M.S. COMPREHENSIVE EXAMINATION

Retake, 23 August 2019

You have four hours, after a 20 minute reading period. You do not need to use the whole time period. You need to answer all three questions, which will be weighted equally.
Watch the time carefully. The logic used to answer each question is important, so be sure to specify your reasoning with full sentences. You may support your answers using diagrams or mathematical derivations where appropriate. If you use graphs, make sure that they are large enough. We are expecting precise and concise answers. Also make sure your writing is legible: if we cannot read it, your answer will be assumed wrong.

I. Purchasing wine of unknown quality

Note: You may find similarities between this question and Question 1 from the June examination. Do not try to replicate the results from the previous examination, it will lead you astray. Instead, concentrate on the present question. If you find a part difficult, take the result for granted and move on to the next part.

We consider the market for wine. There are $\mathcal{M}$ producers. Each of them may supply at most one unit of wine. Producers supply different qualities, and they have different production costs. We consider the stylized case whereby higher-cost producers systematically supply a higher-quality wine (if you wonder why that would be, think of the case where wine grown at higher altitude is higher quality, and it is more costly to farm at higher altitude). This assumption allows us to index producers by a single continuous parameter $\theta \in [0, 1]$ that reflects the cost (and the quality) of the wine produced. That is, the opportunity cost of producing one unit of wine for a seller of type θ is precisely θ. We further assume that θ is uniformly distributed on $[0, 1]$.

On the demand side, there are many buyers, each purchasing at most one unit of wine. There are more buyers than the $\mathcal{M}$ units of wine that can be produced, therefore in equilibrium some buyers drink water (which is available for free). All buyers have identical preferences. They satisfy the expected utility hypothesis with von-Neumann-Morgenstern utility function $u(w) = -\exp(-\alpha w)$, for some $\alpha > 0$. If a buyer purchases water, they get a payoff of zero and therefore their *ex post* utility is $u(0) = -1$. If they purchase a unit of wine of quality θ at price $p > 0$, they get a payoff of $K\theta - p$ and therefore their *ex post* utility is $-e^{\alpha(p-K\theta)}$.

Buyers do not know the quality of wine prior to purchase. Therefore, sellers get the same price $p > 0$ in the market, which they take as given. As usual, producers are willing to sell if the market price exceeds their opportunity cost.

Supply

[0.4] 1. If the market price is equal to $0 < p \leq 1$, how many producers are willing to sell? What is the density of the distribution of sellers selling in that market in terms of the characteristic θ?

[0.2] 2. On a graph with quantity on the horizontal axis and price on the vertical axis, represent the supply curve for wine $S(Q)$. Indicate $\mathcal{M}$ on the horizontal axis and 1 on the vertical axis. Label your graph "Figure 1."

Demand

[0.3] 3. Interpret the parameter α.

[0.3] 4. Are buyers risk-averse? If so, is absolute risk aversion constant, increasing, or decreasing? Justify your answers.

[0.3] 5. How much would a buyer be willing to pay for a unit of wine of *known* quality θ? Your answer, which you can denote $P(\theta)$, should depend only on θ and K. Explain why α does not appear in $P(\theta)$. (Hint: there is no optimization problem here.)

[0.2] 6. Show that as long as $K > 1$, buyers' willingness to pay for wine of *known* quality θ always exceeds the opportunity cost of supplying that particular unit.

From now on, we assume that $K > 1$.

[0.3] 7. Based on your answer to part 6, how many units of wine should be traded in a socially optimal allocation?

Now suppose that the market price is $0 < p \leq 1$. Although buyers cannot tell the quality of a given unit of wine purchased, they are intelligent: they know the answer to part 1 above and can therefore assess the *distribution* of quality in the market.

[0.6] 8. Show that a buyer's willingness to pay for a unit of wine of *unknown* quality when the market price is p is given by the function

$$\tilde{P}(p) = \frac{1}{\alpha} \ln\left(\frac{\alpha K p}{1 - e^{-\alpha K p}}\right).$$

[0.3] 9. Of course, we prefer to express willingness to pay as a function of quantity rather than price. Using part 1 and part 8, express $\tilde{P}$ as a function of Q, the quantity of wine sold in the market, rather than p. You will call that new function $\bar{P}(Q)$.

The rest of the question is devoted to representing the willingness-to-pay function $\bar{P}(Q)$ on Figure 1 to find the market equilibrium. To that effect, we will analyze the behavior of the function $\phi(x) \equiv \ln\left(\frac{x}{1-e^{-x}}\right)$ for $x > 0$.

[0.3] 10. By writing $\phi(x) = -\ln\left(\frac{1-e^{-x}}{x}\right)$, show that $\lim_{x\to 0}\phi(x) = 0$. (Hint: for all $y \in \mathbb{R}$, $e^y \equiv \sum_{k=0}^{+\infty} \frac{y^k}{k!}$.)

[0.3] 11. Show that $\phi'(x) > 0$ for all $x > 0$.

For the rest of the problem you can take for granted that, in addition to the properties shown in parts 10 and 11, $\phi(x)$ is concave and its slope at $x = 0$ approaches $\frac{1}{2}$.

[0.6] 12. Find a set of two inequalities involving α and K that ensure that the willingness-to-pay function intersects the supply function $S(Q)$ at some quantity $0 < Q^e < \mathcal{M}$.

[0.5] 13. Assuming the conditions in part 12 are satisfied, represent $\bar{P}(Q)$ on Figure 1. Represent the equilibrium (Q^e, p^e). Indicate the location of producers unwilling to sell wine, if any.

[0.4] 14. Does the market mechanism always lead to a socially efficient outcome in this market? Justify your answer using answers to the previous parts, and provide some intuition.

II. Linear estimators

This question begins with a typical single-X Ordinary Least Squares regression:

$$y_i = \beta_1 + \beta_2 X_i + u_i, \tag{1}$$

with each u_i an independent $N(0, \sigma^2)$ random variable.

The X_i variables are considered non-random and the model above is considered to be the correct one for each y_i.

1. Prove that the OLS estimator

   $$\hat{\beta}_2 = \frac{\sum_{i=1}^n (X_i - \overline{X})(y_i - \overline{y})}{\sum_{i=1}^n (X_i - \overline{X})^2}$$

   can be written as a linear function of the y_i variables, in the form

   $$\hat{\beta}_2 = \sum_{i=1}^n a_i y_i.$$

   What is a_i?

2. Prove that this estimator is unbiased. (*Hint: either summation notation or matrix notation is fine here and in subsequent parts of this question.*)

3. Explain why the presence of an omitted explanatory variable in your model could cause bias in your $\hat{\beta}_2$. (*Hint: Your explanation should also indicate why bias may not result.*)

4. Explain why your estimator might then be inconsistent.

5. Now you consider another estimator:

   $$\tilde{\beta}_2 = \frac{\sum_{i=1}^n (Z_i - \overline{Z})(y_i - \overline{y})}{\sum_{i=1}^n (Z_i - \overline{Z})(X_i - \overline{X})}$$

   Show that this estimator can also be thought of as a linear function of the y_i variables, of the form

   $$\tilde{\beta}_2 = \sum_{i=1}^n b_i y_i.$$

   What is b_i?

6. Under the assumption that the Z_i variables are non-random, and that

   $$\sum_{i=1}^n (Z_i - \overline{Z})(X_i - \overline{X})$$

   does not equal zero, prove that $\tilde{\beta}_2$ is unbiased.

7. Is it possible for $\tilde{\beta}_2$ to be unbiased, even when $\hat{\beta}_2$ is biased due to the omitted-variables problem?

8. Is it possible for $\tilde{\beta}_2$ to be consistent, even when $\hat{\beta}_2$ is not (again, due to the omitted-variables problem)?

For the rest of this question, you may take the model in Equation (1) to be correctly specified (i.e., the functional form is correct and there is no omitted variable). All other assumptions normally made for OLS may or may not hold, and these are the focus of the rest of this question.

9. Pick any three of the parts below and answer briefly. Each describes a common estimation problem that could arise. Indicate the consequences for OLS and (where relevant) a preferred estimator.

   (i) Each u_i has $E(u_i) = 0$ and $V(u_i) = \sigma^2$, but u_i is not normally distributed.

   (j) Each u_i is distributed as $N(0, \sigma^2 W_i)$.

   (k) A second equation is known to be relevant:

   $$X_i = \gamma_1 + \gamma_2 W_i + v_i$$

   where W_i is a non-random policy variable announced in advance and easily observed by all, and each $v_i \sim N(0, \sigma_v^2)$ is independent of every other variable in the model.

   (l) The previous part applies, except that u_i and v_i are correlated.

   (m) It is believed that, in the original equation, the values of both the intercept and the slope changed midway through the sample.

III. Sugar tax

Many people blame sugar-sweetened beverages (SSBs) for growing obesity and related problems in rich and poor countries alike. Many places have imposed taxes on SSBs in an attempt to remedy these problems.

1. In order to understand the effect of SSB taxes on obesity, it is first important to determine how much of the tax, which is typically imposed on distributors, is passed through to the retail prices faced by consumers. In a 2018 'Research Letter' published in the *Journal of the American Medical Association*, Cawley, Willage and Frisvold use the Philadelphia airport to estimate the pass-through rate of a 1.5 cent/ounce tax on SSBs. The setting of the Philadelphia airport provides a natural experiment for this study since part of the airport is in the city of Philadelphia (subject to the SSB tax) and part of the airport is outside the city boundary (not subject to the SSB tax).

   Using this natural experiment as an example, describe carefully the tradeoffs between internal validity and external validity in empirical research. Discuss the considerations a researcher should take into account when deciding on a research design that balances these tradeoffs.

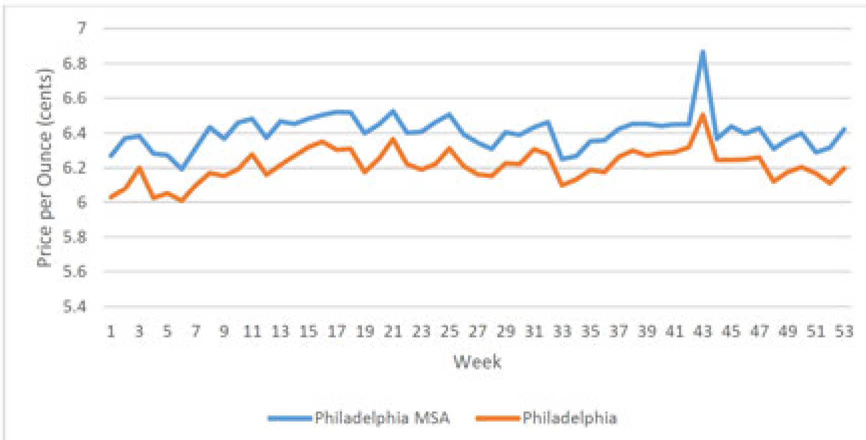
To better understand the effects of the SSB tax in Philadelphia on the retail market for SSBs, Cawley, Frisvold, Hill, and Jones (CFHJ) conduct a broader analysis that uses data collected from more conventional retail settings in the city of Philadelphia and surrounding areas. We will use pieces of this 2018 NBER Working Paper (#24990) for the remainder of this question. CFHJ describe their core Difference-in-Difference (DiD) empirical specification as follows:

> We estimate the effects of the tax using a DiD design, which compares the change in outcomes in stores in Philadelphia to those in stores in comparison (or control) communities. The DiD equations are of the general form:
>
> $$Y_{ist} = \alpha + \beta_1 Post_t + \beta_2 Phila_s + \beta_3 Phila_s * Post_t + \beta_4 S_s + \delta_i + \varepsilon_{ist},$$
>
> where Y_{ist} is either the price per ounce or availability of the beverage i in store s in time period t. *Phila* is a binary variable equal to 1 if the store is in Philadelphia and equal to 0 if the store is in a comparison area. *Post* indicates that an observation occurred after the Philadelphia tax took effect. S_s is a vector of indicators for store type: stand-alone convenience stores, gas stations with convenience stores, small grocery stores, pharmacies, and warehouse stores; large grocery stores are the omitted store type.[6] δ_i is a vector of product fixed effects. β_3 is the coefficient of interest and represents the change in the outcome (price per ounce or availability of beverages) before the tax to after the tax, in Philadelphia relative to the comparison communities. The regression is estimated using ordinary least squares when the outcome is price; it is estimated by logistic regression when the outcome is an indicator variable for product availability. In all cases, we cluster standard errors at the store level to account for correlations between observations within stores.[7]

2. What does it mean to cluster standard errors? Why do CFHJ in this case cluster standard errors at the store level? In this analysis, would you expect standard errors clustered in this way to be larger or smaller than non-clustered standard errors? Justify your reasoning.

3. To justify their DiD identification strategy, CFHJ provide the following appendix figure for the year prior to the implementation of the SSB tax in Philadelphia ("Philadelphia" line in figure). The "Philadelphia MSA" line includes the city of Philadelphia plus all the surrounding areas that are part of the broader Metropolitan Statistical Area.

   Appendix Figure 1: Average Weekly Price per Ounce of Regular Soda, Diet Soda, and Juice Drinks in 2016

   Do you think this figure justifies the authors' use of DiD in this analysis? Be as specific as possible in your evaluation of this figure as it relates to their DiD identification strategy.

4. The authors designed their collection of data for this study in a particular way. Specifically:
   (a) They randomly selected 66 retailers in the city of Philadelphia (taxed). They then used Census data to select a comparison store in the untaxed MSA based on the store type and the percentage of African-American, percentage Hispanic, and percentage in poverty in the neighborhood of the store. Using economic concepts relevant to this research question, discuss the empirical advantage of matching the comparison stores to the taxed stores in this way.
   (b) They collect data from stores on the posted shelf prices of 38 taxed and 8 untaxed products (e.g., juice and water that were not subject to the SSB tax) across a range of beverage types, manufacturers, and container sizes. Why would the authors choose to collect data on untaxed beverages? Explain carefully.
   (c) They collected data at two points in time: Nov-Dec 2016 and Nov-Dec 2017. What is the advantage of collecting data in Nov-Dec for both years?

5. In Table 1 below, the authors report mean prices from their data in cents per ounce, with standard errors in (.) and sample sizes in [.]. What can we infer from these descriptive statistics?

Table 1: Average Price per Ounce of Beverages Before and After the Implementation of the SSB Tax in Philadelphia

	Stores in Philadelphia			Comparison Stores		
	2016	2017	Difference	2016	2017	Difference
Taxed Beverages	7.474 (0.203) [1,253]	9.332 (0.226) [1,052]	1.859 (0.303) [2,305]	7.865 (0.200) [1,562]	8.257 (0.207) [1,366]	0.392 (0.288) [2,928]
Untaxed Beverages	8.718 (0.402) [134]	9.061 (0.404) [158]	0.343 (0.574) [292]	8.025 (0.318) [203]	8.449 (0.344) [187]	0.424 (0.468) [390]

6. The authors report their estimated DiD coefficients in Table 2, where prices are in cents per ounce and each cell shows the DiD estimate (β_3 in the specification above) from a separate regression. Write one paragraph that interprets and discusses these results in the context of the research question and relevant economic concepts.

Table 2: The Impact of the SSB Tax on Prices

	Full Sample
All Taxed Beverages	1.558 (0.172) [5,233]
SSB	1.512 (0.168) [3,463]
Regular Soda	1.552 (0.126) [2,389]
Sports Drink	1.450 (0.299) [374]
Energy Drink	2.071 (1.089) [572]
Juice Drink	1.788 (0.490) [205]
Sweet Tea	0.900 (0.279) [378]
Diet Soda	1.588 (0.128) [1,315]

7. The authors estimate the same DiD specification with product availability rather than price as the outcome variable (i.e., if a given product was available for purchase in a given store (1)

or not (0)). The results from these regressions are available in Table 3 below. What more do we learn from these results that is not evident from the Table 2 results above? Be specific and discuss the practical value-added of these results in the context of this research question.

	Full Sample
All Taxed Beverages	-0.043
	(0.026)
	[10,488]
SSB	-0.046
	(0.029)
	[6,348]
Regular Soda	-0.058
	(0.039)
	[3,312]
Sports Drinks	0.017
	(0.036)
	[1,104]
Energy Drinks	-0.071
	(0.049)
	[1,104]
Juice Drinks	0.004
	(0.043)
	[1,104]
Sweet Tea	-0.021
	(0.040)
	[1,656]
Diet Soda	-0.070
	(0.039)
	[2,208]

8. The authors use a "triple" DiD approach to unpack their DiD estimates and understand heterogeneity in these estimates. To estimate a triple DiD, one simply adds a triple interaction to the basic DiD specification above (i.e., all the terms shown in the simple DiD continue to be included). One dimension of heterogeneity they explore is "travel time to the closest competitor outside the city" (called Time). They do this by adding the triple interaction "$\text{Phila}_s \times \text{Post}_t \times \text{Time}_s$" to the simple DiD. The authors estimate a positive and statistically significant coefficient on this triple interaction. In the context of this problem, what can you infer about how consumers and retailers in the city of Philadelphia responded to the SSB tax?