Guiding Consumers through Lemons and Peaches: Analyzing the Effects of Search Design Activities

Pedro Gardete
(Nova SBE)

Megan Antill
(Stanford University)

NBER IO Program Meeting - Feb 2020

Contribution

- Contribution: Assess the extent to which different information provision strategies affect consumer welfare and seller outcomes
  - Characterize consumer search for multiattribute products
  - Determine the empirical effects of endogenous search frictions in a mature Internet context
  - Determine how strategic consumers are while searching (Gabaix and Laibson 2006, Frazier et. al 2009, Powell 2010, Liang et. al 2017)

Contribution

- Contribution: Assess the extent to which different information provision strategies affect consumer welfare and seller outcomes
  - Characterize consumer search for multiattribute products
  - Determine the empirical effects of endogenous search frictions in a mature Internet context
  - Determine how strategic consumers are while searching (Gabaix and Laibson 2006, Frazier et. al 2009, Powell 2010, Liang et. al 2017)
- Challenge:
  - Assessing the effect of different information scenarios implies incorporating sophisticated consumer behavior
  - While go-to techniques (Weitzman, simultaneous search) are popular because they reduce the burden of estimating structural parameters...
  - ...they impose strict assumptions on consumer behavior, usually easily falsifiable

Research Goals

1. Propose a model of consumer search, which allows for the following innovations:
    - Allow for piecemeal search within each alternative
    - Allow for arbitrary relationships across observable characteristics
    - Allow for flexible search paths (while alleviating the resulting curse of dimensionality)
    - Identify and estimate an unobserved utility component (correlated with remaining characteristics)
    - Allow for forward-looking consumers

Research Goals

1. Propose a model of consumer search, which allows for the following innovations:
   - Allow for piecemeal search within each alternative
   - Allow for arbitrary relationships across observable characteristics
   - Allow for flexible search paths (while alleviating the resulting curse of dimensionality)
   - Identify and estimate an unobserved utility component (correlated with remaining characteristics)
   - Allow for forward-looking consumers
2. Once the consumer behavior fundamentals are recovered, we:
   - Consider various information provision strategies
   - Measure their welfare effects for buyers and seller
   - Pitch our model against a popular myopic learning analogue

Preview of Results

- Parameters
  - Preference; Search costs; Covariance matrix of characteristics
- Counterfactual Analysis:
  - Search frictions hamper conversion rates significantly
  - Information provision policies carry at most modest seller gains
  - Effects of information design are directionally similar for consumers and seller
  - Myopic model provides a poor approximation to the full dynamic problem

Context

SHIFT

- Online startup with unique business model:
  - Online customers can browse vehicles and schedule test drives online
  - Seller takes vehicle to potential buyer's address
  - Buyer takes it or leaves it; no bargaining

Context

- Online startup with unique business model:
  - Online customers can browse vehicles and schedule test drives online
  - Seller takes vehicle to potential buyer's address
  - Buyer takes it or leaves it; no bargaining
- Unique Dataset:

Context

SHIFT

- Online startup with unique business model:
  - Online customers can browse vehicles and schedule test drives online
  - Seller takes vehicle to potential buyer's address
  - Buyer takes it or leaves it; no bargaining
- Unique Dataset:
  - Each alternative has multiple search actions associated with it

Context

- Online startup with unique business model:
  - Online customers can browse vehicles and schedule test drives online
  - Seller takes vehicle to potential buyer's address
  - Buyer takes it or leaves it; no bargaining
- Unique Dataset:
  - Each alternative has multiple search actions associated with it
  - Researchers have access to the data viewed by consumers

Context

SHIFT

- Online startup with unique business model:
  - Online customers can browse vehicles and schedule test drives online
  - Seller takes vehicle to potential buyer's address
  - Buyer takes it or leaves it; no bargaining
- Unique Dataset:
  - Each alternative has multiple search actions associated with it
  - Researchers have access to the data viewed by consumers
  - Access to all consumer clickstream data and test-drive requests

Vehicle Information

Website organized in two tiers: Main listing page and Vehicle detail page

- Main Listing Page:
  - Search Filters
  - Vehicle Picture; Make/Model; Price; Year; Mileage

Vehicle Information

Website organized in two tiers: Main listing page and Vehicle detail page

- Main Listing Page:
  - Search Filters
  - Vehicle Picture; Make/Model; Price; Year; Mileage
- Vehicle Detail Page:
  - Additional vehicle pictures
  - Vehicle history (number of previous accidents; number of previous owners)
  - Inspection reports (number of inspection notes)

Dataset

Dataset

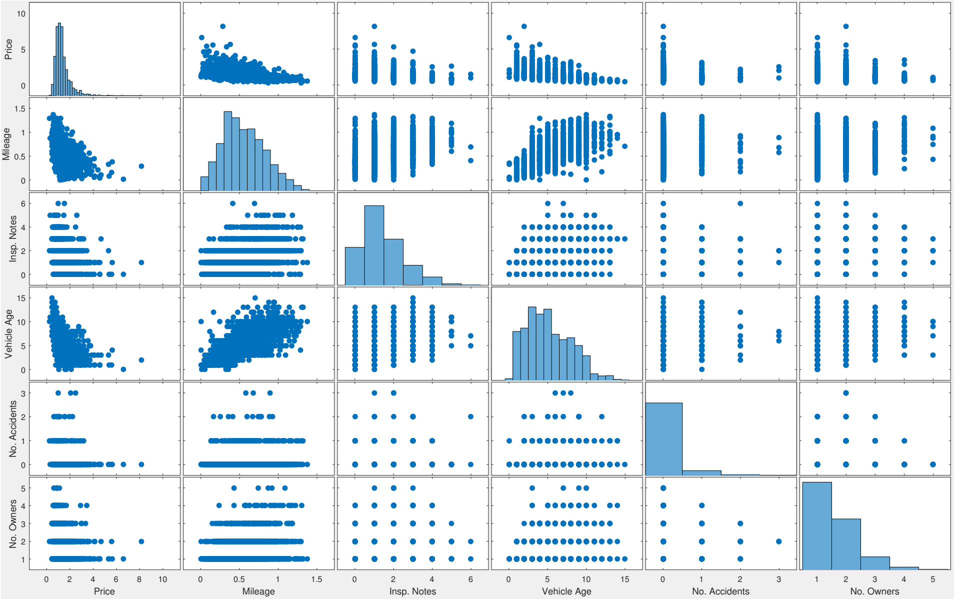
- Click stream data of 24,116 users, browsing 1,573 (sedan) vehicles
- Every event that occurs on the website, February 2016 - September 2016
- All vehicle characteristics, histories, and inspection reports
- Conversion variable: Ordered a test drive

Summary Statistics - User Level

	Mean	Std. Dev.	Min	Max
Vehicle Detail Pages Viewed	2.28	3.47	1	129
Number of Sessions	6.03	21.14	1	762
Photo Sets Browsed	1.63	2.65	0	74
Vehicle Histories Browsed	0.69	2.09	0	101
Inspection Reports Browsed	0.30	1.27	0	53
Average Session Duration (min.)	14.2	13.39	0.1	282.04
Activity Range (days)	18.67	38.5	0	232
Number of users: 24,116				

Note: Statistics are per user. The event "Photo Sets Browsed" equals one for each user-vehicle observation whenever a user browsed more 20 or more pictures of a vehicle.

Histograms and Scatter Plots of Vehicle Characteristics



Model - Utility and Beliefs

Search Model

- Ex-post consumption utility:

$$v_{ijt} = \sum_{k=1}^{K} \beta_k x_{jk} + \nu_{ij} + \epsilon_{ijt} \tag{1}$$

$$v_{i0t} = \epsilon_{i0t} \tag{2}$$

- i: visitor
- j: vehicle
- k: vehicle characteristic
- x_{jk}: level of characteristic k vehicle j
- β_{ik}: consumers preference over characteristics
- ν_{ij}: unobservable utility component
- ϵ_{ijt}: idiosyncratic utility component

Beliefs

- Consumers' beliefs are consistent with the empirical joint distribution of characteristics

$$\mathcal{G}_{make\_model,price,mileage,age,color,owners,accidents,notes} \tag{3}$$

Beliefs

- Consumers' beliefs are consistent with the empirical joint distribution of characteristics

$$\mathcal{G}_{make\_model, price, mileage, age, color, owners, accidents, notes} \tag{3}$$

- Use a statistical copula to model the joint distribution of characteristics flexibly, and provide a first-stage estimate of consumer beliefs:

$$C\left(F_{X_1}\left(x_1\right), \ldots, F_{X_K}\left(x_K\right)\right)$$

- Challenges (in the paper):
  - Make/model and color are not ordinal variables
  - Some variables are discrete
  - Incorporating unobserved utility component ν_{ij}

Model: Search

Mapping between search actions and learned characteristics

Knowledge Level	Characteristic	Utility Component
Observable ex-ante	**Ex-ante Characteristics**	
	Make/Model	$make\_model_j$
	Price	$price_j$
	Mileage	$mileage_j$
	Age	age_j
	Color	$color_j$
Learned upon visiting a vehicle's history	**Vehicle History**	
	# Previous Owners	$owners_j$
	# Accidents	$accidents_j$
Learned upon visiting inspection reports	**Vehicle Status**	
	# Inspection Notes	$notes_j$
Learned after browsing one photo set	**Perceived Quality/Fit**	
	Vehicle Photos	ν_{ij}

- State Variables:
  - Ω_i: What a consumer has learned about vehicle characteristics: x and ν

Search Model

- At each decision point, consumer i solves problem:

$$V\left(\Omega_i, \epsilon_i\right) = \max \left\{ \underbrace{V_0\left(\Omega_i, \epsilon_1\right), ..., V_J\left(\Omega_i, \epsilon_J\right)}_{\textit{Terminal Decisions}}, \underbrace{V_{1,s}\left(\Omega_i, \epsilon_1\right), ..., V_{J,s}\left(\Omega_i, \epsilon_J\right)}_{\textit{Search Decisions}} \Bigg|_{s=1..S} \right\}$$

Search Model

- At each decision point, consumer i solves problem:

$$V\left(\Omega_i, \epsilon_i\right) = \max \left\{ \underbrace{V_0\left(\Omega_i, \epsilon_1\right), ..., V_J\left(\Omega_i, \epsilon_J\right)}_{\textit{Terminal Decisions}}, \underbrace{V_{1,s}\left(\Omega_i, \epsilon_1\right), ..., V_{J,s}\left(\Omega_i, \epsilon_J\right)}_{\textit{Search Decisions}} \Big|_{s=1..S} \right\}$$

- Upon stopping search, consumer earns utility:

$$V_j\left(\Omega_i, \epsilon_i\right) = E\left(v_{ij} \middle| \Omega_i, \epsilon_i\right) \quad (4)$$

$$= E\left(\sum_{k=1}^{K} \beta_k x_{jk} + \nu_{ij} \middle| \Omega_{ij} \right) + \epsilon_{ij} \quad (5)$$

Search Model

- At each decision point, consumer i solves problem:

$$V\left(\Omega_i, \epsilon_i\right) = \max \left\{ \underbrace{V_0\left(\Omega_i, \epsilon_1\right), \ldots, V_J\left(\Omega_i, \epsilon_J\right)}_{Terminal\ Decisions}, \underbrace{V_{1,s}\left(\Omega_i, \epsilon_1\right), \ldots, V_{J,s}\left(\Omega_i, \epsilon_J\right)}_{Search\ Decisions} \Bigg|_{s=1..S} \right\}$$

- Upon stopping search, consumer earns utility:

$$V_j\left(\Omega_i, \epsilon_i\right) = E\left(v_{ij} \middle| \Omega_i, \epsilon_i\right) \tag{4}$$

$$= E\left(\sum_{k=1}^{K} \beta_k x_{jk} + \nu_{ij} \middle| \Omega_{ij}\right) + \epsilon_{ij} \tag{5}$$

- Utility of searching:

$$V_{j,s}\left(\Omega_i, \epsilon_i\right) = -c_s + E_{\epsilon_i', \omega_{js}}\left(V\left(\Omega_i \cup \omega_{js}, \epsilon_i'\right) \middle| \Omega_i\right) + \epsilon_{ij}^s \tag{6}$$

Estimation Approach

- Estimator needs to meet a few challenges:
  - Consumers take multiple actions, so likelihood of interest is of the search sequence
  - Each iteration solves as many value functions as consumers
  - State space is large ($\simeq 7.72 \times 10^{317}$ points; action space itself is 2^{3n})

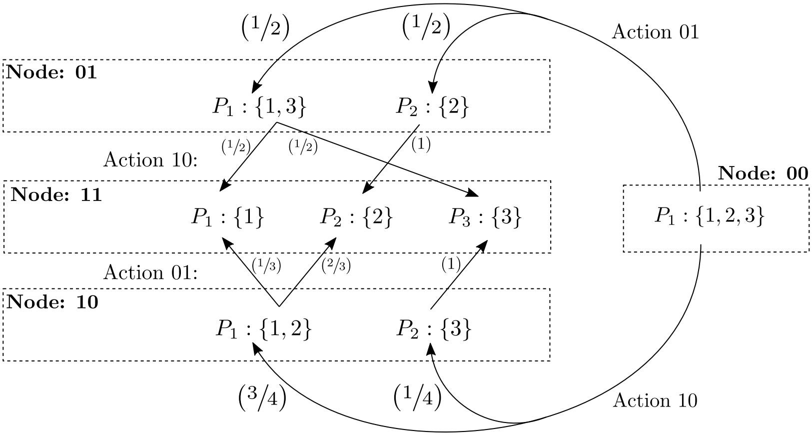
Estimation Approach

- Estimator needs to meet a few challenges:
  - Consumers take multiple actions, so likelihood of interest is of the search sequence
  - Each iteration solves as many value functions as consumers
  - State space is large ($\simeq 7.72 \times 10^{317}$ points; action space itself is 2^{3^n})
- State-Partitioning Estimator:
  - Take draws of characteristics from empirical distribution
  - Then, we assume the agent behaves as if she knew the drawn characteristics; she uses them to form beliefs
  - As agents search, they learn which drawn vehicles remain admissible, given what they have learned so far
  - The resulting tree is sparse and can be solved through backward induction
- Estimator provides a feasible approximation to the complete problem

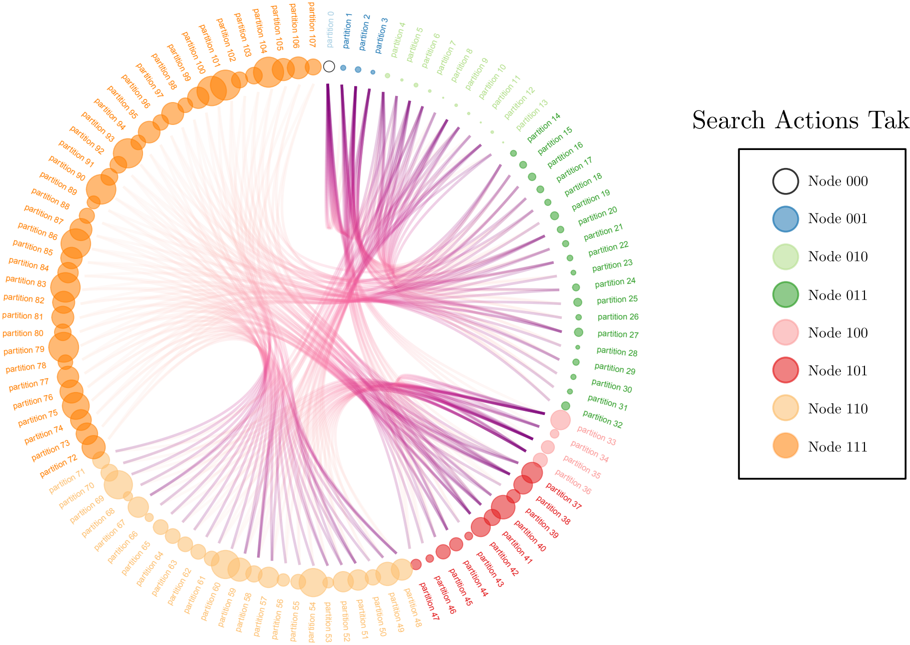
Example of Decision Tree (1 vehicle; 2 actions; 4 simulations)

Simulations

#	x_1	x_2	freq.
1)	1	2	1
2)	1	3	2
3)	2	2	1



Example of Decision Tree from Data (1 vehicle; 50 simulations)



Partitioning Decision Tree

- Advantages
  - Decision tree admits sparse representation when taking action A followed by B leads consumer to same state as taking action B followed by A
  - Can handle arbitrarily correlated unobserved characteristic (support is based on a sparse grid, as in Heiss and Winschel, 2008)

Partitioning Decision Tree

- Advantages
  - Decision tree admits sparse representation when taking action A followed by B leads consumer to same state as taking action B followed by A
  - Can handle arbitrarily correlated unobserved characteristic (support is based on a sparse grid, as in Heiss and Winschel, 2008)
- Limitations
  - The dimensionality curse is alleviated, but not completely mitigated (in our case, $n \leq 8$)

Partitioning Decision Tree

- Advantages
  - Decision tree admits sparse representation when taking action A followed by B leads consumer to same state as taking action B followed by A
  - Can handle arbitrarily correlated unobserved characteristic (support is based on a sparse grid, as in Heiss and Winschel, 2008)
- Limitations
  - The dimensionality curse is alleviated, but not completely mitigated (in our case, $n \leq 8$)
- Implementation aspects (aka, "nerd corner")

Partitioning Decision Tree

- Advantages
  - Decision tree admits sparse representation when taking action A followed by B leads consumer to same state as taking action B followed by A
  - Can handle arbitrarily correlated unobserved characteristic (support is based on a sparse grid, as in Heiss and Winschel, 2008)
- Limitations
  - The dimensionality curse is alleviated, but not completely mitigated (in our case, $n \leq 8$)
- Implementation aspects (aka, "nerd corner")
  - Non-vectorizable and not GPU-friendly; tree is represented by pointers to objects
  - Core decision-tree code: ~3,000 lines in C++ code (~7,000 total)
  - Code is parallelizable

Identification

- Preference parameters
  - Relationship between characteristics and purchase rates

Identification

- Preference parameters
  - Relationship between characteristics and purchase rates
- Search costs
  - Average search intensity of each search option

Identification

- Preference parameters
  - Relationship between characteristics and purchase rates
- Search costs
  - Average search intensity of each search option
- Covariance parameters
  - Variance of unobservable utility component is normalized to one
  - Covariance parameters are identified by search rates conditional on known characteristic levels

Results

Results: Preference Parameters

Table: Model Estimates: Preference Parameters

	Parameter Estimates
Vehicle Characteristics	
price	-0.372**
	(0.000)
mileage	-0.19**
	(0.000)
notes	-0.36**
	(0.000)
age	-0.01
	(0.348)
accidents	0.032
	(0.434)
owners	-1.611**
	(0.000)
Log-Likelihood: -38,937.35	
Make-Model and Color Dummies ✓	
N= 12,887	

Note: Standard errors in parentheses. Significance levels: † p≤0.10, * p≤0.05, ** p≤0.01.

Results: Search Costs

Table: Model Estimates: Search Cost Parameters

	Parameter Estimates
Segment 1	
c_{photos}	2.306**
	(0.031)
$c_{vehicle\ history}$	0.248**
	(0.048)
$c_{insp.report}$	5.255**
	(0.094)
Segment 2	
c_{photos}	0.034*
	(0.02)
$c_{vehicle\ history}$	4.382**
	(0.05)
$c_{insp.report}$	3.309**
	(0.022)
Other Parameters	
α	-0.732**
	(0.01)
Make-Model and Color Dummies	✓
N= 12,887	

Note: Significance levels: † p≤0.10, * p≤0.05, ** p≤0.01. Size of segment 1, implied by $\widehat{\alpha}$, is 67.5%.

Results: Variance-Covariance Parameters

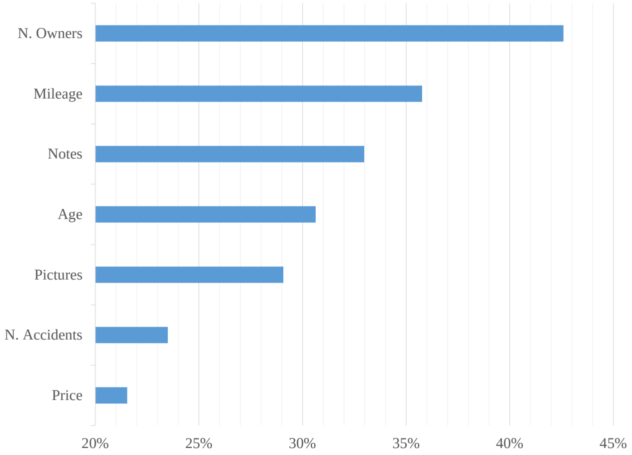
Table: Estimated Variance-Covariance Matrix

$$\widehat{\Sigma}_0 = \begin{pmatrix} 1 & -0.66 & -0.31 & -0.73 & -0.07 & -0.21 & -\mathbf{0.44} \\ -0.66 & 1 & 0.37 & 0.71 & 0.08 & 0.27 & \mathbf{0.76} \\ -0.31 & 0.37 & 1 & 0.41 & 0.07 & 0.14 & \mathbf{0.67} \\ -0.73 & 0.71 & 0.41 & 1 & 0.09 & 0.24 & \mathbf{0.61} \\ -0.07 & 0.08 & 0.07 & 0.09 & 1 & 0.021 & \mathbf{0.1} \\ -0.21 & 0.27 & 0.14 & 0.24 & 0.021 & 1 & \mathbf{0.55} \\ -\mathbf{0.44} & \mathbf{0.76} & \mathbf{0.67} & \mathbf{0.61} & \mathbf{0.1} & \mathbf{0.55} & 1 \end{pmatrix} \Bigg| \begin{pmatrix} \textit{Price} \\ \textit{Mileage} \\ \textit{Insp.notes} \\ \textit{Vehicle age} \\ \textit{No.accidents} \\ \textit{No.owners} \\ \nu \end{pmatrix}$$

Note: In bold, cross-correlation elements induced by estimated parameters $s_1..s_6$. Rightmost vector shows the corresponding characteristics. Non-bold estimates all significant at 1% level. Coefficients in bold: † p$\leq$0.10, * p$\leq$0.05, ** p$\leq$0.01.

% Utility Variance Explained by each Characteristic

Figure: Utility Variance Decomposition



Note: Above, values obtained by simulating $Var_{x_j}\left[E\left(u|\, x_j\right)\right] \div Var\left(u\right)$.

Results: Model Fit

Table: Search Moments from Dataset and Model

	Data				Model Prediction			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Conversion Rate	0.047	0.213	0	1	0.055	0.228	0	1
N. Searches	1.639	3.22	1.082	9	1.433	1.153	0.5	6.054
Vehicle Histories	0.441	0.743	0	4	0.397	0.316	0.108	1.844
Inspection Reports	0.175	0.496	0	4	0.161	0.184	0	1.157
Photo Sets	1.023	0.873	0	4	0.875	0.664	0.268	3.376

N: 12,887 consumers

Counterfactual Analyses

Counterfactual Analyses

1. Effects of different information design policies
2. Comparison with myopic search policy

Counterfactual Analysis: Information Design and Conversions

Table: Conversion Effects of Exchanges of Attribute Visibilities

	Vehicle Detail Page Attribute	
Front Page Attribute	Vehicle History	Inspection Report
Mileage	+0.36% (+6.58%)**	-0.11% (-1.91%)
Age	+0.36% (+6.55%)**	-0.02% (-0.33%)
Price	+0.63% (+11.34%)**	-0.00% (-0.03%)
N: 12,887 consumers		

Note: Above, impact of exchanging attributes in the front page with the ones in the vehicle detail page on mean conversions. Relative changes in parentheses.

Significance levels: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

Counterfactual Analysis: Information Design and Welfare

Table: Welfare Effects of Exchanges of Attribute Visibilities

	Vehicle Detail Page Attribute	
Front Page Attribute	Vehicle History	Inspection Report
Mileage	+1.65% (+1.73%)**	-0.39% (-0.41%)**
Age	+1.76% (+1.85%)**	-0.11% (-0.12%)
Price	+2.35% (+2.48%)**	+0.02% (+0.01%)*
N: 12,887 consumers		

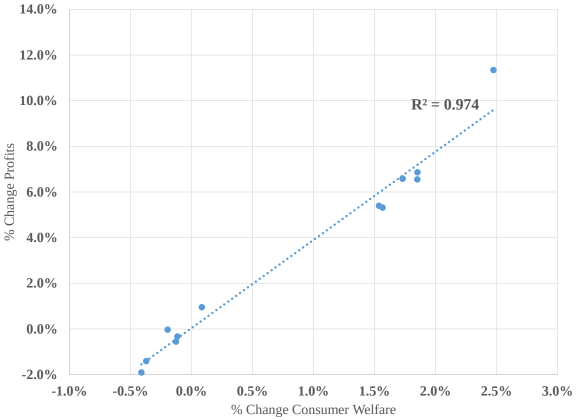
Note: Above, impact of exchanging attributes in the front page with the ones in the vehicle detail page on the number of searches. Relative changes in parentheses. Significance levels: † $p \leq 0.10$, * $p \leq 0.05$, ** $p \leq 0.01$.

Discussion

- Conversions can increase moderately through information design
- Consumer welfare effects are low, but are often in line with the seller's own welfare
- Contrasts with ideas related to price obfuscation activities:
  - *...Internet tools will help consumers to find and to process information, but retailers may simultaneously harness the power of the Internet to make information processing problems more formidable...* - Ellison and Ellison, 2012
- We may be at a point where prices are nonetheless competitive, and so there is little incentive to not provide better matches, from the seller's point of view

Alignment between Consumer and Firm preferences

Figure: Information Design Effects: Consumer Welfare and Profits



Note: High correlation found between consumer and seller welfares

Myopic Policy

- Knowledge gradient approach (Frazier et. al 2009, Powell 2010); Myopic policy (Liang et. al 2017)
  - Decision makers are assumed to look 'one period' ahead
  - They search only if the immediate option value is high enough
  - Otherwise, pick a terminal action

Myopic Policy

- Knowledge gradient approach (Frazier et. al 2009, Powell 2010); Myopic policy (Liang et. al 2017)
  - Decision makers are assumed to look 'one period' ahead
  - They search only if the immediate option value is high enough
  - Otherwise, pick a terminal action

- Model:

$$V_j\left(\Omega_i, \epsilon_i\right) = E\left(v_{ij} \middle| \Omega_i, \epsilon_i\right) \tag{7}$$

$$\begin{aligned} V_{j,s}\left(\Omega_i, \epsilon_i\right) = & -c_s + \\ & E_{\epsilon_i^{'}, \omega_{js}}\left[\max\left\{v_{i0}, V_1\left(\Omega_i, \epsilon_i^{'}\right), .., V_j\left(\Omega_i \cup \omega_{js}, \epsilon_i^{'}\right), .., V_J\left(\Omega_i, \epsilon_i^{'}\right)\right\} \middle| \Omega_i\right] + \epsilon_{ij}^{s} \end{aligned} \tag{8}$$

Assumption of Strategic Behavior Fits Data Best

Table: Consumer Behavior Statistics Compared with Myopic Model

	Data				Knowledge Gradient			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Conversion Rate	0.047	0.213	0	1	0.057	0.233	0	1
N. Searches	1.639	3.22	1.082	9	1.16	0.65	0.46	4.319
Vehicle Histories	0.441	0.743	0	4	0.279	0.163	0.057	1.143
Inspection Reports	0.175	0.496	0	4	0.119	0.084	0	0.694
Photo Sets	1.023	0.873	0	4	0.762	0.428	0.257	2.699

N: 12,887 consumers

- Vuong (1989) non-nested test rejects that both models are "at the same distance" from true model, $p < 0.01$

Conclusion

- Propose a flexible model of search, which allows for the following distinctive features:
  - Incremental search over multiple characteristics
  - Rich joint distribution of characteristics
  - Flexible search paths
  - Unobserved utility component
  - Forward-looking consumers

Conclusion

- Propose a flexible model of search, which allows for the following distinctive features:
  - Incremental search over multiple characteristics
  - Rich joint distribution of characteristics
  - Flexible search paths
  - Unobserved utility component
  - Forward-looking consumers
- Assessed implications of search frictions and search design
  - Search design has moderate conversion effects ($\simeq$6%)
  - Search design tends to make sellers more better off, when compared to buyers (influence takes place on the margin)
  - In our setting, search design has similar directional effects for seller and consumers, in contrast with early predictions (e.g., Ellison and Ellison 2012)

Thank you