Economics Letters 149 (2016) 75–78

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

A theory of dynamic investment in education in response to accountability pressure

Tom Ahn

Department of Economics, University of Kentucky, 225D BE, Lexington, KY 40506, United States

HIGHLIGHTS

- I propose a dynamic investment model of schools in an environment with accountability.
- Scores continually degrade as costly investments are made to coincide with sanctions.
- Simulations show that blindly setting thresholds or rewards leads to lower scores.
- RD analysis with data from NC, which had a merit-pay system, corroborates the model.

ARTICLE INFO

Article history:
Received 19 July 2016
Received in revised form
24 September 2016
Accepted 9 October 2016
Available online 24 October 2016

JEL classification:
I2

Keywords:
Education
Accountability
Dynamic investment

ABSTRACT

While well-implemented accountability systems are effective in inducing sharp test scores gains after intervention, it remains a mystery why such schools with the technical capacity to improve would allow productivity to decline to the point of sanction in the first place. We present a theory of dynamic investment where schools look forward and rationally choose the timing of reforms to increase achievement at the point of sanctions. Theory shows that policy makers must select the strength of sanctions carefully to maximize education production. Regression discontinuity analysis of a merit-pay system in North Carolina corroborates the theory.

1. Introduction

The economics of education literature has established that accountability systems can elicit modest improvements in student academic outcomes (Clotfelter et al., 2004; Chiang, 2009). However, beyond the estimated treatment effect, not much is understood about how schools evolve to require accountability intervention and what schools do to increase academic achievement once treated.

Academic achievement, which often determines a school's accountability status, does not dramatically change from year to year. Many schools experience gradual performance decreases until they are sanctioned. After sanctions are applied, performance increases sharply and then slowly declines again. This pattern is difficult to reconcile with a model of rational agents (schools) that should correct their behavior and increase test scores prior to failing. Empirical work (in particular, in the quasi-experimental treatment effects literature) implies that schools may be myopic (and are shocked into action by sanctions).[1] Given that schools seem to have the expertise and sophistication to respond in meaningful ways to improve academic outcomes when pressured, it remains an open question why schools are unresponsive (or unaware) until actual sanctions are levied.

We present a theory of targeting, which replicates the observed saw-tooth pattern of gradual declines in academic achievement punctuated by sharp increases at the point of sanction, while maintaining the assumption that schools look forward in time and make rational decisions. Schools allow performance to decline and eventually fail because targeting is costly and must be undertaken as infrequently as possible to maximize its twin objectives:

[1] As most accountability results are reported online, it seems unlikely that principals and teachers would be caught off-guard by failure. In addition, that we see similar responses in schools that have had previous accountability failures is surprising. In fact, previous research has shown that experienced principals have the savvy and ability to respond to pressures by changing the school environment and altering recruiting strategies (Ahn and Vigdor, 2016a; Ahn, forthcoming).

E-mail address: thomas.ahn@uky.edu.

http://dx.doi.org/10.1016/j.econlet.2016.10.013
0165-1765/

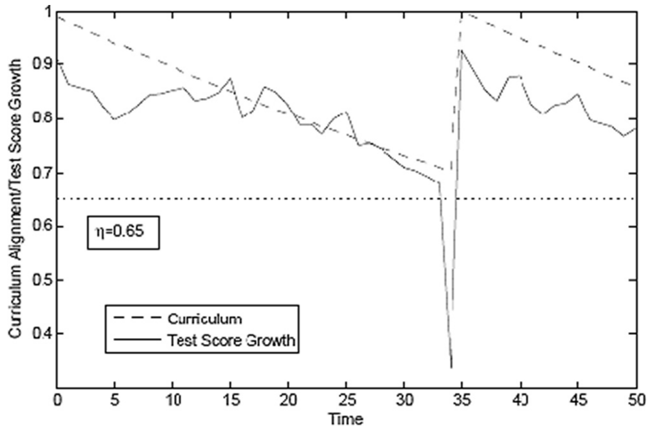

Fig. 1. $\delta = 0.99, \beta = 0.70, \mu = 1, \rho = 0.95, \sigma = 0.1, \alpha = 0.75, \eta = 0.65, \gamma = 0.35, \lambda = 0.50, \overline{W} = 1$. Autoregressive transition probability matrix estimated at 50 points. First 100 periods discarded for initial conditions reasons.

increasing achievement and avoiding sanctions. The point of failure serves as the most opportune time to make investments in reforms. The model makes predictions about the pattern of changes in test scores close to the accountability threshold and the targeting behavior of schools.

We then show using North Carolina data that test score growth in response to a merit-pay system and variables that proxy for targeting both exhibit the saw-tooth pattern predicted by the model. There is a positive jump in test scores and the proxy variable at the point of sanction. Trends both to the left and to the right of the discontinuity yield negative estimates.

2. Theory model

We present a simple model of targeting, where effort costs associated with the adjustment process means that the school's academic productivity is allowed to decay until sanctions are levied (for example, a bonus payment is lost), after which the school to invests in reforms. A school's education production function is:

$$f(\varepsilon, W) = \varepsilon W^{\alpha}$$

where $f(\cdot)$ is growth of test score, and W represents academic productivity. For example, W could be a measure of how well the school's curriculum is aligned to the state exam. The term ε is a productivity shock (in student ability). We assume $0 < \alpha < 1$ to ensure score growth is concave in input. We restrict $\underline{W} \leq W \leq \overline{W}$, with $W = \overline{W}$ implying that curriculum is perfectly tuned to the exam, and $W = \underline{W}$ indicating that curriculum has no relationship to the exam. Over time, the curriculum becomes misaligned:

$$W' = \delta W$$

where $\delta < 1$, and prime($'$) indicates next period. Misalignment can be due to many factors, such as changes to the exam, teacher transfers resulting in loss of pedagogic know-how, declining lesson preparation by teachers, etc. To account for retention of cohorts of students for more than one year, ε is modeled as an AR(1) process. Productivity follows an AR(1) process: $\varepsilon' = \mu + \rho\varepsilon + \nu$ with $0 < \rho < 1$ and $\nu \sim N(0, \sigma_{\nu}^{2})$. Given ability ε and productivity W, a school invests resources (teacher/administrative efforts) to maximize achievement.

The school, operating in a merit-pay system, attempts to maximize achievement and qualify for the bonus:

$$u(\varepsilon, W, \gamma, \eta) = f(\varepsilon, W) + \gamma I[f(\varepsilon, W) > \eta]$$

where $I[\cdot]$ is an indicator that equals one when $f(\varepsilon, W) > \eta$. That is, in addition to education output, if scores growth is greater than a defined cutoff η, the school receives additional utility γ (bonus amount, normalized to how the school values achievement).

A school has the choice to align its curriculum to perfectly follow the exam each period.[2] If the school invests, it must devote resources away from education production this year. We model this by discounting production by $\lambda < 1$. Therefore, the current period utility of a school that invests is:

$$u(\varepsilon, W, \gamma, \eta, \lambda) = \lambda f(\varepsilon, W) + \gamma I\left[f(\varepsilon, W) > \frac{\eta}{\lambda}\right].$$

Note not only is production discounted, the probability of bonus receipt also declines.[3]

The only reason to invest in the current period is for higher growth in subsequent periods.[4] As schools are now forward-looking agents, we use a value function representation. A school in each period has the option of maintaining the status-quo or investing. The value of maintaining the status-quo is:

$$V^{s} = \varepsilon W^{\alpha} + \gamma I[\varepsilon W^{\alpha} > \eta] + \beta E_{\varepsilon'|\varepsilon} V(\varepsilon', \delta W)$$

and the value of re-aligning is:

$$V^{*} = \lambda\varepsilon W^{\alpha} + \gamma I\left[\varepsilon W^{\alpha} > \frac{\eta}{\lambda}\right] + \beta E_{\varepsilon'|\varepsilon} V(\varepsilon', \overline{W})$$

where β is an inter-temporal discount rate. The trade-off in investment is the current reduction in output (and reduced likelihood of bonus receipt) against increased output (and increased likelihood of receiving the bonus) in subsequent periods. The value function is:

$$V(\varepsilon, W) = \max\left\{V^{s}, V^{*}\right\} \quad \forall(\varepsilon, W)$$

The model is solved numerically via value function iteration.

Fig. 1 presents the Monte Carlo results of the model. There are two characteristics worthy of note. First, after a school realigns its curriculum, scores sharply increase and then gradually decline the further away a school is from the bonus threshold, as it chooses to let the alignment decrease instead of making continual costly adjustments. Second, corresponding measures of alignment mirrors changes in growth at the threshold.

The model vividly shows that the policy planner should be careful in setting standards. Fig. 2 presents policy simulation results as the threshold value that determines whether schools qualified under the accountability system is adjusted (η).[5] Too high a threshold forces schools to adjust constantly, which predictably results in score declines and an increase in the variance of student outcomes year-to-year. Surprisingly, too low a threshold also induces frequent readjustment, leading to lower achievement. This behavior arises because with the threshold so low, the school does not have to "pay the cost" of forfeiting the bonus during the adjustment year.

Fig. 3 demonstrates that policy makers also need to be careful in setting the amount of reward (γ). Too low a reward results in low test scores, as schools experience no pressure from accountability. Achievement is allowed to lapse to low values before re-targeting the curriculum, resulting in large variance in student outcomes between cohorts. However, too generous a reward also results in declining achievement as schools "chase the bonus", choosing to adjust more often than is optimal for educational output. In the next section, we describe the data and the policy environment used to corroborate the theory model.

2 Qualitative results remain unchanged if a school can choose degree of alignment (at increasing cost for more drastic changes).

3 Investment in re-alignment is modeled as becoming effective the next period. Qualitative results remain unchanged if investment becomes effective immediately.

4 Alignment adjustment is *costless* in the model (besides the loss in education achievement due to allocation of effort). Introducing adjustment cost would induce schools to delay realignment.

5 We run the model 1000 times (with new error draws) and report the average value of mean and variance in achievement and mean alignment across 100 periods.

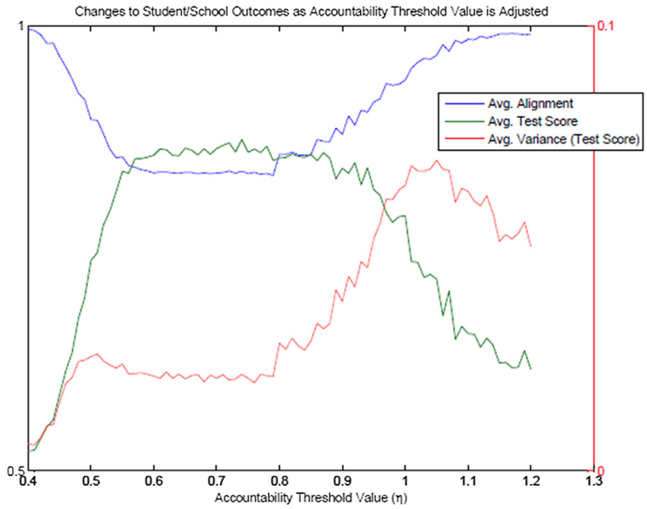

Fig. 2. Policy simulation results for changes to threshold value (η). Averaged across 100 periods.

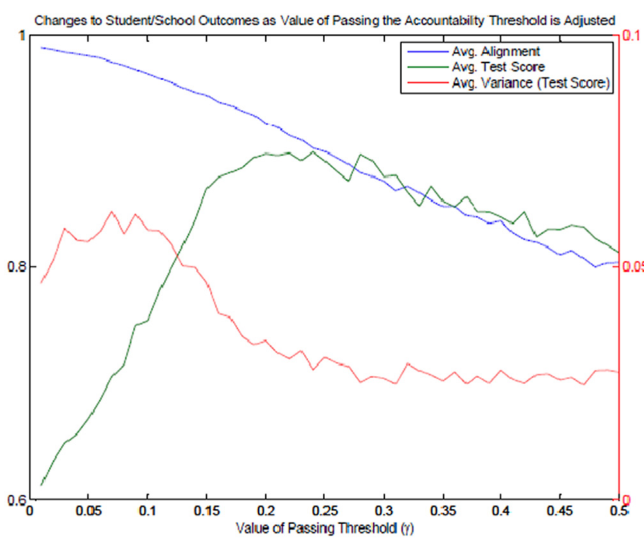

Fig. 3. Policy simulation results for changes to value of passing threshold (γ). Averaged across 100 periods.

Table 1
Parametric RD estimates of the impact of failing.

	Math score growth		Change in Avg. Yrs. experience	
Treatment point	0.0386*** (0.0137)	0.0409** (0.0203)	0.4882* (0.2821)	1.0838** (0.4433)
Trend to the left	−0.4462*** (0.0791)	−0.5600** (0.0233)	−5.5679* (3.0523)	−10.1285** (5.1463)
(Trend to the left)^2		−0.3960 (0.9518)		−34.2102* (18.5924)
Trend to the right	−0.2447* (0.0133)	−0.6577* (0.3526)	−2.6521* (1.6072)	−17.3678** (8.4497)
(Trend to the right)^2		2.4664 (1.5497)		77.9736* (40.1130)

Note: Standard errors in parentheses. Covariates include % poverty, % female, % minority.
* Denotes significance at 10% level.
** Denotes significance at 5% level.
*** Denotes significance at 1% level.

3. Empirical corroboration

We use administrative data on public schools from North Carolina and the "composite growth index" data from the Department of Public Instruction. The normalized index ranges from −0.45 to 0.66. Summary statistics from our sample, which consists of students and teachers in grades 3–5 in 2006–2007, are presented in the online appendix.[6]

During our sample year, North Carolina had in place a merit-pay system which paid $750 to all teachers at schools that achieved 'expected' performance, by scoring above zero on the normalized index. Schools were assessed by calculating year-over-year growth in math and reading standardized test scores.[7] Approximately 65% of schools qualified for the bonus. The bonus system was administered simultaneously with the federal No Child Left Behind (NCLB) system. As NCLB sanctions were not based on test score growth, there is low correlation between the bonus and NCLB outcomes.[8]

To corroborate the saw-tooth pattern in the theory model, we estimate a parametric regression discontinuity (RD) regression (Lee and Lemieux, 2010).[9] Table 1 and Fig. 4 present empirical results for linear and quadratic functions of the assignment variable. Trends are allowed to differ on either side of the discontinuity. RD results for the "treatment effect" of the bonus is shown for math test score growth and average years of teacher experience at the school.[10] Test score growth results shows the sharp increase in achievement at the sanction point. The corresponding sharp changes in teacher experience level show the evolution in curriculum alignment. Both regressions also have slopes to the left and right of the threshold that are negative, showing that both test scores and curriculum alignments naturally degrade through time.

4. Discussion

A growing body of literature has shown the positive impact judicious use of accountability systems could have on education production. However, accountability is not a panacea, instantly raising test scores by introducing "market pressure". Our understanding of how successful accountability system work may be less than complete. In particular, the efficacy of the "sanction treatment" in the quasi-experimental literature may imply that schools cannot foresee future negative outcomes to avoid it, yet have the technical capacity to achieve extra-normal gains when pressured.

Our model attempted to re-introduce schools as rational and forward-looking agents that respond to the accountability pressure optimally. The saw-tooth patterns in achievement and measures of targeting predicted in the theory model is corroborated by

6 See sites.google.com/site/tomsyahn/. The sample contains ∼350,000 students and 87,000 teachers. While student data contains grades 3–5 (as only tested students are in the data), teachers may be assigned to kindergarten through grade 5.

7 There was also $1500 payments for 'exceptional' performance. About 1/6 of schools qualified for the $1500 bonus. See Ahn and Vigdor (2016b) for a complete description.

8 Table 2 in the online appendix shows a cross-tabulation of NCLB and bonus system status. More than 30% of the sample fails under one accountability regime yet passes under the other.

9 As a robustness check, a non-parametric RD analysis was also done. See online appendix. Additionally, all standard robustness checks for RD is presented in the online appendix.

10 Average years of teacher experience declines because experienced teachers that transfer/retire are often replaced with newly-minted teachers. The assumption here is that it takes effort to recruit effective teachers with experience. See Ahn (forthcoming).

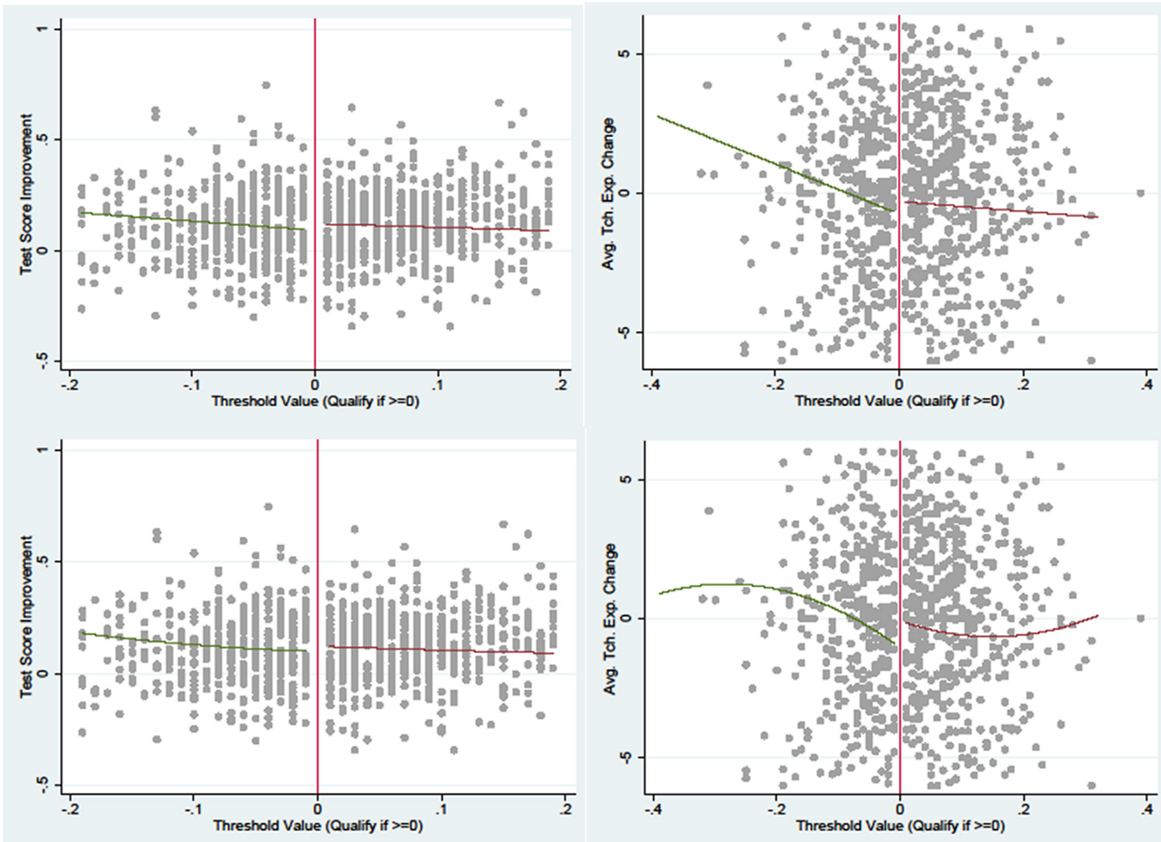

Fig. 4. RD graph of impact of bonus loss on math score growth and change in average years of teacher experience. Top two graphs show linear model. Bottom two graphs show quadratic model.

RD analysis using data and a bonus system from North Carolina. The theory also shows that simplistic attempts to induce higher achievement by increasing market pressure (via increasing bonus payouts) or raising the threshold value may actually result in lower average achievement and cause greater harm to students.

Our incomplete understanding of what induces schools to change behavior in the presence of accountability system means naïve implementation of such programs could have unintended consequences, and more nuanced research into what motivates schools and how they respond to market pressure is required.

References

Ahn, T., 2016. Strategic matching of schools and teachers with (and without) accountability pressure. Educ. Financ. Policy (forthcoming).

Ahn, T., Vigdor, J.L., 2016a. Opening the Black Box: Behavioral Responses of Teachers and Principals to Pay-for-Performance Incentive Programs Working Paper.

Ahn, T., Vigdor, J.L., 2016b. When Incentives Matter Too Much: Explaining Significant Responses to Irrelevant Information. NBER Working Paper #20321.

Chiang, H., 2009. How accountability pressure on failing schools affects student achievement. J. Public Econ. 93 (9), 1045–1057.

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., Aliaga, R., 2004. Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers? J. Policy Anal. Manag. 23 (2), 251–271.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econom. Lit. 48 (2), 281–355.