Problem Set 1
11.126/11.249/14.48
Due February 27, 2007

1. (25 points) Suppose we have a data set with information on N men containing their earnings, their ages, their fathers' earnings, and their fathers' ages. We estimate a regression of the form

$$\ln y_i^s = \alpha + \beta \ln y_i^f + \gamma_1 a_i^s + \gamma_2 a_i^{s2} + \delta_1 a_i^f + \delta_2 a_i^{f2} + \epsilon_i$$

where y_i^s is son's earnings averaged over several years, y_i^f is father's earnings averaged over several years, a_i^s is son's age, and a_i^f is father's age.

(a) (5 points) What is the interpretation of β? Try to be precise.

Answer: β is the percentage increase in our best prediction of y_i^s for a one percent increase in y_i^f, holding the son's and father's ages constant. Alternatively, β is the elasticity of our best prediction of y_i^s with respect to y_i^f, again holding son's and father's ages constant.

(b) (5 points) Is β likely to be a good estimate of the effects of a pay raise on the earnings of a man's son? Explain why or why not.

Answer: Probably not. Regression estimates tell us how to predict one variable by using another, but they can't always be interpreted as causal. Here, we could estimate $\beta > 0$ because high IQ fathers have high IQ sons, and IQ increases wages. So the father's wages are a good predictor of the son's wages (the father's wages are correlated with son's IQ), but it doesn't follow that if we increased the father's wages without changing the son's IQ, then his son's earnings would rise. Alternatively, father's wages could be a proxy for the "culture" in which the son grew up, and some cultures may be more conducive to high earnings than others. But it doesn't follow that if we increase the father's earnings, then we will change the cultural upbringing of the son.

Note that the question was ambiguous as to whether we're increasing the father's earnings while the son is still a child or later on, once the son has grown up. In the latter case, it's very unlikely that we can interpret our estimate of β as causal; but the causal interpretation is problematic even in the former case.

(c) (5 points) If we estimate this regression because we are interested in determining the degree of mobility in the US economy, does the answer to (b) matter? Why or why not?

Answer: This depends a bit on how we interpret "mobility." If high mobility means that poverty *per se* doesn't diminish someone's life prospects, then β is relevant to mobility if and only if it is causal (i.e.,

the answer to (b) is yes). But a more natural interpretation of mobility is that growing up in a poor household doesn't mean you have poor economic prospects. In that case, we don't really care whether β is large because money during childhood itself affects earnings later or β is large because it proxies for genetic factors, culture, etc. (and these other factors affect son's earnings in adulthood). All we care about is whether boys who grow up in poor households end up poor themselves, and β answers that question whether or not it is causal. This is all quite subtle but important. Whenever you see a regression, you should ask two questions: 1) Can I interpret this regression coefficient as an estimate of a causal effect (and not just a correlation in the data), and (if not) 2) Is the author's argument presuming a causal interpretation? Here the regression probably isn't causal but that's not necessarily a big problem. Often it is unclear whether a regression estimate can be taken as causal, and the author definitely IS assuming a causal interpretation.

Some people misread the question to say, "does β matter?" The answer to this question is obviously yes.

(d) (10 points) Suppose we compare two economies with the same β. In the first economy, the standard deviation of ϵ_i is very low (most people's ϵ_i is close to 0), while in the second economy the standard deviation is very high (large negative and large positive values of ϵ_i are common).

   i. Given the same y_i^f, a_i^s and a_i^f, is our best guess of y_i^s the same or different in the two economies? Explain why.
Answer: Our best guess of $\ln y_i^s$ is the same: $\alpha + \beta \ln y_i^f + \gamma_1 a_i^s + \gamma_2 a_i^{s2} + \delta_1 a_i^f + \delta_2 a_i^{f2}$.
One person pointed out that it doesn't follow that our best guess of y_i^s is the same, because ln is a concave function; this is a technical and unintended complication, so don't worry about it.

  ii. In each economy, we can rank the sons and the fathers by their earnings, with 1 being the lowest rank and N being the highest. Suppose we want to know the probability that a son's rank exceeds his father's. Will this probability be the same or different in the two economies? Explain why.
Answer: The probabilities will be different in general. If the father had a low rank, then the probability that the son's rank will exceed his father's will be higher in the high-variance economy; if the father had a high rank, the opposite will be true. To see this, imagine the limiting case where $\epsilon_i = 0$ for everyone and all sons and fathers have the same age; in that case, everyone has the same rank as their father. If we introduce some randomness (increase the standard deviation of ϵ_i), then clearly I have at least some chance of exceeding my father's rank. If we make the standard deviation of ϵ_i approach infinity, then everyone has the

2

same chance of ending up in a given rank, because the differences in ϵ_i swamp differences in $\beta \ln y_i^f$. That means that I have a $1/N$ chance of ending up at any rank, no matter where my father was, and if my father's rank was r then I have a $(N-r)/N$ chance of exceeding his rank.

Alternatively, imagine a world with 2 people in each generation. If the low-variance economy, the son of the poor father is very likely to remain the poor man of his generation. In the high-variance economy, he may either become very, very poor or very rich. In the former case he stays at the same rank (1), while in the latter case he has a good shot at surpassing the son of the rich man.

The lesson: depending on what we mean by "mobility", we may care about more than the value of β.

This question was hard–don't worry if you didn't get it right.

2. (25 points) Write a short essay (a couple of paragraphs) exploring the patterns in Figure 4 of "The Consequences of Increasing the Nation's Supply of College Graduates." In particular, you may want to touch on a) why male enrollment dipped in the 1970's and then grew slowly until the mid-1990's, and b) why the pattern for women is different.

Answer: Answers will vary, and there isn't a "right" answer here. Some points that one could discuss:

- Male enrollment dipped in the 1970's. There are several possible reasons for this dip, including

  - Lower wage premia for college education in the 1970's, as Freeman discusses.
  - The Vietnam war. Several people talked about the role of the war in delaying or preventing college education for young men in the 1970's. The conventional wisdom is instead that the war encouraged college enrollment in the late 1960's as a means to avoid the draft. The draft ended in 1973, so that there was now less incentive to go to college as an alternative to Vietnam.
  - The baby boom. If the supply of college spots is not perfectly elastic, then large cohorts will tend to have lower enrollment rates than small cohorts. The baby boom generation started to hit college age roughly around the time the enrollment rate of young men fell.
  - Increased competition for college spots from women.

- Male enrollment increased from about 1980 onward as the college wage premium rose, but it never surpassed the levels of the late 1960's, even though the college wage premium did. The reasons for this anomaly are unclear, but some possibilities are 1) a natural limit in the proportion of the population with the cognitive skills to

3

complete college and 2) the increase in the proportion of Black and Hispanic men in the population, who typically have had lower college attendance rates.

- Women's enrollment rates increased steadily with barely a blip in the 1970's. This is most likely due to women's expanding presence in the labor force–homemaking doesn't usually require a college degree.

  I gave "good" answers grades in the low 20's, so 25 means an exceptional essay. Clarity (not just content) counts for something.

3. (20 points) Read the article by Leon Feinstein on the reading list.

   (a) (6 points) The article shows that cognitive development at young ages predicts educational outcomes in adulthood. Give at least two stories where this does *not* imply that an intervention that improves childhood test scores will have a causal effect on adult education.

   Answer: Answers will vary, but some possibilities:

   - High childhood test scores are a proxy for a home environment that is supportive of education (e.g., parents read to their children). Children from this type of home environment will be much more attracted to school, so childhood test scores successfully predict schooling later in life. However, improving the test scores without changing the home environment won't affect schooling levels. More generally, childhood test scores could be correlated with all sorts of environmental characteristics, and these omitted variables might be the ones that really have a causal effect on later schooling. (Remember omitted variables bias in regression?) Feinstein controls for a crude measure of SES, but that might not be enough.

   - Young children who score high on tests are naturally bright, and naturally bright children eventually get a lot of schooling. However, an intervention that improves test scores need not make kids any brighter; it might just improve the skills required for the test. More generally, childhood test scores could be correlated with various "innate" characteristics of children, and these omitted variables might be the ones that really have a causal effect on schooling.

   - We can imagine interventions improving test scores in two ways: by teaching broad skills that are applicable to many situations, or by teaching a narrow set of skills that are useful for the test but mostly useless everywhere else. We probably wouldn't expect an intensive month-long, full-time course in cube-stacking to have an effect on kids' educational attainment.

   (b) (7 points) A skeptic looks at Figures 2 and 3 and complains, "These figures use too crude a measure of SES by splitting children into only

4

two groups. We can't tell from looking at these figures whether early childhood test scores really do predict later test scores among people with the same family background." Does the skeptic's criticism make sense? Explain why or why not.

Answer: Yes. Feinstein's Figures use two categories of SES, high and low (putting aside the omitted middle category); suppose that the low group really contains two sublevels, 1 and 2. Suppose also that a person's precise SES level does predict educational attainment, but among families at the same SES level test scores do not. However, higher SES is correlated with higher test scores. When we look within the low category of SES, the higher testing kids will be predominantly from SES 2 while the low test score kids will be predominantly from SES 1. Thus we will observe that the high test score kids get more education later on than the low test score kids, *within the broad SES category "low"*. This is true even though high test score kids do not get more education *within the same narrow SES level* (1 or 2). The same pattern could hold for the high SES category.

The lesson: if we control for some factor by using a crude measure, then we haven't completely controlled for it.

This was a hard question, so don't worry too much if you didn't get it.

(c) (7 points) Another skeptic complains, "The tests at different ages are not comparable at all–one is measuring things like cube stacking and another is measuring things like reading. We can't infer from Figures 2 and 3 that "ability" at early ages predicts "ability" later in any meaningful way." Does the skeptic's criticism make sense? Explain why or why not.

Answer: No, probably not. We can think of each test as measuring a weighted average of "general ability" g and test-specific ability s. The critic is saying that the overall test score probably puts a lot of weight on s and not much on g. That might be true, but it will make it harder for one test score to predict another–if there was no weight on g at all, we wouldn't expect the test scores to be correlated at all. So this criticism should lead us to conclude that the correlations among test scores understate the persistence of "ability" rather than overstating it. To put it another way: if the critic were right that the tests are measuring totally different things, we wouldn't observe the patterns in Figures 2 and 3 at all.

Again, this was a hard question, so don't worry too much if you didn't get it.

4. (30 points) This problem is designed to help you think about how the rate of return to education is determined. Consider a model where people choose their level of education in order to maximize the present value of their lifetime earnings, i.e., education is a purely an investment. Person i is

in school from time $t = 0$ until time $t = s_i$ (s_i is the level of schooling they achieve), and during this time they earn no income. Assume, however, that there are no tuition costs. After time s_i, a person enters the labor force and earns an amount $y(s_i, t)$ until $t = \infty$ (we abstract away from retirement), where $\ln y(s_i) = \alpha + \beta_1 s_i + \gamma(t)$. All people face the same earnings function $y(s_i, t)$, and they all discount the future at the interest rate r. In continuous time, the present value of a person's lifetime earnings is

$$PV_i = \int_{s_i}^{\infty} e^{-rt} y(s_i) dt$$

(a) (15 points) How much schooling does each person choose to get in this model as a function of β_1? Explain why the result makes sense in words.

Answer: This problem is a little tricky with a general function $\gamma(t)$; apologies for that, I didn't vet a last-minute edit to the problem well enough. However, the basic approach to solving this problem is straightforward. Take the derivative of PV_i with respect to s_i:

$$
\begin{aligned}
\frac{dPV_i}{ds} &= \int_{s_i}^{\infty} e^{-rt} \frac{\partial y(s,t)}{\partial s} dt - e^{-rs} y(s,s) \\
&= \int_{s_i}^{\infty} e^{-rt} \beta_1 e^{\alpha + \beta_1 s + \gamma(t)} dt - e^{-rs} y(s,s) \\
&= \frac{\beta_1}{-r + \gamma'(t)} e^{-rt} e^{\alpha + \beta_1 s + \gamma(t)} |_s^{\infty} - e^{-rs} y(s,s) \\
&= \frac{\beta_1}{r - \gamma'(s)} e^{-rs} y(s,s) - e^{-rs} y(s,s) \\
&= e^{-rs} y(s,s) [\frac{\beta_1}{r - \gamma'(s)} - 1]
\end{aligned}
$$

where $y(s,t)$ indicates earnings at time t with schooling s, and $y(s,t) = e^{\alpha + \beta_1 s + \gamma(t)}$. We do need to assume that the integrand converges at infinity (i.e., $\gamma(t)$ doesn't grow too fast)–otherwise PV_i can be infinite.

This expression has an intuitive interpretation, by the way: the first term is the marginal benefit from schooling, which is the increase in wages over the rest of life, and the second term is the marginal cost, which is the earnings you would get today by working instead of going to school. (Remember, the model is abstracting from intuition.)

6

OK, now we know that if this derivative is positive, then the person will want to increase schooling; if the derivative is negative, the person will want to decrease schooling; and if the derivative is 0, the person is happy with the level of schooling they're at. Note that $e^{-rs}y(s,s) > 0$, so we can ignore that part, and everything depends on the sign of $\frac{\beta_1}{r-\gamma'(s)} - 1$.

If $\gamma(t)$ is constant, as you were free to assume, then the derivative is 0 iff $\beta_1 = r$. Note that the person can't control β_1 or r, so this means that 1) if $\beta_1 > r$, the person wants infinite schooling; 2) if $\beta_1 < r$, the person wants no schooling, and 3) if $\beta_1 = r$, the person is indifferent across all levels of schooling. We get such a stark result because the return to schooling is linear–it's like an investment that we can make to any degree that we want. If the investment has higher returns than the interest rate ($\beta_1 > r$), we want to borrow (and invest) as much as possible. If it has lower returns than the interest rate, it's a raw deal and we should pass.

If you don't assume that $\gamma(t)$ is a constant, it's a bit hairier. If the function is linear, $\gamma(t) = \gamma t$, then we get the same kind of solution as above but with $r - \gamma$ in place of r. If $\gamma(t)$ is convex, then there will be a unique level of schooling that maximizes PV_i; at some point as earnings rise, the marginal cost of schooling induces the person to drop out. If $\gamma(t)$ is concave, then setting $\frac{dPV_i}{ds} = 0$ finds a minimum, and the person will always want either infinite or zero schooling.

(b) (5 points) Suppose that the return to schooling β_1 is a decreasing function of the average level of schooling in the population $\bar{s}$, $\beta_1(\bar{s})$. Draw a supply and demand diagram with $\bar{s}$ on the horizontal axis and β_1 on the vertical axis.

Answer: If we take the case where $\gamma(t) = \gamma t$, then the supply curve is just a horizontal curve at $\beta_1 = r - \gamma$. Supply is perfectly elastic: if the return to education ever rose above $r - \gamma$, people would stay in school forever, and if it ever fell lower, no one would go to school.

The demand curve is just $\beta_1(\bar{s})$; nothing more to say.

(c) (5 points) In this model, what is the equilibrium return to schooling $\beta_1(\bar{s})$? What is the equilibrium effect of an increase in demand on the return to schooling?

The equilibrium return is always $\beta_1 = r - \gamma$, or $\beta_1 = r$ is we make $\gamma(t)$ a constant function. That means that shifts in demand have NO effect on the return to schooling!

Is this model completely crazy? In the short run, yes; the supply of educated labor is probably closer to perfectly inelastic in the short run. In the long run, it is much more elastic. If we push the Human Capital model to its limit and see education as a pure investment that anyone can make–so who makes the investment doesn't affect the return, just like buying a stock–then we find that the return to

schooling will be perfectly elastic like this.

More generally, this exercise gives us ballpark number for what return to schooling to expect. Insofar as schooling is an investment, it should pay returns similar to those paid by other investments (like stocks, bonds, opening a business, etc.).

(d) (5 points) Compare your answer to (c) with the college/high school earnings differentials in Table 1 of "The Consequences of Increasing the Nation's Supply of College Graduates." Are the observed earnings differentials higher or lower than you would expect based on this model?

Answer: If we think of r as the real rate of interest on a safe asset (like government bonds), then the college/high school earnings differentials are surely too high, especially in recent years. (With $\gamma(t)$ a constant, the model predicts that the earnings differential will be $(1+r)^4 - 1 \approx 4r$. With $\gamma'(t) > 0$, the model predicts an even lower equilibrium return to education.) Of course, education is a risky asset, so perhaps we should think of r as the return to something like stocks–perhaps on the order of an 8% real return. This still seems lower than the differentials in Figure 1. Adding tuition costs to the model would raise the expected college/high school earnings gap too, but most college students attend public universities where tuition costs are low relative to forgone earnings. Accounting for risk and tuition costs might get us to the earnings differential seen from 1950-1980, but it is hard to see how a model like this will generate the very high differentials seen in 1990 and 2000. The question, then, is why aren't more people making (highly profitable) investments in education?