Search with Learning in the Retail Gasoline Market

Xiaosong (Andy) Wu — University of Melbourne
Matthew S. Lewis — Clemson University
Frank A. Wolak — Stanford University

DSE
17 December 2022

Standard search models

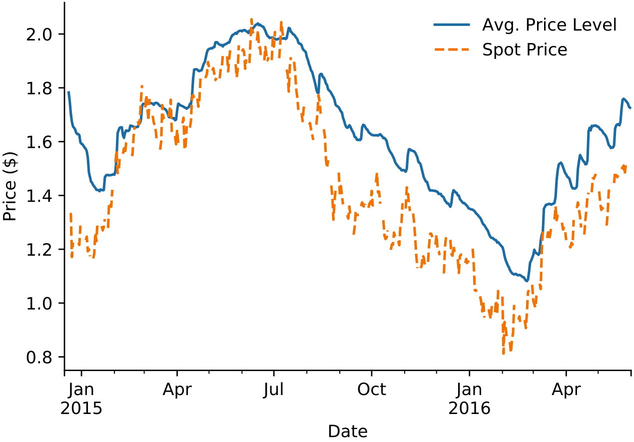
- Since Stigler 1961, consumer search models have shaped how we think about consumer behaviors and competition in many markets.
- A large empirical literature identifies search costs based on a key assumption: Consumers **know** the price distribution (e.g. Hortacsu and Syverson 2004, Hong and Shum 2006, De los Santos et al. 2012, Wildenbeest 2011, Honka 2014).
- The **known** price distribution is convenient because it allows econometricians to derive search benefits from the empirical price distribution, but it is a strong assumption.
    - Assumes consumers have a correct initial price belief.
    - Assumes consumers do not update their beliefs with new information.

Prior beliefs / learning

- Consumers are unlikely to know the price distribution with any uncertainty if consumers are unfamiliar with the market or the market conditions are changing rapidly.
- By eliciting consumer price beliefs, experimental papers find that consumers have prior beliefs different from the actual price distribution and update their beliefs as they obtain new information (i.e. Matsumoto and Spence 2016 and Jindal and Aribarg 2021).
- Ignore learning can lead to biased search cost estimates (i.e. Koulayev 2013 and De los Santos et al. 2017).
- Learning affects search behaviors and market outcomes, as the benefit and the cost of search together determine search behaviors.
  - Lewis (2011) finds reduced-form evidence that the observed price patterns in retail gasoline market are consistent with a model where consumers have price beliefs different from the actual price distribution.

Price beliefs and asymmetric cost pass-through

Figure: Average City Gasoline Price and Wholesale Cost



Estimating search with learning

- We relax the known price distribution assumption and estimate a search with learning model (Rothschild 1974) in the retail gasoline market.
- The learning process has 2 components:
    1. Consumers are allowed to formulate prior price beliefs based on past observed prices.
    2. Consumers are allowed to update their beliefs as they search along their driving trips before purchase decisions.
- We estimate parameters governing the search and learning process using a novel panel dataset of station-specific prices and quantities combined data that describe the traffic flows in the city.

Estimating the learning process

- Ideally, identifying learning requires $P(Search_{i,t}|p_{i,t}, p_{i,t-1}, ...)$.
- When individual search histories are not observed, integration of all unobserved search sequences can cause the curse of dimensionality problem.
- We overcome this challenge by leveraging a crucial observation in the retail gasoline market: price search is not random but **ordered**.
  - Use traffic data to simulate search sequences.
  - Total daily gasoline sales at each station is an aggregation of the purchase decisions of individuals searching and learning along different driving trips.
- Given these search routes, the joint variation of the sequence of prices and market shares identifies learning.

Preview of results

- Learning is a crucial component of price search in the retail gasoline market.
- Consumers place a significant weight on past prices when formulating their prior beliefs. The average absolute difference between consumer price beliefs and the actual price level is 2.7 cpg.
- Consumers are uncertain about their prior beliefs and learn quickly. One (two) new price observation(s) can reduce the prior bias by 77% (87%).
- Ordered search generates more realistic substitution patterns across stations that depend on shared traffic.
- Demand elasticities change over time with fluctuations in price beliefs, consistent with the rockets and feathers phenomenon.
- Learning can results in negative cross-price elasticities for some station pairs.

Data

- The market is a mid-sized city in the US, spanning from December 14, 2014 to May 31, 2016
  - Web-scraped daily gasoline price for all of the 46 gas stations in the city.
  - Obtained daily total gasoline transaction for 33 of the 46 stations from a major credit card company.
- Observe the empirical distribution of traffic flows
  - Obtained the Origin and Destination Table from the Department of Transportation.
  - Generated search routes (an ordered set of stations).

Figure: A Travel Route

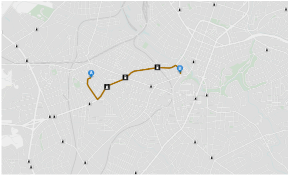


Table: Search Routes

Search Route	No. Drivers
(1, 4, 2)	2500
(4, 2)	800
...	...

Data

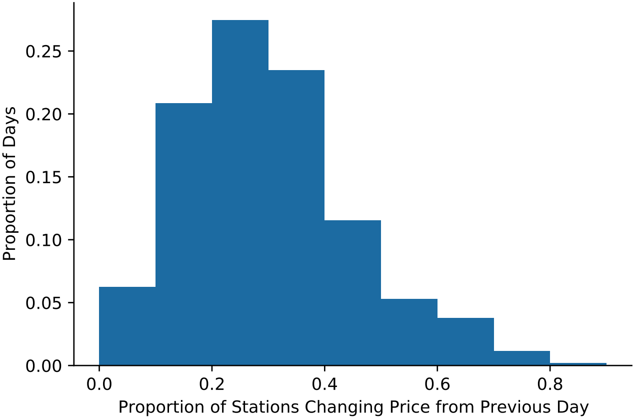
Table: Summary Statistics of the Station Characteristics

	Obs.	Mean	SD	25%	50%	75%
Panel (a): All Stations						
Avg. Price ($)	46	1.61	0.06	1.58	1.59	1.68
Avg. Quantity (gl.)	33	978.09	1207.78	235.76	396.09	1525.02
Major Brands	46	0.37	0.49	0.00	0.00	1.00
Number of Islands	46	3.59	1.73	2.00	3.00	5.00
Easy Left-Turns	46	0.26	0.44	0.00	0.00	0.75
No Left-Turns	46	0.28	0.46	0.00	0.00	1.00
Direct Traffic (1,000s)	46	11.52	4.99	8.32	10.66	15.09

- 10% increase in traffic volume corresponds to a 12.6% increase in sales.
- An average driver sees 3.5 prices along a travel route.

Sources of Price Uncertainty

Figure: Proportion of Stations Changing Price From Previous Day



Sources of price uncertainty

Isolate different sources contribute to the uncertainty,

$$p_{jt} = \sum_{j=2}^{J} \psi_j Station_j + \sum_{t=1}^{T} \gamma_t Day_t + \nu_{jt},$$

- ψ_j measures the persistent price differences across stations
- γ_t is the price change common to all station
- $\nu_{j,t}$ is the price change specific to a station on a day
  - Chandra and Tappata (2011) show that price rankings vary significantly over time
- $\tilde{p}_{j,t} = \gamma_t + \nu_{j,t}$ time-variant price

Table: Summary Statistics of Relative Price and Price Level

	Obs.	Mean	SD	Min	Max
Relative Price Changes					
ν_{jt}	23732	0.000	0.032	-0.167	0.169
Price Level Changes					
$abs(\Delta\gamma_t)$	528	0.008	0.012	0.000	0.108

Descriptive evidence on price distribution uncertainty

	(1)	(2)	(3)
ln(Own Price)	-2.632***	-2.961***	-2.942***
	(0.147)	(0.160)	(0.163)
ln(Past Price Level)	1.220***	1.132***	1.133***
	(0.108)	(0.109)	(0.116)
ln(1st Neighbor Price)	1.464***	1.171***	1.124***
	(0.152)	(0.162)	(0.163)
ln(2nd Neighbor Price)		0.757***	0.786***
		(0.144)	(0.146)
ln(Total Sales 1-Day Ago)			0.234***
			(0.043)
ln(Total Sales 2-Day Ago)			-0.076*
			(0.042)
ln(Total Sales 3-Day Ago)			-0.057
			(0.042)
ln(Total Sales 7-Day Ago)			-0.060
			(0.038)
R^2	0.928	0.928	0.929
Observations	15985	15985	14960

- The positive and significant coefficient on past price level suggests that consumers formulate their prior beliefs using past prices.
- Past prices influence current demand through prior beliefs but not through purchase timing.

Model: utility

Consumers' indirect utility for one unit of gasoline (10 gallons) at station j on day t is equal to

$$u_{j,t} = X_j \beta - p_{j,t}$$

where X_j are station j's non-price characteristics, and $p_{j,t}$ is the unit price of gasoline. Divide the utility into 2 parts,

$$u_{j,t} = \underbrace{X_j \beta - \psi_j}_{\text{Known: } V_j} - \underbrace{(\gamma_t + \nu_{j,t})}_{\text{Unknown: } \tilde{p}_{j,t}}$$

We make the following assumption,

- Consumers know the mean utility, V_j prior to search.
- Consumers search to realize $\tilde{p}_{j,t} = \gamma_t + \nu_{j,t}$ and **Bayesian update** their beliefs on the distribution of $\tilde{p}$.

Model: learning (Bayesian updating)

- Time-variant prices $\tilde{p}_{j,t} \sim N(\gamma_t, \sigma^2)$.
- Consumers are uncertain about the price level, but know σ^2.
- Prior beliefs about the average price level are normal

$$N(\mu_0, \frac{\sigma^2}{\alpha_0}),$$

  where μ_0 is the prior mean and $\frac{\sigma^2}{\alpha_0}$ is the prior uncertainty and α_0 is commonly known as the prior weight.
- After observing n prices $(x_1, x_2, ..., x_n)$, a consumer's posterior belief is $N(\mu_n, \frac{\sigma^2}{\alpha_0+n})$, where

$$\mu_n = \frac{\alpha_0}{\alpha_0 + n}\mu_0 + \frac{1}{\alpha_0 + n}\sum_{k=1}^{n} x_k.$$

Model: learning (Bayesian updating)

Posterior mean depends on the prior mean and prior weight,
$\mu_n = \frac{\alpha_0}{\alpha_0+n}\mu_0 + \frac{1}{\alpha_0+n}\sum_{k=1}^{n} x_k$.

- μ_0 may not equal to γ_t
- α_0 determines the speed of learning.
- The larger the α_0, the slower the update
  - Consumers believe that the observed price change is more likely to be specific to that station
  - When $\alpha_0 = \infty$, then consumes do not learning
- The smaller the α_0, the faster the update
  - Consumers believe that the observed price change is more likely to be common to the market

Model: ordered search

- Consumers are forward-looking.
- On day t, consumers are assigned to a search route based on the empirical distribution of search routes
- Consumers each search and learn along their search routes
- Search is ordered with no recall
- Sampling prices along a search route is assumed to be free
- There is a cost to postpone purchase to a future trip
  - Postponement cost is consumer specific
- Consumers decide at which station to purchase gas by comparing the realized utility at a station with the continuation value of search

Model: ordered search

- We introduce for a turn cost, τ_{rn}, the ex-ante known mean utility is $V_{rn} = V_{r(n)} - \tau_{rn}$.
- At station $n < N_r$, the value function after observing its price x is

$$W_{irn}\left(\mu_{rn}, x_{rn}\right) = \max\Bigg\{V_{rn} - x_{rn},$$
$$\underbrace{\int W_{irn+1}\left(h\left(\mu_{rn}, n+1, x_{rn+1}\right), x_{rn+1}\right) \cdot dF_{rn}\left(x_{rn+1}\right)}_{z_{rn}(c_i|\theta) - \mu_{rn}}\Bigg\},$$

- At the final station N_r,

$$W_{irN_r}\left(\mu_{rN_r}, x_{rN_r}\right) = \max\Bigg\{V_{rN_r} - x_{rN_r},$$
$$\underbrace{-c_i + \sum_{r' \in \mathbb{R}} \lambda_{r'} \cdot \int W_{ir'1}\left(h\left(\mu_{rN_r}, 1, x_{r'1}\right), x_{r'1}\right) \cdot dF_{rN_r}\left(x_{r'1}\right)}_{z_0(c_i|\theta) - \mu_{rN_r}}\Bigg\}.$$

- We can numerically solve for $z_{rn}(c|\theta)$, $\forall r, n$ using policy function iteration.

Model: purchase decision

Postponement costs and purchase probability

- Consumers purchase at different stations along a route because they have heterogeneous postponement costs.
- We can solve for a critical postponement cost c^*_{rn}, where $z_{rn}(c^*_{rn}|\theta) - \mu_{rn} = V_{rn} - x_{rn}$.
- Postponement cost lower bound: purchase from the nth station if $c_i \geq c^*_{rn}$.
- Postponement cost upper bound: purchase from the nth station if consumer i has not purchased from any stations before n, $c_i < c^{**}_{rn} = \min(c^*_{r1}, ..., c^*_{rn-1})$.
- Assume postponement cost distribution has a probability mass of $1 - \eta_t$ at 0. Positive postponement costs follow a log-normal distribution, purchase probability conditional on a route is

$$q_{rnt} = \begin{cases} \eta_t \left(\Phi(lnc^{**}_{rnt} - \mu_c) - \Phi(lnc^*_{rnt} - \mu_c) \right) & \text{if } c^{**}_{rnt} > c^*_{rnt} \\ 0 & \text{otherwise.} \end{cases}$$

Estimation

Estimation assumption: purchases are made by a new group of drivers each day.

Prior belief

- $\mu_{0t} = \pi\gamma_{t-7} + (1-\pi)\gamma_t$, where π is the weight on past prices.
- We estimate the weight on past prices (π) and prior weight (α_0), allowing us to empirically test the different models.
- if $\pi = 0$, correct belief, $\mu_{0t} = \gamma_t$.
- if $\pi = 1$ and $\alpha_0 = \infty$ reference price search model (Lewis 2011).
- if $\pi = 0$ and $\alpha_0 = \infty$, standard search model with a known price distribution.

Non-linear least squares

- Sum up the purchase probability at a station over all the routes it is on to obtain the model predicted market share.
- Estimate the parameters that minimize the squared distance between the model predicted and observed market share.

Results

Table: Estimation Results

	Learning		No Learning	
	Coeff.	SE	Coeff.	SE
Prior				
Bias (π)	0.587	(0.061)	0.027	(0.039)
Learning (α_0)	0.304	(0.116)		
Station Attributes				
Major Brand 1	0.505	(0.041)	-0.065	(0.115)
Retail Brand 1	-0.131	(0.041)	-0.168	(0.043)
Retail Brand 2	-0.064	(0.053)	-0.322	(0.094)
Small-Sized Station	-1.173	(0.043)	-1.719	(0.099)
Large-Format Station	0.662	(0.058)	0.904	(0.052)
Left-Turn Cost	1.106	(0.060)	1.767	(0.152)
Postponement Cost				
Constant (μ_c)	-0.574	(0.063)	-0.288	(0.061)
Standard Deviation (σ_c)	1.055	(0.092)	1.250	(0.095)
Pseudo-R^2	0.888		0.870	

Results

- The weight placed on past prices (0.59) suggests that the average abs. difference between the estimated prior mean and the actual price level is 2.7 cpg.
  - Approximately 3.3 times the size of the average abs. day-to-day changes in price level.
- A small prior weight (0.3) suggests fast learning. One new price observation reduces the initial difference by 77%.
  - The magnitude of the estimated prior uncertainty is similar to the short term variation in price levels.
  - On average, consumers have a decent understanding of the volatility of price level variation in the market.
- The median postponement cost is 56 cents for 10 gallons of gas purchased.
  - Ignore learning overestimates median postponement cost by 33%.

Own-price elasticities

Table: Summary of the Station Average Own-Price Elasticity Estimates

	Obs.	Mean	SD	Min	50%	Max
Panel (a): Learning						
Price ($\tilde{p}_{j,t}$)	37	-8.39	4.68	-24.40	-7.66	-2.67
Price Reputation (ψ_j)	37	-24.40	12.71	-60.74	-20.99	-8.70

- An average station's own-price elasticity with respect to a change in the time-variant price is -8.
- An average station's own-price elasticity with respect to a change in the price reputation is -24.
- The elasticities are comparable with the literature (Wang 2009).

Spatial competition and cross-price elasticities

Table: Regression Results of Estimated Cross-Price Elasticities on Distance Measures Between Stations

	(1)	(2)	(3)
Driving Distance	-0.067	0.001	-0.000
	(0.015)	(0.005)	(0.006)
Abs. Mean Utility Distance	-0.061	-0.082	-0.093
	(0.024)	(0.021)	(0.021)
Common Traffic		4.286	
		(0.945)	
Common Traffic Easy Access			5.144
			(1.167)
Common Traffic Costly Left-Turn			1.748
			(0.616)
Constant	0.574	0.008	0.026
	(0.106)	(0.054)	(0.057)
R^2	0.03	0.11	0.13
Observations	1665	1665	1665

Negative cross-price elasticities

Table: Summary Statistics on Cross-Price Elasticities

	Obs.	Mean	SD	2.5%	10%	50%	90%	97.5%
Cross-Elasticity	1665	0.148	1.221	-0.229	-0.003	0.000	0.163	1.529

Figure: Competing Stations Along a Hypothetical Travel Route

$$\xrightarrow{\qquad\qquad\qquad\qquad\qquad\qquad}$$

1 2 3

- Consider an example where all 3 stations have a positive share of consumers, $+\infty > c_1^* > c_2^* > c_3^*$.
- $q_1 = 1 - G(c_1^*)$, $q_2 = G(c_1^*) - G(c_2^*)$, and $q_3 = G(c_2^*) - G(c_3^*)$, where $c_1^* = z_1^{-1}(V_1 - \frac{\alpha_0}{\alpha_0+1}x_1 + \frac{\alpha_0}{\alpha_0+1}\mu_0)$, $c_2^* = z_2^{-1}(V_2 - \frac{\alpha_0+1}{\alpha_0+2}x_2 + \frac{1}{\alpha_0+2}x_1 + \frac{\alpha_0}{\alpha_0+2}\mu_0)$, and $c_3^* = z_3^{-1}(V_3 - \frac{\alpha_0+2}{\alpha_0+3}x_3 + \frac{1}{\alpha_0+3}(x_1 + x_2) + \frac{\alpha_0}{\alpha_0+3}\mu_0)$.
- Substitution: $x_1 \uparrow$, $c_1^* \uparrow$, $G(c_1^*) \uparrow \Rightarrow q_1 \downarrow$ and $q_2 \uparrow$
- Learning: $x_1 \uparrow, c_2^* \downarrow$, and $c_3^* \downarrow, G(c_2^*) \downarrow, G(c_3^*) \downarrow, \Rightarrow q_2 \uparrow$ and q_3?
- Negative cross-price elasticities arise when two stations are never immediate competitors but are 3.7 stations away from each other.

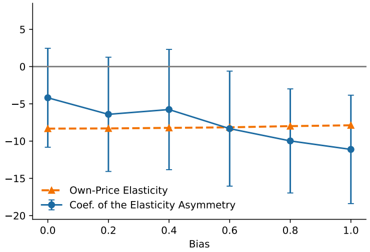
Biased priors and asymmetric search behaviors

Table: Search Behavior and Demand and Past Prices

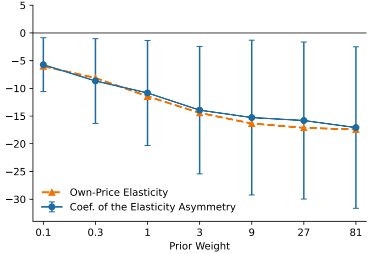
	Share of Purchase (1)	No. Station Searched (2)	Demand Elasticity (3)
$\Delta\phi_{t-7}$	-0.272	1.909	-8.619
	(0.008)	(0.052)	(3.971)
Constant	0.607	2.624	
	(0.000)	(0.003)	
R^2	0.74	0.69	0.23
Observations	522	522	18726

- Sharper increases or decreases in price level can cause prior beliefs to be more different from actual prices, potentially leading to larger asymmetry in search behaviors and demand elasticities.
- We construct two measures of search intensity: the share of searching consumers who buy and number of stations searched conditional on purchase.
- Fluctuations in own-price elasticities drive fluctuations in profit margins associated with asymmetric cost pass-through.
- When prices increase, consumers search more and demand become more elastic.

Counterfactuals - priors and asymmetric demand elasticities



(a) Prior Bias and Elasticity Asymmetry



(b) Prior Uncertainty and Elasticity Asymmetry

- Examine the relationship between the learning primitives and the demand asymmetry.
- Panel (a) suggests the larger prior bias results in more asymmetric demand elasticity.
- Panel (b) suggests that conditional a biased prior, the slower learning results in more asymmetric demand elasticity.

Conclusion

- Estimate a model of optimal search by consumers who are uncertain about the price distribution.
- Leverage traffic data to simulate search orders.
- Ordered search can identify the learning process using only aggregated data in a consumer search framework.
- Show learning is an important component of price search in the retail gasoline market.
- Establish an important relationship between learning primitives and search and demand asymmetry.