ECON 441 - Group Project

Maura Coughlin

Spring 2021

Admin Details:

Topic: Group Project - Platform Design in Online Shopping
Due Date: April 29, 2021 at 11:59PM Houston Time
Submission Guidelines:

- All group members are expected to participate and contribute.
- Your group should compile a single file of the group's answers, written in your own group's words (keep this in mind if you talk about the project with other groups) and a single Matlab file. Answers should be typed up.
- This assignment will be graded much more closely for content and accuracy than the homework (which to some extent prioritize effort and correct approaches rather than correct conclusions). As such, do not rush through the written answers - they are arguably the main component of the project.
- Per usual, do not distribute any material without my explicit permission. This assignment is based on Dinerstein et al (2018).
- To submit, collect your Matlab files and typed answers (as a PDF) into a single zip file, named 'ECON441_GroupProject.zip' and email it to me at maura.coughlin@rice.edu by the deadline with subject 'ECON441 Project'. Cc all group members, as well as Rajarshi.

Assignment Notes:

The Matlab portion of this project does not include any data. Rather you will gain experience with another common coding practice among IO economists - coding up the details of a simplified model and solve the model by evaluating differing parameters yourself. This practice is distinct from solving a model to learn empirical results, but is very helpful for understanding the intuition of a model. Note some of the loops you writ may take longer to compute than you are used to. You will also be asked about an empirical application, but will not be coding that portion.

Part 1: Motivation

In recent years (and even more so in the last year), the prevalence of online shopping has grown immensely. Online retailers such as Amazon account for a growing share of all retail purchases. Online shopping provides many conveniences for consumers but also introduces additional aspects to the process of selecting a good to purchase. Imagine you are looking to purchase more Bounty paper towels and you turn to Amazon. If you search for "Bounty paper towel" you will receive well over 100 search results which include products that relate to your search term in varying degrees. How Amazon presents those results to you may greatly impact the purchase you ultimately make.

1. Why would an economist expect online shopping platforms to lead to less price dispersion for comparable products? Why might that not have occurred in practice?
2. Does it seem likely that consumers evaluate and compare the utility of all available relevant products when making an online purchase? Why or why not? Provide a discussion of markets where you think this is more likely to be true and markets where you think it is less likely to be true.
3. Explain at least two ways in which the only platform design can impact the ultimate purchase decision of the consumer (meaning what other than consumer preferences and product prices and characteristics might drive the purchase choice).
4. There are three main parties involved in such online shopping: the sellers of the goods, the online platform, and the consumers. Explain the economic objectives of each of these agents.
5. Detailed browsing data from eBay indicates that consumers often only evaluate a small number of products, even in cases where there are many available offerings. Does this surprise you? Why or why not? How does this sort of behavior impact our standard approach to estimating demand for differentiated products?

Part 2: Equilibrium Model of Consumer Search and Price Competition

Products and sellers (we treat them interchangably) are denoted by subscript $j = 1, .., J$ (seller j offers product j). Each product j is defined by a fixed vector of attributes x_j and is offered on eBay for a seller-chosen price, p_j. Each seller faces cost c_j of producing product j.

Consumers, indexed by subscript i, have characteristics $\zeta_i \sim F(\cdot)$ and make the decision of which single unit (if any) to buy and receive utility $u_{ij} = u(x_j, p_j; \zeta_i)$.

The model is based on eBay's online platform and reflects the role of the platform in impacting both consumer search and sellers' pricing incentives. **The platform underwent an interesting algorithm redesign.** Prior to the change, eBay would show consumers individual offers from a larger set of offers potentially matching their search term using an algorithm that ranked the offerings by relevance. After the redesign, there were two steps. In the first, after entering a search term consumers picked the exact product of interest. In the second step, eBay compared listings of that product ranked mostly by price.

1. Consider the algorithm redesign.

   (a) Explain how the redesign impacts the consumer search process. What are the likely effects of the redesign on consumer purchasing behavior relative to before?

   (b) Explain how the redesign impacts the incentives of sellers in the online market. What would you expect to be the effect of the policy on the number of different offerings and posted prices?

   To formalize the role of the platform in this otherwise standard demand set up, we introduce the notion of consumer attention/awareness of different products. We define $a_j(p_j, x_j; p_{-j}, x_{-j}) \in [0, 1]$ as the awareness/visibility function of product j for all consumers on the platform. You can think of a_j as the probability a consumer is presented with product j in their search results. The platform charges the seller a fixed per transaction fee T and retains a fraction t of the selling price. Assume the platform does not otherwise enter into consumer utility.

2. A consumer makes a choice of product from a choice set L that may not include all of the available products j. We can define demand in terms of the choice set L by $D_j(p_j, p_{-j}; L) = \int \mathbb{1}(u(x_j, p_j; \zeta_i) \geq u(x_k, p_k; \zeta_i) \forall k \in L) dF(\zeta_i)$. Letting a_L be the probability of a given choice set L arising, write total demand for product j, $D_j(p_j, p_{-j})$, using the above defined $D_j(p_j, p_{-j}; L)$.

3. Further assume that the probability event of any given product appearing in choice set L is an independent event. Using this assumption, write a_L as a function of individual product level a_j's.

4. Using the relevant notational details given, define firm j's profit maximization problem, take the first order condition, and solve for the profit maximizing price.

5. Note that $\frac{\partial D_j(p_j, p_{-j})}{\partial p_j} = \sum_L a_L \frac{\partial D_j(p_j, p_{-j}; L)}{\partial p_j} + \sum_L \frac{\partial a_L}{\partial p_j} D_j(p_j, p_{-j}; L)$. Interpret the two right hand side effects and relate the interpretation to how platform design can incentivize price decreases. How does this relate to the roles of online platforms?

   For estimation, consider the following utility specification (where "TRS" is in dummy variable indicated whether product j is sold by a top-rated seller and q_j is a measure of product quality that vertically distinguishes products) for a targeted product. A "targeted product" is the focal product of the market (e.g. the product that corresponds to the user search term), such as Bounty paper towels in the introduction of this project document. Assume that the error term is an iid draw from a type 1 extreme value distribution.

$$u_{ij} = \alpha_0 + \alpha_1 p_j + \alpha_2 TRS_j + \alpha_3 p_j TRS_j + \alpha_4 q_j + \epsilon_{ij}$$

6. Describe what each term in targeted product utility captures and where appropriate, how they are distinct from the other terms in utility.

For the non-targeted products, utility of listing m is given by the following. λ refers to the degree of horizontal differentiation of various non-targeted products.

$$u_{im} = \delta + \lambda\epsilon_{im}$$

7. Discuss the interpretation of all of the non-targeted utility components.

Denote the consumer's consideration set (the set of products they consider when deciding on a purchase) as J_i, including J_i^J and J_i^M representing the targeted and non-targeted listings in the consideration set. The model includes a stochastic model of how listings are selected onto a page displaying results. Each product is associated with a sampling weight determining its likelihood of appearing in the displayed results. Before the policy change, sampling weights were set to be equal to measured quality, $\omega_j = q_j$. After the redesign, the weight became a function of price, $\omega_j = \exp\left[-\gamma\left(\frac{p_j - \min_{k\in\mathcal{J}_i^J}(p_k)}{std_{k\in\mathcal{J}_i^J}(p_k)}\right)\right]$. Pricing is assumed as a standard Nash Equilibrium

Part 3: Simplified Toy Model and Matlab Exercise

To illustrate the key features of the model, consider a highly simplified example with just two products ($J = 2$). Each product is associated with a product quality $q_1 > q_2$, prices p_1, p_2, and production costs c_1, c_2. Normalize $q_2 = 0$ and $c_2 = 0$ such that $q_1 = q$ and $c_1 = c$. Consumer utility is defined in this example as $u_{ij} = \zeta_i + q_j - \alpha p_j$ and $u_{i0} = 0$. Consumer characteristics ζ_i is distributed uniformly on $[0, 1]$. For simplicity, assume the platform shows a single product to consumers and the probabilities of showing each product (which is a function of prices and quality) are $a_1 \in [0, 1]$ and $a_2 = 1 - a_1$.

Sellers set prices to maximize profits and see demand as a function of consumer preferences and the platform strategy:

$$D_j(p_j, p_{-j}) = \begin{cases} a_j(p_j, p_{-j}) & \text{if } p_j < q_j/\alpha, \\ a_j(p_j, p_{-j})(1 + q_j - \alpha p_j) & \text{if } p_j \in [q_j/\alpha, (1 + q_j)/\alpha], \\ 0 & \text{if } p_j > (1 + q_j)/\alpha \end{cases}$$

1. Explain with a combination of math and words why demand is this three-piece demand function. It is most helpful to think of $j = 1$.

   Assume that the platform bases the visibility parameters on the prices and quality (estimated by the platform with noise, σ) in the following form:

$$a_1 = \frac{[\exp(q - \beta p_1)]^{1/\sigma}}{[\exp(q - \beta p_1)]^{1/\sigma} + [\exp(-\beta p_2)]^{1/\sigma}}$$

   and $a_2 = 1 - a_1$.

2. Discuss the role of β in the platform strategy.

3. In your Matlab code (using the HelpCode as an initial starting point), recreate the six figures in the figure below, "Comparative Statics," to illustrate the tradeoffs associated with different platform strategies (values of β).

   (a) The HelpCode provides details, such as which and how many parameter values to consider. Please refer to that code.

   (b) To do so, you fill fix the other parameters:

   $$\alpha = .5, \sigma = 1, q = 1, c = .5, T = .3, t = .1$$

   Then, at different possible values of β, you will use the included functions func_find_prices.m and func_foc_costs.m to solve for the resulting equilibrium prices and resulting profits and consumer surplus.

   (c) You will then use the resulting prices to solve for the corresponding visibility parameters, labeled below as $A1$ and $A2$.

4. For each of the six subfigures, provide an economic interpretation of the plot. (Note, if you are unable to recreate the subfigures, you can still answer this question based on the include figure). Be sure to also note any competing effects present across different aspects of this simplified model.

5. Similarly, in your Matlab code, recreate the four figures in the figure below "Optimal Betas" that display the optimal value of β, denoted β^* varies with the assumed parameter values of α, c, q, and σ. Note optimal refers to the value of the parameter that maximizes consumer surplus.

   (a) The HelpCode provides details, such as which and how many parameter values to consider. Please refer to that code.

   (b) To do so, you will vary only one parameter at a time and follow similar steps as before for solving the equilibrium value of β^*.

6. For each of the four subfigures, provide an economic interpretation of what the figure shows.

Part 4: Empirical Results and Main Findings

In this final section, you will be asked to interpret estimates of the main model (not the simplified toy model) and the effects of the platform redesign. In some instances (which are specified), analysis focuses on a single product market - that for the highly specific video game of Halo Reach.

1. Below is a table "Summary Statistics" that includes select statistics for transactions in certain product categories before and after the platform redesign. Using information in this table, describe any patterns that stand out to you, both about consumer behavior/attention and the potential effect of the policy change.

   Consider now the specific product searches for Halo Reach.

2. The figure below titled "Consideration Effects" documents consumer search behavior before and after the algorithm redesign. Interpret the effects of the redesign on consumer consideration. How does this agree or disagree with your expectations based on the model?

3. The table below "Estimation Results" presents the main model estimation results (standard errors in parentheses). The "Before" column refers to the model estimates based on the data prior to the platform redesign. The "Predicted after" column uses the model primitives estimated in the before analysis, applied to the new platform structure.

   (a) Comment on the estimates of price elasticities, both before the redesign and the change in elasticities resulting from the redesign.

   (b) Interpret and comment on the estimates of the utility parameters.

   (c) Comment on the change in the platform parameter estimates from the redesign.

   (d) Comment on the change in the purchase rates from the redesign.

4. The predicted effects of the redesign (based on the estimates from before the redesign) can be compared with the observed data after the redesign. The figure "Price Distribution" below compares the distribution of list prices for Halo Reach before the redesign and after (both predicted and observed). Based on this figure, what do you think about the model's suitability to evaluate the effect of the redesign? Comment on the impact of the redesign on the price distribution and what this might mean for consumer welfare.

5. These estimates presented are for the specific product category of the Halo Reach video game. How would you expect the results to change if instead the model was estimated on a product category with more product differentiation, such as cell phones?

6. Although the role of the platform to steer consumers to see and consider certain product offerings is rather clear, such forces also exist outside of the online retail setting. Discuss such possible influences on consumer purchases of in-store purchases. Do you think these forces help or harm consumers?

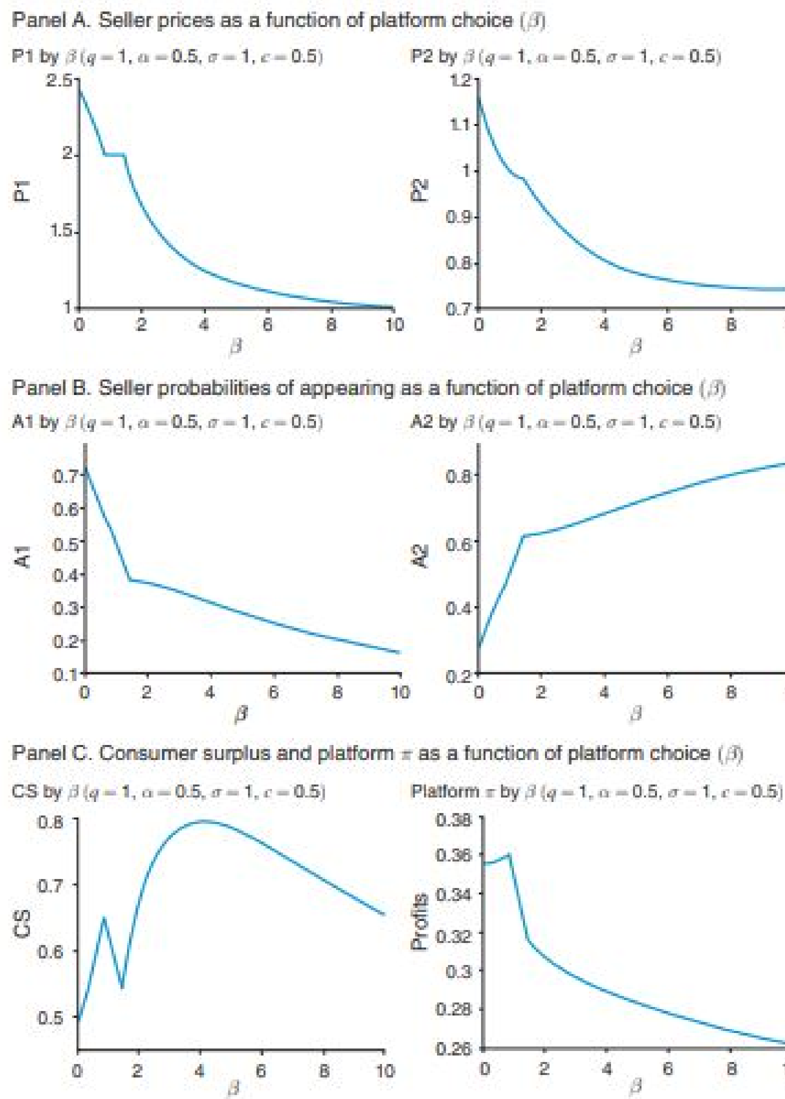

Figure 1: Comparative Statics

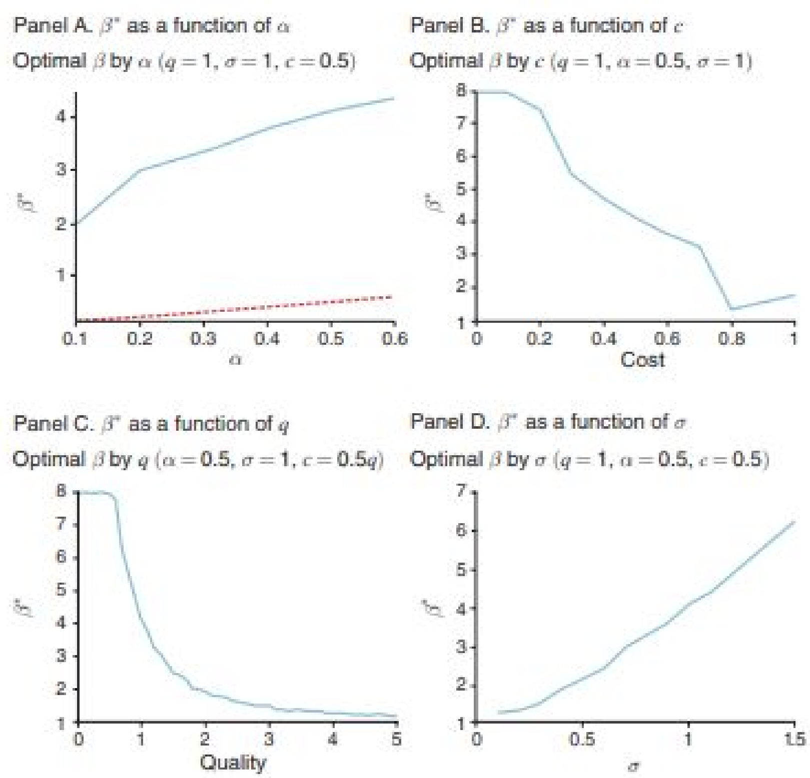

Figure 2: Optimal Betas

		Cell Phones	Digital Cameras	Textbooks	Video Game Systems	Video Games	iPhone 4
Average number of active listings	Before	23.33	40.64	16.27	35.41	38.93	29.49
75th/25th percentile of							
Posted prices	Before	1.22	1.32	1.61	1.39	1.47	1.28
Transacted prices	Before	204.10	1.31	1.25	1.29	1.29	138.44
Average price percentile of bought items	Before	40.11	31.85	29.82	17.34	27.82	37.40
	After	37.79	24.94	20.75	19.72	19.22	33.07
	Change	−2.32	−6.90	−9.07	2.38	−8.61	−4.33
		(0.98)	(1.21)	(1.76)	(0.84)	(0.61)	(0.62)
Posted prices ($, mean)	Before	562.45	1,418.31	67.98	290.00	48.60	749.25
	After	462.88	1,170.81	63.11	285.08	48.14	567.74
	Change	−99.57	−247.50	−4.86	−4.92	−0.45	−181.50
		(11.47)	(13.14)	(2.61)	(5.32)	(0.84)	(12.04)
Transacted prices ($, mean)	Before	412.30	1,162.51	51.03	222.62	45.71	676.72
	After	403.75	980.06	42.89	257.16	42.17	554.82
	Change	−8.55	−182.46	−8.14	34.53	−3.53	−121.90
		(11.26)	(10.78)	(0.59)	(3.47)	(0.40)	(6.83)
Number of transactions	Before	1,762	650	482	1,045	3,873	2,605
	After	3,594	3,108	3,941	1,666	2,537	4,346
	Change	1,832	2,458	3,459	621	−1,336	1,741
TRS share of transactions (%)	Before	43.87	70.92	27.39	43.92	27.11	36.62
	After	39.12	78.93	45.27	42.62	44.42	40.52
	Change	−4.75	8.00	17.88	−1.31	17.31	3.90
		(1.44)	(1.93)	(2.18)	(1.96)	(1.22)	(1.20)

Figure 3: Summary Statistics

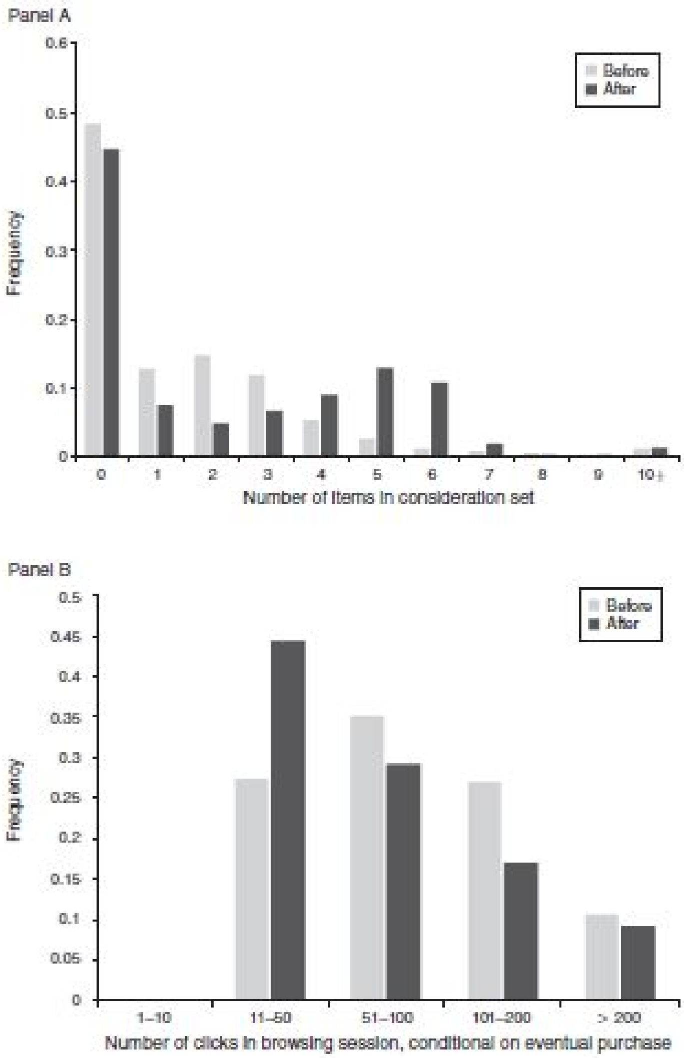

Figure 4: Consideration Effects

	Before	Predicted after
Platform Parameters		
Average number of listings on the site	21	28
Average number of TRS listings on the site	3	10
Prob. of a single-targeted item consideration set	0.13	0.08
Estimated gamma		0.80
		(0.14)
Median prob. of appearing in a search		0.08
Prob. of appearing if lower price by 10%		0.11
Demand		
Constant (*Halo Reach* new fixed price)	3.72	
	(1.07)	
Price	-0.24	
	(0.03)	
Top-rated seller (TRS)	4.13	
	(2.75)	
Price × TRS	−0.10	
	(0.08)	
Quality (0 to 1)	0.64	
	(0.36)	
Constant (Other listings)	−8.37	
	(0.41)	
Size of epsilon (Other listings)	1.70	
	(0.14)	
Implied Price Elasticities		
Average own–price elasticity	−10.64	−13.53
Average own–price TRS elasticity (TRS)	−14.16	−17.09
Average own–price TRS elasticity (Non-TRS)	−9.95	−12.84
Supply		
Median price − cost (TRS)	$2.35	$1.92
Median margin (percent of P) (TRS)	0.06	0.05
Median price − cost (Non-TRS)	$3.48	$2.61
Median margin (percent of P) (Non-TRS)	0.09	0.07
Purchase Rates (percent)		
(*Halo Reach*) observed	1.03	1.23
(*Halo Reach*) predicted	0.93	1.47
(Other) observed	1.96	2.63
(Other) predicted	1.27	1.85
(Buy Box, conditional on HR purchase) observed		64
(Buy Box, conditional on HR purchase) predicted		65

Figure 5: Estimation Results

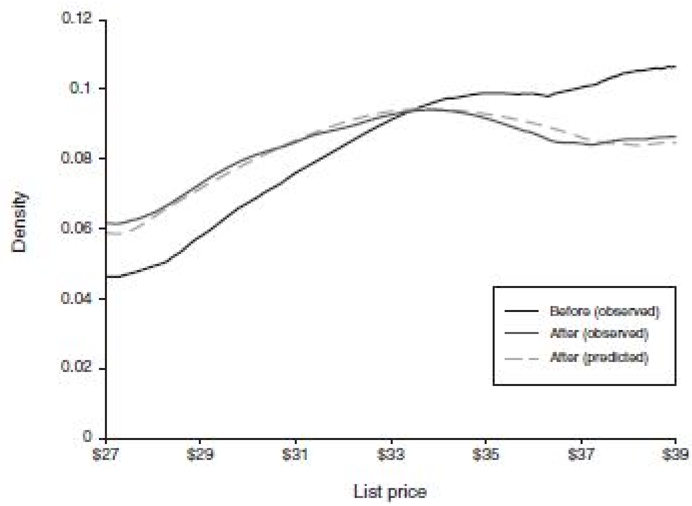

Figure 6: Price Distribution