University of California, Davis

Department of Agricultural and Resource Economics

M.S. Comprehensive Exam, July 2018

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 33.3% of the total exam score.

Question 2 is worth 33.3% of the total exam score.

Question 3 is worth 33.3% 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.
Question 1.
Farmers in the Pajaro Valley of California grow annual vegetables. Let $Q$ denote quantity of a particular crop, e.g., melons, and assume that melons are produced with variable inputs labor, fertilizer, and water pumped from an underground aquifer. Let $L$, $F$, and $W$ denote quantities of labor, fertilizer, and water applied by a representative farmer. Land area is a fixed input, $A$. We will normalize the output price of melons to be 1.0, and define the normalized prices of labor, fertilizer, and water as $P_L$, $P_F$, $P_W$, respectively. Labor and fertilizer are purchased in the open market, and $P_W$ is determined by the marginal cost of pumping: $P_W = C(\Omega)$, $C' > 0$, $\Omega = \sum W_i$, i.e. $\Omega$ is the total volume of groundwater pumped in Pajaro. Each individual farmer treats $P_W$ as fixed (given).

a. Assume the production process is Cobb-Douglas.
   i. Write down the optimization problem for a representative melon producer.
   ii. Take the first-order conditions to identify the farmer’s optimal application of labor, fertilizer, and groundwater. Be careful to make sure your work is consistent with the information given in the problem.

b. What condition(s) on your Cobb-Douglas production function must be satisfied for the melon producer’s optimization problem to be bounded?

c. Suppose that 50 farmers operate in the Pajaro Valley and pump groundwater. Draw a graph to show how the equilibrium pumping cost, $P_W^*$ and volume, $\Omega^*$, is determined in this valley. Label your graph carefully and be sure to define and explain any curves or functions you include in the graph.

d. Indicate whether the following statement is true or false and explain your answer:
   "Groundwater in the Pajaro Valley is a common resource and its utilization is subject to the problem known as "the tragedy of the commons."

Due in part to California’s prolonged drought, many aquifers in the state are being pumped faster than they are replenished by groundwater recharge. California recently passed a law requiring the state’s aquifers to be managed in a sustainable way. Suppose in response to the law each farmer in the Valley is given a fixed allocation, $\bar{W}_i$, of water that she may pump in any given year.

e. Rewrite your optimization problem from part (a) using the Lagrange multiplier method to reflect the presence of this constraint and re-derive the first-order conditions. Finally, give an economic interpretation of the value of the Lagrange multiplier at the farmer’s constrained optimum.

f. Indicate whether the following statement is true or false and explain your answer: "The melon farmer’s profit will be reduced by the presence of the fixed water allocation, $\bar{W}_i$."

2 of 8
Suppose there are two crops grown in the Pajaro valley, melons, and sweet corn and also suppose that, given the fixed allocations of groundwater, the marginal value product of groundwater used to produce sweet corn is greater than the marginal value product of groundwater used to producer melons.

g. Given these conditions, could a market for groundwater in Pajaro increase economic welfare and the profits of sweet corn and melon growers? Explain your answer, being careful to provide as much detail as you can as to how a groundwater market would work (or not work) in this setting.
Question 2.
A local strawberry grower hires workers for harvests that occur on Thursdays and Fridays, in
order to supply weekend farmers markets. At the time they are hired, workers are offered two
alternative payment schemes. Under the first, workers are paid according to a “piece rate”
model: the berries they pick are weighed and the worker receives a fixed payment per pound.
Under the second, workers are paid by the hour.

You may assume that the workers treat both the weight of total berries picked and the number of
hours they will be offered as random variables.

The grower provides you with the total earnings of each of the workers hired for a given week.

a. How would you test the hypothesis that a worker’s expected earnings are unaffected by
the type of payment scheme?

Be specific about the assumptions you make that support your approach. To demonstrate
your interpretation of the hypothesis, provide a picture that describes the data-generating
process when the hypothesis is correct and another that describes the data-generating
process when the hypothesis is not correct.

b. State assumptions about expected earnings for the two groups that allow you to draw
pictures that illustrate the following items:
   1. The probability of a Type-I error
   2. The probability of a Type-II error
   3. The power of your test

(You can vary your assumptions in ways that let you illustrate each part, and you are not
bound by these choices for subsequent parts of this question.)

c. Does it matter for the interpretation of your test results whether workers choose their own
payment scheme, as opposed to being randomly assigned to one scheme or the other?
Why or why not? Your answer should be consistent with your assumptions in part (a).

Harvesting strawberries is physically demanding, so a model that predicts earnings should take
into account some measure of ability to perform the work. Suppose you are reviewing a study
in which data were collected for both types of workers, and the following model was estimated:

\[ Y_{it} = \beta_1 + \beta_2 X_{it} + \epsilon_{it} \]

The dependent variable is earnings per day; \( i = 1 \) for piece-rate workers and \( i = 2 \) for hourly-wage
workers. The \( t \) index is used to identify individual employees within each set. The independent
variable is the age of each worker, in years.

d. How would you interpret the estimated \( \beta_2 \), and does your interpretation suggest any
modifications of the functional form for this model? (You may ignore those
modifications in subsequent parts of this question.)
e. How would you interpret the estimated $\beta_1$?

f. The study reported some surprising results: the estimated value of $\beta_1$ is positive and statistically significant and the estimated value of $\beta_2$ is negative and statistically significant. The opposite results were obtained in a model estimated using data for piece-rate workers only: $\beta_1$ is negative and statistically significant and the estimated value of $\beta_2$ is positive and statistically significant.

Draw a picture that explains how to reconcile these results. (Hint: it may be helpful to think about the ages you are likely to observe for the two groups.)

g. The estimates reported in the study were obtained using Ordinary Least Squares. There are two hypotheses of interest to you. Each pertains to the choice of equation and technique for estimation. One hypothesis pertains to whether the expected value of $Y_{it}$ given $X_{it}$ is the same for $i=1$ and $i=2$. The second pertains to whether the variance of each $u_{it}$, denoted by $V(u_{it})=\sigma_1^2$, equals the variance of each $u_{2t}$, denoted by $V(u_{2t})=\sigma_2^2$.

i. Explain how you might test the first hypothesis, assuming that the second one is correct.

ii. Now explain how you might test the second hypothesis, assuming that the first one is not correct.

h. Does the fact that workers are allowed to choose their type of compensation (i.e., piece-rate vs. hourly wage) affect how you will approach the hypothesis tests in part (g)?

i. Suppose you conclude that $V(u_{it})$ is proportional to some observed $Z_{it}$. What changes, if any, would you make to the empirical work reported in the study?

j. Suppose instead that you had no $Z_{it}$ variable, but you believe that the variance of $u_{it}$ is not constant. Explain what you would do if you think the variance differs only for $i=1$ and $i=2$ and then instead when you think the variance is not constant within each group.

k. For this part of the question, you may ignore all of the previous parts and return to the specification given before part c. You may also ignore data on workers who earn an hourly wage, and focus only on the piece-rate employees.

Suppose for this part that there is a variable called ability, $A_{it}$, that directly measures ability or productivity in strawberry harvests. This variable is only partially correlated with age. If the correct predictor of earnings should be ability $A_{it}$, what are the consequences of instead using age ($X_{it}$) in the regression model for earnings? It is sufficient to consider the effects on the estimate of $\beta_2$.­
Question 3.
Understanding the determinants and impacts of technology adoption is a common research objective for applied economists. In 1958, Willard Cochrane postulated a “treadmill” theory to describe how the diffusion of a new and improved production technology affects producers. This theory suggests that producers who adopt the technology early reap higher profits for a time, while later adopters may not benefit at all and non-adopters may actually be hurt by the introduction of the new production technology.

a. The predictions of this theory rely on several assumptions. One key assumption pertains to the output market into which producers sell their goods or services. Carefully describe this output market assumption and explain why it is central to the treadmill theory. (Hint: For the purposes of this question, assume the new technology reduces production costs for adopters.)

Harou et al. (2017) test whether the notion of “better late than never” suggested by the Cochrane treadmill theory holds in the case of pineapple cultivation in Ghana. Farmers in this area began growing pineapple – which requires a longer-term and higher upfront investment than their traditional crops – in the early 1990s as improvements in transportation infrastructure and logistics made export to Europe more viable. Harou et al. use a panel data set with two rounds (1997 and 2009) in this analysis.

b. Using panel data with so many years between rounds often raises concerns of attrition (i.e., some households cannot be found 12 years later), which might result in biased estimates. Explain (either in general or in this specific case) how attrition could introduce bias.

c. The researchers observe the year a given pineapple farmer began to cultivate pineapple and the change in household wealth between 1997 and 2009 as measured by an asset index.

   i. Propose a specification that uses these variables and a vector of control variables to test the predictions of the treadmill theory.

   ii. Using the notation from your specification, write the specific hypothesis that would test the predictions of the treadmill theory.

   iii. Given an example of one control variable you think should belong in this specification (without worrying about whether the variable is available in their dataset).

   iv. If the control variables in this specification are not used in the specific hypothesis test in c(ii) above, then why do they belong in the specification at all? Explain carefully.

   v. Estimating the coefficients for this hypothesis test with OLS will yield biased and inconsistent estimates. Explain.
d. As an alternative to OLS, the researchers use 2SLS (IV) estimation with soil characteristics – soil organic matter (SOM) and soil acidity – as instruments because pineapple grows well in soils with high SOM and high acidity.
   i. Explain why this form of estimation is referred to as “2 Stage Least Squares.”
   ii. Explain carefully how 2SLS can, unlike OLS, provide consistent estimates of this specification.
   iii. In this case, what must be true about these two soil instruments in order for them to be valid (i.e., in order for the resulting estimates to be consistent)?

The results of this 2SLS estimation are shown in Table 4 below. Note that the authors chose to represent duration of pineapple cultivation relative to a 2004 market shock that dramatically lowered pineapple prices for these Ghanaian farmers.

e. Pretend this is your paper. Write a thoughtful, concise paragraph that interprets these results in a way that addresses this question: Are these results consistent with the predictions of the treadmill theory?

f. Compare the OLS estimate on “Duration pre 1/2004 (years)” to the estimate in column 1, which uses instrumented “Duration pre 1/2004 (years).” What does this comparison suggest about nature of the inconsistency of OLS in this context?

g. Evaluate whether the output market assumption you described in part (a) is relevant to the case of small farmers adopting a high value crop for export (e.g., pineapple in Ghana). Based on your evaluation, how appropriate is the treadmill theory for understanding the effects of a new technology on farmers in this case?
Table 4
Effect of instrumented duration pre-shock, pre-shock squared, and nonadopter on Δ comprehensive asset index

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration pre 1/2004 (years)</strong></td>
<td>0.120*</td>
<td>0.121*</td>
<td>0.119*</td>
<td>0.164</td>
<td>0.226</td>
<td>0.042*</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.059)</td>
<td>(0.058)</td>
<td>(0.511)</td>
<td>(5.853)</td>
<td>(0.019)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>−0.695***</td>
<td>−0.733***</td>
<td>−0.721***</td>
<td>−0.640</td>
<td>−1.573</td>
<td>−1.021***</td>
</tr>
<tr>
<td></td>
<td>(0.187)</td>
<td>(0.222)</td>
<td>(0.211)</td>
<td>(0.930)</td>
<td>(26.820)</td>
<td>(0.145)</td>
</tr>
<tr>
<td><strong>Age of respondent</strong></td>
<td>−0.014</td>
<td>0.018</td>
<td>0.015</td>
<td>0.021</td>
<td>−0.038</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.043)</td>
<td>(0.043)</td>
<td>(0.156)</td>
<td>(2.333)</td>
<td>(0.035)</td>
</tr>
<tr>
<td><strong>Age, squared</strong></td>
<td>0.000</td>
<td>−0.000</td>
<td>−0.000</td>
<td>−0.000</td>
<td>0.000</td>
<td>−0.001</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.002)</td>
<td>(0.022)</td>
<td>(0.000)</td>
</tr>
<tr>
<td><strong>Δ household size</strong></td>
<td>0.014</td>
<td>0.014</td>
<td>0.016</td>
<td>0.007</td>
<td>−0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.034)</td>
<td>(0.043)</td>
<td>(0.713)</td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td><strong>Δ household members of working age (18–50)</strong></td>
<td>−0.057</td>
<td>−0.056</td>
<td>−0.062</td>
<td>−0.052</td>
<td>−0.044</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.041)</td>
<td>(0.100)</td>
<td>(1.082)</td>
<td>(0.032)</td>
<td></td>
</tr>
<tr>
<td><strong>Highest level of schooling: high school</strong></td>
<td>0.249</td>
<td>0.239</td>
<td>0.191</td>
<td>0.304</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.289)</td>
<td>(0.334)</td>
<td>(3.360)</td>
<td>(0.272)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Duration pre 1/2004, squared</strong></td>
<td></td>
<td></td>
<td>−0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.036)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nonadopter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(78.271)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>1.325</td>
<td>0.749</td>
<td>0.779</td>
<td>0.564</td>
<td>0.965</td>
<td>0.725</td>
</tr>
<tr>
<td></td>
<td>(0.892)</td>
<td>(0.864)</td>
<td>(0.892)</td>
<td>(3.718)</td>
<td>(22.050)</td>
<td>(0.761)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>157</td>
<td>142</td>
<td>142</td>
<td>142</td>
<td>142</td>
<td>142</td>
</tr>
<tr>
<td><strong>Adjusted R²</strong></td>
<td>0.23</td>
<td>0.22</td>
<td>0.22</td>
<td>0.20</td>
<td>0.10</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Notes: Δ Variable = Variable (2009) − Variable (1998). The difference in age squared is included as a control but is not statistically significant from zero, so is omitted from the table above. Standard errors reported in parentheses are bootstrapped. 
*P < 0.1; **P < 0.05; ***P < 0.01.