

Problem Set 4 Solutions

Economics of Education
11.126/11.249/14.48

Frank Levy
Spring 2007

Due in Section Friday, May 11, 2007

1. (35 points) Consider the following model of the labor market for teachers: In each location (e.g., a state or a school district), schools require a fixed number of teachers. That is, we're going to ignore any benefits from reduced class sizes. However, voters have preferences over teacher wages (which affect taxation levels) and average teacher quality (which affects student test scores), and this yields a demand function for teacher quality $Q_D(w, Z_D)$, where w is teachers' wage and Z_D is anything else that increases demand for quality. Similarly, college graduates decide whether to enter teaching or an alternative field, and the average quality of people in the teaching profession can be expressed in a quality supply function $Q_S(w, Z_S)$, where Z_S is anything other than the teacher wage that makes "high quality" workers want to become teachers.
 - (a) (6 points) Draw a graph that shows how average teacher quality and teachers' wages are determined in the long run. Do you expect the supply function Q_S to be more elastic or less elastic in the short run? Explain why.

Answer: Here is the email clarification I sent out:

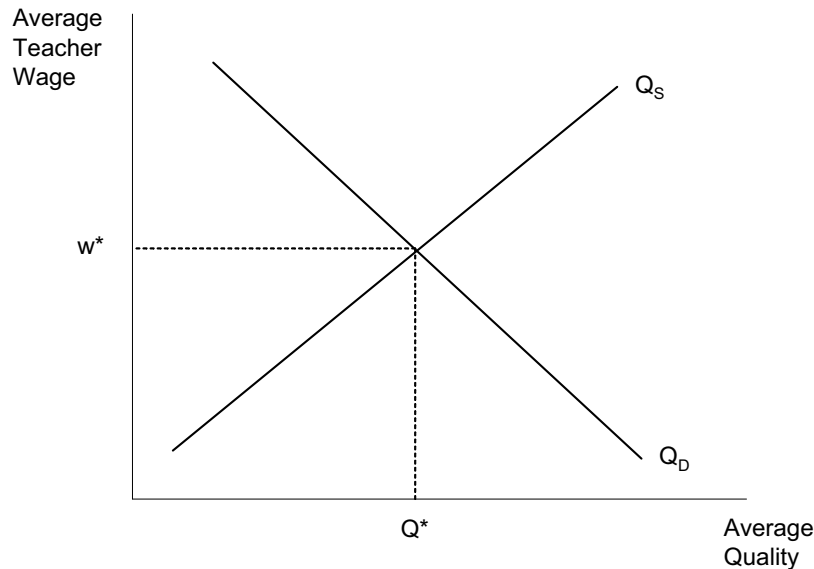
It was pointed out to me that this question is confusing, because Q_D is not a demand function in the normal sense. There are a few ways to make the problem clearer and simpler, but here is the most straightforward way to think about it: Imagine the market being at the level of the state. The average quality of people becoming teachers depends on the wage that teachers receive. (This gives the Q_S function.) Within the state, decision-making bodies (schools or school districts) have to trade off different kinds of expenditure. When teacher wages are higher, something has to go—taxes have to be raised, renovations have to be delayed, or the school has to hire cheaper (and presumably lower quality) teachers from the available pool. Think of w as the average teacher wage, but there is variation around it with better teachers typically getting higher pay. Q_D is a reflection of how the school's desired level of teacher quality responds when average teacher wages go up. If you've done this problem already and felt like it made sense, you don't need to go over it again. The interpretation of Q_D doesn't really affect parts (c) and (d), and it's not fundamental for (a) or (b).

I graded this part liberally. The basic idea is that teacher wages and teacher quality are both outcomes of the interaction between two factors: 1) The effect of wages on the willingness of different types

of people to become teachers, and 2) the preferences of the people hiring the teachers (i.e., their willingness to trade off teacher quality against other uses of funds).

We expect Q_S to be upward-sloping; low teacher wages draw disproportionately from the pool of college graduates with few alternative opportunities (and some people who are dedicated to teaching). Higher wages allow schools to pull more college graduates from the top of the class. It's not necessary that Q_D be downward-sloping, and this depends partly on how you interpreted the question. Within the context of the suggestion in my email, however, Q_D is very likely to slope downward. Higher wages for all teachers mean that schools have less money to work with. Something has to go: they can raise local taxes to pay the higher wages, they can cut non-teacher expenditures, or they can hire cheaper teachers from within the available pool. On the assumption that cheaper teachers are not as good on average, teacher quality will go down if they respond in part with the last option.

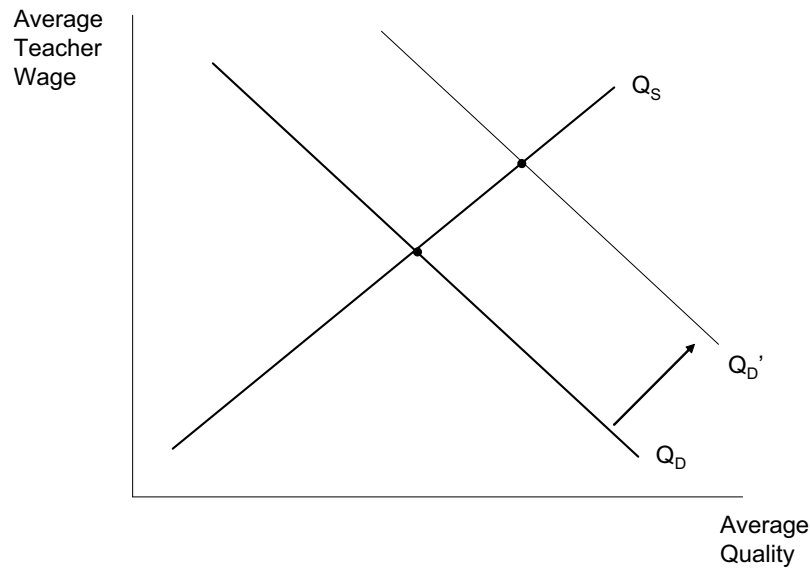
We expect Q_S to be less elastic in the short run for the standard reasons. The set of teachers working in a specific area is relatively fixed from one year to the next. With time, however, higher-quality college graduates can be pulled into teaching, and higher-quality teachers can be induced to move from other locations.



- (b) (7 points) When studies (like Loeb & Page) try to estimate the relationship between teacher quality and wages, are they trying to estimate Q_D , Q_S , or some combination of the two? In the context of

this bare-bones model, do they want to use changes in Z_D or changes in Z_S to estimate this function?

Answer: They are trying to estimate Q_S ; the question is whether schools could attract higher-quality teachers if they chose to pay higher wages. In order to estimate Q_S , they want to use changes in Z_D . Z_D shifts around Q_D , so that the points observed trace out Q_S . (See below.)

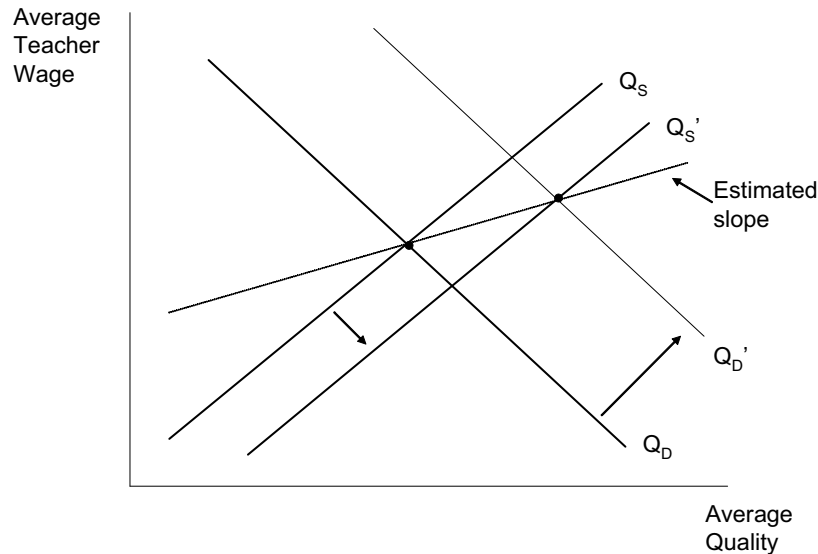


- (c) (12 points) Loeb & Page take several different approaches to the data, but their preferred approach (in Tables 4 and 5) is to regress a state's 10-year change in student outcome on the state's 10-year change in teacher wages, controlling for changes in the wages of female college graduates in the state. To what extent do you think they are picking up changes in Z_D ? To what extent do you think they are picking up changes in Z_S ? Insofar as they are picking up the "wrong" kind of variation (i.e., not what you specified in part (b)), in which direction will their estimates be biased? Explain your answers.

Answer: The control for alternative wages of female college graduates (the denominator of their "relative wage" variable) is meant to hold Z_S fixed. Some of the additional controls they add in Model III may be intended to fix Z_S as well. (However, many of these controls are probably best interpreted as factors that would affect dropout rates even if teacher quality stayed the same; this is the problem addressed in part (d).) At least in Models I and II, there are no explicit controls for Z_D . This means that observed changes in teacher wages are surely picking up changes in Z_D and may or may not be picking up

changes in Z_S . Insofar as alternative wages are the only important factor affecting the supply of quality, Loeb & Page's estimates will be correct. On the other hand, it is possible that states have experienced changes in the quality of their teacher training programs, in the social desirability of teaching as a profession, or in the overall skill level of their populations. Any of these factors might increase Z_S , even controlling for wage rates in alternative occupations.

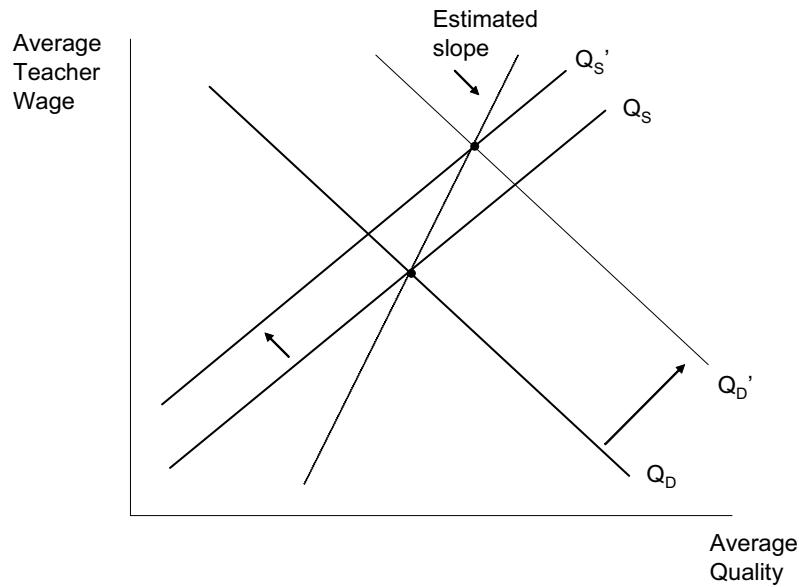
The direction of bias will depend on how shifts in Z_S are correlated with shifts in Z_D . If increases in Z_S tend to accompany increases in Z_D , we have the following picture, and the estimated effect of wages on teacher quality is too strong.



If decreases in Z_S tend to accompany increases in Z_D , then the estimated effect of wages is too weak:

- (d) (8 points) Loeb & Page do not measure teacher quality directly; instead they measure student outcomes (which are presumably affected by teacher quality but are also affected by other factors). Suppose they were able to use the ideal, pure source of variation you identified in part (b). Would you still have concerns that the estimates in Tables 4 and 5 might be biased? Explain why or why not.

Answer: You might. When Z_D is high, decision-makers are putting a high value on teacher quality. We might think that when parents are pushing harder for teacher quality, the average home environment is better. That is, parents may be teaching their kids at home, giving them better preparation for school, or generally supporting their studies more. These factors will tend to lower dropout and raise col-



lege enrollment even if teacher quality doesn't increase substantially.

2. (30 points) Write short (one or two paragraph) essays on each of the following:

(a) (15 points) Computerized scoring software like Criterion can be used in two different ways: it can be used pedagogically to give students feedback on their writing (when the teacher might have insufficient time to read every essay), or it can be used evaluatively to assign grades on students' coursework. What problems might arise when software is used evaluatively that would not be as severe when it is used pedagogically? Illustrate with examples from Criterion.

Answer: Broadly, the issue is that students have much stronger incentives to exploit or game the system when it is being used evaluatively. Scoring software will typically base scores on a list of *proxies* for good answers, because it is difficult to write a program that can recognize the same qualities of a good answer that a human grader would emphasize. As long as students' goal is to produce high quality work, these proxies may correspond very closely to the true characteristics of good answers. If students deliberately try to maximize the proxies, however, they will not necessarily be producing high-quality work.

There are several examples you could use to illustrate. For instance,

- Criterion has an easier time evaluating essay form than essay content. Although the software can detect whether the answer uses words from the question prompt, it is difficult to automate

evaluations of logical validity and appropriateness of examples. Even if students do not deliberately write "garbage" in a high-stakes setting, they might spend their time focusing on form and settle for mediocre content.

- Criterion has restrictive standards for form: it expects a standard five-paragraph essay. This could cause problems pedagogically insofar as the teacher wants students to learn about other writing styles, but it is likely to cause more problems in evaluation (when a good but non-standard essay is given a low score).
- At least in older versions, Criterion based scores partly on length. Length may be a reasonable predictor of quality, but adding unnecessary fluff to an essay surely decreases its quality. In high-stakes settings, students might respond to incentives by padding their writing.

(b) (15 points) The paper by Rouse, Krueger & Markman on evaluating computerized instruction suggests disappointing results. What are some drawbacks of computerized instruction methods relative to more traditional approaches? Do you think that these drawbacks will fade as computer technology becomes more advanced?

Answer: Answers will vary, but here are some possible points:

- Computers have a harder time training and testing open-ended questions that are more likely to lead to the "expert thinking" described earlier in the course.
- Humans can more easily notice particular problems in a student's learning and respond appropriately. Computers can be programmed to adapt, but usually only within a restricted number of ways.
- Computers may teach a narrower set of skills than human instructors. (This is a reasonable interpretation of the Rouse, Krueger & Markman results.) Computers are very efficient at presenting students with similar tasks again and again, but programming costs become high if they are used to present students with a wide variety of tasks.
- Computerized instruction involves less personal interaction, both between students and the teacher and among students. This might hamper students' development of social skills.
- Computers can be a distraction if students play games when the instructor isn't looking.

The extent to which these drawbacks will fade seems debatable. Faster hardware *per se* probably won't help, but more advanced software could alleviate some of these issues. For example, software might learn about a student's strengths and weaknesses based on inputs and use this information to tailor the learning program. Even a simple system where students progress through "levels" does this

to some extent, but improved software might be able to form more nuanced assessments of student skills. Other issues are not likely to fade at all with advances in computerization—indeed, the loss of personal interaction and the problem with distraction could become worse as computer environments become more immersive.

3. (35 points) The article by Gordon Winston talks about peer effects (among other things); this is the idea that how much you learn depends in part on who your fellow students are. Consider the following model: There are two types of students, with ability $a = 1$ and $a = 0$ respectively; each type is $\frac{1}{2}$ of the cohort of entering freshmen. There are two colleges, each with an equal number of spots available. When a person with ability a goes to a college where the average ability is \bar{a} , they leave with human capital

$$H = \alpha + \beta a + \gamma \bar{a} + \delta a \bar{a}$$

and this is also their lifetime productivity once they finish college.

- (a) (10 points) Suppose you are the education czar of the United States, and you get to command each college to admit a certain proportion of high ability students. Your goal is to maximize GDP. How do you want to divide the high ability students between the two colleges? (Hint: The answer will depend on some of the parameters $\alpha, \beta, \gamma, \delta$. The two colleges are equivalent, so you can just say that college 1 will be the "better" college with average ability $\bar{a}_1 \geq \frac{1}{2}$. Then college 2 must have average ability $\bar{a}_2 = 1 - \bar{a}_1$ because the total number of high ability students is fixed. Write down an expression for this cohort's average productivity, and maximize it with respect to \bar{a}_1 . If you don't need to do the math in order to come up with the answer, that's fine, but be sure to explain why your answer is correct well.)

Answer: The pedantic way of solving this problem is to note that there are four types of student: high ability students at college 1 get human capital H_1^1 ; high ability students at college 2 get H_2^1 ; low ability students at college 1 get H_1^0 ; and low ability students at college 2 get H_2^0 . Note that the proportion of the population that consists of high ability students at college 1 must be $\frac{1}{2}\bar{a}_1$, while the rest of the high ability people, $\frac{1}{2}(1 - \bar{a}_1)$, go to college 2. Similarly, a proportion $\frac{1}{2}(1 - \bar{a}_1)$ of the people are low ability at college 1 and a proportion $\frac{1}{2}\bar{a}_1$ of the people are low ability at college 2. So average

human capital for this cohort, as a function of \bar{a}_1 , is

$$\begin{aligned}
 \bar{H} &= \frac{1}{2}\bar{a}_1 \cdot H_1^1 + \frac{1}{2}(1 - \bar{a}_1) \cdot H_2^1 + \frac{1}{2}(1 - \bar{a}_1) \cdot H_1^0 + \frac{1}{2}\bar{a}_1 \cdot H_2^0 \\
 &= \frac{1}{2}\bar{a}_1 \cdot [\alpha + \beta + \gamma\bar{a}_1 + \delta\bar{a}_1] + \frac{1}{2}(1 - \bar{a}_1) \cdot [\alpha + \beta + \gamma(1 - \bar{a}_1) + \delta(1 - \bar{a}_1)] \\
 &\quad + \frac{1}{2}(1 - \bar{a}_1) \cdot [\alpha + \gamma\bar{a}_1] + \frac{1}{2}\bar{a}_1 \cdot [\alpha + \gamma(1 - \bar{a}_1)] \\
 &= \alpha + \frac{1}{2}\beta + \frac{1}{2}[\bar{a}_1^2 + 2\bar{a}_1(1 - \bar{a}_1) + (1 - \bar{a}_1)^2]\gamma + \frac{1}{2}[\bar{a}_1^2 + (1 - \bar{a}_1)^2]\delta \\
 &= \alpha + \frac{1}{2}\beta + \frac{1}{2}[\bar{a}_1^2 + (1 - \bar{a}_1)^2]\delta
 \end{aligned}$$

Taking the derivative with respect to \bar{a}_1 gives

$$\begin{aligned}
 \frac{d\bar{H}}{d\bar{a}_1} &= [\bar{a}_1 - (1 - \bar{a}_1)]\delta \\
 &= [2\bar{a}_1 - 1]\delta
 \end{aligned}$$

Remember that we can declare that school 1 is the "better" school, so that $\bar{a}_1 \geq \frac{1}{2}$. This means that if $\delta > 0$, then $\frac{d\bar{H}}{d\bar{a}_1}$ is always positive and it is optimal to set $\bar{a}_1 = 1$. If $\delta < 0$, then $\frac{d\bar{H}}{d\bar{a}_1}$ is always negative and it is optimal to set $\bar{a}_1 = \frac{1}{2}$. (A different way to look at this is to note that $\frac{d\bar{H}}{d\bar{a}_1} = 0$ when $\bar{a}_1 = \frac{1}{2}$. But this is a maximum when $\delta < 0$ and a minimum when $\delta > 0$. When $\bar{a}_1 = \frac{1}{2}$, either $\bar{a}_1 = 0$ or $\bar{a}_1 = 1$ are equivalent maximums, and we can just say that $\bar{a}_1 = 1$.) There is a knife-edge case when $\delta = 0$; in that case, GDP is the same no matter what we do.

The idea here is that what matters is the relative value of high-ability peers to high- and low-ability students. The absolute value of peer quality (γ) isn't important; all we care about is which group benefits more from a high quality peer environment. If high ability people do ($\delta > 0$), then we try to maximize the exposure of high ability students to other high ability students by grouping them all in the same college. We do the opposite when $\delta < 0$.

- (b) (10 points) You should have found two cases in part (a) depending on the values of the parameters. Try to come up with at least one reason why each of the cases might be true in practice. Which of the cases do you think is more likely to hold for college education? Would your answer change if we were talking about primary or secondary schooling instead?

Answer: The distinction between the two cases is not in how disruptive low-ability students are or in how beneficial high-ability students

are; that gets captured by γ . The issue is instead whether high- or low-ability students benefit *more* from being with high-ability peers. The answer might be that they do ($\delta > 0$) because curricula can be designed more appropriately when they are targeting a specific type of student. When classes contain a mix of students, teachers must design courses so that both types learn; separating students allows some teachers to specialize in teaching high-ability types and others to specialize in teaching low-ability types. The answer might be that they do not ($\delta < 0$) because it is very important for low-ability types to observe the behaviors and thought processes of high-ability peers. Because high-ability types already know what it takes to be high-ability, they do not benefit from this observation as much. (Other answers are possible.)

The U.S. educational system seems to stratify ability types more at the post-secondary level than at the primary and secondary levels. This suggests a belief that δ is larger (more positive) in a university setting than in a high school setting. Whether this is optimal seems debatable, but one justification might be that young, low-ability types' attitudes and thought processes are still flexible. They will benefit a lot from exposure to high-ability peers, but by the time they hit university age it is too late to mold them.

- (c) (10 points) Suppose that the labor market is perfectly competitive, so workers earn their productivity. Think about the maximum tuition each student would be willing to pay to get into a given college. How would this maximum willingness to pay vary with \bar{a} for a low ability student? How would it vary with \bar{a} for a high ability student?

Now imagine that we had a free and perfectly competitive market in education and that colleges just tried to maximize "profits" (which they can spend on architectural curiosities or elite sports teams). Each college can charge one price to high ability students and a different (higher) price to low ability students. Do you think a perfect free market will maximize GDP? (Hint: This turns on the observations about willingness to pay above.)

Answer: A low-ability student would pay some constant plus $\gamma\bar{a}$ to attend a college where average ability is \bar{a} . A high ability student would pay a constant plus $(\gamma + \delta)\bar{a}$.

There is more than one way to think about this problem, but here is the cleanest: Imagine each student engaging in two transactions with her college. First, she buys a spot at the college by paying tuition, and the amount of tuition will be increasing in \bar{a} . Secondly, if she is high ability then she sells her services as a peer to the university (i.e., receives scholarships or other aid), and the amount the university is willing to pay might depend on \bar{a} too.

Suppose there are two colleges, one with \bar{a}_1 and one with $\bar{a}_2 < \bar{a}_1$. If $\delta > 0$, then two things are true: high-ability students are willing

to outbid low-ability students for a spot at college 1, and college 1 is willing to pay more in financial aid to high-ability students than college 2 is. (College 1 is willing to pay more in financial aid because attracting a high-ability student adds more value to its institution.) Both of these forces mean that college 1 will pull high-ability students away from college 2, and this will continue until the colleges are completely specialized (with either high- or low-ability students, but not a mix).

If $\delta < 0$, then the opposite is true: low-ability students are willing to outbid high-ability students for a spot at college 1, and college 2 is willing to pay a higher financial aid premium to students with high ability. Both forces mean that high-ability students will go from college 1 to college 2, and this will continue until the colleges have the same levels of peer quality.

These outcomes are the same as what you should have found in part (a). Thus a perfectly-functioning free market will allocate students across colleges to maximize GDP (or to maximize economic efficiency more generally).

- (d) (5 points) Suggest one reason why the sorting of high and low ability students across colleges might be different in a real world higher education system than in the idealized market of part (c). Explain your answer.

Answer: Multiple answers are possible. One factor is that students may be unable to borrow an unlimited amount to go to college. If $\delta < 0$, then the perfect market equilibrium entails low-ability students paying very high tuitions to go to school with high-ability peers, while the high-ability students are subsidized. A low ability student who cannot borrow might instead pay lower tuition to go to school with low-ability peers, even if this is a less profitable investment in the long run.

Another factor is that colleges are typically not profit-maximizing institutions. This could play out in different ways, but you might think that colleges would be willing to sacrifice some profit in exchange for greater prestige (i.e., a high-ability student body). Of course, not every college can end up with the high-ability students, but colleges that can afford to lose money (because they have a large endowment) might end up with all high-ability students even if a more even sorting were optimal.