University of California, Davis

Department of Agricultural and Resource Economics

M.S. Comprehensive Exam, June 2020

You have THREE hours for this exam after a 15 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, which will be weighted equally.

Watch the time carefully. The logic used to answer each question is important, so be sure to clearly specify your reasoning using full sentences. 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 to be wrong.

QUESTION 1.

This question asks you to explore various dimensions of one key shock to food markets due to the coronavirus pandemic. Meat processing (pork, beef, chicken) was severely disrupted this spring when outbreaks of the virus among processing plant workers caused the temporary shutdown of several U.S. processing plants, resulting in higher consumer prices (P^R) and lower prices for live cattle (P^F). This increase in the farm-to-retail price spread, P^R-P^F, has led to calls from farm groups and politicians to investigate meat packers for violations of U.S. antitrust laws, i.e., laws to prevent unlawful exercise of market power.

1. Suppose you were asked to investigate the impact of the retail price changes on consumer welfare. To limit our scope, let us focus on a single meat product, pork. How would you propose to compute the welfare losses to U.S. pork consumers from this price increase? Focus on the relevant economic theory here and *not* on empirical measurement issues. In other words, assume you will be able from the available data to know the magnitude of price increase for pork products and how long the price increase persisted.

2. For many politicians and commentators, it did not make sense that consumer prices were higher after the plant closures, but farm prices were lower. This is what led to suspicion that the meat processors were manipulating the market. Suppose, however, that meat markets are perfectly competitive. Focusing again on a single meat, pork, use one or more graphs, given our assumption of perfect competition, to show how higher P^R and lower P^F could emerge as equilibrium responses to the processing plant closures.

3. Obviously, the quantities, Q, of pork sold by pork processors declined as a result of the plant closures. However, I am going to propose the following proposition: *sales revenues earned by pork processors actually* ***increased*** *as a result of the plant closures.* Your job is to discuss this proposition. You might agree or disagree, but the key to a good answer is to explain the economic factors that would go into determining whether my proposition is true or not. Graph(s) may be helpful. Define carefully any notation you introduce.

4. Pork processing in the U.S. is highly concentrated, i.e., a few firms procure most of the live hogs sold in the U.S., and sell most of the fresh pork products in the U.S. Suppose you were building an economic model of the U.S. pork industry and wanted to allow for the possibility that pork processors exercised market power in (i) acquiring live hogs from farmers and (ii) selling processed pork products downstream to food service and retailers. Show how you would model pork processors' behavior to allow for the exercise of market power. A good answer here will involve setting up an objective function for a representative processor and deriving first-order conditions. Hint: You can always choose your units of measurement (i.e. normalize) so that one unit of live hog produces one unit of processed pork and thereby have a single quantity variable for each.

QUESTION 2.
Suppose that you would like to investigate the cumulative effects of long-term smoking on the infection fatality rate (IFR) of the new respiratory disease, COVID-19. You have data from one country for all individuals who tested positive for the novel coronavirus. For simplicity, assume that every individual infected with the virus either dies or fully recovers within 14 days after being infected and your sample only includes people who were tested positive at least 14 days ago. Suppose you observe age of the individuals, A_i, years of smoking, S_i, and whether they recovered or passed away, Y_i (i.e., if a person recovers $Y_i = 0$, otherwise $Y_i = 1$). Let $D_i = 1$ if $S_i > 0$ and $D_i = 0$ if $S_i = 0$.

1. The first model specification that you try is the following:

$$P\{Y_i = 1\} = \beta_0 + \beta_1 D_i + \beta_2 D_i S_i + \gamma_1 A_i + \gamma_2 A_i^2. \tag{1}$$

   (a) Provide an interpretation for coefficients β_1 and β_2.

   (b) Why is it important for estimation of β_2 to include term $\beta_1 D_i$ in (1)? (Hint: People with known underlying lung conditions may avoid smoking in the first place and thus represent a different population.)

   (c) Write down a linear regression model corresponding to (1).

   (d) Suppose that the data (Y_i, S_i, A_i) is i.i.d. Is the linear regression model corresponding to (1) homoscedastic or heteroscedastic? Does your answer affect consistency of the OLS estimates of β_1 and β_2?

   (e) It is well known that respiratory diseases tend to have more severe effects on older populations. What would be the likely direction of the bias in β_2 if you were to omit the A_i terms in (1)?

   (f) Discuss drawbacks and limitations of the linear model specification (1) in the context of this research context.

2. Next, you come up with the following latent index model:

$$Y_i^* = \beta_0 + \beta_1 D_i + \beta_2 D_i S_i + \gamma_1 A_i + \gamma_2 A_i^2 + \varepsilon_i, \tag{2}$$

   where Y_i^* is an unobservable index that measures health condition of the lungs and ε_i represents all unobserved factors affecting the lungs condition. Healthy lungs correspond to low values of Y_i^*. The patient dies if Y_i^* surpasses a certain threshold, $\bar{Y}$. Suppose that $\varepsilon_i \sim N(0,1)$. The CDF and PDF of $N(0,1)$ are denoted $\Phi(z)$ and $\phi(z)$, correspondingly. You do not need to use the explicit forms of $\Phi(z)$ and $\phi(z)$ to answer the questions below.

   (a) What is the name of this model? (No need to describe its properties.)

   (b) Derive the IFR, which is the conditional probability of death given S_i and A_i.

   (c) What is the average effect of having smoked in the past on IFR? Does it depend on A_i?

   (d) What is the marginal effect of an additional year of smoking on the IFR? Does it depend on A_i?

3. Finally, you realize that only symptomatic patients get tested and end up in the sample. You assume that symptoms only occur in patients with $Y_i^* \geq \underline{Y}$. Suppose that $\underline{Y}$ is known and that $\underline{Y} < \bar{Y}$.

   (a) Discuss how $\underline{Y}$ would affect the IFR estimated without taking into account the sample selection.

(b) What is the probability of being tested conditional on S_i and A_i and being infected?

(c) What is the probability of death conditional on S_i and A_i and being infected and tested?

(d) Write down a likelihood function for the version of model (2) that takes into account sample selection.

QUESTION 3.

Economists are generally obsessed by markets, including where they emerge, how they function, when and why they fail to function well, and how they shape welfare outcomes. Many important market questions hinge on how markets interact across physical distance and whether they are spatially integrated.

1. Adam Smith argued that "The division of labor is limited by the extent of the market." Although he didn't articulate it explicitly or as an algebraic model, there is nonetheless a model behind this claim.
   (a) Describe two assumptions of the implicit model that supports this claim.

   (b) Briefly discuss the welfare implications of this claim.

2. The introduction of mobile phones has had important impacts on markets around the world. These impacts have been especially pronounced in places that previously lacked good communication or transportation infrastructure. Muto and Yamano (2009) study the impact of mobile phone coverage on market participation and other outcomes for farmers in Uganda. Consider key features of the conceptual framework they use in their analysis.

> 3. CONCEPTUAL FRAMEWORK
>
> Suppose that the farm-gate price of farmer i at time t of commodity j is defined as $p_{itj}^{FG} = p_{tj}^{M} - \gamma_j(I_t)\tau_i^2$. p_{tj}^{M} is the price of commodity j in the nearest market to the household i at time t. τ_i is the distance between the market and farmer i. We assume that $\gamma_j(I_t)$ is the sensitivity of the output price of commodity j with regard to the distance to the market and that the farm-gate price has a quadratic functional form with the distance to the market. As the information, I_t, increases by one unit, $\gamma_j(I_t)$ decreases: $\frac{\partial \gamma_j(I_t)}{\partial I_t} < 0$.

   (a) Use the subscripts as they appear on the four right-hand side variables in this simple farm-gate price model to explicitly state the assumptions the authors make.

   (b) The authors further assume that

$$\left|\frac{\partial \gamma_P(I_t)}{\partial I_t}\right| > \left|\frac{\partial \gamma_N(I_t)}{\partial I_t}\right|$$

   where P denotes a perishable commodity and N denotes a non-perishable commodity. Explain in words what this assumption means and provide a brief justification for it using a specific example that involves the introduction of mobile phones, which increases I_t.

(c) Take the 2nd cross-partial derivative of p_{itj}^{FG} with respect to I_t and then τ_i. Interpret this derivative as specifically as possible in the context of this research setting.

(d) Farmers throughout Uganda produce maize and bananas, much of which they consume in their own households. If they choose to sell any of their production in the market, they generally rely on traders. These traders buy from many farmers at p_{itj}^{FG} then deliver the goods to larger markets in the hopes of selling for a profit.

Without access to the mobile phone network, most farmers (especially farmers with limited land holdings (i.e., smallholder farmers)) have no reliable means of communicating with traders, who typically visit farmers in an area unannounced and try to fill their truck as quickly as possible with a specific crop produced by local farmers. This creates a coordination problem that delays the delivery to market and reduces the quality of a perishable good like bananas. In this setting, traders can exploit the fact that they have much better information than farmers.

i. True or False: "The costs associated with this coordination problem are not technically considered transaction costs." Justify your answer.
ii. True or False: "Asymmetric information between farmers and traders creates costs that are not technically considered transaction costs." Justify your answer.

(e) In a related article, Jensen (2007) ["The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector" *The Quarterly Journal of Economics* 122(3): 879-924] studies the impact of mobile phones on fishermen and fish markets in India. In contrast to the Ugandan farmers in Muto and Yamano (2009) who do not directly market what they produce, these fishermen both produce (i.e., capture) and market fish.

i. Carefully explain how reliance on traders may change the expected economic impacts of mobile phones on Ugandan farmers.
ii. Does reliance on traders alter the impacts you would expect to see of expanding the mobile network coverage on those who do not have access to their own mobile phone? Be specific.

3. Muto and Yamano (2009) use a panel dataset based on two large household surveys conducted in 2003 and 2005 in their analysis. They leverage the fact that the mobile phone network expanded from 46% of the population in 2003 to 70% in 2005 to estimate the impact of mobile phones on banana farmers.

   (a) Their identification strategy relies on the expansion of mobile phone coverage to some new villages between these two survey rounds. Specifically, in one part of their analysis they compare market-level outcomes in these new villages to those in villages that had coverage or lacked coverage in both years.
      i. This identification strategy relies on one key assumption. Explain what this assumption is and why it is central to their strategy.
      ii. Provide an example of a violation of this assumption that would clearly undermine causal identification in their case.

   (b) The authors also estimate market participation at the farmer-level using specifications that take as dependent variables the probability that a farmer sells banana and the ratio of sales to production (i.e., the share of total banana production that he sells) and use household fixed effects (FE) to account for time invariant household characteristics. They report their results in Table 7 (we focus here only on the results for Pr(selling banana).

Table 7. *Market participation and mobile coverage/possession: banana*

Variables	Pr(selling banana)			Ratio of sales quantity to production		
	FE	FE	FE-IV	FE	FE	FE-IV
Household mobile phone possession dummy[a]	0.203* (2.35)	0.209* (2.44)	0.151 (0.19)	0.118* (2.39)	0.121* (2.45)	−0.062 (0.14)
Community mobile phone coverage dummy	0.055 (1.57)	−0.095 (1.41)	−0.094 (1.11)	0.054** (2.68)	−0.011 (0.29)	0.000 (0.02)
Distance to district center (miles) × mobile coverage		0.008** (2.62)	0.007** (2.53)		0.003* (1.99)	0.003* (1.81)
Year 2005(=1)	0.481** (17.73)	0.502** (17.84)	0.505** (10.85)	0.174** (11.18)	0.183** (11.31)	0.191** (7.04)
F-stat on IVs			2.58			2.58
# of observations	1,161	1,161	1,151	1,161	1,161	1,151

Note: Numbers in parentheses are absolute *t*-values. The distance to the district center is not included in the models because it is fixed over time.
[a] Instrumental variables: The household mobile possession (HHmobile) is instrumented by the four interaction terms between the mobile coverage dummy and the four household characteristics: log of farm equipments value, age of household head, education of male adult, and education of female adult. These IVs together passed the over-identification test at the 1% significance level.
* Indicates significance at the 5% level.
** Indicates significance at the 1% level.

      i. Carefully compare the results in the first and second column of results in Table 7. Write a concise paragraph that interprets and evaluates these two columns of results. Pay particular attention to what we learn in the context of the research question as we move from the first to the second column.
      ii. In column FE-IV, the authors instrument for "Household mobile phone possession dummy" with the instruments described in footnote 'a.' Explain how the "*F*-stat on IVs" is computed. Why does it matter? How does the value of this F-stat help you to interpret these FE-IV results?