E C O N O M I C S B U L L E T I N

A theoretical model of the distribution of teacher attention under benchmark testing.

Justin Ross
School of Public Environmental Affairs - Indiana University, Bloomington

Abstract

This paper provides a simple theoretical model designed to capture the targeting incentives created by benchmark testing. Under high-stakes benchmark testing, schools and teachers are judged on the fraction of students that meet some given level of educational attainment. The incentive for teachers is then to allocate their resources towards students who are on the margin of the pass/fail level of educational attainment. This behavior has some empirical support and the aim of the model is to provide a formal means of developing hypotheses for future research.

Citation: Ross, Justin, (2008) "A theoretical model of the distribution of teacher attention under benchmark testing.." *Economics Bulletin,* Vol. 9, No. 29 pp. 1-8
Submitted: December 9, 2008. **Accepted:** December 16, 2008.
URL: http://economicsbulletin.vanderbilt.edu/2008/volume9/EB-08I20014A.pdf

1 Introduction

The debate over public school performance has largely focused on incorporating accountability, best observed in the *No Child Left Behind Act of 2001* (NCLB) with standardized testing. As part of the act, schools must demonstrate proficiency in select subject areas, and those who are deficient must prove to be making progress on an annual basis. The benchmark for these assessments are made based on the fraction of the school's students who pass some minimum level of proficiency on their respective annual standardized tests. Failure to have the minimum proportion of students attain the proficient level can result in withheld federal funds. The focus on passing a specific fraction of students provides incentives for schools and, by extension, their teachers to focus resources more heavily on the distribution of student achievement gains. In order to maximize the number of students that pass the proficiency test, the optimal strategy would focus on directing attention away from students whose outcomes are less dependent on the teacher's attention. High-aptitude students that can pass the test with little input from the teacher, and likewise the low-aptitude students who would not likely pass the test even with a great deal of attention from the teacher would be the best choices for teacher neglect under this incentive scheme created by these high-stakes proficiency testing. The purpose of this paper is to provide a baseline theoretical model to complement and motivate future empirical research on this behavior, as well as serve as a starting point for more complete models of the education system.

This targeting behavior under high-stakes testing has been gaining empirical support as well as national attention. TIME Magazine recently ran a cover story about the unintended detrimental consequences NCLB has had on the brightest students (Cloud, 2007). Jacob (2002) found evidence that teachers substituted away from untested or low-stakes subject matter to the high-stakes subjects their performance was judged on. Additionally, schools began retaining students earlier in their school career as a means of improving the share that pass the exam. At the school level, Jacob (2002) found an increase in special placements who typically are not included in the testable student body.

Wilson et al. (2004) conducted interviews with 22 English headmasters about the changes their school have made in response to high-stakes testing. Eight of these headteachers indicated they have strategies to target students whose grades are at the C/D borderline in order to raise their schools' respective ranking, while two others said they had attempted such strategies in the past but found them ineffective. Burgess et al. (2005) found that, with the implementation of high-stakes testing in the United Kingdom, educational gains for low-ability students declined when their schools experienced increases in the number of students who were near the pass/fail margin.

Using student-level panel data from the Texas Assessment of Academic Skills, Reback (2006) found evidence of schools redistributing resources to focus on those borderline pass/fail students. Reback used a two-stage probability model based on the probability of the student achieving a passing grade and the probability this student contributed to improving the school's rating. Reback found that students who were likely to contribute to a school's rating, were more likely to make higher than expected gains in their test performance in the following year. These students tended to be the lower-performing students, as the high ability students tended to perform according to expectations.

The following section will provide a preliminary theoretical model with simple assumptions to explain some of these existing empirical results.

2 The Model

The assumed institutional structure is the common end-of-the-year proficiency test. A school receives approximately nine-months with students, at the end of which a proficiency exam is given where each student who passes a predetermined benchmark score according to his or her grade level is considered to have "passed" the exam. The objective of the school then is to maximize the percentage of students who achieve a passing score. Since enrollment at public schools is taken as a given, this objective is equivalent to maximizing the number of students who pass the exam. The targeting behavior to be modeled is observed at the teacher level, who can discriminate their teaching efforts among the different students in their classroom. An exogenously given number of students with varying levels of aptitude are assigned to the teacher. Aptitude, a, is defined as a composite of the students' abilities, willingness to learn, parental support, prior knowledge, and any other factors that may create heterogeneity among the students to absorb the teacher's instruction and improve their ability to pass a proficiency test. It is assumed that teachers are able to perfectly observe their students' aptitudes and rank them in order from lowest to highest over the range $[p, q]$. As a result of this ranking, students can be uniquely identified according to their aptitude.

The teacher's tools for improving the educational attainment of each student is to provide them with instructional attention, which can be divided into lecturing to the class or providing individual attention to a student. The major difference is that lecture (L) is consumed as a pure-public good by the students while individual attention $(t(a))$ is purely private. Students are assumed to experience diminishing marginal returns to their education in both of these inputs. Brown and Saks (1975; 1980) used the same set of assumptions when modeling a teacher's choice, with a semantic difference in describing individual attention as private tutoring. Similarly, Guest (2001) and Epstein and Spiegel (1996) use a similar modeling of a university instructor's resources with lecture as a public good and office hours as a private good. Note that by defining individual attention as a pure private good, the possibility of peer effects is eliminated. The teacher is assumed not to value his or her lecture time any differently from time spent on individual attention, so that the budget constraint may be specified as

$$T = L + \int_p^q t(a)da. \qquad (1)$$

Where T is total time the teacher has for instruction.

The students' educational attainment can be defined as a function of their aptitude, the amount of time the teacher lectures, and the amount of individual attention they receive. As mentioned previously, students have diminishing returns to both inputs. A teacher will receive some positive and finite level of utility (j) for bringing a student to a passable level. The teacher is assumed to have perfect information about the students' educational attainment relative to the benchmark standard they need to have in order to pass the exam. The teacher will receive zero utility (V) from a student who does not pass the benchmark standard. Letting $e()$ be the education function of the students and $\bar{e}$ be the benchmark level, the optimization problem of the teacher is

$$Max\ V = \int_p^q V(e)f(a)da$$

where

$$V(e) \begin{cases} j & \text{if } e(L, t(a), a) \geq \bar{e} \\ 0 & \text{else} \end{cases}$$

where $\infty > j > 0$ and $f(a)$ is the distribution of the students according to aptitude level in the classroom.

For simplicity a uniform distribution is assumed so that $f(a)$ can be ignored. The teacher simultaneously determines the total amount of lecture and the distribution of individual time to each student. While the decision is simultaneous, it can be thought of occurring in two stages. In the first stage the teacher chooses how much time to spend lecturing, and in the second stage they choose how to allocate individual time with different students. To solve I begin with the second stage, where the distribution of individual attention is conditional on the amount of lecture performed in the first stage. Letting $\hat{e}(\cdot)$ represent the difference function between the benchmark level and the educational attainment with only lecture, it implicitly defines the optimal allocation of individual attention among the students of different aptitudes ($\hat{t}(a)$):

$$\bar{e} - e(L, 0, a) = \hat{e}(L, \hat{t}(a), a). \tag{2}$$

Since the teacher will only be able to increase utility by moving students from failing the benchmark standard to passing, equation (2) is only relevant to a subgroup of students. This subgroup will be defined according to aptitude ranges $[\underline{a}, \bar{a}]$, where $\bar{a}$ is the lowest aptitude student who can pass the benchmark with only the lecture input the teacher performed in the first stage and $\underline{a}$ being the lowest aptitude student to have received individual attention and pass the test. For additional clarity, Figure 1 illustrates this intuition for a given level of lecture. The process can be thought of as a teacher having already lectured for L amount of time, and begins providing individual attention to the student with the highest aptitude who fails the exam with only lecture (just below $\bar{a}$) until their educational attainment reaches $\bar{e}$. Once this is achieved, they move on to the next lowest aptitude student and provide them with individual attention until they can reach a passable level. This process repeats until they exhaust their total available time at student of aptitude $\underline{a}$. In this context, the problem of maximizing the number of students who can pass the benchmark level is equivalent to minimizing the position of $\underline{a}$ (i.e. moving $\underline{a}$ as close to p as possible).

As already stated, equation (2) defines $\hat{t}(a)$ between $\bar{a}$ and $\underline{a}$ while being equal to zero outside this range. It implicitly defines $\hat{t}$ as a function of lecture, aptitude, and the benchmark educational attainment of $\bar{e}$. Since we have lecture and total time given, both $\bar{a}$ and $\underline{a}$ are known and also become functions of lecture. In the general form, $\bar{a}$ can be implicitly defined from the following:

$$\bar{e} = e(L, 0, \bar{a}(L)) \tag{3}$$

By totally differentiating Equation (3) with respect to L, it can be seen that $\bar{a}'(L)$ will be strictly negative:

$$0 = e_L + e_{\bar{a}} \frac{d\bar{a}}{dL}$$

$$\frac{d\bar{a}}{dL} = -\frac{e_L}{e_{\bar{a}}} < 0$$

This demonstrates that the more the teacher lectures, the more students will pass with only lecture. This however does not necessarily mean that it will maximize the total number of students who pass the exam, as there are diminishing returns for all students to lecture and that individual attention time is sacrificed. Again, the best strategy for the teacher is to begin with the student of the highest aptitude that will fail the test with only lecture. After prescribing the minimal amount of individual attention necessary to get that student to $\bar{e}$, the teacher will then begin with the student of the next highest ability that is still failing the test and repeat the process until they run out of time. As a result of this process, the following equality must hold:

$$\int_{\underline{a}(L)}^{\bar{a}(L)} \hat{t}(L, a, \bar{e}) da = T - L \tag{4}$$

From this equality, $\underline{a}(L)$ can be solved for and differentiated. The position of $\underline{a}$, the lowest aptitude student to receive individual attention and pass benchmark $\bar{e}$, is expected to be a convex function of lecture. A teacher who lectures too little will begin providing individual attention at too high of level of $\bar{a}$, and a teacher who provides too much lecture will not have enough time left over to provide individual attention to reach the minimum possible $\underline{a}$ student. Using Leibniz Integral Rule, equation (4) can be differentiated with respect to L as:

$$\int_{\underline{a}(L)}^{\bar{a}(L)} \frac{\partial \hat{t}(L, a, \bar{e})}{\partial L} da + \hat{t}|_{\bar{a}} \frac{\partial \bar{a}}{\partial L} - \hat{t}|_{\underline{a}} \frac{\partial \underline{a}}{\partial L} = -1 \tag{5}$$

Notice now that $\hat{t}|_{\bar{a}} = 0$ because they are the lowest aptitude students to pass the test with only lecture, and hence do not receive individual attention. Substituting this into the above equation and solving for $\frac{\partial \underline{a}}{\partial L}$ arrives at:

$$\frac{\partial \underline{a}(L)}{\partial L} = \frac{\int_{\underline{a}(L)}^{\bar{a}(L)} \frac{\partial \hat{t}(L,a,\bar{e})}{\partial L} da + 1}{\hat{t}|_{\underline{a}}} \tag{6}$$

Setting this equation equal to zero yields the $\underline{a}$ minimizing level of lecture, which in turn is the utility maximizing condition for the teacher:

$$-\int_{\underline{a}(L)}^{\bar{a}(L)} \frac{\partial \hat{t}(L, a, \bar{e})}{\partial L} da = 1 \tag{7}$$

Since the budget constraint mandates an inverse relationship between $\hat{t}$ and L, the partial derivative $\partial \hat{t}/\partial L$ is negative and so is its integration over $[\underline{a}, \bar{a}]$. Graphically, the negative of this integral represents the diminished vertical distance in $\hat{t}$ from an increase in lecture and intuitively it represents the marginal benefit of increasing lecture. Increasing lecture increases the number of students between $[\bar{a}, q]$ and it reduces the amount of individual attention required for students between $[\underline{a}, \bar{a}]$ to reach the benchmark level. This is set against its marginal cost, which is just one unit of time.

It can be seen from equation (7) that increasing $\bar{e}$, the only policy variable in this model, will result in the aptitude range $[\underline{a}, \bar{a}]$ covering a range of higher aptitude students who offer higher returns to lecture. Holding L fixed then would leave the marginal benefit greater than marginal cost, suggesting that an increase in the benchmark level will increase the amount of lecture the teacher performs.

It is difficult to confirm at this point that the function $\underline{a}$ is convex using calculus. It is likely that at low levels of lecture there are still high returns and $\underline{a}$ will be a student of relatively high aptitude. By increasing the lecture, the position of $\underline{a}$ will move to the left at the expense of reduced individual attention. Conversely, lecturing too much will result in too few students receiving enough individual attention to achieve $\bar{e}$, and $\underline{a}$ will begin moving to the right.

A couple of interesting features emerge from studying this model and provide some testable hypotheses for empirical researchers. One is that structural break in educational attainment that occurs around the middle aptitude students. Another comes from observing that neither p nor q enter into the teacher's optimization problem at any point. This implies that adding another student to the classroom whose aptitude is greater than $\bar{a}$ or lower than $\underline{a}$ will not change the allocation of time, but only influence the resulting pass-fail rate.

Also interesting is considering what will occur if the benchmark level ($\bar{e}$) is raised. Holding, lecture constant, $\bar{a}$ would move to the right and the teachers would begin taking attention away from the students closer to $\underline{a}$ and redistribute it to the higher aptitude students. Unambiguously this lowers the pass rate for the teacher, but it has this interesting effect on the distribution of educational gains within the system, particularly it should mean greater educational gains for the higher-aptitude students.

3 Conclusion

The purpose of this paper has been to provide a theoretical model to support previous empirical findings and motivate future research on the distribution of educational attainment among students. The distinct feature of the model is the incentive of teachers to divert attention away from the highest and lowest aptitude students to those towards the middle near the benchmark level of educational attainment. The formalization of this behavior in the model allows for several empirically testable hypotheses in addition to this targeting behavior.

The model also seems to predict that if this targeting behavior is occurring, then adding a new student with very high- or low-aptitude will not have an impact on the distribution of students' educational attainment. If instead we add a mid-level student who is close enough to the pass/fail margin, then we will see a change in those achievement distributions. If researchers obtain data that contains observations of students who transfer to different schools, they may be able to study this effect.

This model serves as a starting point as it has many simplified assumptions. Attempting a different production function, providing alternative non-teaching uses of time, and incorporating peer effects are all possible extensions for the model. This would also be useful in developing a holistic model of the education system, as it could serve as a counterpart to a model that formalizes the incentives of administrators determining class composition and size, as well as parents choosing school districts with students of high or low aptitude dominating the student body.

References

Brown, Byron W. and Daniel H. Saks, "The Production and Distribution of Cognitive Skills within Schools," *Journal of Political Economy*, 1975, *83* (3), 571–93.

_ **and** _ , "Production Technologies and Resource Allocations within Classrooms and Schools: Theory and Measurement," in R. Dreeben and J.A. Thomas, eds., *The Analysis of Educational Productivity*, Cambridge, MA: Ballinger, 1980, pp. 53–117.

Burgess, S, C Propper, H Slater, and D Wilson, "Who Wins and Who Loses from School Accountability? The Distribution of Educational Gain in English Secondary Schools," Working Paper 05/128, University of Bristol 2005.

Cloud, John, "Are We Failing Our Geniuses?," *TIME Magazine*, August 27 2007, *170* (9).

Epstein, Gil S. and Uriel Spiegel, "A Lecturer's Optimal Time Allocation Policy," *Education Economics*, 1996.

Guest, Ross, "The Instructor's Optimal Mix of Teaching Methods," *Education Economics*, 2001, *9* (3), 313–326.

Jacob, Brian A., "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools," Working Paper 8968, National Bureau of Economic Research 2002.

Reback, Randall, "Teaching to the Rating: School Accountability and the Distribution of Student Achievement," SSRN Working Paper 0602, Barnard College, Department of Economics May 2006.

Wilson, D., B. Croxson, and A. Atkinson, "What Gets Measured Gets Done: Headteachers' Responses to the English Secondary School Performance Management System," Technical Report 04/107, University of Bristol 2004.

Figure 1: Theoretical Allocation of Individual Attention with Benchmark Testing.