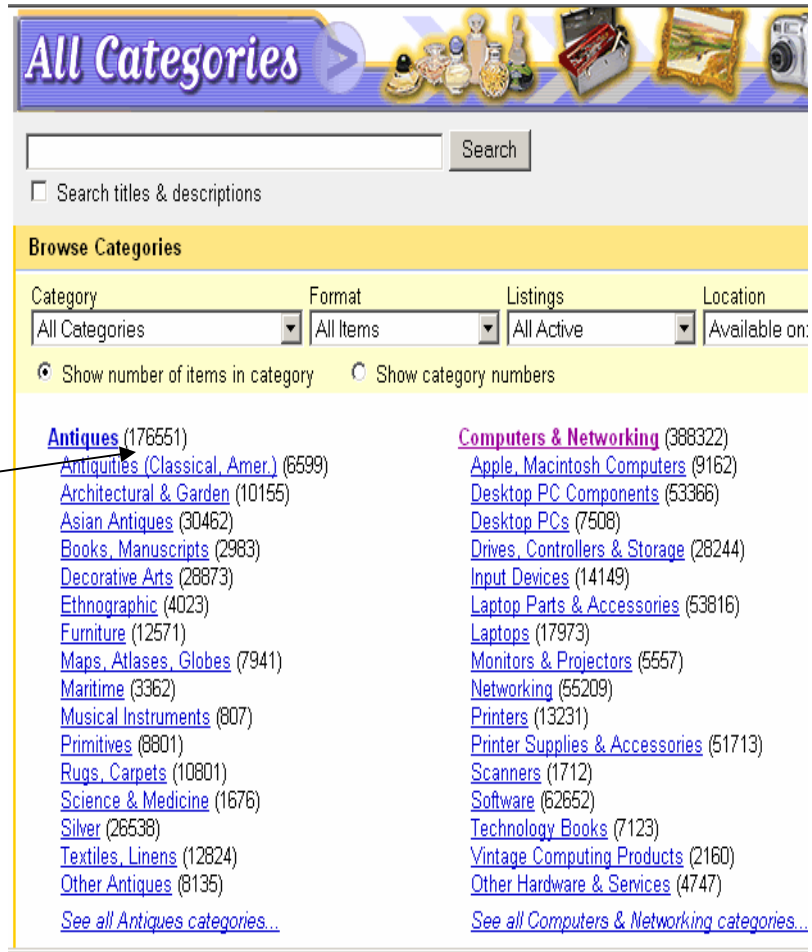
- Sampling is an issue
- Visualization is a necessity
- Auction dynamics are important

Sampling Auction Data

- Data collection via web agents
  - Get list of data pages (use website search / other)
  - Collect info from all data pages
- Non-sampling errors
  - Measurement bias (data refreshing, search algorithm, unrecorded data)
  - Selection bias (untruthful, poor website design, web-agent interference, robot restrictions)
- Sampling costs, errors, and issues
  - Fixed (programming) + very negligent per-item
  - Some items/categories harder to collect (legal/technical)
- Sample/Population
  - SRS assumption (is there a sampling frame?)
  - Stratified sampling: increase accuracy (estimate strata size, cat overlap)
  - Cluster sampling: account for concurrency, longitudinal
  - Goal of study (exploratory, descriptive, analytical), parameter of interest (difference?) tolerable error rates

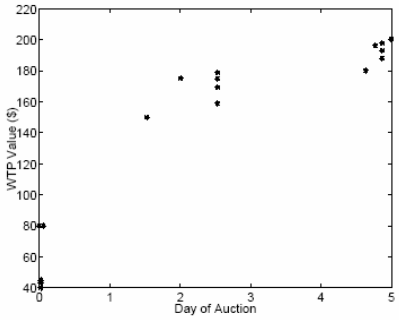
Example: What % of eBay auctions transact?

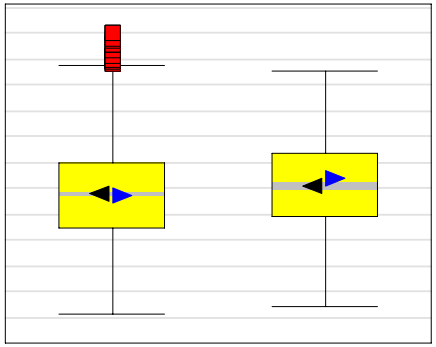
- Define population (when, where, type of transactions)
- Collect Sample
  - Stratify by category (top level)
  - Estimate sub-pop size from current category size
  - Look at all closed auctions, record % transacting, sample size
- Compute estimate + error

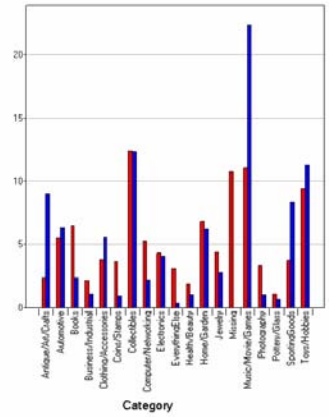


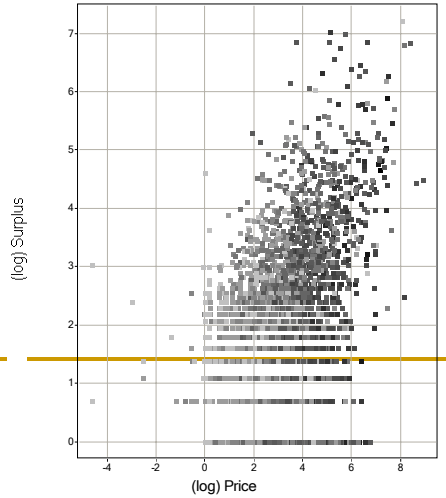
Visualization

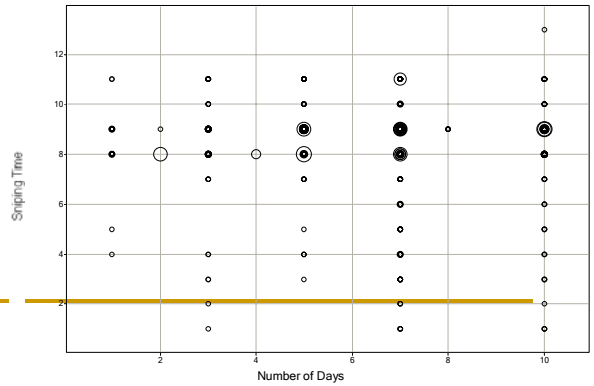
- Auction rules
- Skewed variables (price, rating, #bids…)
- Validation
- Interplay of multiple factors
- Discrete/cont



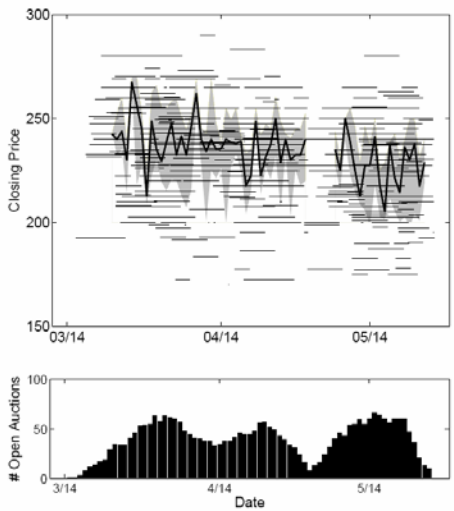


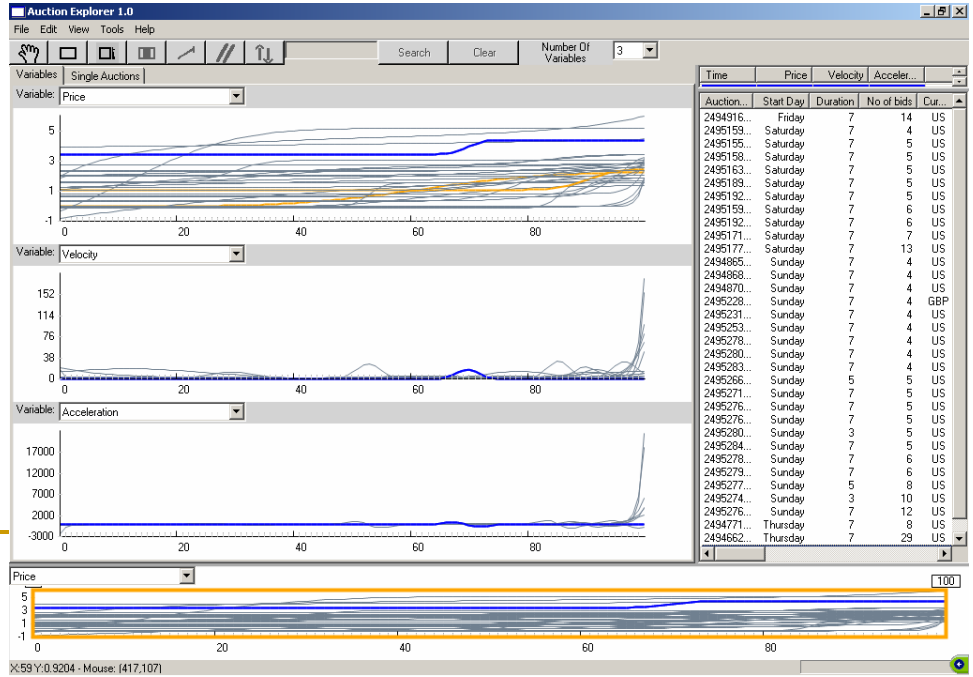






Visualization: special data structures





Information in Auction Data

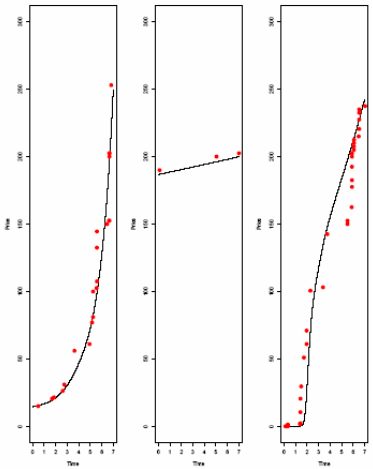
- Current research uses mostly a snapshot

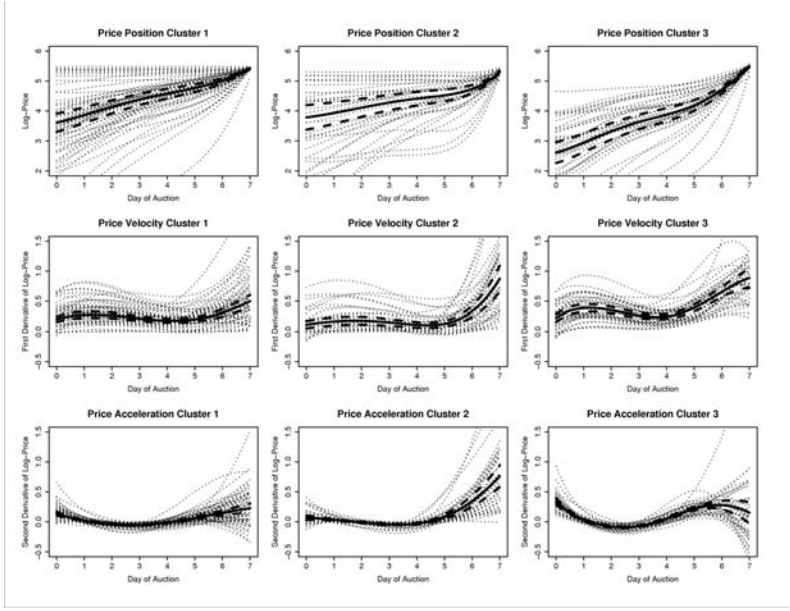
AuctionID	Price	# Bids	# Bidders	Open Bid	Duration	Category	winnerId	winnerRating	sellerId	sellerRating	Currency
2242934096	860	50	16	0.01	10	Tickets & Travel	meigsmike	4	alphapenguin	119	US
8104666657	9.57	7	6	0.5	10	Collectibles	firebird16	66	jscomics	5703	US
2243064894	62.99	14	11	0.99	10	Collectibles	moto26	853	sf.bay.trader	2760	US
2243273716	389	18	12	50	10	Collectibles	beancounte	32	morphling88	325	US
3723066321	6.99	1	1	6.99	7	Collectibles	wadeltons	5	prime-star	25433	US
2243514515	11.5	1	1	11.5	7	Collectibles	sullee-socc	125	prime-star	25433	US
2243538083	11.5	7	2	3.99	7	Collectibles	mminnie	98	mickeyjj77	81	CA
2243580975	6.5	3	2	2.99	7	Collectibles	sandyreser	177	rankens-toys	3385	US
3815779306	5.5	2	2	2.99	7	Collectibles	sandyreser	177	rankens-toys	3385	GBP
2243581144	5.5	2	2	2.99	7	Collectibles	sandyreser	177	rankens-toys	3385	US
3912166297	2.99	1	1	2.99	7	Collectibles	djleeds	56	rankens-toys	3385	AU
2243609083	74.75	12	8	5	7	Collectibles	junk-sales	136	deb*bee	387	US
4012045841	36	7	3	5	7	Collectibles	ufantastik%	33	deb*bee	387	US

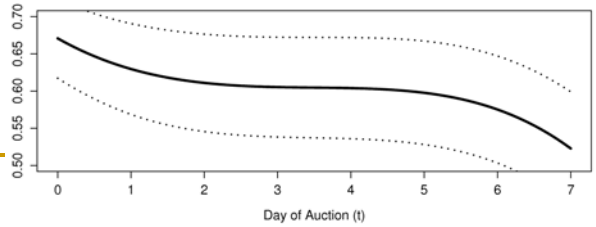
- Some of these change during the auction

Auction Dynamics

- What goes on *during* the auction is informative
- Price process and price dynamics
- Effect of various factors (open bid,…)
- Economic theory?
- Functional Data Analysis!







Final words

- Special data structures allow more advanced techniques, new Qs
- New findings can add directions for developing new (online) theory
- Exploratory data analysis is foundation of formal modeling