University of California, Davis

Department of Agricultural and Resource Economics

M.S. Comprehensive Exam, June 2017

You have four hours for this exam after a 20 minute reading period. You do not need to use the whole time period. This exam consists of three questions. You must answer all three questions.

Question 1 is worth 32.5% of the total exam score.

Question 2 is worth 15% of the total exam score.

Question 3 is worth 32.5% of the total exam score.

Question 4 is worth 20% of the total exam score.

Watch the time carefully. The logic used to answer each question is important, so be sure to clearly specify your reasoning, with full sentences. Please support your answers as rigorously as possible – e.g., using diagrams or equations. If you use graphs, make sure they are clearly labeled and large enough to read easily. This is not the time to economize on paper, but keep your responses clear and concise. Make sure your writing is legible; if we can't read it, it will be assumed wrong.

1. **Cheese, Carbon, and Climate Change:** California pursues its own climate policy that is intended to reduce carbon emissions. Among the industries included under the climate policy are processors of dairy products. Consider a processing plant that produces cheese (ignore the value of any byproducts generated in the production of cheese). Because we can choose units freely to measure quantities, we assume without loss of generality that under the current input mix, one unit of farm milk produces one unit of cheese and one unit of carbon emissions. A processor combines milk, M, with other inputs, **X**, to produce cheese, Q. (Bold letters denote vectors and normal font denotes scalers.) Although California is a large cheese producer, cheese is produced many places in the world and the market is global in geographic scope. Thus, we assume California processors are price takers, selling cheese at the global price of P. For the moment, we will also assume that each cheese processor is a price taker in acquiring raw milk from farmers at price F.

a. The following is the production function for producing cheese:

$$(1) \qquad Q = \min\{M, L^{\beta_1} E^{\beta_2} K^{\beta_3}\},$$

where $\sum_i \beta_i < 1$, and L, E, and K denote labor, energy, and capital inputs, respectively. Provide a verbal interpretation of this production function.

b. Consider a cheese processor's profit-maximizing choice of Q. First, can we be assured that the profit is bounded, given the information provided so far? Explain why or why not. Then depict a processor's short-run profit-maximization problem graphically. Use the same notation as introduced in the problem and make certain that your graph is consistent with the information provided.

c. So far we have not accounted for the cheese processors' carbon emissions. One proposal for regulating carbon emissions is a carbon tax. Suppose California imposes a tax, T, per unit of carbon emitted. Using a new graph, show how a carbon tax will impact price received for cheese and cheese production for our processor.

d. Next we will consider what impact, if any, the carbon tax will have on the price of milk (F) received by dairy farmers. Farm production is not presently subject to California's climate regulations. Remembering your concepts from basic production theory, derive a representative cheese processor's unconditional input demand for farm milk and show how the carbon tax will affect the processor's demand. You need not draw a graph for this part, but you may if it helps in developing your answer.

e. Cheese of course is not the only consumer product manufactured from farm milk, but it is a major use of California farm milk. For simplicity, we will assume that all California farm milk is bought and processed in state, i.e., by California processors. Given your answer in part (d) show graphically how the carbon tax will influence price and output for farm milk.[1]

f. Farm milk is highly perishable and costly to transport. A plausible alternative to assuming that processors are price takers in acquiring farm milk is to think that some might be monopsonists. Let us consider the case of a cheese processor who is the only buyer of farm milk in its procurement area. Let the supply of milk to the processor be M(F), where $M' > 0$. On a single graph show how the monopsony processor's optimal volume of milk to purchase and the price paid are influenced by the carbon tax T.

g. Carbon emissions are a function of the energy intensity of a production process. In addition to reducing production of products that are significant emitters of carbon, another goal of a carbon tax is to incentivize firms to adopt less carbon intensive production processes. To see how this might work, we will focus on the firm's use of inputs E and L, treating K as fixed. First draw an isoquant in the (L, E) space that is consistent with equation (1). Then show how imposing a carbon tax (which for this purpose you can treat as a tax on the E input) can cause the firm to alter its input mix so as to emit less carbon.

h. As everyone knows, climate change is a global issue in the sense that it is global carbon emissions that influence climate. A major concern with local jurisdictions such as California imposing climate policies is *leakage*. Leakage occurs when production and emissions from the regulated jurisdiction (California in our case) leak to unregulated jurisdictions (the rest of the U.S. and most of the rest of the world in our case). Given the assumptions made in this question regarding the market for cheese, is leakage a significant concern in this case? Explain carefully. Graphs may be very helpful in your answer, but are not essential.

[1] Anyone with a deep familiarity with milk pricing in California will know that the farm price of milk is heavily regulated, with the final price representing a blend of the values of farm milk for its different final-product uses—e.g., fluid milk, cheese, butter, powder. We will ignore these complications for answering this question.

2. **Bernoulli Random Variable Estimator:** Suppose you have *N* independent observations of a random variable that takes on a value of 0 or 1, (note this is called a Bernoulli random variable). x takes on the value 1 with a probability θ and the value 0 with the probability $(1-\theta)$.

   You have two candidate estimators for θ:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{N}$$

$$\tilde{\theta} = \frac{1000 + \sum_{i=1}^{n} x_i}{2000 + N}$$

To help you evaluate which estimator you might decide to use, your extremely talented research assistant (not the clown you'll meet in problem 4 below) produced a Monte Carlo analysis. Results are displayed in Figure 1 for the two estimators for two cases:

Case 1: $\theta = 0.5$

Case 2: $\theta = 0.1$

   a. When the true probability of success $\theta = 0.5$, which estimator performs better? Explain which numbers in the output make your point. When the true probability of success is 0.1, which estimator performs better? Explain which numbers in the Monte Carlo output make your point.
   b. If you were ignorant about the true value of θ, which estimator would you recommend and why?
   c. Is $\hat{\theta}$ a consistent estimator for θ? Is $\tilde{\theta}$ a consistent estimator for θ? Sketch a proof of your claims. How might you use the simulation program provided to show your results?

```
-----------------------------------------------------------------------------
--------------------
      name:  <unnamed>
  log type:  text
 opened on:  19 Jun 2017, 19:49:39

.
. program define final, rclass
  1.         syntax [, obs(integer 25) theta(real 0.5)]
  2.         /* the above line creates local macros obs and theta they are
only
>        set to 25 and 0.5 by default if other values aren't specified */
.        drop _all
  3.         set obs `obs'
  4.         gen x = uniform() <= `theta'
  5.         summarize x
  6.         local s=r(sum)
  7.         local thetahat=`s'/`obs'
  8.         local thatatilde=(`s' + 1000)/(`obs' +2000)
  9.         return scalar thatatilde=`thatatilde'
 10.         return scalar thetahat=`thetahat'
 11. end

.
. simulate thetahat =r(thetahat) thatatilde =r(thatatilde), reps(10000)
nodots: final, obs(25) th
> eta(0.5)

      command:  final, obs(25) theta(0.5)
     thetahat:  r(thetahat)
   thatatilde:  r(thatatilde)


. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    thetahat |     10,000     .499448     .099176        .12        .84
  thatatilde |     10,000    .4999932    .0012244   .4953086   .5041975

.
. simulate thetahat =r(thetahat) thatatilde =r(thatatilde), reps(10000)
nodots: final, obs(25) th
> eta(0.1)

      command:  final, obs(25) theta(0.1)
     thetahat:  r(thetahat)
   thatatilde:  r(thatatilde)


. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    thetahat |     10,000     .099756    .0607315          0        .36
  thatatilde |     10,000    .4950587    .0007498   .4938272   .4982716

.
. log close
      name:  <unnamed>
```

Figure 1 Simulation Code

3. **Where's the Snow? Asymmetric Information and a Day on the Slopes:** Information shapes the actions and interactions of market agents and resulting market outcomes in many important ways.

a. Several models in economics highlight how asymmetric information affects market actions and outcomes. Describe briefly what economists mean by "asymmetric information."

b. Provide an example of asymmetric information that is common and can be observed in Davis, CA. Discuss how a standard model of asymmetric information applies in your example, including what the model predicts about the behavior of market agents and the market outcomes.

Now consider the case of snow skiing. Suppose that demand for skiing is given by

$$q_{it} = b_i(100 + N_{it} - p_{it})$$

where q_{it} is the total number of lift tickets sold at resort *i* on day *t*, b_i is a time-invariant demand parameter specific to resort *i*, N_{it} is the number of inches of new snow received in the past 24 hours, and p_{it} is the price of a lift ticket at resort *i* on day *t*. Assume that the daily cost of operating ski resort *i* on day *t* is given by $f_i + c_i N_{it}$, where f_i is the fixed cost of operation and c_{it} is the cost of preparing the ski runs for skiers, which increases with each additional inch of new snow.

c. Write out the resort's profit maximization problem. Assuming that N_{it} is a decision variable (e.g., all snow is man-made), provide the FOC that determines the optimal N_{it}^* and solve for N_{it}^*.

d. In practice, demand for skiing is more responsive to new snow on the weekend than during the week. Using the indicator notation $1(i = weekend) = \{0,1\}$, revise the simple demand equation to reflect this fact. Provide the FOC that determines the optimal N_{it}^* and solve for N_{it}^*.

e. Now assume that resorts have no control over N_{it} (i.e., no snow is man-made), but can choose how much new snow to report on their website, denoted by $\widehat{N_{it}}$. Write out the modified profit maximization problem. What is the optimal snow report $\widehat{N_{it}^*}$, given by this problem. Reconcile this result with reality by explaining what is missing from the model.

Zinman and Zitzewitz (ZZ) (2016) contribute to the literature on deceptive advertising with an analysis of snow reports by ski resorts, which can be compared to high resolution snow reports from the US National Weather Service (called SNODAS).

f. Using different notation than above, ZZ first estimate this specification by OLS

$$(1) \qquad s_{rt} = \beta \times w_t + \mathbf{a}_w + \mathbf{n}_r + e_{rt},$$

where s_{rt} is natural new (or "fresh") snowfall reported by resort r on day t, w_t is an indicator variable for whether t is a weekend day, $\mathbf{a}_w$ is a vector of fixed effects for calendar weeks (Wednesday–Tuesday), $\mathbf{n}_r$ is a vector of resort fixed effects, and e_{rt} is an error term. The fixed effects control for any bias arising from the proportion of snow reports on weekends varying between more and less snowy weeks of the year (e.g., if resorts were open only on weekends at the beginning and end of the season) or between more and less snowy resorts.[25]

i. In order to avoid serious bias in estimated coefficients, one rarely omits a constant from regression specifications. What must be true about equation (1) for it to effectively include a constant term?

ii. Explain carefully how equation (1) provides a test that resorts exaggerate their snow reports with reference to your modified profit maximization in part (e) above.

Results from this and additional regressions are provide in Table 2 below.

TABLE 2—WEEKEND EFFECT REGRESSIONS, WITH AND WITHOUT CONTROLLING FOR ACTUAL SNOWFALL

Dependent variable: Inches of new natural snowfall reported by resort					
Observations include	All observations	w/SNODAS			
	(1)	(2)	(3)	(4)	(5)
Weekend (Sat. and Sun.)	0.233**	0.242**	0.174*	0.175*	0.183**
	(0.099)	(0.117)	(0.101)	(0.091)	(0.089)
SNODAS ($t+1$)					0.359***
					(0.025)
SNODAS (t)			0.716***		0.621***
			(0.042)		(0.029)
SNODAS ($t-1$)					0.104***
					(0.020)
Weighted SNODAS				0.967***	
				(0.054)	
Observations	56,402	39,920	39,920	39,920	39,920
Unique days	752	707	707	707	707
R^2	0.123	0.141	0.309	0.341	0.351

Notes: Each specification is estimated by OLS with fixed effects for weeks (Wed.–Tues.) and resort. Weighted SNODAS averages the SNODAS observations for days t and $t+1$, weighting by the number of hours in the 7AM-to-7AM local time window that overlap with the SNODAS observation window in question (weights are 0.75 on t and 0.25 on $t+1$ for the Eastern time zone; and 0.625 and 0.375, respectively, for the Pacific time zone). Columns 2–5 are restricted to observations with SNODAS data for days $t-1$ to $t+1$. Standard errors allow for clustering within both day and resort.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

g. Pretend that this is your paper. Write a thoughtful one paragraph interpretation of these results. In the paragraph, compare and contrast the results in the five different specifications.

h. ZZ cluster standard errors “within both day and resort”. What does it mean to cluster standard errors “within day” and why do the authors do this? What does it mean to cluster standard errors “within resort” and why do the authors do this?

i. The results in Table 2 are potentially biased by selective reporting. Selective reporting refers to the potential practice by resorts to actively update their snow reports when they receive new snow but to let old (favorable) reports sit on their website during periods without snow. Explain how such a practice would affect the results in Table 2. Describe a robustness test the authors could run to show how big of a problem such a bias is like to be in this case.

j. Skiers (and, yes, snowboarders too) live for “epic” powder days. Suppose that demand for lift tickets is especially high once reported new snow is above 12”. Provide a modified specification that would test whether this threshold translates into an incentive to exaggerate snow reports.

4. **Who Is Going to Pass the MS Exam?** The GAC is interested in studying how the GRE quantitative score predicts success in the MS program. Success in the program is defined as passing a high stakes test, which students can either pass $(y = 1)$ or fail $(y = 0)$. Pass/Fail is based entirely on results this single examination. You are provided with information on whether the students passed or failed the examination and their quantitative GRE score.

a. Write down an econometric model of this problem.

b. Propose two different estimation approaches for analyzing this data and discuss their relative strengths and weaknesses.

You are subsequently provided with students' actual scores on the exam. Upon further examination of the scores, you notice that an important proportion of the students obtained the maximum possible score on the exam. Figure 2 displays a histogram of the scores.

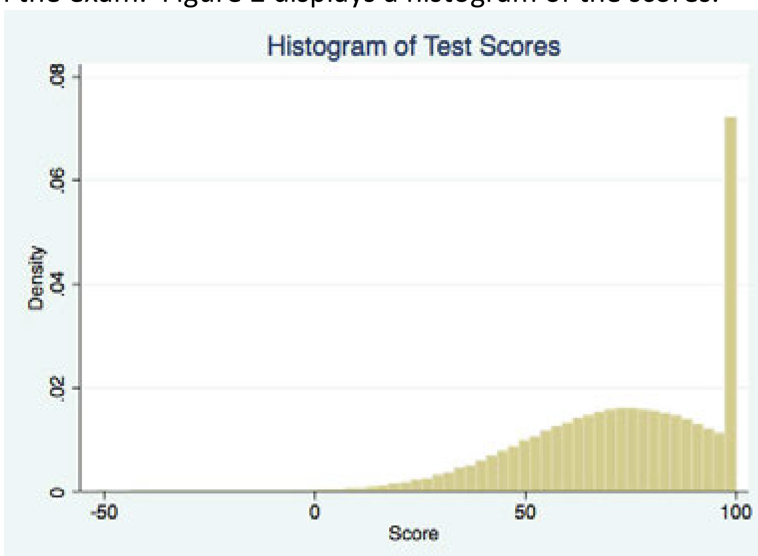

Figure 2 Test Scores

c. Write down an econometric model of this problem. Propose an econometric approach for analyzing this data and discuss its strengths and weaknesses.

Whoops, you discover the code used to generate Figure 1. had an important mistake. Your intrepid RA had coded missing values as 999, which Stata kindly rounded down to 100 in the above histogram. It turns out that that students who performed exceptionally well in the most important courses in the MS Program (256A & 256B) did not have to write the final exam. In other words, you do not observe test results for students who performed well in 256A & 256B because they didn't have to write the test.

d. How would you analyze a sample of the entire class (i.e. a sample which includes students who were exempted and not exempted)? Propose an econometric approach for analyzing this data and discuss its strengths and weaknesses.