## ECO410H: Practice Questions 6

- 1. Read Hausman and Leonard (2002): start by reading the abstract and then browsing through the sections headings and tables to get an overall sense and then read it. Keep going until the end even if you feel confused. Next are some detailed questions to help guide you through. *The solutions either refer you to the article or contain an outline of the answer, but not detailed explanations.* You are expected to read the article: it is well-written and provides many detailed and intuitive explanations.
  - (a) Consider Equation 3 on page 245.
    - i. How many observations (rows) are in the data used to estimate this equation?
    - ii. How many times is this equation estimated?
    - iii. What is  $\alpha_i$ ? How many variables does this term add to this equation? What kind of variables are they?
    - iv. What is  $W_i$ ? How many variables does this term add to this equation? What kind of variables are they?
    - v. What is  $I_{it}$ ? How many variables does this term add to this equation? What kind of variables are they?
    - vi. What is the interpretation of the estimated parameters  $\delta_1$ ?
    - vii. What is the estimate of the  $\delta_1$  for Angel Soft? Is it statistically significant?
    - viii. What is the estimate of the  $\delta_1$  for Scot Tissue? Is it statistically significant?
    - ix. Can you use the estimated parameters  $\delta_1$  to measure a causal relationship?
    - x. What if you found out that on average prices in all markets had been declining prior to the introduction of KBT and that the trend simply continued? Could this explain why the authors obtain a negative value for  $\delta_1$ ?
    - xi. What if you found out that Kleenex introduced KBT earliest in markets where it expected the strongest demand (hence making it a strong competitor and forcing its rivals to lower prices a lot)? Could this explain why the authors obtain a negative value for  $\delta_1$ ?
    - xii. What is the economic and econometric importance of the fact that Kleenex rolled out its product in waves? Could  $\delta_1$  be estimated if KBT introduced in all cities simultaneously? Explain.
    - xiii. What if firms manufacturing competing brands of bath tissue lowered their prices at the moment KBT is introduced but then raised them back up for all subsequent weeks. If this is the pattern in the data, what would the estimated parameter  $\delta_1$  be?
    - xiv. What is  $M_{it}$ ? How many variables does this term add to this equation? What kind of variables are they?
    - xv. What is the approximate value of  $\sum_{i=1}^{30} \sum_{t=1}^{196} M_{it}$ ?
    - xvi. What is the interpretation of the estimated parameter  $\delta_2$ ?
    - xvii. What is the estimate of the  $\delta_2$  for Angel Soft? Is it statistically significant?
    - xviii. What is the estimate of the  $\delta_2$  for Scot Tissue? Is it statistically significant?
    - xix. Does it make sense to add up  $\delta_1$  and  $\delta_2$ ? What would be the interpretation?

- xx. Does this equation seek to estimate the direct or indirect effect of KBT on competitors' prices? Explain.
- (b) What percent of variation in log prices is not explained by variation in  $W_t$ ,  $I_{it}$ , and  $M_{it}$  for Angel Soft brand bath tissue? For which brands is most of the variation in log prices explained by variation in  $W_t$ ,  $I_{it}$ , and  $M_{it}$ ? Give an intuitive explanation based on what the article says about these brands.
- (c) Consider Equation 5 on page 248. [Hint: To answer these questions you may find it useful to look at the actual estimation results reported in Appendix Table 1 (may make it seem more concrete and less abstract).]
  - i. How many observations (rows) are in the data that could possibly be used to estimate this equation? (Give an upper bound estimate as exact one is time consuming to compute.)
  - ii. How many observations (rows) did Hausman and Leonard actually use? (Note: There is a typo on page 250 of this article it should say "For the 17 wave ONE cities" not "wave THREE cities.")
  - iii. Why didn't they use all of the available data?
  - iv. How many times is this equation estimated?
  - v. Given that expenditure shares over the seven brands must sum to one, there must be one brand serving as the omitted category in estimation. Which did the authors specify as the omitted brand?
  - vi. How many variables does the term  $\alpha_{in}$  add to this equation?
  - vii. How many variables does the term  $Z_{nt}$  add to this equation? Are they all dummy variables?
  - viii. What is the expected sign of the  $\gamma_{ij}$  parameters when i = j and  $i \neq j$ ? Are the estimation results consistent with expectations?
- (d) Look at Appendix Table II. Is there evidence that without month-year dummies there is an endogeneity bias? (Hint: Compare own-price elasticity estimates and think about the direction of bias caused by unobserved demand shifters.)
- (e) Using Equation 9 on page 250 calculate the own-price elasticity for Kleenex in the city of Charlotte. Do a rough check of your answer with that reported in Table III on page 251 (which is a rough check because Table reports average over all cities). [Hint: To do this calculation you will need to use estimates in Appendix Table I and Table I (page 244) and refer back to Equations 5, 6, and 7.]
- (f) Using Equation 9 on page 250 calculate the cross-price elasticity between Kleenex and Cottonelle in the city of Charlotte. [Hint: To do this calculation you will need to use estimates in Appendix Table I and Table I (page 244) and refer back to Equations 5, 6, and 7.]
- (g) What is the meaning of the "virtual prices" reported in Table IV and how are the useful in estimating the consumer benefit from the increased variety created by a new good?
- (h) In Table IV (page 255) the estimate of 3.5% that appears in the lower left corner is an estimate of what?

- (i) In Table VII (page 262) what is the interpretation of the 7.3% estimate?
- (j) When backing out costs using the Bertrand NE assumption, are the authors assuming equal constant marginal costs across products (c) or do they allow for differences  $(c_i)$ ?
  - i. Why would assuming constant and EQUAL marginal costs across brands be a bad idea in the bath tissue industry?
- 2. Now that you have worked through many detailed questions, consider these bigger-picture questions about Hausman and Leonard (2002).
  - (a) Describe three important facts about their data.
  - (b) How do they obtain the marginal costs of producing Kleenex bath tissue? Explain.
  - (c) Which specific model of firm behavior do they use? What are the choice variables of the firms? What equilibrium do they assume the market is in?
  - (d) Do the authors need to assume a particular model of firm behavior to estimate the price effect of the KBT introduction?
  - (e) What is the overall purpose of estimating the two equations below?

$$s_{int} = \alpha_{in} + \beta_i \log Y_{nt} / P_{nt} + \sum_{j=1}^{I} \gamma_{ij} \log p_{jnt} + Z_{nt} \theta_i + \varepsilon_{int}$$
$$\log u_{nt} = \mu_n + \lambda \log X_{nt} + \delta \log P_{nt} + Z_{nt} \phi + \eta_{nt}$$

(f) What is the overall purpose of estimating the equation below?

$$logp_{it} = \alpha_i + W_t + I_{it}\delta_1 + M_{it}\delta_2 + \varepsilon_{it}$$

(g) Considering the equation below, generally speaking what do the  $\gamma$  parameters reflect? Do the estimates of the  $\gamma$ 's affect own and cross price elasticity estimates? Explain.

$$s_{int} = \alpha_{in} + \beta_i \log Y_{nt} / P_{nt} + \sum_{j=1}^{I} \gamma_{ij} \log p_{jnt} + Z_{nt} \theta_i + \varepsilon_{int}$$

(h) Considering the equation below, what are  $c_j$ ,  $s_j$ ,  $p_j$ , and  $e_{jt}$ ? Where do they come from? (For example, are they in Hausman and Leonard's data?)

$$s_i(p_1,...,p_N) + \sum_{j=1}^n \frac{p_j - c_j}{p_j} s_j(p_1,...,p_N) e_{jt}(p_1,...,p_N) = 0$$

- (i) Consider Table VII entitled "Total Consumer Welfare Effect." What are the causes of the reported positive total consumer welfare effects? For the city with the smallest effect, is the effect statistically significant?
- This question gives you excerpts and tables from a scholarly peer-reviewed empirical IO paper: Goeree (2008) "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica* 76(5): 1017 - 1074.

Summary: Goeree (2008) studies the U.S. personal computer market using a static Bertrand model of firm behavior and a discrete-choice model of consumer choice. The data contain 15 major firms that sell 723 distinct models of personal computers to home users. Goeree observes sales of each model, prices, three non-price characteristics (CPU speed, whether it has a Pentium processor and whether it is a laptop), and some consumer demographics.

(a) Use Table 1 (reproduced below) to find a lower bound estimate of the HHI. Remark on concentration.

	Percentage Dollar		Average Annual			Median Percentage Markup, Home Sector		
	Hom	e Marke	Share	Ad	Ad-to-Sales	Median Price	Over Marginal	Including
Manufacturer	1996	1997	1998	Expend	Ratio	Home Sector	Costs	Ad Costs
Industry					3.4%	\$2239	15%	10%
Top 6 firms	65.67	68.31	75.26	\$469	9.1%	\$2172	17%	12%
Acer	6.20	6.02	4.37	\$117	5.4%	\$1708	11%	9%
Apple	6.66	5.79	9.16	\$161	5.3%	\$1859	16%	9%
AST	3.08	1.53					13%	
Compaq	11.89	16.29	16.43	\$208	2.4%	\$2070	23%	16%
Dell	2.46	2.87	2.57	\$150	2.1%	\$2297	10%	
Gateway	8.94	11.77	16.43	\$277	5.6%	\$2767	12%	10%
Hewlett-Packard	4.02	5.52	10.05	\$651	17.7%	\$2203	16%	10%
IBM	8.49	7.42	6.85	\$1189	20.1%	\$2565	16%	10%
Micron	3.26	4.05	1.68				7%	
NEC	3.22							
Packard–Bell	23.48							
Packard–Bell								
NEC		21.02	16.33	\$327	7.2%	\$2075	16%	11%
Texas								
Instruments	1.40						7%	
15 included	83.11	82.27	83.88					

 TABLE I

 Summary Statistics for Market Shares, Advertising, Prices, and Markups<sup>a</sup>

<sup>a</sup>Notes: Others in the 15 included are ATT (NCR), DEC, and Epson, each of which held less than 1% of the home (and total) market shares in 1996 and 1997. AST and Micron held less than 1% total market shares on average. In 1997 three mergers occurred: Packard Bell, NEC, ZDS; Acer, Texas Instruments; Gateway, Advanced Logic Research. Ad expenditures (in \$M) and ad-to-sales ratios are annual averages from LNA and include all sectors (home, business, education, government). Percentage markups are the median (price – marginal costs)/price across all products. The last column is percentage total markups per unit after including advertising. These are determined from estimated markups and estimated effective product advertising in the home sector.

(b) Use Table 3 (reproduced below) to answer: (i) which are dummy variables, (ii) What is the meaning of the number 0.443 (fourth number down in the first column of numbers)? (iii) What is the meaning of the number 2.54 (fifth number down in the second column of numbers)?

	Sample		Population	
Variable Description	Mean	Std. Dev.	Mean	Std. Dev.
Male	0.663	0.474	0.661	0.473
White	0.881	0.324	0.881	0.324
Age (years)	47.38	15.68	46.87	15.13
$30 \text{ to } 50 \ (=1 \text{ if } 30 < \text{age} < 50)$	0.443	0.497	0.449	0.497
Education (years)	13.98	2.54	14.00	2.35
Married	0.564	0.496	0.572	0.495
Household size	2.633	1.429	2.631	1.428
Employed	0.695	0.460	0.693	0.461
Income (\$)	56,745	45,246	56,340	44,465
Inclow (=1 if income $<$ \$60,000)	0.667	0.471	0.669	0.471
Inchigh (= 1 if income > $100,000$ )	0.107	0.309	0.106	0.308
Own PC (= 1  if  own a PC)	0.466	0.499	0.470	0.499
PCnew (=1 if PC bought in last 12 months)	0.113	0.317	0.112	0.316
Media Exposure	Mean	Std. Dev.	Min	Max
Cable (=1 if receive cable)	0.749	0.434	0	1
Hours cable (per week) cable (per week)	3.607	2.201	0	7
Hours noncable (per week)	3.003	2.105	0	6.2
Hours radio (per day)	2.554	2.244	0	6.5
Magazine (=1 if read last quarter)	0.954	0.170	0	1
Number magazines (read last quarter)	6.870	6.141	0	95
Weekend newspaper (=1 if read last quarter)	0.819	0.318	0	1
Weekday newspaper (=1 if read last quarter)	0.574	0.346	0	1

## TABLE II Descriptive Statistics for Simmons Data<sup>a</sup>

<sup>a</sup>Notes: Unless units are specified, variable is a dummy. Number of observations in survey is 39,931. Sample size is 13,400. Media exposure summary statistics are based on reports published by Simmons Market Research.

## (c) Consider this excerpt and Table D3 (reproduced below) from Goeree (2008).

Finally, I estimate a logit model to show that the instruments I use in the full model address the endogeneity issues. The first column reports OLS results. As expected, the price coefficient is negative but small in magnitude. Instruments appear to address the endogeneity of price issue and result in estimates for the price coefficient that are significantly higher in absolute value.

	OLS	OLS	IV	IV	IV	IV
Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Price	$-0.05^{*}$	$-0.04^{*}$	-0.71**	-1.07**	-2.41**	-2.46**
	(0.04)	(0.04)	(0.13)	(0.17)	(0.51)	(0.41)
CPU Speed	0.16**	0.17**	0.07	0.06	0.17*	0.07*
-	(0.08)	(0.08)	(0.08)	(0.08)	(0.14)	(0.11)
Pentium	-0.36**	0.04	-0.68**	-0.85**	1.51**	2.04**
	(0.09)	(0.14)	(0.11)	(0.19)	(0.29)	(0.38)
Laptop	-1.27**	-1.25**	1.99**	2.35**	3.82**	3.83**
	(0.09)	(0.09)	(0.16)	(0.20)	(0.58)	(0.47)
Acer		1.64**		1.65**		1.67**
		(0.16)		(0.17)		(0.20)
Apple		1.81**		0.62**		0.98*
11		(0.20)		(0.29)		(0.56)
Compaq		1.77**		1.72**		1.66**
1 1		(0.16)		(0.17)		(0.21)
Dell		0.72**		0.43**		0.06
		(0.16)		(0.17)		(0.25)
Gateway		1.98**		1.43**		0.69**
		(0.17)		(0.21)		(0.33)
HP		-0.13		-0.26*		-0.43**
		(0.17)		(0.18)		(0.22)
IBM		1.01**		0.64**		0.13
		(0.16)		(0.19)		(0.28)
Micron		0.55**		-0.71**		-2.40**
		(0.17)		(0.28)		(0.56)
Pbell		2.24**		2.18**		2.10**
		(0.17)		(0.19)		(0.23)
Constant	-12.64**	-13.83**	-12.83**	-13.99**	-13.33**	-14.20**
	(0.89)	(0.89)	(0.90)	(0.90)	(1.32)	(1.21)
First Stage						
Adjusted $R^2$	0.13	0.30	0.42	0.53	0.33	0.49
<i>E</i> -statistic	0.12	0.00	126.47	112.7	210.69	143.76
Prob > F			0.00	0.00	0.00	0.00
Instruments						_
BLP 95 series approvimation			x	x		
BLP 99 direct approximation			Δ	Δ	x	x
			1	1		

## TABLE D.III Logit IV Results<sup>a</sup>

<sup>a</sup>Notes: The dependent variable is  $\ln(s_j) - \ln(s_0)$  based on 2112 observations. All regressions include time dummies. Asymptotically robust standard errors are in parentheses. \*\* indicates *t*-stat > 2; \* indicates *t*-stat > 1. BLP series approximation IV are the sum of the values of the same characteristics of other products offered by the same firm, the sum of the values of the same characteristics of all products offered by rival firms, the number of ownfirm products, and the number of rival firm products. The more direct approximation IV are based on BLP 1999. These are used in the full model and described in the paper.

- i. The assumed functional form for demand is logit, which is a different functional form from others you have seen such as linear, constant elasticity and even the AIDS functional form in Hausman and Leonard (2002). This explains the dependent variable. Given what you can glean from the given information and tables, which approach does Goeree use to address price endogeneity in demand estimation?
- ii. Is the effect of price on consumer choice statistically different from zero in the column "OLS (i)"? In the column "IV (iii)"? Even before seeing the estimation results why should Goeree expect price to be endogenous? Be specific.

- (d) Goeree (2008) estimates the price elasticity of the IBM Thinkpad is -6.9745. What does -6.9745 mean exactly? Is this estimate plausible?
- (e) Explain what Goeree (2008) means by "I use information from the first-order conditions to estimate marginal cots, which allows me to calculate markups." Did Hausman and Leonard (2002) do the same thing?