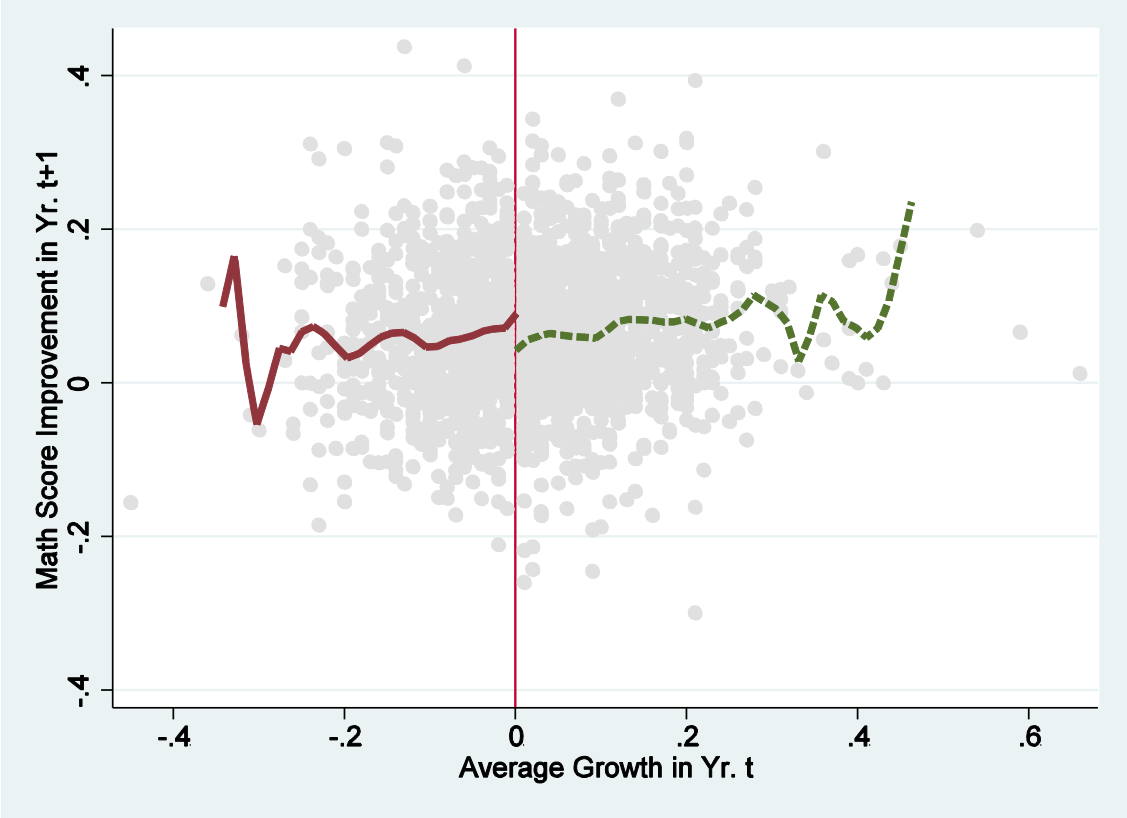
How Salient are Performance Incentives in Education? Evidence from North Carolina

Tom Ahn & Jacob Vigdor

10-3-2012

Big Picture



- Do teachers/schools respond to cash incentives?
- How do they respond? (Not as simple an answer as one would think.)

Preview of Results

- Teachers respond to loss of cash bonus this year by increasing test score next year.
- Teachers may be responding to the incentives in a fashion that is difficult to reconcile with a simple rational utility model.
- RD effects are particularly prominent among schools with a bad track record of:
  - failing to consistently qualify for bonuses.
  - failing to make adequate yearly progress, a metric that is UNRELATED to the bonus incentives.
- Implies a need to look beyond simple expected utility maximizing model, perhaps at behavioral models.

Data

North Carolina Education Data Set (NCEDS)

- NC Education data set of all public school students and teachers.
- Data on elementary school students in grades 3 - 5 from 2005/06 - 2006/07 used.
- Approximately 570,000 observations.

ABC Incentive system

- NC students in Grs. 3 - 8 take End-of-Grade (EOG) exams in reading and mathematics.
- A school's avg. score for year t is compared to the same students' scores from year $t-1$.
- Teachers are paid a cash bonus based on school-level growth in test scores.
- Exists simultaneously with No Child Left Behind.

Summary Stats

Table : Summary Statistics

Variable	Mean (Std. Dev.)
Δ math score	0.0617 (0.4555)
Δ reading score	-0.0348 (0.6362)
math proficiency level	2.8408 (0.8439)
reading proficiency level	3.2976 (0.7838)
Δ math proficiency level	0.0454 (0.6251)
Δ reading proficiency level	-0.0327 (0.7516)
% minority	0.3959 (0.4891)
% poverty	0.4582 (0.4982)
Years since last bonus	0.6524 (0.5001)
Number of no bonus years in last 5 years	1.2267 (1.2530)
Years since AYP made	0.5558 (0.9146)
Number of AYP failed since 2002-03	1.0547 (1.0776)
Observations	569,808

Checks 1

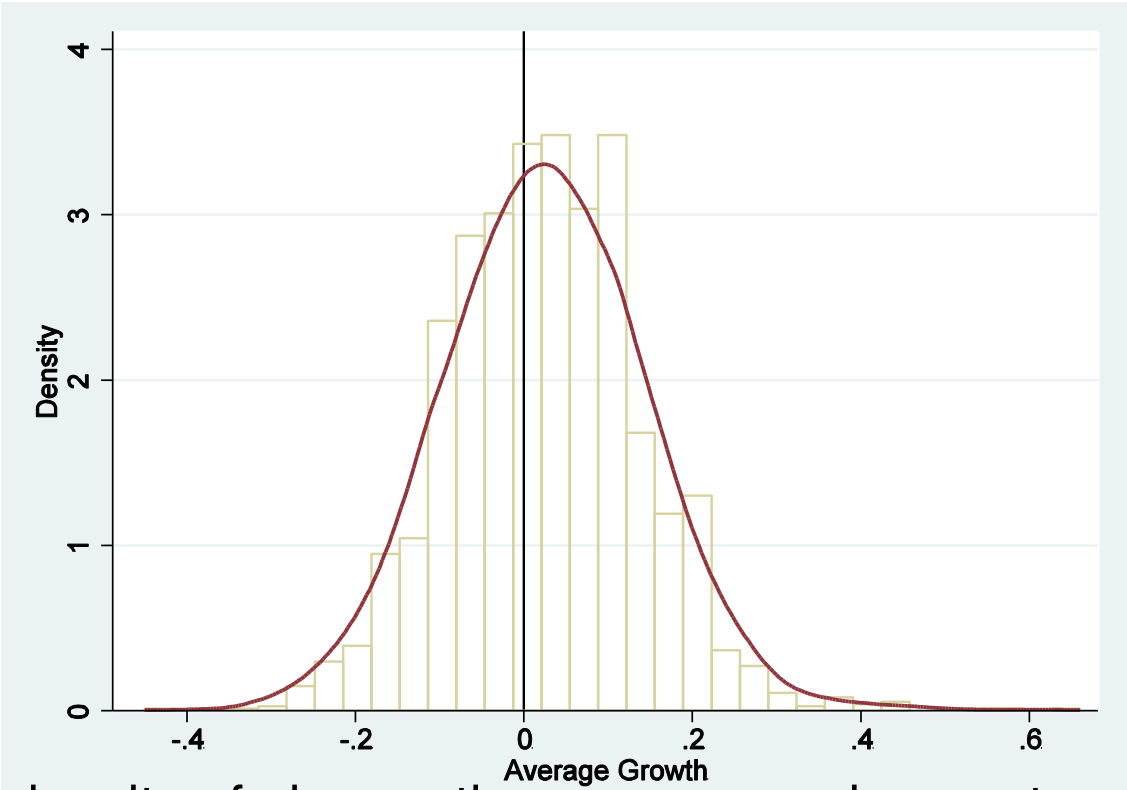


Figure : density of observations across assignment variable

Checks 2

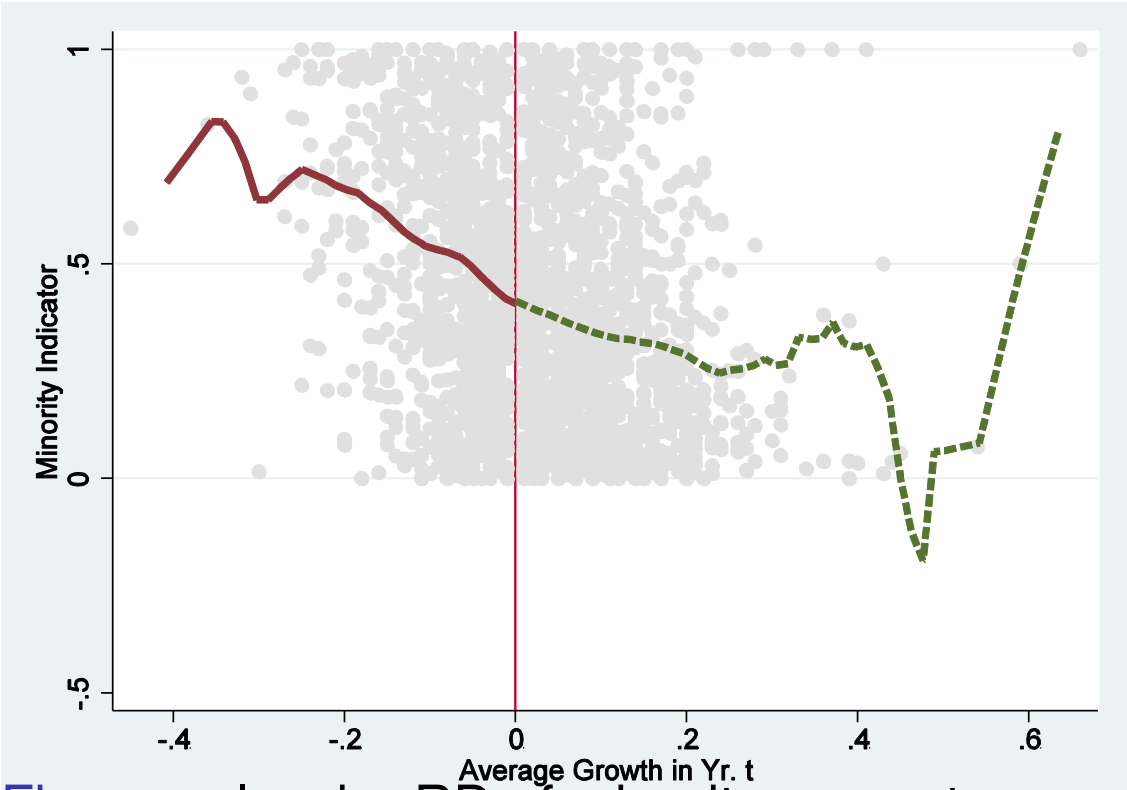


Figure : placebo RD of minority percentage

Checks 3

Figure : placebo RD of female percentage

Checks 4

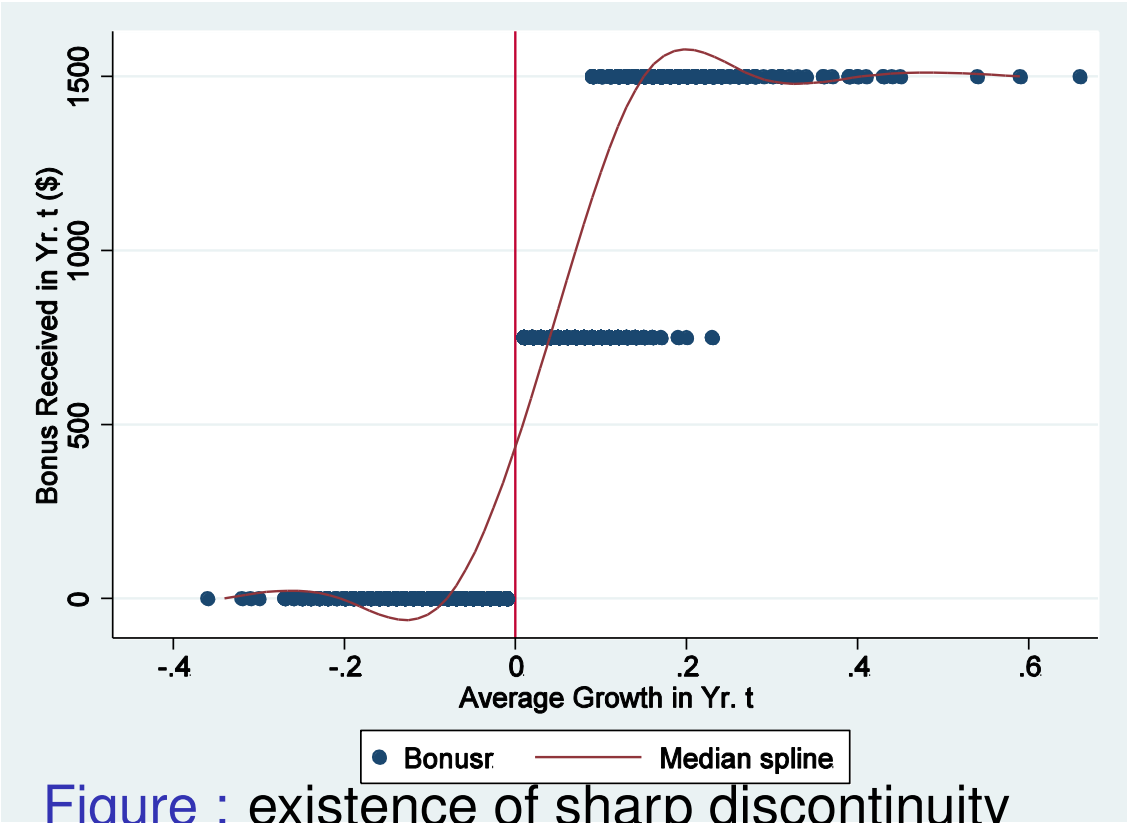


Figure : existence of sharp discontinuity

Results 1

- Just failing to qualify for the bonus spurs extra-normal gains in the next period.
- Just qualifying for the bonus results in a slight dip in scores the next period.
- Rational utility model should imply that there should be NO break at the discontinuity.

Table : Regression Discontinuity Results for Bonus Receipt: Entire Sample

Outcome Measure	RD Effect (Std. Err.)	Bandwidth
Δ math score	-0.0188 (0.0032)***	0.1195
	-0.0200 (0.0049)***	0.0597
	-0.0175 (0.0024)***	0.2390
Δ reading score	-0.0114 (0.0064)*	0.0829
	-0.0325 (0.0104)***	0.0415
	-0.0050 (0.0045)	0.1659

Are Teachers Irrational??

- No.
- Asymmetric responses for just failing vs. just passing. Why?
- Implies a ‘finish-line’ effect, in which schools that discover that they are just short of the bonus make strong efforts to qualify, while those that succeed slack off in the next year.
- Behavior seems consistent with confusion about how well or how poorly schools/teachers are doing.

Results 2

- Schools have a track record of success develop a sense of complacency.
- Confusing/multiple accountability systems may induce the 'wrong' schools.

Table : RD Results for Bonus Receipt: Math Score Only, By Accountability History

Accountability History	RD Effect (Std. Err.)	Bandwidth
No bonus more than 2 out of last 5 years	-0.0813(0.0107)***	0.0735
	-0.0495 (0.0168)***	0.0367
	-0.0673 (0.0075)***	0.1469
Bonus in 3 or more of the last 5 years	-0.0093 (0.0053)*	0.0593
	-0.0252 (0.0106)	0.0296
	-0.0092 (0.0035)***	0.1186
Failed to make AYP for the last 2 or more years running	-0.1282 (0.0186)***	0.0374
	-0.1036 (0.0116)***	0.0187
	-0.1037 (0.0107)***	0.0748
Made AYP every year since 2003	-0.0033 (0.0056)	0.0540
	-0.0146 (0.0106)	0.0270
	-0.0074 (0.0036)**	0.1081

Results 3

- Teachers can improve the performance of traditionally disadvantaged students when they feel ABC pressure...
- ... but why?

Table : RD Results for Bonus Receipt: Math Score Only, By Demographic Subsamples

Subsample	RD Effect (Std. Err.)	Bandwidth
Minority	-0.0783 (0.0123)***	0.0612
	-0.0821 (0.0223)***	0.0306
	-0.0612 (0.0081)***	0.1224
Non-minority	-0.0019 (0.0073)	0.0977
	0.0038 (0.0120)	0.0489
	-0.0156 (0.0054)***	0.1955
Poverty	-0.0476 (0.0124)***	0.0571
	-0.0787 (0.0244)***	0.0285
	-0.0352 (0.0081)***	0.1151
Non-poverty	-0.0271 (0.0062)***	0.1419
	-0.0169 (0.0091)*	0.0710
	-0.0275 (0.0049)***	0.2839

Results 4

- Teachers are successful in raising the test scores of students close to the cut off for AYP when they feel ABC pressure...
- ... but why?

Table : RD Results for Bonus Receipt: Math Score Only, By Proficiency Level

Level	RD Effect (Std. Err.)	Bandwidth
I	-0.0052(0.0228)	0.0979
	-0.0020 (0.0372)	0.0490
	-0.0037 (0.0168)	0.1958
II	-0.0338 (0.0205)*	0.0469
	-0.1001 (0.0342)***	0.0234
	-0.0389 (0.0127)***	0.0937
III	-0.0405 (0.0064)***	0.1523
	-0.0467 (0.0092)***	0.0762
	-0.0317 (0.0051)***	0.3047
IV	-0.0150 (0.0116)	0.0996
	-0.0025 (0.0192)	0.0498
	-0.0216 (0.0085)**	0.1992

Conclusion

What did we learn?

- Some optimism for efficacy of accountability systems:
  - Teachers are capable of extra-normal exertion to improve student performance and they can be induced to do so.
  - Teachers can improve the performance of traditionally disadvantaged students.
- Some pessimism (or at least caveats) for efficacy of accountability systems:
  - Schools that are 'close' to the finish line are capable of exerting extra-normal effort to push themselves across (next year), but complacency sets in immediately afterward.
  - Schools have a track record of success develop a sense of complacency.
  - Confusing/multiple accountability systems may induce the 'wrong' teachers to focus efforts on the 'wrong' students.

Recommendations:

- Make accountability system easier to understand and simpler to evaluate teachers/school performance and,
- Make the standards tougher to attain so that schools find it difficult to have a consistent track record of bonus receipt.