Second Midterm Review Sheet, Part I

Ec240a – Second Half, Fall 2016

In preparing for the exam review your lecture notes, assigned course readings, problem sets and prepare answers to the questions on the review sheet(s). You may bring to the exam a single 8.5×11 inch sheet of paper with notes on it. No calculation aides are allowed in the exam (e.g., calculators, phones, laptops etc.). You may work in pencil or pen, however I advise the use of pencil. Please be sure to bring sufficient blue books with you to the exam if possible.

[1] Consider the following joint probability density function for x and y

$$f_{X,Y}(x,y) = \begin{cases} x^2 + \frac{8}{3}y^3, & (x,y) \in \mathbb{I} \times \mathbb{I} \\ 0 & (x,y) \notin \mathbb{I} \times \mathbb{I} \end{cases}$$

where $\mathbb{I}$ denotes the closed unit interval $[0,1]$.

[a] Show that $f_{X,Y}(x,y)$ is a valid probability density function.

[b] Compute the conditional density $f_{Y|X}(y|x)$.

[c] Compute the conditional expectation function $\mathbb{E}[Y|X=x]$.

[d] Compute the linear predictor $\mathbb{E}^*[Y|X=x]$.

[e] Consider a joint distribution with a conditional that coincides with the one derived in part [b] above, but where the marginal distribution of X is uniform on $\mathbb{I}$. Redo parts [c] and [d] using this alternative joint distribution. Explain your findings.

[2] Consider the following linear predictor

$$\mathbb{E}^*[Y|X] = \alpha_0 + \beta_0 X.$$

Let $X^* = X + U$ with U independent of (X,Y) and mean zero.

[a] Show that $\mathbb{C}(X^*, Y) = \mathbb{C}(X, Y)$.

[b] Show that $V(X^*) = \mathbb{V}(X) + \mathbb{V}(U)$.

[c] Let $\mathbb{E}[Y|X^*] = a_0 + b_0 X^*$. Derive expressions for a_0 and b_0 in terms of α_0, β_0, $\mathbb{V}(X)$ and $\mathbb{V}(U)$.

[d] Now let

$$\begin{pmatrix} Y \\ X \\ U \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_Y \\ \mu_X \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_Y^2 & \sigma_{YX} & \sigma_{YU} \\ \sigma_{YX} & \sigma_X^2 & \sigma_{XU} \\ \sigma_{YU} & \sigma_{XU} & \sigma_U^2 \end{pmatrix} \right).$$

Compute $\mathbb{E}[Y|X^*]$ for this new setup. HINT: First calculate the joint distribution of Y and X^* using the properties of the multivariate normal discussed in lecture.

[e] Comment on any implications of your results for linear regression analysis in the presence of regressor measurement error.

[3] Consider the statistical model

$$\begin{aligned} Y &= \alpha_0 + \beta_0 X + U \\ X &= \eta_0 + Z'\pi_0 + V \end{aligned}$$

with

$$\left(\begin{array}{c} U \\ V \end{array}\right)\Bigg| Z \sim \mathcal{N}\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_U^2 & \rho\sigma_V\sigma_U \\ \rho\sigma_V\sigma_U & \sigma_V^2 \end{array}\right)\right).$$

[a] Let $\mu^2 = \pi_0^{'}\mathbb{V}(Z)\pi_0/\mathbb{V}(V)$. Consider the (mean squared error minimizing) linear predictor of Y given X. Let b_0 be the coefficient on X in this linear predictor. Show that

$$b_0 = \beta_0 + \rho\frac{\sigma_U}{\sigma_V}\frac{1}{\mu^2+1}.$$

Comment on your result. Why does $b_0 \neq \beta_0$ for $\rho \neq 0$? Discuss the above statistical model in light of the Card and Krueger (1996) model of schooling discussed in the readings.

[b] Assume that $\pi_0 \neq 0$ and show that

$$\mathbb{E}^*\left[Y|\, X, V\right] = \alpha_0 + \beta_0 X + \rho\frac{\sigma_U}{\sigma_V}V.$$

Provide an intuitive explanation for why additionally conditioning on V ensures that the LP coefficient on X is equal to β_0. Explain why this linear predictor is not well-defined if $\pi_0 = 0$.

[c] Using a sample of $N = 32,587$ Honduran males aged 45 to 51 in 1988 we compute the least squares fit of the logarithm of monthly earnings (LogEarnings) onto a constant and years of completed schooling (YrsSch):

$$\text{LogEarnings} = \underset{(0.0064)}{5.2067} + \underset{(0.0010)}{0.1324}\ \text{YrsSch}. \tag{1}$$

In Honduras compulsory schooling for the above cohorts began in the first February after turning seven (i.e., the school year begins in February and you must attend school if you are seven years old). Let D_{FMA} be a dummy variable taking a value of one if an individual was born in February, March or April and zero otherwise, D_{MJJ} a dummy for being born in May, June or July and D_{ASO} a dummy for being born in August, September or October. A least squares of fit of YrsSch onto a constant and these three dummy variables using the same sample yields

$$\text{YrsSch} = \underset{(0.0539)}{4.1127} - \underset{(0.0742)}{0.3372}\ \mathrm{D_{FMA}} - \underset{(0.0734)}{0.2711}\ D_{MJJ} + \underset{(0.0769)}{0.0406}\ \mathrm{D_{ASO}}. \tag{2}$$

Let $\hat{V}$ be the fitted least squares residual associated with (2) above. The least squares fit of LogEarnings onto a constant, YrsSch of $\hat{V}$ is

$$\text{LogEarnings} = \underset{(0.1197)}{5.1765} + \underset{(0.0302)}{0.1400}\ \text{YrsSch} - \underset{(0.0302)}{0.0076}\ \hat{V}. \tag{3}$$

[i] The estimated asymptotic variance-covariance matrix (divided by the sample size) of the least

squares coefficient estimates reported in (2) above is

$$\begin{pmatrix} 0.0029 & & & \\ -0.0029 & 0.0055 & & \\ -0.0029 & 0.0029 & 0.0054 & \\ -0.0029 & 0.0029 & 0.0029 & 0.0059 \end{pmatrix}.$$

Test the hypothesis, at the $\alpha = 0.05$ level, that years of completed schooling cannot be predicted by quarter of birth. Provide a precise statement of this hypothesis in terms of the population analogs of the estimated coefficients reported in (2), construct a test statistic, and compare it to the appropriate critical value. For your reference the 0.95 quantiles of χ^2 random variables with parameters 1, 2 and 3 are, respectively, 3.84, 5.99 and 7.81. Why might quarter-of-birth predict years of completed schooling?

[ii] Test the hypothesis, at the $\alpha = 0.05$ level, that the coefficients on YrsSch in (1) and (3) coincide. Explain yourself.

[4] 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 $\left(U_i^S, U_i^D\right) \overset{i.i.d}{\sim} F$ 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 $\mathbb{C}\left(U^D, U^S\right) = 0$. Evaluate the coefficient on $\ln(P)$, does it coincide with an economically interpretable parameter? Assume that $\mathbb{V}\left(U_i^S\right) / \left(\mathbb{V}\left(U_i^S\right) + \mathbb{V}\left(U_i^D\right)\right) \approx 1$, does your answer change? Why?

[5] Let $Y \in \mathbb{Y} = \{0, 1, 2, 3, \ldots\}$ be a scalar 'count' variable, X a $K \times 1$ vector of continuously-valued regressors (here X *does not* include a constant), and A unobserved unit-specific heterogeneity. Assume that

$$\begin{aligned} \mathbb{E}[Y|X, A] &= \exp(\alpha_0 + X'\beta_0 + A) \\ X &= (I_K \otimes W)' \pi_0 + V \end{aligned}$$

with

$$\left.\begin{pmatrix} A \\ V \end{pmatrix}\right| W \sim N\left(\begin{pmatrix} \mu_A \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_A^2 & \Sigma_{AV} \\ \Sigma_{AV}' & \Sigma_{VV} \end{pmatrix}\right).$$

[a] Let $m^{\text{ASF}}(x) = \mathbb{E}_A\left[\mathbb{E}\left[Y|\, X = x, A\right]\right]$ denote the average structural function (ASF). Interpret this function. Under what conditions will $\mathbb{E}\left[Y|\, X = x\right] = m^{\text{ASF}}(x)$? Are these conditions satisfied for any configuration of parameter values in the family of models described above? Be precise and explain (in words).

[b] Show that

$$\mathbb{E}\left[Y|\, X, V\right] = \exp\left(\alpha_0^* + X^{'}\beta_0 + \Sigma_{AV}\Sigma_{VV}^{-1}V\right).$$

Provide a closed formed expression for α_0^*. Interpret $\Sigma_{AV}\Sigma_{VV}^{-1}V$; in what sense is this a 'proxy' for A? HINT: Recall that for $U \sim \mathcal{N}\left(\mu, \sigma^2\right)$ we have $\mathbb{E}\left[\exp\left(U\right)\right] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$.

[6] Consider the following joint probability density function for x and y

$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{11}\left(x^2 + y\right), & (x,y) \in [0,2] \times [0,1] \\ 0 & (x,y) \notin [0,2] \times [0,1] \end{cases}$$

[a] Show that $f_{X,Y}(x,y)$ is a valid probability density function.

[b] Compute the conditional density $f_{Y|X}(y|\, x)$.

[c] Compute the conditional expectation function $\mathbb{E}\left[Y|\, X = x\right]$.

[d] Compute the linear predictor $\mathbb{E}^*\left[Y|\, X = x\right]$.

[e] Consider a joint distribution with a conditional that coincides with the one derived in part [b] above, but where the marginal distribution of X is uniform on $[0, 2]$. Explain, qualitatively, how this change would affect your answers in parts [c] and [d] above?

[7] Consider the population of married men. Let Y denote log earnings for a generic random draw from this population, X his years of completed schooling and W the schooling of his spouse. Assume that the conditional mean of own log earnings given own and spouse's schooling is

$$\mathbb{E}\left[Y|\, X, W\right] = \alpha_0 + \beta_0 X + \gamma_0 W,$$

while the best linear predictor of spouse's schooling given own schooling is

$$\mathbb{E}^*\left[W|\, X\right] = \delta_0 + \zeta_0 X.$$

You may assume that the joint distribution of (W, X, Y) is such that these objects are well-defined.

[a] Show that $\zeta_0 = \rho_{WX}\frac{\sigma_W}{\sigma_X}$, with ρ_{WX} the correlation of W with X, and σ_W and σ_X respectively the standard deviation of W and X. Further show that $\delta_0 = \mu_W - \rho_{WX}\frac{\sigma_W}{\sigma_X}\mu_X$ with μ_W and μ_X denoting the population means of W and X.

[b] Using your answers in [a] above, as well as the form of $\mathbb{E}\left[Y|\, W, X\right]$, provide an expression for $\mathbb{E}^*\left[Y|\, X\right]$.

[c] Consider another population of married men where $F_{Y|W,X}(y|\, W = w, X = x)$, $F_W(w)$ and $F_X(x)$ coincide with those for the population described above, but where $F_{W,X}(w,x)$ differs. Assume that in this alternative population $\rho_{WX} = 0$. Solve for $\mathbb{E}\left[Y|\, X, W\right]$, $\mathbb{E}^*\left[W|\, X\right]$ and $\mathbb{E}^*\left[Y|\, X\right]$. Use the notation established in parts [a] and [b] to formulate your answer.

[d] Assume that $F_W(w)$ and $F_X(x)$ are identical and that marriage is homogamous in terms of education so that $W = X$ for all couples (i.e., individuals choose partners with identical levels of education). Show

that in this world $\rho_{WX} = 1$. Solve for $\mathbb{E}[Y|X, W]$, $\mathbb{E}^*[W|X]$ and $\mathbb{E}^*[Y|X]$. Use the notation established in parts [a] and [b] to formulate your answer.

[e] Compare the form of $\mathbb{E}^*[Y|X]$ in the original population with that in the two alternative populations of parts [c] and [d]. In which population does log earnings rise most steeply with years of schooling? Provide some intuition for your answer (5 sentences).

[f] Assume that schooling is binary valued, taking on the values 0,1. Let R_W be a 2×1 vector equal to $(1,0)'$ if $W = 0$ and $(0,1)'$ if $W = 1$. Let S_X be the analogous 2×1 vector defined using X. Let $T_{WX} = (R_W \otimes S_X)$ and

$$\mathbb{E}^*[Y|T_{WX}] = T'_{WX}\pi,$$

where a constant is not included and $\pi = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})'$. Show that

$$\pi_{jk} = \mathbb{E}[Y|W = j, X = k].$$

[g] Consider the null hypothesis that $\mathbb{E}[Y|W, X] = \alpha_0 + \beta_0 X + \gamma_0 W$. Maintaining this null find an explicit expression for each component of π in terms of α_0, β_0 and γ_0. Express this null in the form $C\pi = c$ for some matrix of constants C and vector of constants c.

[h] Let $W = 1$ if a wife has completed primary school and zero otherwise, let $X = 1$ if a husband has completed primary school and zero otherwise. A least squares fit, loosely based on data from Brazil, of log husband's earnings on T_{WX} as defined in [f] using a random sample of size $N = 50,000$ yields point estimate of

$$\hat{\pi} = \begin{pmatrix} 5.50 \\ 6.00 \\ 5.00 \\ 7.00 \end{pmatrix}$$

with an estimated asymptotic variance-covariance matrix of

$$\hat{\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Can you reject the null hypothesis (at the $\alpha = 0.05$ level) formulated in part [g] on the basis of this sample? For your reference the 0.95 quantiles of χ^2 random variables with parameters 1, 2 and 3 are, respectively, 3.84, 5.99 and 7.81.

[8] Let Y denote log-earnings and X years of completed schooling for a cohort of workers. Assume a random sample of size N is available from this population. Let $D_x = 1$ if $X = x$ and zero otherwise. Assume that $X \in \{0, \ldots, 16\}$ with positive probability attached to each support point.

[a] Let

$$\mathbb{E}^*[Y|D_1, \ldots, D_L] = \alpha_0 + \sum_{l=1}^{16} \gamma_{0l} D_l.$$

What is the relationship between this linear predictor and $\mathbb{E}[Y|X = x]$?

[b] Assume that $\Pr(X = 6) = 0$. Is the linear predictor defined in part [a] still well-defined? Why or why

not?

[c] You hypothesize that $\mathbb{E}[Y|X=x]$ is linear in x. Consider the linear predictor in part [a] and let $\beta = (\alpha, \gamma_1, \ldots, \gamma_{16})'$. Show how your hypothesis may be equivalently expressed as set of linear restrictions of the form $C\beta_0 = c$. Provide explicit expressions for C and c. Describe how you would construct a test statistics for your hypothesis. What is the asymptotic sampling distribution of your statistics under the null? Assume that your have a consistent estimate $\hat{\Lambda}$ of the asymptotic variance-covariance matrix of $\sqrt{N}\left(\hat{\beta} - \beta\right)$, with $\hat{\beta}$ the least squares estimate.

[d] After attending the labor lunch you now believe that $\mathbb{E}[Y|X=x]$ is linear in x but with discrete jumps at $X = 12$ and $X = 16$. Describe, in detail, how you would evaluate this new hypothesis?

[9] Let $A \in \{a_l, a_h\}$ denote an individual's unobserved 'entrepreneurial acumen', and X be a binary indicator taking a value of one if an individual completed an undergraduate degree and zero otherwise. Let Y equal annual earnings. The following table gives the conditional mean of Y for each of the four possible 'entrepreneurial acumen' and schooling combinations

	$A = a_h$	$A = a_l$
$X = 1$	\$45,000	\$35,000
$X = 0$	\$50,000	\$15,000

Assume that $m(x,a) = \mathbb{E}[Y|X=x, A=a]$ is a structural function in the following sense: in subpopulations homogenous in 'entrepreneurial acumen', $m(x,a)$, traces out how average earnings would change with external manipulations in college completion behavior. The population frequency of each of the four schooling and 'entrepreneurial acumen' combinations is

	$A = a_h$	$A = a_l$
$X = 1$	0.20	0.10
$X = 0$	0.05	0.65

[a] While on the elevator in Evans Hall you heard a grumpy individual (possibly a professor) claim "the best students should just start a tech firm in their parents' garages, we can train the rest to become corporate lawyers". Comment with reference to the population described above.

[b] Calculate the average annual earnings level in this economy, $\mathbb{E}[Y]$, and the averages conditional on college completion, $\mathbb{E}[Y|X=1]$, and not, $\mathbb{E}[Y|X=0]$.

[c] Calculate average earnings in a counterfactual world where $\Pr(X=1|A=a_l) = 1$ and $\Pr(X=1|A=a_h) = 0$.

[d] What is the expected earnings gain associated with college completion for a random draw from the population?

[e] Let $W = 1$ if an individual operated a lemonade stand at some point during childhood and zero otherwise. Assume that (i) $0 < \Pr(X=1|W=w) < 1$ for $w \in \{0,1\}$ and (ii) that X is conditionally independent of A given W. Show that for $q(x,w) = \mathbb{E}[Y|X=x, W=w]$ we have $\mathbb{E}[q(x,W)] = \mathbb{E}[m(x,A)]$. How would your answer change if $\Pr(X=1|W=1)$ were equal to one?

[f] Maintaining the assumptions of part [e] above show that

$$\mathbb{E}\left[\frac{\mathbf{1}(X=x)Y}{\Pr(X=x|W)}\right]=\mathbb{E}[m(x,A)].$$

Provide an intuitive discussion of this result.

[g] Available is a random sample of size N from the population of high school graduates. For each unit we observe $Z=(W,X,Y)'$. Let

$$R_1=(\mathbf{1}(X=0)\mathbf{1}(W=0),\mathbf{1}(X=0)\mathbf{1}(W=1),\mathbf{1}(X=1)\mathbf{1}(W=0),\mathbf{1}(X=1)\mathbf{1}(W=1))'$$

where $\mathbf{1}(\bullet)$ denotes the indicator function and

$$\mathbf{S}=\begin{pmatrix}Y\\W\\X\\WX\end{pmatrix},\mathbf{R}=\begin{pmatrix}R_1' & \underline{0}_3'\\ \underline{0}_3\underline{0}_3' & I_3\end{pmatrix}$$

with $\underline{0}_k$ a $k\times 1$ vector of zeros and I_k a $k\times k$ identity matrix. Establish the following notation: $\mu_{xw}=q(x,w)$, $\sigma_{xw}^2=\mathbb{V}(Y|X=x,W=w)$, $p_x=\Pr(X=x)$, $q_w=\Pr(W=w)$, and $r_{xw}=\Pr(X=x,W=w)$. Assume (i) σ_{xw}^2 is finite and (ii) that $r_{xw}>0$ for all four x and w combinations.

Consider the estimate

$$\hat{\beta}=\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_i'\mathbf{R}_i\right]^{-1}\times\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_i'\mathbf{S}_i\right].$$

[i] Show that $\hat{\beta}\to\beta_0$ and provide an expression for each of the seven elements of β_0 in terms of the notation established above (i.e. in terms of μ_{xw}, σ_{xw}^2 etc.). How would your analysis change if it were the case that $r_{11}=0$?

[ii] Show that $\sqrt{N}\left(\hat{\beta}-\beta_0\right)$ converges in distribution to a normal random variable. Provide an explicit expression for the covariance matrix of this normal distribution in terms of the notation established above (i.e. in terms of μ_{xw}, σ_{xw}^2 etc.).

[iii] Using the elements of $\hat{\beta}$ construct estimates of $\mathbb{E}(q(1,W))$ and $\mathbb{E}(q(0,W))$. Establish the consistency of these estimates.

[iv] You are interested in the joint hypothesis that $\mu_{10}=\mu_{00}$ and $\mu_{11}=\mu_{01}$. Discuss the substance of this hypothesis in light of the empirical set-up developed in parts [a] to [f] above. Show that this hypothesis may be represented as a restriction of the form $C\beta_0=c$ for some matrix C and column vector c.

[v] Your professor provides you will the following (consistent) estimates: $\hat{\mu}_{00}=10,000$, $\hat{\mu}_{01}=20,000$, $\hat{\mu}_{10}=50,000$, $\hat{\mu}_{11}=10,000$, $\hat{\sigma}_{00}^2=1,000$, $\hat{\sigma}_{01}^2=2,000$, $\hat{\sigma}_{10}^2=500$, $\hat{\sigma}_{11}^2=1,000$, $\hat{r}_{00}=0.2$, $\hat{r}_{01}=0.3$, and $\hat{r}_{10}=0.25$. Construct a test statistic for the hypothesis described in part [iii] above and compare it to the appropriate critical value. For your reference the 0.95 quantiles of χ^2 random variables with parameters 1, 2 and 3 are, respectively, 3.84, 5.99 and 7.81.

[10] Consider the following statistical model for the earnings of Berkeley students

$$Y=\alpha+\beta G+2A+U,\ \mathbb{E}^*[U|G,A]=0,$$

where G equals 1 if the student graduated and zero if they dropped out; A equals the sum of the math and verbal SAT scores.

[a] Let $\mathbb{E}^*[A|G] = \delta_0 + \delta_1 G$; calculate $\mathbb{E}^*[Y|G]$ (symbolically, at least in part)

[b] You read in the Oakland Tribune newspaper that Berkeley graduates earn an average of \$65,000 per year nationwide, while the earnings of dropouts average only \$35,000. Relate this population earnings differential to the parameters of the statistical model given above.

[c] The same article reports that average SAT scores for Berkeley graduates equal 1300, while for dropouts they equal 1100. Use this information to calculate δ_0 and δ_1.

[d] You are considering dropping out of Cal to spend more time on Telegraph Avenue. What is the expected earnings cost associated with this decision? Explain. [2 - 3 sentences]

[e] You move to Oakland upon graduation, your neighbor to the left tells you that he dropped out of Berkeley during the Free Speech Movement, your neighbor to the right tells you that he graduated from Berkeley about the same time. What is your expectation of the annual earnings of your two neighbors. Explain. [2 - 3 sentences]

[11] Let $C_t = 1$ if an individual (son) went to college and zero otherwise. Let $C_{t-1} = 1$ if the corresponding individual's father 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 son's of college graduates, what fraction go on to complete college themselves? Among son's of non-graduates, what fraction go on to complete college themselves?

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

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

	$C_t = 0$	$C_t = 1$
$C_{t-1} = 0$	\$7,000	\$60,000 .
$C_{t-1} = 1$	\$28,000	\$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 $\pi_{c_{t-1}} = \Pr\left(C_{t-1} = c_{t-1}|\, C_t = 1\right)$. Consider the estimand

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

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 asked to predict the effects of the new policy. Jerry estimates that after the community college expansion the distribution of college attendance in California will look like

	$C_t = 0$	$C_t = 1$
$C_{t-1} = 0$	0.50	0.30
$C_{t-1} = 1$	0.10	0.10

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?

[f] Did you use β as defined in part (d) above for your analysis in part (e)? Why or why not? [3 to 8 sentences].

[12] Let $D = 1$ if an individual graduated from college and zero if they stopped their education after completing high schools (high school dropouts are excluded from the population under consideration), let be X a vector of respondent attributes measured at, or prior to, high school graduation, and Y log-earnings as an adult. Assume that potential earnings are given by

$$Y(d) = \alpha + \beta D + U,$$

for $d = \{0, 1\}$. Assume that $U \perp D | X$.

[a] Consider the alternative response function

$$Y(d) = \alpha + \beta D + U(d),$$

with $(U(0), U(1)) \perp D | X$. In what sense is this model less restrictive? Which model is more appropriate for evaluating the returns to college attendance? [4 to 6 sentences]

[b] Let $e(x) = \Pr(D = 1 | X = x)$ denote the conditional probability of treatment given $X = x$. Using the original (more restrictive) model show that

$$\mathbb{E}[Y(D - e(X))] = \beta \mathbb{E}[D(D - e(X))].$$

[c] Consider the estimator

$$\hat{\beta} = \frac{\frac{1}{N}\sum_{i=1}^{N} Y_i (D_i - \hat{e}(X_i))}{\frac{1}{N}\sum_{i=1}^{N} D_i (D_i - \hat{e}(X_i))},$$

with $\hat{e}(X_i)$ an estimate of $e(X_i)$. Make an (informal) argument for the consistency of this estimate for β. Can you provide an intuitive explanation for it? [3 to 5 sentences]

[d] Assume that

$$\mathbb{E}[U | X] = X'\gamma.$$

Consider the linear regression of Y onto a constant, D and X. Show that the coefficient on D coincides with β under the maintained assumptions.

[e] Consider the following ratio of expectations

$$\frac{\mathbb{E}\left[\left(Y - X'\gamma^*\right)\left(D - e^*\left(X\right)\right)\right]}{\mathbb{E}\left[D\left(D - e^*\left(X\right)\right)\right]}.$$

Show that this expectation equals β when (i) $\gamma^* = \gamma$ and $e^*(X) \neq e(X)$, (ii) $\gamma^* \neq \gamma$ and $e^*(X) = e(X)$, and (i) $\gamma^* = \gamma$ and $e^*(X) = e(X)$ with γ and $e(X)$ as defined in parts (b) and (d) above. Suggest an estimator based on your analysis and discuss some of its possible advantages relative to those featured in parts (c) and (d) above. [8 to 15 sentences].

[13] 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

$$S = \gamma_0 + \delta_0 W + V$$

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. 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

$$\mathbb{E}^*[\log Y|\, S, V] = \alpha_0 + \beta_0 S + \lambda V. \tag{4}$$

[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 (4) 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.