Problem Set 1

Health Economics II
Maghfira Ramadhani[*]

February 15, 2024

1 Logit Demand

We are given data on over-the-counter (OTC) headache medicine. The data is at the store-week level for four brands and three package sizes. A brand-size pair is a product. Consider the utility function for product j in market t for consumer i:

$$\begin{aligned} u_{ijt} &= X_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \varepsilon_{ijt} \\ &= \delta_{jt} + \varepsilon_{ijt}. \end{aligned}$$

where ε_{ijt} is an i.i.d. draw from type I extreme value distribution, X_{jt} is a vector of product characteristics, p_{jt} is the price of product j in market t, and ξ_{jt} is an unobserved characteristic of the product.

1. Table 1 shows summary statistics that provides the mean of each of the following for each brand- size pair: market share of sales, unit price, price/50 tab, and wholesale price.

Table 1: Summary Statistics

	Advil			Bayer			Store Brand		Tylenol		
	25	50	100	25	50	100	50	100	25	50	100
Market share of sales	0.120	0.077	0.036	0.043	0.035	0.082	0.093	0.071	0.149	0.178	0.116
Unit price	0.119	0.103	0.082	0.107	0.072	0.040	0.039	0.044	0.137	0.099	0.070
Price per 50 tab	5.927	5.145	4.080	5.346	3.607	1.983	1.928	2.224	6.841	4.942	3.508
Wholesale price	2.030	3.623	6.091	1.847	2.422	3.712	0.910	1.919	2.182	3.672	5.755

If we assume that ε_{ijt} is distributed type I extreme value which have the density f and cumulative distribution F as follows

$$\begin{aligned} f(\varepsilon_{ijt}) &= e^{-\varepsilon_{ijt}} e^{e^{-\varepsilon_{ijt}}} \\ F(\varepsilon_{ijt}) &= e^{-e^{-\varepsilon_{ijt}}} \end{aligned}$$

Consequently, the probability that consumer i chooses product j in market t is given by

$$\begin{aligned} s_{jt} &= \Pr(\delta_{jt} + \varepsilon_{ijt} > \delta_{kt} + \varepsilon_{ikt} \quad \forall k \neq j) \\ \Leftrightarrow s_{jt} &= \Pr(\varepsilon_{ikt} < \varepsilon_{ijt} + \delta_{jt} - \delta_{kt} \quad \forall k \neq j) \\ \Leftrightarrow s_{jt} &= \int \left(\prod_{k \neq j} F(\varepsilon_{ijt} + \delta_{jt} - \delta_{kt}) \right) f(\varepsilon_{ijt}) d\varepsilon_{ijt} \\ \Leftrightarrow s_{jt} &= \frac{\exp(\delta_{jt})}{\sum_k \exp(\delta_{kt})}. \end{aligned} \tag{1}$$

[*]Replication codes available at https://github.com/maghfiraer/health-ii-ps

If we assume that the mean utility of the outside good is zero, then the market share of sales for product j in market t is given by

$$s_{0t} = \frac{1}{\sum_k \exp(\delta_{kt})}. \tag{2}$$

Dividing equation (1) by equation (2) yields

$$\begin{aligned} &\frac{s_{jt}}{s_{0t}} = \exp(\delta_{jt}) \\ &\Leftrightarrow \ln(s_{jt}) - \ln(s_{0t}) = X_{jt}\beta - \alpha p_{jt} + \xi_{jt} \\ &\Leftrightarrow \ln(s_{jt}) = \ln(s_{0t}) + X_{jt}\beta - \alpha p_{jt} + \xi_{jt}. \end{aligned} \tag{3}$$

We can estimate equation (3) with log market share of sales as the outcome variable and price and product characteristics as the explanatory variables using OLS, or IV to solve the price endogeneity problem.

2. The demand model estimates by OLS using price and promotion as product characteristics is shown in column (1) of Table 2.

3. The demand model estimates by OLS using price and promotion as product characteristics including product dummies is shown in column (2) of Table 2.

4. The demand model estimates by OLS using price and promotion as product characteristics, including store-product (the interaction of store and product) dummies is shown in column (3) of Table 2.

Table 2: Demand model OLS estimates

	OLS		
	(1)	(2)	(3)
Price	-0.049	-0.277	-0.297
	(0.002)	(0.009)	(0.009)
Promotions	0.154	0.287	0.277
	(0.016)	(0.012)	(0.011)
Fixed Effect	-	Product	Store-Product
Adjusted-R^2	0.014	0.048	0.036
Observations	38544	38544	38544

5. The demand estimates using wholesale price as an instrument is shown in column (1) - (3) of Table 3.

6. The demand estimates using Hausman instrument (i.e. the average price of the same product in other stores) is shown in column (4) - (6) of Table 3.

Table 3: Demand model IV estimates

	IV: Wholesale Price			IV: Hausman Instrument		
	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.010	0.016	0.008	-0.051	-0.501	-0.501
	(0.003)	(0.018)	(0.018)	(0.002)	(0.012)	(0.012)
Promotions	0.176	0.376	0.371	0.154	0.219	0.215
	(0.016)	(0.013)	(0.012)	(0.016)	(0.012)	(0.011)
Fixed Effect	-	Product	Store-Product	-	Product	Store-Product
Montiel-Pflueger F-statistics	276954	14225	14970	2498670	50394	64374
Adjusted-R^2	0.008	0.024	0.008	0.014	0.034	0.023
Observations	38544	38544	38544	38544	38544	38544

7. The IV estimates when instrumenting for price using wholesale price is shown in column (1) - (3) of Table 3. The coefficient for prices are practically zero. This is because the wholesale price is not a good instrument for price. I suspect that wholesale price is highly correlated with product price but fail to satisfy the exclusion restriction, this could happen when the store collect the same portion of markup from the price of product sold. The result from using Hausman instrument in column (1) - (3) of Table 3 show that it is a better instrument for price. The specification in column (6) is the best model in addressing endogeneity issues since it control for unobserved store-product characteristics as well as using a good instrument for price.

8. The own-price elasticities for products j in the market t is given by

$$\frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}} = -\alpha p_{jt}(1 - s_{jt}).$$

From the data, the mean own-price elasticity for each product j is given by

$$\left. \frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}} \right|_t = -\alpha \frac{1}{T} \sum_t p_{jt}(1 - s_{jt}).$$

The mean own-price elasticity for each product j is shown in Table 4. The own-price elasticity estimates from specification (1) to (3) are not reliable as the estimates suffers from weak instrument issue. Result from specification (5) and (6) do not differ much, but the preferred specification (6) is the best model in addressing endogeneity issues. From the preferred specification (6), the own-price elasticity for each product is negative, which is consistent with the law of demand. The own-price elasticity for Advil with 100 tabs is the most elastic, while the own-price elasticity for Store Brand with 50 tabs is the least elastic.

Table 4: Mean own price elasticity for different product across specifications

	Advil			Bayer			Store Brand		Tylenol		
	25	50	100	25	50	100	50	100	25	50	100
Specification (1)	-0.027	-0.050	-0.082	-0.027	-0.036	-0.038	-0.018	-0.043	-0.030	-0.042	-0.065
Specification (2)	0.042	0.076	0.126	0.041	0.056	0.059	0.028	0.067	0.047	0.065	0.100
Specification (3)	0.020	0.037	0.061	0.020	0.027	0.028	0.014	0.032	0.023	0.032	0.048
Specification (4)	-0.132	-0.240	-0.398	-0.129	-0.176	-0.184	-0.089	-0.209	-0.147	-0.206	-0.313
Specification (5)	-1.306	-2.381	-3.939	-1.280	-1.742	-1.824	-0.879	-2.073	-1.458	-2.037	-3.106
Specification (6)	-1.307	-2.382	-3.941	-1.281	-1.743	-1.825	-0.879	-2.074	-1.459	-2.037	-3.107

9. Let be $\mathcal{J}$ the choice sets of available products. Under the assumptions of the logit model, the expected consumer surplus that an individuals receives from getting to choose the OTC headache medicine that maximizes their utility is

$$\begin{aligned} \mathrm{E}[CS_i | j \in \mathcal{J}] &= \frac{1}{\alpha} \mathrm{E}[\max_j (\delta_{jt} + \varepsilon_{ijt})] \\ &= \frac{1}{\alpha} \ln \left(\sum_{j \in \mathcal{J}} \exp(\delta_{jt}) \right) \end{aligned}$$

When a product l is not offered in the market, the expected consumer surplus that an individuals receives from getting to choose the OTC headache medicine that maximizes their utility is

$$\begin{aligned} \mathrm{E}[CS_i | j \in \mathcal{J} \setminus l] &= \frac{1}{\alpha} \ln \left(\sum_{j \in \mathcal{J} \setminus l} \exp(\delta_{jt}) \right) \\ &= \frac{1}{\alpha} \ln \left(\sum_{j \in \mathcal{J}} \exp(\delta_{jt}) - \exp(\delta_{lt}) \right) \end{aligned}$$

The change in consumer surplus from removing product l from the choice set is given by

$$\begin{aligned}
\mathrm{E}[\Delta CS_i] &= \frac{1}{\alpha}\ln\left(\frac{\sum_{j\in g}\exp(\delta_{jt}) - \exp\left(\delta_{lt}\right)}{\sum_{j\in g}\exp\left(\delta_{jt}\right)}\right) \\
&= \frac{1}{\alpha}\ln\left(1 - \frac{\exp\left(\delta_{lt}\right)}{\sum_{j\in g}\exp\left(\delta_{jt}\right)}\right) \\
&= \frac{1}{\alpha}\ln\left(1 - s_{lt}\right)
\end{aligned}$$

Using the data, the mean change in consumer surplus from removing product l from the choice set is given by

$$\Delta CS_i|_t = \frac{1}{\alpha}\frac{1}{T}\sum_t \ln\left(1 - s_{lt}\right)$$

The mean change in consumer surplus from removing a specific product from a market is shown in Table 5. Again comparing specification (1) -(3) is not useful as they suffer from weak instrument issue. Looking only at the preferred specification (6), the mean change in consumer surplus from removing a specific product from the market is negative, which is consistent with the intuition that removing a product from the market will reduce the consumer surplus. The mean change in consumer surplus from removing Tylenol with 50 tabs is the highest, while the mean change in consumer surplus from removing Bayer with 25 tabs is the lowest. These value as we can see from the derivation are related to its share of sales of the product in the market. The higher the share of sales, the higher the mean change in consumer surplus from removing the product from the market.

Table 5: Mean change in consumer surplus from removing specific product

	Advil			Bayer			Store Brand		Tylenol		
	25	50	100	25	50	100	50	100	25	50	100
Specification (1)	-12.391	-7.840	-3.549	-4.269	-3.413	-8.255	-9.557	-7.299	-15.615	-18.977	-12.001
Specification (2)	8.046	5.091	2.305	2.772	2.216	5.360	6.206	4.739	10.139	12.323	7.793
Specification (3)	16.686	10.558	4.779	5.749	4.596	11.116	12.869	9.829	21.027	25.555	16.161
Specification (4)	-2.558	-1.619	-0.733	-0.881	-0.705	-1.704	-1.973	-1.507	-3.224	-3.918	-2.478
Specification (5)	-0.258	-0.163	-0.074	-0.089	-0.071	-0.172	-0.199	-0.152	-0.325	-0.395	-0.250
Specification (6)	-0.258	-0.163	-0.074	-0.089	-0.071	-0.172	-0.199	-0.152	-0.325	-0.395	-0.250

2 Adverse selection and (selection on) moral hazard

Define each of the following terms and describe the data and identifying variation necessary to estimate their effects in health insurance markets: adverse selection, moral hazard, and selection on moral hazard. What approaches have the papers discussed in class taken to estimate these effects? What have these papers found and concluded?

Adverse selection

Adverse selection happens when willingness to pay is positively correlated. The source of asymetric information because the insurance company do not observe the individual health status. Under this condition, individual who have the highest willingness to pay are individual with the highest expected cost and consequently the average cost is always higher than the marginal cost causing underinsurance.

1. **Einav et al. (2010)** use the data on health insurance options, choices, and medical insurance claims of Alcoa, Inc in 2004 in which they focused on subsample of 3,779 employes who chose one of the two modal insurance plans, contracts H and L.

- They use **price variation** in which due to Alcoa's organization structure, employees doing similar jobs are exposed to different the price of insurance for the same sets of coverage options across different section of the company.
- Using this variation, they **estimate** the marginal cost and demand curve, and **find** that the marginal cost is increasing in price (decreasing in quantity) which is consistent with adverse selection.

Moral hazard

Under moral hazard, individual who choose more comprehensive coverage have less incentive to reduce claim probability, i.e. being more risky. The source of asymetric information is due to the change in behaviour of the individual after coverage.

1. In the paper discussing adverse selection, **Einav et al. (2010)** use the same data and price variation to estimate the effect of moral hazard by computing marginal cost curve for contracts H and L.

- They define moral hazard as the difference in the expected cost of the individual who choose the high coverage and the individual who choose the low coverage.
- They **find** that they can not reject the null of no moral hazard.

2. **Lin and Sacks (2019)** use the data from the RAND Health Insurance Experiment (HIE) in which family are randomly enrolled in either free healthcare or healtcare with some cost-sharing up to an OOP maximum.

- They use the **variation in monthly spending between those with and without cost-sharing** to estimate the effect of moral hazard. Individual with cost-sharing plan, who had coinsurance up until they hit the maximum dollar expenditure (MDE), will have an effective price of zero after hitting this MDE.
- To test for intertemporal substitution by comparing health care demand in cost-sharing and in free care plans in the final month of a coverage year, they **estimate** monthly demand in these plans, adjusting for differences in site-by-start date and for general trends.
- They **find** that individual respond to both current price and expected future price when choosing care, i.e. individual who already passed their MDE at the end of the year will consume more care compared to the very first month of the year after.

3. **Brot-Goldberg et al. (2017)** use proprietary data set from a large self-insured firm which spans from 2006 to 2015 containing individual-level line-by-line health claims providing granular information on medical spending, medical diagnoses, and patient-provider relationships. The data also contain employee and dependent demographic and employment characteristics as well as the linked benefit decisions of HSA elections and 401(k) contributions.

- They leverage a **natural experiment** at this firm that required all of its employees to switch from an insurance plan that provided free health care to a nonlinear, high-deductible health plan (HDHP).
- They show that the impact of increased cost sharing on overall medical spending and for specific type of patient and procedure using **difference in differences** and **find** that the switch caused a spending reduction between 11.8% and 13.8% of total firm-wide health spending.
- They then **decompose the overall change in spending** from the required switch to the HDHP into consumer price shopping, outright quantity reductions, and quantity substitutions to lower-cost procedures and **find** that spending reductions are entirely due to outright reductions in quantity. They **find** that consumers respond heavily to spot prices at the time of care,reducing their spending by 42% when under the deductible, conditional on their true expected end-of-year price and their prior year end-of-year marginal price.

Selection on moral hazard

Selection on moral hazard is when individual selecting insurance coverage based on their underlying health risk or status and their behavior response after coverage.

1. **Einav et al. (2013)** use the data from Alcoa (the same company as in Einav et al. (2010)) from 2003 and 2004, in some analysis the data is extended through 2006. The data contain the menu of health insurance options available to each employee, the employee's coverage choices, and detailed claim-level information of the employee (and any covered dependents') utilization and expenditure for the year. It also contain information on demographics of employee, union affiliation, whether theey are hourly or salary type employee, age, race, gender, annual earnings, job tenure, number and ages of insuren family members.

- To identify and estimate moral hazard they use **variation in the health insurance options** offered to different group of workers at different points of time in the imposition of the new PPO plans.
- They introduce a theoretical model of coverage choice and utilization and **use Markov Chain Monte Carlo Gibbs sampling to estimate the model parameters** that explain the joint distribution of health status F_λ, moral hazard ω, and risk aversion ψ.
- They **find** significant heterogeneity in moral hazard and selection on moral hazard. Individual who has a higher behaviour response to coverage are more likely to select the more generous coverage. They **find** that for determining the choice between the two plans, selection on moral hazard is roughly as important as selection on health risk, and considerably more important than selection on risk aversion. Selection on moral hazard implies that those with the lowest moral hazard effects of insurance are those who have the lowest willingness to pay for incremental coverage and are therefore the first (as the price of coverage increases) to switch from higher to lower coverage.

References

Brot-Goldberg, Z. C., A. Chandra, B. R. Handel, and J. T. Kolstad (2017). What does a deductible do? the impact of cost-sharing on health care prices, quantities, and spending dynamics. *The Quarterly Journal of Economics 132*(3), 1261–1318.

Einav, L., A. Finkelstein, and M. R. Cullen (2010). Estimating welfare in insurance markets using variation in prices. *The Quarterly Journal of Economics 125*(3), 877–921.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on moral hazard in health insurance. *American Economic Review 103*(1), 178–219.

Lin, H. and D. W. Sacks (2019). Intertemporal substitution in health care demand: Evidence from the rand health insurance experiment. *Journal of Public Economics 175*, 29–43.