University of California, Berkeley
Department of Economics
Field Exam

Labor Economics

Summer 2020

There are three questions on this exam. **Please answer all three.** You should plan to spend about one hour per question. Using notes and other references is allowed, but do not communicate with anyone else about the exam.

Question I

Postel-Vinay and Robin (2002, EMA) – henceforth PVR – proposed a "sequential auction" model of employer competition where firms compete for workers by offering them the lowest wage necessary to poach them from rivals. In the PVR model, the output from matching a worker of type ϵ with a firm of productivity p is εp. Workers search on the job and encounter firm types drawn from a distribution $F(\cdot)$ of employer productivity. When an employed worker encounters another firm, the incumbent and new firm simultaneously submit wage offers to the worker. The worker accepts whichever offer yields the highest expected utility. In equilibrium, the winning offer will always be submitted by the more productive employer.

Denote the wage necessary to compel a worker of type ϵ to leave a firm with productivity q to join a firm of productivity $p > q$ by $\phi(\epsilon, p, q)$. PVR show that this "poaching wage" must obey the equation

$$U(\phi(\epsilon, p, q)) = U(\epsilon q) - \kappa \int_q^p \bar{F}(x) U'(\epsilon x) \epsilon dx$$

where $U(\cdot)$ is the worker's utility function, $\bar{F}(x)$ gives the probability of encountering an employer with productivity greater than level x, and $\kappa > 0$ is a constant.

a) Derive an expression for the log poaching wage when utility is logarithmic (i.e., when $U(\phi) = \ln \phi$).

b) Provide an economic interpretation to this equation.

c) Show that the log poaching wage is additively separable in the worker type ϵ, the productivity p of the hiring firm, and the productivity q of the poached firm. (Hint: use the fundamental theorem of calculus)

d) Compare and contrast the expression for log poaching wages with the usual AKM specification of log wages.

e) In your answer to part c), denote by $\lambda(q)$ the additive component of log poaching wages involving q and by $\psi(p)$ the component involving p. Show that $\lambda'(q) > 0$ while $\psi'(p) < 0$. Provide intuition for this result.

f) Suppose that for a set of firms $\mathcal{J}$, you managed to estimate all of the $\{\psi(p_j), \lambda(p_j)\}_{j\in\mathcal{J}}$ as fixed effects. How could you use these fixed effects to recover an estimate of the underlying log productivities $\{\ln p_j\}_{j\in\mathcal{J}}$?

g) Which component of poaching wages should be more variable across firms: $\psi(p_j)$ or $\lambda(p_j)$? Provide intuition for your answer.

h) Suppose you find that your fixed effect estimates $\left\{\hat{\psi}(p_j)\right\}_{j\in\mathcal{J}}$ are positively correlated with independent proxies of firm productivity (such as value added or firm size). What might explain this violation of the PVR model prediction that $\psi'(p) < 0$?

Question II

You are studying a multi-site randomized job training experiment. Treated subjects in the experiment received a voucher allowing them to participate in job training for free. The experiment was run separately at each of S experimental sites. At site $s \in \{1.....S\}$, experimental participants were randomly and independently assigned to treatment with probability $\pi_s \in (0,1)$. Let $Z_{is} \in \{0,1\}$ indicate whether individual i at site s was assigned treatment, let Y_{is} denote this participant's subsequent earnings, and let N_s denote the total number of participants at site s. You receive data $\{Y_{is}, Z_{is}\}_{i=1}^{N_s}$ for each site s.

1. Let $Y_{is}(1)$ and $Y_{is}(0)$ denote potential earnings for individual i at site s with and without the voucher. Define $\beta_s \equiv E[Y_{is}(1) - Y_{is}(0)]$ to be the average treatment effect at site s. Show that β_s is identified and propose an unbiased estimator $\hat{\beta}_s$. What feature of the experiment allows you to identify β_s?

2. Suppose the average treatment effect is constant across sites: $\beta_s = \bar{\beta}\ \forall s$. Consider the following ordinary least squares (OLS) regression pooling data across the sites:

$$Y_{is} = \alpha_s + \beta_{OLS} Z_{is} + \epsilon_{is},$$

where α_s is an experimental site fixed effect and $Cov(Z_{is}, \epsilon_{is}) = 0$ by definition. Explain why the OLS slope coefficient $\hat{\beta}_{OLS}$ is an attractive estimator of $\bar{\beta}$.

3. Now suppose that the average treatment effects β_s may differ across sites. Provide an expression for the OLS slope coefficient β_{OLS} in terms of the β_s, π_s, and N_s. Explain why $\hat{\beta}_{OLS}$ might still be a useful statistic to look at even if treatment effects are not constant.

4. A classmate suggests that you cluster your standard errors for the OLS regression from part (2) by experimental site. When you do this, you discover that the standard error for $\hat{\beta}_{OLS}$ increases substantially. Interpret this finding. Discuss arguments for and against clustering your standard errors in this case.

5. Suppose your estimator from part (1) is normally distributed and centered at the truth: $\hat{\beta}_s \sim N(\beta_s, \tau_s^2)$. Suppose you are also willing to assume the average treatment effects β_s are normally distributed across sites: $\beta_s \sim iid\ N(\mu, \sigma^2)$. Derive an expression for the posterior mean for the treatment effect at site s given your estimate, given by $\beta_s^* = E[\beta_s | \hat{\beta}_s]$. (You can treat τ_s^2, μ, and σ^2 as known.)

6. Suppose you are only interested in the average treatment effect at site 1, labeled β_1. Treating β_1 as a fixed and unknown parameter, provide expressions for the mean squared error (MSE) of $\hat{\beta}_1$ and β_1^* as estimators of β_1. On MSE grounds, when would you prefer to use each estimator?

7. You learn that many individuals in the control group who were denied the experimental voucher participated in job training anyway by paying for it themselves, while some members of the treatment group did not use the voucher. Let $D_{is} \in \{0,1\}$ indicate whether individual i at site s actually received job training. Propose a strategy for using data $\{Y_{is}, Z_{is}, D_{is}\}_{i=1}^{N_s}$ to estimate the effect of receiving job training at site s. What additional assumption(s) do you need for your strategy to work? Are there reasons to worry that these assumptions might be violated? Provide a brief discussion.

Question III

Note that there are 2 parts to this question.

1. Consider a model where each individual (indexed by i) has to choose an occupation $j \in \{1, 2, 3...J\}$. Suppose that the utility person i will get from occupation j is:

$$u_{ij} = v_j + \sigma\epsilon_{ij} \tag{1}$$

where v_j is a shared value, and $\sigma\epsilon_{ij}$ is an idiosyncratic component. We will assume that each of the $\epsilon_{ij}' s$ is distributed as Extreme Value Type 1 (EV-1). In this case we know that the average probability of choice j is

$$p_j = \frac{exp(v_j/\sigma)}{\sum_{k=1}^{J} exp(v_k/\sigma)}. \tag{2}$$

a) Explain why we have to impose a "normalization assumption" on the $v_j' s$.
Hint: show what happens if we transform the $v_j' s$ by adding some constant a to each valuation.

b) Assume from now on that to resolve the normalization problem we agree to set $v_1 = 0$. What is the interpretation of v_k for $k > 1$?

c) Consider the case where $\sigma \to 0$. What happens to the p_j's?

d) Consider the case where $\sigma \to \infty$. What happens to the $p_j' s$?

e) A well-known fact about models like (1) is that the expected value of the idiosyncratic error for choice j, conditional on choice j being the best one for a given individual (i.e., that $u_{ij} \geq u_{ij'}\ \forall\, j'$) is that:

$$E[\sigma\epsilon_{ij} | u_{ij} \geq u_{ij'}\ \forall\, j'] = \sigma(\gamma - \ln p_j)$$

where γ is Euler's constant.

(i). Use this expression to give an expression for $Emax(u_{ij})$ the expected maximum utility that individual i can attain if she chooses her occupation after she learns the values of $\{\epsilon_{i1}, \epsilon_{i2}, ...\epsilon_{iJ}\}$.

(ii) Show that $Emax(u_{ij})$ is increasing in σ. Explain why.

(iii) Let $v^{max} = \max_j v_j$. Develop an expression for $maxE[u_{ij}]$, the maximum expected value of u_{ij} that individual i can attain if she has to make a choice of occupations before she learns the values of $\{\epsilon_{i1}, \epsilon_{i2}, ...\epsilon_{iJ}\}$. (Hint: the mean of an EV-1 random variable is γ).

(iv) Define the option value (OV) of being able to delay making a choice until the $\epsilon_{ij}' s$ are known:

$$OV = Emax(u_{ij}) - maxE[u_{ij}]$$

Show how this is related to the set of relative probabilities $\{p_j/p_1\}$.

2. Suppose that we have a multinomial choice problem as in question 1, except now we assume that the utility that individual i attaches to choice j is

$$u_{ij} = x_i'\beta_j + \epsilon_{ij} \tag{3}$$

where x_i is set of observed characteristics of individual i, β_j is a choice-specific vector of coefficients, and ϵ_{ij} is an EV-1 error term. As in part 1 there is normalization issue, so we will assume that $\beta_1 = 0$ (so $x_i'\beta_1 = 0$ for all i). In this case we know that the multinomial logit probability of choice j is:

$$p_j = \frac{exp(x_i\beta_j)}{\sum_{k=1}^{J} exp(x_i\beta_j)}. \tag{4}$$

Suppose you were to select all the individuals who make either choice 1 or choice k, and within that sample estimate a simple logit model for the probability that i selected choice k rather than choice 1.(This is sometimes called a "conditional logit model"). Specifically suppose you assume that q_k, the probability of selecting choice k, conditional on selecting either choice 1 or choice k, has the standard logit form:

$$q_k = \frac{exp(x_i\delta_k)}{1 + exp(x_i\delta_k)} \tag{5}$$

a) Prove that if (4) is correct than $\delta_k = \beta_k$.

b) Explain how you can estimate the "multinomial logit" coefficients $\{\beta_k\}_{k=2}^{J}$ from a series of simple logits estimated on the conditional samples.