



# Can failure succeed? Using racial subgroup rules to analyze the effect of school accountability failure on student performance

David P. Sims\*

*Economics Department, Brigham Young University, 130 FOB, Provo, UT 84602, United States*

## ARTICLE INFO

### Article history:

Received 20 April 2012

Received in revised form 20 November 2012

Accepted 5 December 2012

### JEL classification:

I28

### Keywords:

No Child Left Behind

Stigma

School accountability

Student achievement

## ABSTRACT

Many school accountability programs are built on the premise that the sanctions attached to failure will produce higher future student achievement. Furthermore, such programs often include subgroup achievement rules that attempt to hold schools accountable for the performance of all demographic classes of students. This paper looks at two issues: the degree to which such rules increase the likelihood of school failure, and the effect of failure to meet a performance standard on subsequent student achievement. Using data from California's state accountability program, I find that subgroup rules lead to otherwise similar schools having different probabilities of failure. I also find that subgroup induced failure leads to lower future student achievement under both the state's system and its' implementation of No Child Left Behind. This implies that small demographic differences play a large role in how schools are judged and how they perform under current accountability systems.

© 2012 Elsevier Ltd. All rights reserved.

## 1. Introduction

The implementation of the federal No Child Left Behind Act of 2001 (NCLB) punctuated a decade of school accountability reforms. At the heart of this agenda is the belief that additional incentives for school employees will lead to better outcomes from students. This policy of (most often) punishing schools for failure to meet performance standards has been controversial in education circles. On the one hand, a number of studies suggest that both self-adopted state accountability regimes and those implemented to meet the federal mandate of No Child Left Behind have led to test score improvements for at least some groups of students. However, critics have charged that these systems fail to recognize the diversity of challenges facing schools and claim that recent accountability regimes induce widespread gaming of the system in ways that may negatively affect students.

One of these concerns, raised by Kane and Staiger (2002, 2003), is that the reliance of accountability systems on racial subgroup rules lead to different achievement requirements for otherwise similar schools. This paper expands upon their investigation to show how the racial subgroup rules in the California accountability system lead to vastly different failure probabilities for otherwise similar schools. Furthermore, this paper shows that these failure probabilities translate into future reductions in student achievement at schools that fail because they have an extra subgroup. As the subgroup thresholds, conditional on smooth race and demographic controls, induce a discontinuous change in failure probabilities I argue that these estimates may be interpreted as a causal effect of school failure on future student achievement.

The remainder of the paper proceeds as follows: Section 2 briefly introduces the school accountability literature and gives some background information on the California programs, Section 3 discusses the data, Section 4 explains the empirical strategy, Section 5 presents the results, and Section 6 concludes.

\* Tel.: +1 801 422 1554; fax: +1 801 422 0194.

E-mail address: [davesims@byu.edu](mailto:davesims@byu.edu).

## 2. Background

### 2.1. Previous research on accountability

While the literature studying the empirical effects of recent U.S. school accountability programs is too large to consider in its entirety, there are two broad threads that are extremely important to the context of the present study. First, there is a large and growing body of literature suggesting that both state accountability programs of the 1990s and the NCLB reforms led to an improvement in average test scores for some affected students (e.g. Carnoy & Loeb, 2002; Dee & Jacob, 2009; Hanushek & Raymond, 2005; Neal & Whitmore Schanzenbach, 2010; Reback, 2008). While this literature finds some test score effects both on the low stakes, nationally standard NAEP, as well as on individually designed state exams, it also provides evidence that the gains are not shared across all subjects or all types of students. In particular, positive effects in math scores are not accompanied by positive language score effects (Dee & Jacob, 2009), while students who are in the middle of the ability distribution, and thus most likely to play a marginal role in the school's accountability outcome, may account for all of the achievement gains (Neal & Whitmore Schanzenbach, 2010; Reback, 2008). This narrow nature of measured gains leaves open the possibility that these test score results may overstate the actual social value of the accountability programs.

These papers manage to overcome some daunting empirical challenges. Evaluating accountability programs that punish underperforming schools is inherently difficult because of the lack of a meaningful control group. Schools that have test scores low enough to fail are different in many salient ways than those with higher scores. Furthermore, because schools that failed have low test scores and test scores are noisy, some scores will likely improve due to mean reversion even in the absence of educational policy reform. This suggests that simple regression models may fail to capture the true effect of an accountability program on student learning. For example, in a reexamination of the Florida accountability program, Figlio and Rouse (2006) find that most of the effect of the program on achievement disappears when lagged test scores and more stringent controls are added.

Thus, the school achievement literature has largely adopted quasi-experimental methods to ascertain achievement effects. Partly because of this empirical framework, the above studies have focused on identifying the reduced form effect of the passage of a particular accountability program on student achievement. However, some papers have emphasized particular mechanisms that might induce schools to improve student achievement, such as fear of public stigma (Figlio & Rouse, 2006; Ladd & Glennie, 2001) and threats of school choice (Greene, 2001; West & Peterson, 2006), as well as the actual availability of school choice (Hastings & Weinstein, 2007). These mechanisms emerge logically from the premise that schools respond to threats of failure by improving students' test scores to avoid it, and most studies cite this accountability pressure as a main consideration in motivating their results (Chiang, 2009).

The effect of school failure itself receives somewhat less academic attention. Schools that actually fail are usually provided with some sort of supplemental student services or funding but face a schedule of increasing sanctions with each additional failure. Clearly, the structure of these laws suggests that punishing schools for failing is central to accountability, and that such punishments are expected to improve future performance. Indeed, estimates of the effect of the failure itself on future academic outcomes are the cornerstone of Rouse, Hannaway, Goldhaber, and Figlio's (2007) study of the Florida report card system. Their findings of positive achievement effects due to school failure have generally been accepted and seen as a gateway to the more important policy mechanism questions such as vouchers.

However, the recent work of Hemelt (2011) suggests that a re-examination of failure effects is necessary. Using a regression discontinuity approach in Maryland data, he finds that schools which barely fail to meet the academic standards suffer decreased future test scores relative to those that barely meet the standard. The present study uses a different identification strategy and also finds negative effects of failure to meet a performance standard on future test scores.

The second relevant strand of the school accountability literature concerns the unintended consequences of school accountability programs. These consequences have proved to be numerous and have reached into many facets of school operations including disability classifications (Figlio & Getzler, 2006), school lunch menus (Figlio & Winicki, 2005), and school calendars (Sims, 2008). More disturbing is the evidence that high stakes accountability programs led to an increased prevalence of teachers cheating on behalf of students to raise their exam scores (Jacob & Levitt, 2003). Another commonly studied consequence is the possibility that incentives for schools to focus on marginal students might adversely impact those students at either of the extremes of the ability distribution. The evidence on this is mixed, suggesting in some instances that the accountability system may lower the scores of these students (Ballou & Springer, 2008; Krieg, 2008), while implying in other cases infra-marginal students are simply excluded from the test score gains provided by the accountability system (Neal & Whitmore Schanzenbach, 2010; Reback, 2008).

In this paper, I examine another unintended consequence of accountability program design, the possibility that racial subgroup rules produce distorted accountability results.

Under most state accountability regimes and later NCLB, schools with a certain number of eligible students from a particular demographic subgroup were required to apply the academic standards separately to the average performance of students from that significant subgroup. Failure by the subgroup to reach the standard was usually sufficient to fail the entire school. Groups of eligible students were those of various race or ethnicity groups: blacks, Hispanics, Asians, and so forth, as well as socioeconomically disadvantaged students, special education students and English learners. Each state set the  $n$ -size or threshold level for the number of students in a school

Download English Version:

<https://daneshyari.com/en/article/354557>

Download Persian Version:

<https://daneshyari.com/article/354557>

[Daneshyari.com](https://daneshyari.com)