Ec142, Spring 2018

Professor Bryan Graham

Review Sheet 2

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, April 26th.

[1] You are given a random sample from South Africa in the late 1980s. Each record in this sample includes, Y, an individual's log income at age 40, X the log permanent income of their parents, and D a binary indicator equaling 1 if the respondent is White and zero if they are Black. Let the best linear predictor of own log income at age forty given parents' log permanent income and own race be

$$\mathbb{E}^*[Y|X,D] = \alpha_0 + \beta_0 X + \gamma_0 D.$$

[a] Let $Q = \Pr(D=1)$, assume that $\mathbb{V}(X|D=1) = \mathbb{V}(X|D=0) = \sigma^2$ and recall the analysis of variance formula $\mathbb{V}(X) = \mathbb{V}(\mathbb{E}[X|D]) + \mathbb{E}[\mathbb{V}(X|D)]$. Show that

$$\mathbb{V}(X) = Q(1-Q)\{\mathbb{E}[X|D=1] - \mathbb{E}[X|D=0]\}^2 + \sigma^2.$$

[b] Let $\mathbb{E}^*[D|X] = \kappa + \lambda X$. Show that

$$\lambda = \frac{Q(1-Q)\{\mathbb{E}[X|D=1] - \mathbb{E}[X|D=0]\}}{Q(1-Q)\{\mathbb{E}[X|D=1] - \mathbb{E}[X|D=0]\}^2 + \sigma^2}.$$

[c] Assume that $\beta_0 = 0$. Show that in this case $\gamma_0 = \mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0]$.

[d] Let $\mathbb{E}^*[Y|X] = a + bX$. Maintaining the assumption that $\beta_0 = 0$ show that

$$b = \frac{Q(1-Q)\{\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0]\}\{\mathbb{E}[X|D=1] - \mathbb{E}[X|D=0]\}}{Q(1-Q)\{\mathbb{E}[X|D=1] - \mathbb{E}[X|D=0]\}^2 + \sigma^2}.$$

[e] Let $Q(1-Q) = 1/10$, $\sigma^2 = 3/10$ and $\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0] = \mathbb{E}[X|D=1] - \mathbb{E}[X|D=0] = 3$. Provide a numerical value for $\mathbb{V}(X)$ and b.

[f] On the basis of β_0 a member of the National Party argues that South Africa is a highly mobile society. One the basis of b a member of the African National Congress argues that it is a highly immobile one. Comment on the relative merits of these two assertions.

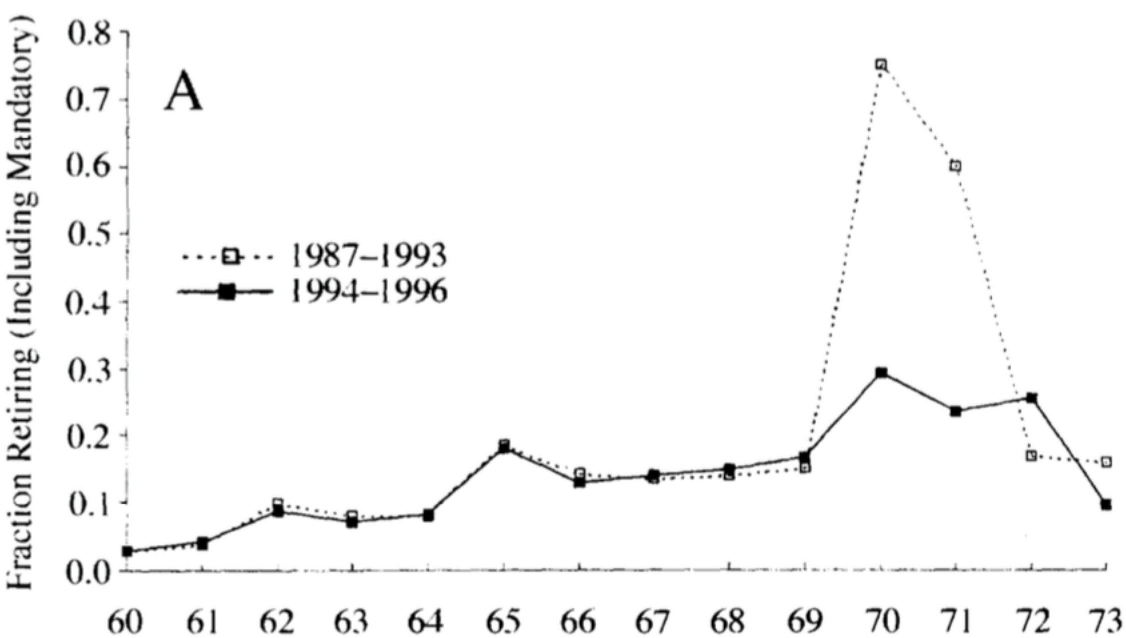
[2] Prior to 1994 colleges and universities in the United States were exempt from laws prohibiting mandatory retirement, consequently many institutions forced faculty to retire at age 70. After 1994 mandatory retirement rules were prohibited by Congress. Ashenfelter and Card (*AER*, 2002) study the effects of this exemption expiration on faculty retirement behavior in a sample of 104 colleges and universities. Let Y^* equal age of retirement, let C equal the age at which a faculty-member is lost to follow-up, $D = 1(Y^* \leq C)$ be a censoring indicator, and $Y = \min(Y^*, C)$ be the observed age at exit from the sample. Let T_y (for $y = 60, 61 \ldots, 72, 73$) equal the calendar year during which an individual was age y. So, for example, an individual who turned sixty in 1992 would have $T_{60} = 1992$, while one who did so in 1999 would have $T_{60} = 1999$.

[a] Let $X = 0$ if $T_{70} < 1994$ and $X = 1$ if $T_{70} \geq 1994$. Assume that

$$\lambda(y; X) = \Pr(Y^* = y | Y^* \geq y, X) = \frac{\exp(\alpha_y + X\beta_0 + X \times \mathbf{1}(y \geq 70)\gamma_0)}{1 + \exp(\alpha_y + X\beta_0 + X \times \mathbf{1}(y \geq 70)\gamma_0)}. \quad (1)$$

In the context of the Ashenfelter and Card (*AER*, 2002) study interpret the hazard function $\lambda(y; X)$ when $X = 0$ and when $X = 1$. Interpret β_0 and γ_0 in terms of the hazard function.

[b] Reference the figure below when answering the following questions (justify your answers). What signs do you expect β_0 and γ_0 to take? Does the evidence appear consistent with the hypothesis that $\beta_0 = 0$ and $\gamma_0 < 0$?



[c] Assume that $D \perp Y^* | X$. Interpret this assumption. Describe how it could be

violated.

[d] Assume the first four lines of the Ashenfelter and Card (*AER*, 2002) dataset equal

	Y	D	X
1	65	0	0
2	72	1	0
3	61	1	1
4	70	0	1

What are these units' contributions to the corresponding "person-period" dataset (in the sense described in the Singer and Willett book)? Write out the corresponding rows. Describe, in detail, how you could use this person period dataset to construct estimates of $\alpha_{60}, \alpha_{61}, \ldots, \alpha_{73}$, β_0 and γ_0 .

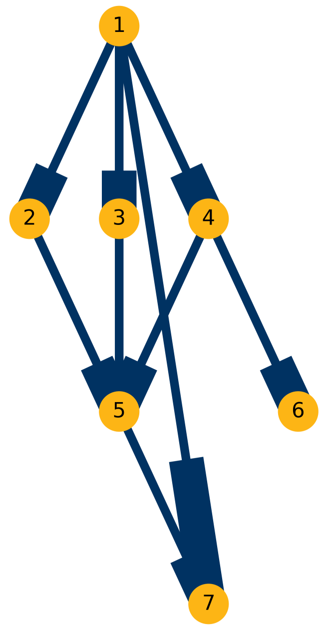
[e] Let $S(y; X) = \Pr(Y^* > y|\, X)$. Does $\Pr(Y > y|\, X) = S(y; X)$? If not, does $\Pr(Y > y|\, X) > S(y; X)$ or $\Pr(Y > y|\, X) < S(y; X)$? Why? Describe a method for constructing an estimate of $S(y; X)$. Describe, in detail, how you could use this estimate to compute the effect of the end of mandatory retirement on median retirement age. Use the information in the table below to implement your procedure. Specifically construct an estimate of $S(y; 1)$ and $S(y; 0)$ for $y = 60, 61...., 72$. Use your results to construct an estimate of the median retirement age before and after 1994.

TABLE 2—AGE-SPECIFIC RETIREMENT RATES, BEFORE AND AFTER 1994

			Average retirement rate		Change in retirement rate	
Age	Number of observations	Percentage post-1994	1987–1993	1994–1996	Unadjusted	Adjusted from logit
60	7,343	31.8	3.3	3.0	−0.3	−0.2
			(0.3)	(0.4)	(0.4)	(0.5)
61	7,027	32.4	4.1	4.4	0.3	0.3
			(0.3)	(0.4)	(0.5)	(0.5)
62	6,665	32.9	10.3	8.9	−1.4	−1.4
			(0.5)	(0.6)	(0.8)	(0.8)
63	5,838	34.5	8.5	7.3	−1.3	−1.1
			(0.5)	(0.6)	(0.7)	(0.8)
64	5,222	35.4	8.4	8.5	0.1	0.1
			(0.5)	(0.7)	(0.8)	(0.8)
65	4,650	35.1	19.3	18.1	−1.2	−1.4
			(0.7)	(1.0)	(1.2)	(1.3)
66	3,653	35.1	14.7	13.0	−1.7	−1.9
			(0.7)	(0.9)	(1.2)	(1.3)
67	2,969	34.2	13.8	14.0	0.1	−0.1
			(0.8)	(1.1)	(1.3)	(1.4)
68	2,453	34.2	14.3	14.6	0.4	0.7
			(0.9)	(1.2)	(1.5)	(1.5)
69	2,004	33.7	15.4	16.7	1.3	0.6
			(1.0)	(1.4)	(1.7)	(1.7)
70	1,598	35.1	75.6	29.1	−46.5	−43.7
			(1.3)	(2.0)	(2.4)	(2.5)
71	502	58.6	60.6	23.8	−36.8	−32.2
			(3.4)	(2.5)	(4.2)	(4.0)
72	182	67.0	16.7	25.4	8.7	−3.7
			(4.9)	(4.0)	(6.3)	(7.2)

Notes: Retirement rates expressed as percent per year. Estimated standard errors are in parentheses. An individual's retirement age is measured as of September 1 following the date of retirement. The adjusted change in retirement rates is the normalized regression coefficient from a logit model for the event of retirement, fit by age and including a total of 19 covariates: gender, Ph.D., nonwhite race, region (three dummies), Carnegie classification and public/private status of institution, and six department dummies.

[3] The figure below depicts a (hypothetical) supply chain. For example firm 1 sells inputs to firms 2, 3, 4 and 7; firm 6 sells inputs to firm 5 and so on.



[a] Let $\mathbf{D} = [D_{ij}]$ where

$$D_{ij} = \begin{cases} 1, & \{i,j\} \in \mathcal{E}(G) \\ 0, & \text{otherwise} \end{cases} . \tag{2}$$

That is $D_{ij} = 1$ if firm i "sells" to firm j (and zero otherwise). Fill in the table below to construct the adjacency matrix $\mathbf{D}$ for the depicted supplier-buyer network.

		Buyers						
		1	**2**	**3**	**4**	**5**	**6**	**7**
Suppliers	**1**	0						
	2		0					
	3			0				
	4				0			
	5					0		
	6						0	
	7							0

[b] Let A_i equal the productivity of firm i. Assume that

$$A_i = \alpha_0 + \beta_0 \left\{ \frac{\sum_{j=1}^{N} D_{ji} A_j}{\max\left(1, \sum_{j=1}^{N} D_{ji}\right)} \right\} + V_i \tag{3}$$

with V_i independently and identically distributed across agents with mean zero and variance σ^2. Interpret, in words, equation (3). Why might the productivity of a firm vary with that of its suppliers?

[c] Assume firm 1 experiences a shock to V_1 of σ. What is the effect of this shock on the productivity of firm 1's direct customers, firms 2, 3, and 4?

[d] What is the effect of the shock on the productivity of the customers of firm 1's customers? On the customers of those firms customers?

[e] What is the social multiplier associated with this shock to firm 1's productivity?

[f] What is the social multiplier associated with a one standard deviation shock to V_6? How does your answer differ from the one given in [e] above? Why?

[4] 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 conditionally mean zero given W. 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

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

[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.