Ec142, Spring 2017

Professor Bryan Graham

Review Sheet 1

This review sheet is designed to assist you in your exam preparations. I suggest preparing written answers to each question. You may find it useful to study with your classmates. In the exam you may bring in a single 8.5 x 11 sheet of notes. No calculators or other aides will be permitted. Please bring blue books to the exam. The midterm exam will occur in class on Thursday, March 23rd.

[1] For $s \in \mathbb{S}$, a hypothetical years-of-schooling level, let an individual's potential earnings be given by $\log Y(s) = \alpha_0 + \beta_0 s + U$. Here U captures unobserved heterogeneity in labor market ability and other non-school determinants of earnings. Let the total cost of s years of schooling be given by $(\delta_0^* W + V^*) s + \frac{\kappa}{2} s^2$. Here W is an observable variable which shifts the marginal cost of schooling and V^* is unobserved heterogeneity. You may assume that both U and V^* are mean zero. Agents choose years of completed schooling to maximize expected utility

$$S = \arg\max_{s \in \mathbb{S}} \mathbb{E}\left[\log Y(s) - (\delta_0^* W + V^*) s - \frac{\kappa}{2} s^2 \middle| W, V\right].$$

[a] Show that observed schooling is given by

$$S = \gamma_0 + \delta_0 W + V, \quad \mathbb{E}[V | W] = 0$$

for $\gamma_0 = \beta_0/\kappa$, $\delta_0 = -\delta^*/\kappa$, and $V = -V^*/\kappa$.

[b] Assume that W measures commute time to the closest four year college from a respondent's home during adolescence. What sign do you expect δ_0 to have? Explain.

[c] Assume that $\mathbb{E}[U | W, V] = \mathbb{E}[U | V] = \lambda V$. Restate this assumption in words (HINT: Think about V as a latent variable/attribute). What sign do you expect λ to have? Briefly argue for and against this assumption?

[d] Let $\log Y = \log Y(S)$ denote actual earnings. Show that

$$\mathbb{E}^*[\log Y | S, V] = \alpha_0 + \beta_0 S + \lambda V. \tag{1}$$

[e] What determines variation in S conditional on $V = v$? What is the relationship between this variation and the unobserved determinants of log earnings? Use your answers to provide an intuitive explanation (i.e., use words) for why the coefficient on schooling in (1) equals β_0.

[f] The random sample $\{(Y_i, S_i, W_i)\}_{i=1}^N$ is available. Suggest a procedure for consistently estimating β_0.

[g] Let

$$\mathbb{E}^*[\log Y | S] = a_0 + b_0 S.$$

From you analysis in part [f] you learn that $\lambda \approx 0$. Guess what value b_0 takes. Justify your answer.

[2] Let $C_t = 1$ if an individual (child) went to college and zero otherwise. Let $C_{t-1} = 1$ if the corresponding individual's parent went to college and zero otherwise. The following table gives the joint distribution of father and sons' college attendance:

	$C_t = 0$	$C_t = 1$
$C_{t-1} = 0$	0.60	0.20
$C_{t-1} = 1$	0.10	0.10

For example 20% percent of the population consists of pairs with a father who did not attend college, but a son who did.

[a] Among children of college graduates, what fraction go on to complete college themselves? Among children of non-graduates, what fraction go on to complete college themselves?

[b] Let $\mathbb{E}^*\left[C_t \middle| C_{t-1}\right] = a + bC_{t-1}$; calculate a and b.

[c] The following table gives child's adult earnings, Y_t, for each of the four subpopulations introduced above

	$C_t = 0$	$C_t = 1$
$C_{t-1} = 0$	\$8,000	\$60,000
$C_{t-1} = 1$	\$14,000	\$30,000

What is the average earnings level of college graduates in this economy? What is the average earnings of non-college graduates? What is the overall average earnings level? Express your answers symbolically using the notation of (conditional) expectations and also provide a numerical answer.

[d] Let $\pi_{c_{t-1}} = \Pr\left(C_{t-1} = c_{t-1} \middle| C_t = 1\right)$. Consider the estimand

$$\beta = \sum_{c_{t-1}=0,1} \left\{\mathbb{E}\left[Y \middle| C_t = 1, C_{t-1} = c_{t-1}\right] - \mathbb{E}\left[Y \middle| C_t = 0, C_{t-1} = c_{t-1}\right]\right\} \pi_{c_{t-1}}.$$

In what sense does β adjust for "covariate differences" between college and non-college graduates [4 - 5 sentences]? Evaluate β and compare your numerical answer with the raw college - non-college earnings gap you calculated in part (c). Why are these two numbers different [2 to 4 sentences]?

[e] Jerry Brown is considering a community college expansion policy. You have been tasked to asked to predict the earnings gain associated with completing a college degree. Jerry estimates that after the community college expansion the distribution of college attendance in California will look like

	$C_t = 0$	$C_t = 1$
$C_{t-1} = 0$	0.40	0.40
$C_{t-1} = 1$	0.05	0.15

Calculate average earnings in this new economy (you may assume that the mapping from background and education into earnings introduced in part (c) remains the same)? Assume a state tax rate of 10 percent. What is the long run predicted increase in annual tax revenue from the community college expansion? Treat this revenue as a perpetuity and assume a discount rate of 0.05. What is the present value of the increase in tax revenue that is expected to be generated by the community college expansion?

[3] For a random draw from the population of US workers, let Y equal log earnings and X be a binary indicator taking a value of one if the worker is female and zero otherwise. Let $\{(Y_i, X_i)\}_{i=1}^{N}$ be a random sample of size N. Let N_1 denote the number of sampled units that are women (i.e., $X = 1$) and $N_0 = N - N_1$

the number that are male. Assume that

$$Y_i = \alpha_0 + \beta_0 X_i + U_i$$

with

$$Q_{U|X}\left(1/2|\, X\right) = 0.$$

Let a and b be a candidate values for 'the truth' (i.e., α_0 and β_0). Let $u^1_{1/2}(a,b)$ be the median of $U(a,b) = Y - a - bX$ given $X = 1$. Let $u^0_{1/2}(a,b)$ be the corresponding median given $X = 0$. Let $R_1(a,b), \ldots, R_{N_1}(a,b)$ denote the N_1 order statistics of $U(a,b)$ in the $X_i = 1$ subsample. Let $S_1(a,b), \ldots, S_{N_0}(a,b)$ denote the N_0 corresponding statistics from the $X_i = 0$ subsample.

[a] Interpret (in words) the parameters α_0 and β_0. What is true about the distribution of male versus female earnings if $\beta_0 = 0$?

[b] What is the median of $U(\alpha_0, \beta_0)$ given, respectively, $X = 1$ and $X = 0$?

[c] Assume that $a = \alpha_0$ and $b = \beta_0$. Let $j/(N_1+1) < 1/2 \leq (j+1)/(N_1+1)$. Before looking at your sample you are asked to guess the value of $(R_j(a,b) + R_{j+1}(a,b))/2$. What is your guess? Justify your answer.

[d] Let $N_1 = 3$ and $N_0 = 3$ (for this part of the problem only). Consider the order statistic intervals $[R_1(a,b), R_3(a,b)]$ and $[S_1(a,b), S_3(a,b)]$. Assume $a = \alpha_0$ and $b = \beta_0$; what is the ex ante probability that each of these intervals contain zero? Be sure to explain your work.

[e] Let a and b be some candidate intercept and slope values. Describe, in detail, an estimate of $u^0_{1/2}(a,b)$ and $u^1_{1/2}(a,b)$? Denote these estimates by, respectively, $\hat{u}^1_{1/2}(a,b)$ and $\hat{u}^0_{1/2}(a,b)$.

[f] Describe how to construct an approximate 95 percent confidence interval for $u^0_{1/2}(a,b)$ and $u^1_{1/2}(a,b)$?

[g] Describe how to construct an estimate of the asymptotic sampling variances of $\sqrt{N}\left(\hat{u}^1_{1/2}(a,b) - u^1_{1/2}(a,b)\right)$ and $\sqrt{N}\left(\hat{u}^0_{1/2}(a,b) - u^0_{1/2}(a,b)\right)$?

[h] Using your estimates from part (e) and sampling variance from part (g) sketch a procedure for testing the joint null hypothesis $H_0: \alpha_0 = a,\ \beta_0 = b$.

[4] Let Y equal tons of banana's harvested in a given season for a randomly sampled Honduran banana planation. Output is produced using labor and land according to $Y = AL^{\alpha_0}D^{1-\alpha_0}$, where L is the number of employed workers and D is the size of the plantation in acres and we assume that $0 < \alpha_0 < 1$. The price of a unit of output is P, while that of a unit of labor is W. These prices may vary across plantations (e.g., due to transportation costs, labor market segmentation etc.). We will treat D as a fixed factor; A captures sources of plantation-level differences in farm productivity due to unobserved differences in, for example, soil quality and managerial capacity. Plantation owners choose the level of employed labor to maximize profits. The observed values of L are therefore solutions to the optimization problem:

$$L = \arg\max_{l} P \cdot A l^{\alpha_0} D^{1-\alpha_0} - W \cdot l.$$

[a] Show that the amount of employed labor is given by

$$L = \left\{\alpha_0 \frac{P}{W} A\right\}^{\frac{1}{1-\alpha_0}} D. \tag{2}$$

[b] Let $a_0 = \frac{1}{1-\alpha_0} \ln \alpha_0 + \frac{1}{1-\alpha_0} \mathbb{E}[\ln A]$, $b_0 = \frac{1}{1-\alpha_0}$, and $V = \frac{1}{1-\alpha_0} \{\ln A - \mathbb{E}[\ln A]\}$. Show that the log of the labor-land ratio is given by

$$\ln\left(\frac{L}{D}\right) = a_0 + b_0 \ln\left(\frac{P}{W}\right) + V \tag{3}$$

and that, letting $c_0 = \mathbb{E}[\ln A]$ and $U = \ln A - \mathbb{E}[\ln A]$, the log of planation yield (output per unit of land) is given by

$$\ln\left(\frac{Y}{D}\right) = c_0 + \alpha_0 \ln\left(\frac{L}{D}\right) + U. \tag{4}$$

[c] Briefly discuss the content and plausibility of the restriction

$$\mathbb{E}[\ln A | \ln(P/W)] = \mathbb{E}[\ln A]. \tag{5}$$

[d] Using (3), (4) and (5) show that the coefficient on $\ln(L/D)$ in $\mathbb{E}^*[\ln(Y/D) | \ln(L/D)]$ equals

$$\alpha_0 + (1-\alpha_0) \frac{\mathbb{V}(\ln A)}{\mathbb{V}(\ln A) + \mathbb{V}(\ln(P/W))}.$$

Provide some economic intuition for this result.

[e] Using (3), (4) and (5) show that the coefficient on $\ln(L/D)$ in $\mathbb{E}^*[\ln(Y/D) | \ln(L/D), V]$ equals α_0. Provide some economic intuition for this result.

[f] Assume that all plantations face the same output price (P) and labor cost (W). What value does the coefficient on $\ln(L/D)$ in $\mathbb{E}^*[\ln(Y/D) | \ln(L/D)]$ equal now? Why?

[5] Consider the following model of supply and demand:

$$\begin{aligned} \ln Q_i^D(p) &= \alpha_1 + \alpha_2 \ln(p) + U_i^D \\ \ln Q_i^S(p) &= \beta_1 + \beta_2 \ln(p) + U_i^S, \end{aligned}$$

with i indexing a generic random draw from a population of 'markets' ; U_i^D and U_i^S are market-specific demand and supply shocks. We assume that $(U_i^S, U_i^D) \overset{i.i.d}{\sim} F$ for $i = 1, 2, \ldots, N$. In each market the observed price and quantity pair (P_i, Q_i) coincides with the solution to market clearing condition

$$Q_i^D(P_i) = Q_i^S(P_i) = Q_i.$$

[a] Provide an economic interpretation of the parameters α_2 and β_2. What signs do you expect them to take? Why?

[b] Depict the market equilibrium graphically. Solve for the equilibrium values of $\ln Q_i$ and $\ln P_i$ algebraically. How is the market price and quantity related to the demand and supply shocks, U_i^D and U_i^S? Provide some economic content for your answer. Can you use a figure to illustrate it?

[c] Calculate $\mathbb{E}^*[\ln Q | \ln P]$. You may assume that $\mathbb{C}(U^D, U^S) = 0$. Evaluate the coefficient on $\ln(P)$, does it coincide with an economically interpretable parameter? Assume that $\mathbb{V}(U_i^S) / (\mathbb{V}(U_i^S) + \mathbb{V}(U_i^D)) \approx 1$, does your answer change? Why?