

‘Membership Has Its Privileges’: Status Incentives and Categorical Inequality in Education

Thurston Domina,^a Andrew M. Penner,^b Emily K. Penner^c

a) University of North Carolina, Chapel Hill; b) University of California, Irvine; c) Stanford University

Abstract: Prizes – formal systems that publicly allocate rewards for exemplary behavior – play an increasingly important role in a wide array of social settings, including education. In this paper, we evaluate a prize system designed to boost achievement at two high schools by assigning students color-coded ID cards based on a previously low stakes test. Average student achievement on this test increased in the ID card schools beyond what one would expect from contemporaneous changes in neighboring schools. However, regression discontinuity analyses indicate that the program created new inequalities between students who received low-status and high-status ID cards. These findings indicate that status-based incentives create categorical inequalities between prize winners and others even as they reorient behavior toward the goals they reward.

Keywords: prizes; categorical inequality; status incentives; education

PRIZES – formal systems that publicly allocate rewards for exemplary behavior – organize actors in a wide array of settings. From the Nobel Prize to the Honor Roll, prizes articulate a vision of excellence and encourage people to pursue that vision. Prize structures have proliferated in contemporary society (Best 2008), and they appear to be particularly prominent in loosely-coupled systems where they allow prize-givers to organize the behavior of disparate actors without exercising direct authority. Today, prizes play well-documented roles in athletics (c.f. Brown 2011; Ehrenberg and Bognanno 1990), arts and culture (c.f. Goode 1978; Rossman and Schilke 2014), science and letters (Merton 1968; Xie 2014), and business and technology (Anand and Watson 2004; Bourdreau and Lacetera 2011). But even as prize structures create incentives, they also construct social inequalities. Prizes create meaningful new social categories by designating winners and losers and providing opportunities for winners to acquire status, power, and resources (Bourdieu 1991; Frey 2006).

Prizes are particularly important in contemporary American schools. Long viewed as paradigmatic examples of loosely-coupled organizations (c.f. Meyer and Rowan 1975; Weick 1976), American public schools are in the midst of a decades-long organizational transformation. National, state, and local accountability policies have created a system of prizes and sanctions that aim to encourage educators to make progress toward explicitly articulated educational goals (McGuinn 2006; Rhodes 2012). The first generation of school accountability policies leaned heavily on sanctions to establish tighter links between incentives and outcomes (Booher-Jennings 2005; Hallett 2010). Recently, however, educational policy-makers and practitioners have begun to turn toward prize structures. For example, the U.S. Department of Education’s Race to the Top competition attempts to drive school

Citation: Thurston Domina, Andrew M. Penner, and Emily K. Penner. 2016. “Membership Has Its Privileges’: Status Incentives and Categorical Inequality in Education.” *Sociological Science* 3: 190-221.


Received: June 23, 2015

Accepted: August 17, 2015

Published: March 14, 2016

Editor(s): Olav Sorenson

DOI: 10.15195/v3.a10

Copyright: © 2016 The Author(s). This open-access article has been published under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited. 

reform policy-making at the state and district level. Similarly, several states and large public school districts use “performance pay” systems, which provide pecuniary rewards and public recognition for teachers and schools that meet goals for student achievement gains (c.f. Dee and Wycoff 2015; Jackson, Rockoff, and Staiger 2014; Springer 2009). Meanwhile, student-focused prize systems – including merit aid programs, school honor rolls, and token rewards – reward student academic achievement and behavior (Maehr and Midgley 1996; Thistlethwaite and Campbell 1960). Advocates argue that educational prizes provide incentives for students and educators to invest time and energy in schooling (Allen and Fryer 2011; Bishop 2006). Detractors, on the other hand, worry that competitive pressures distract from fundamental educational tasks and decrease intrinsic motivation; they are also concerned that prize structures may broaden existing inequalities and generate new ones (Deci, Koester, and Ryan 1999; Kohn 1993).

Prize programs as a double-edged sword

This paper describes and evaluates a controversial prize program in which schools gave students color-coded ID cards based on their performance on the California Standards Tests (CST). These end-of-course standardized tests, which are high-stakes for California schools and teachers, previously had few consequences for students. In an effort to align school, teacher, and student incentives, two schools, which we pseudonymously refer to as Live Oak High School and Mann High School, awarded students who performed in the highest proficiency bracket on every CST “platinum” ID cards, and students who scored in one of the two highest brackets on all tests “gold” ID cards. With a few exceptions, all other students received plain white ID cards. The schools allocated the cards in school-wide awards ceremonies and required students to display their cards at all times; students were also required to carry matching color-coded homework planners to classes. In addition, the schools granted privileges to platinum and gold ID card holders, including a separate “express” lunch line in the school cafeteria.

The color-coded ID card program was controversial, and district administrators discontinued it after two years. However, the program shares key elements with the wide array of prize systems that organize behavior in loosely-coupled systems by allocating status rewards to exemplary performers (c.f. Rossman and Schilke 2014). Further, it is an example of an increasingly popular approach to student management that attempts to use visual cues to orient within-school status systems around academic achievement (c.f. Taylor 2015; Walen 2012). The color-coded ID card program extends the logic of educational accountability that defines contemporary public education by creating a prize system for students that mirror the accountability system in which teachers and administrators work.

More broadly, as a social structure implemented with explicit rules in a data-rich environment, the ID card program provides a unique opportunity to closely observe the operation of status incentives and the construction of social inequality. Administrators at Live Oak and Mann viewed the color-coded ID card program as a simple and efficient way to coordinate effort in a loosely coupled system. Our analyses indicate that the program worked as intended in this regard. The introduction of

the ID card program changed the incentives facing students, providing them with a set of primarily status-based incentives determined by a clearly articulated set of rules that encouraged effort on a previously un-incentivized set of tests. In the process, the program boosted student achievement, particularly on the tests that the program incentivized.

But we argue that the incentives and information that prize systems produce come at a cost. The color-coded ID card program, like other prize programs, created new status-based identities and inequalities, and may have reinforced existing inequalities. The ID card program created new social categories by drawing categorical distinctions between students based on their achievement on end-of-course tests, and reinforced these categorical distinctions by marking students with ID cards. Building on Tilly's (1998) theory of categorical inequality, we hypothesize that the construction of these new social categories created new inequalities between students who received higher-status platinum and gold ID cards and their peers who received lower-status white ID cards. Laboratory-based research in social psychology and behavioral economics (c.f. Ashforth 1989; Ball et al. 2004; Goette, Huffman, and Meier 2006) indicates that people readily adopt and enact categorical identities even if they know they are based on arbitrary distinctions, and that these status distinctions influence subsequent behavior (Kalkoff and Thye 2006). Our research builds on this work, rigorously documenting the emergence of categorical inequality outside of the laboratory.

Literature review and hypotheses

We draw from literatures in sociology, social psychology, and economics that address the roles of prize structures and incentives in changing behavior to generate hypotheses about the potential impacts of the ID card policy. We hypothesize that the promise of winning and the threat of losing motivate responses to the program's prize structure. However, individuals' responses to these incentives are likely to vary systematically according to their perceived likelihood of winning. Furthermore, we hypothesize that color-coded IDs create meaningful social categories in the school setting, influencing student behavior by constructing new academic identities and reinforcing existing ones.

IDs as incentives

When they designed the color-coded ID card program, administrators at Live Oak and Mann hoped that rewarding achievement would motivate students, improving academic achievement on a wide range of indicators. Implicitly they viewed students as "adolescent econometricians" (Manski 1993) who estimate the economic and other returns to schooling as well as their likelihood of reaping those returns, and use these estimates to formulate plans and make educational choices. Arguing that many students lack sufficient short-term incentives to invest time and effort in schooling (c.f. Breen and Yaish 2006; Morgan 2005), and on the CST in particular, Live Oak and Mann administrators thus sought to boost student achievement on

the CST by increasing and diversifying the returns to educational effort. In doing so, the program echoes several interventions that have experimented with pecuniary incentives for students. This research shows, for example, that cash rewards for passing exam scores influence student course-taking behavior, improve SAT scores, and boost college enrollment (Behrman, Parker, Todd, and Wolpin 2015; Jackson 2010; but see also Angrist and Lavy 2009). Of particular note, experimental evidence suggests that immediate cash incentives boost student performance on tests that otherwise have no stakes for students (Braun, Kirsch, and Yamamoto 2011).¹

Unlike these well-studied financial student incentive programs, the color-coded ID card program primarily utilizes status incentives. While the program does offer tangible rewards, including free admission to athletic and other school events, the ID card program is notable for the public recognition it provides for high achieving students (cf. Walen 2012). Most strikingly, the program gave gold and platinum card holders access to a specially demarcated “express” line in the high school cafeteria. By marking students based on their test scores, the ID card program attempted to build new academically-oriented status hierarchies at Live Oak and Mann. To date, relatively little empirical research has considered status incentives and their consequences (but see Tran and Zeckhauser 2012).

Assuming these incentives appeal to students, we hypothesize that exposure to the color-coded ID card program’s incentives boosts student performance on the tests used to make ID card assignments. Less clear, however, is the extent to which these positive effects ought to generalize beyond CSTs to other measures of achievement. As noted by the Live Oak High School principal, who helped to design the color-coded ID card program, the CST is otherwise a uniquely low-stakes test for high school students:²

It means everything to the school, but nothing to the students. A pop quiz in home economics has more incentives for students. I have had teachers work incredibly hard to prepare their students only to watch them bubble pictures on their answer sheet or take a nap.

According to this educator, the lack of incentives biases student performance on this exam. Research in educational assessment suggests that this concern is valid, indicating that students perform as much as 0.6 standard deviations lower on low-stakes tests than on high-stakes tests (Wise and DeMars 2005; Braun, Kirsh, and Yamamoto 2011).

Accordingly, we expect students to respond to the introduction of incentives associated with the CST by investing more energy on the test. While this effort might spill over into other educational domains, boosting scores on a test that already has high stakes for students, grades, and other measures of student learning and academic behavior, such spill-over is by no means assured. Therefore, we hypothesize that:

Hypothesis 1: Students exposed to the color-coded ID card program will score higher on the CST than students who were not exposed to the color-coded ID card program. The differences between the test scores of students who were and were not exposed to the ID card program will be largest on the CSTs, and smaller for other measures of student achievement or effort.

Contemporary applications of rational choice theory to educational settings typically acknowledge that students face a great deal of uncertainty in estimating the returns to education and their likelihood of realizing educational goals. This uncertainty likely influences student estimates of the costs and benefits of educational investments (Bénabou and Tirole 2003). Put simply, students who believe that their chances of educational success are small may not invest in education, even if they believe that the returns to education are large (c.f. Altonji 1993). Conversely, students who are certain that they will succeed regardless of their effort may also invest little energy in schooling (Rosenbaum 2001).

If students calibrate their responses to academic incentives according to their perceived odds of success, we would expect student responses to the ID card program to vary by prior achievement. The program articulated a clear set of eligibility criteria based on tests that students take and receive scores from annually. If students use information from prior tests to assess their odds of earning high-status ID cards in the future, students who previously scored close to the ID card assignment thresholds may believe that increased effort is worthwhile. However, students who score well below the threshold may question the link between effort and reward. We thus hypothesize that

Hypothesis 2: Differences between students who were and were not exposed to the ID card program will be greater among students whose prior test scores are close to an ID card assignment threshold than among students whose prior test scores are far below thresholds.

Stereotype threat and stereotype lift

Stereotypes about academic performance may also influence the relationship between the ID card program and student achievement. Research related to stereotype threat and stereotype lift suggests that making race (Steele and Aronson 1995), class (Croizet and Claire 1998), or gender (Spencer, Steele, and Quinn 1999) more salient, or raising the anxiety associated with a test, can inhibit the academic achievement of students in negatively stereotyped groups (but see Stoet and Geary 2012), and boost achievement for students in positively stereotyped groups (Walton and Cohen 2003). Of particular relevance to this study, previous research suggests that in contexts of stereotype threat, Latino students score worse than white students (Gonzales, Blanton, and Williams 2002) while white students score worse than Asian students (Aronson et al. 1999).³

The ID card program might activate or exacerbate racial stereotypes regarding the academic achievement through one of several mechanisms: First, the introduction of social incentives might negatively affect the achievement of students who belong to traditionally low-scoring groups through increasing anxiety about scoring well on this test (Ben-Zeev, Fein, and Inzlicht 2005; O'Brien and Crandall 2003). Second, as some research suggests that stereotype threat effects operate only among students who value a particular domain (cf. Aronson et al. 1999; Davies, Spencer and Steele 2005), the ID card policy could induce stereotype threat and stereotype lift effects by giving students additional incentives to care about their achievement on this test. Third, the program could cue stereotypes by providing students with

information about how their peers perform (cf. Davies et al 2002; Inzlicht and Ben-Zeev 2000). That is, in much the same way that informing students that a particular test does not have race or gender differences can eliminate stereotype threat effects (cf. O'Brien and Crandall 2003), visually reinforcing these stereotypes might serve to exacerbate their effects. If any of these mechanisms operate, negative academic stereotypes may undermine the ID card program's incentives for Hispanic students, while positive academic stereotypes may exaggerate incentive effects for Asians. Research on the effects of stereotypes on academic achievement thus suggests that:

Hypothesis 3: Exposure to the ID card program will be associated with larger gains for Asian students than for white students, and smaller gains for Hispanic students than white or Asian students.

Producing categorical inequality

In addition to changing the salience of existing group memberships, the introduction of color-coded ID cards might also create new status distinctions and identities among students. Work on the minimal group paradigm – much of which is conducted with young adults near the age and developmental stage of U.S. high school students – suggests that identity-formation processes can imbue even trivial differences with meaning (Ashforth 1989; Harvey et al. 1961). These differences can have important consequences. Lovaglia et al. (1998) document that attaching status differences to otherwise irrelevant categories can induce differences in students' achievement test scores, and Ball et al. (2004) demonstrate that categorical assignments affect behavior even in cases in which students know that status assignments are random. Papay, Murnane, and Willett (forthcoming) provide provocative data to suggest that these labeling processes are widespread in American education.

Anecdotal information suggests that the introduction of the color-coded ID cards created salient identities with strong, clearly delineated social boundaries. One white card holder at Mann noted that "It makes you feel dumb, like your school is putting you down." Another student recounted hearing a platinum card honors student tell a white card honors student that "Nobody with a white card should be in the honors program." These social distinctions were actively encouraged by the administration: a Mann parent related an incident in which an administrator told a group of girls to aspire to attend prom with platinum card holders.

We view these categorical assignments as a Bourdieuan "rite of initiation," in which ID card assignment "manages to produce discontinuity out of continuity" (Bourdieu 1991:120). That is, the program capitalizes on small differences in achievement test scores to construct mutually exclusive, highly salient social categories with "bright" boundaries (c.f. Alba 2005). We hypothesize that these categorical assignments influence student experiences at Live Oak and Mann, ultimately creating new inequalities that are especially salient at the ID card placement thresholds. We imagine three mechanisms through which such effects might occur. First, the program might send students who receive a white ID card discouraging messages about their academic potential even as it sends encouraging messages to students who receive gold and platinum ID cards. If so, white ID card assignment might depress student academic engagement and achievement (c.f. Wang and Eccles 2013).

Second, the program might influence students' peer interactions. Prior research (c.f. Goette, Huffman, and Meier 2006) indicates that people are more likely to cooperate effectively when they share a group identity. Such a phenomenon might lead to more positive relationships with higher achieving peers for students with gold and platinum ID cards than for students with white ID cards. Third, status-based identities can influence external evaluations of an individual's competence (Ridgeway and Correll 2006). If teachers perceive an association between a student's ID card and his or her competence, they might grade students with white ID cards more harshly and hold lower expectations for these students, further lowering student engagement and learning (Eden and Ravid 1982; Rosenthal and Jacobson 1968).

While we are unable to adjudicate between these three mechanisms, each is consistent with the hypothesis that:

Hypothesis 4: Assignment to a low-status white card (rather than a high-status gold card) will have a negative effect on student test scores and grades.

Data and methods

Program description

In the fall of the 2010-2011 school year, Live Oak and Mann High Schools issued color-coded ID cards to all students based on their performance on the prior year's CSTs. While administrators at the other Sudden Valley high schools were aware of the program, they did not participate in it. Live Oak's principal deemed the program "an instant success" and both schools replicated the program in the 2011-2012 school year. Under the California Schools Accountability Act (1999) students take CSTs each spring in math, English language arts (ELA), social studies, and science through eleventh grade.⁴ Scores are then reported in a series of five performance bands: advanced, proficient, basic, below basic, and far below basic. Aggregated student scores on these tests are publicly reported and linked to a series of sanctions and incentives for schools under California accountability policies and the federal No Child Left Behind Act (2001).

The ID card program was designed to pass these incentives on to students. The program issued platinum ID cards, as well as matching homework planners, to students who scored "advanced" on each of the CSTs that they took. It issued gold cards and planners to students who scored either "proficient" or "advanced" on all of their CSTs. Other students received white cards and planners. The program offered platinum and gold ID card holders free or discounted tickets to home sporting events and school dances; coupons for local businesses; and entrance to lotteries for school event tickets, class rings, and yearbooks. However, the program's most important incentives were social. Schools required students to display both their ID cards and their planners at all times on campus, and they reinforced these identities by creating an "express" line at the school cafeteria for gold and platinum ID card holders. The online supplement provides more details about program design and implementation.

Data

Sudden Valley officials provided our research team access to student-level demographic, test score, and transcript data covering all students enrolled in each of the district's high schools between fall 2008 and spring 2012, a span that includes the two years of program implementation as well as a pre-implementation year and a transition year.

Setting

Live Oak and Mann are large, diverse high schools located in an inner-ring California suburb. Both schools are situated on sprawling campuses along busy commercial strips in middle class neighborhoods. When the ID card program was in place, administrators adorned breezeways and other public spaces on both campuses with banners and signs that borrowed phrases from credit card marketing campaigns (e.g. "The Gold Card: Never leave home without it" and "The Platinum Card: Membership has its privileges").

Table 1 provides a demographic snapshot of Live Oak and Mann, as well as the other high schools in their district and the state of California. Both Live Oak and Mann enroll an ethnically diverse, predominately middle class student body. Asians are the largest ethnic group at Live Oak and Mann, accounting for 37 percent of the student body at both schools. While Mann enrolls a somewhat poorer and more Hispanic student body than Live Oak, students at both schools are advantaged compared to their peers at other high schools in the district. Supplementary analyses suggest that Live Oak and Mann students were less likely to leave the district or change high schools within the district than their peers in other Sudden Valley schools, both before and after the ID program's implementation. While white card holders have slightly higher rates of school attrition than gold and platinum card holders, there is no significant discontinuity in odds of attrition or changing schools across the ID assignment threshold.

In spite of their relatively high performance compared with other Sudden Valley schools, principals at both schools felt a great deal of pressure to demonstrate academic improvement. The Academic Performance Index scores reported in Table 1 are a composite of test scores prepared by the California Department of Education for use in the state's accountability policy. Administrators at Live Oak and Mann speak frequently of their school's API scores, comparing them to the scores of a nearby charter school with selective admissions, as well as scores in a more affluent nearby suburb.

Table 2 summarizes the rate of white, gold, and platinum card attainment for students enrolled in Live Oak and Mann in the second program year. This table shows that less than half of the student body in these schools received the desirable gold or platinum cards: 59 percent of students received white ID cards, 29 percent received gold ID cards, and 13 percent received platinum ID cards. Consistent with the API data reported above, the proportion of Live Oak students who received the platinum card is higher than the comparable proportion for Mann students (16 percent vs. 10 percent), while Mann has a higher proportion of students who received gold cards than Live Oak.

Table 1: Comparison of student demographics in Live Oak and Mann High Schools with other Sudden Valley Unified High Schools, spring 2012

	Live Oak HS	Mann HS	Other Sudden Valley Unified HS	California, Grades 9–12
Hispanic (%)	22	33	75	50
Asian (%)	37	37	13	12
White (%)	36	24	9	28
African-American/Other race (%)	5	6	3	10
Male (%)	52	49	51	51
Female (%)	48	50	49	49
Free/reduced price lunch (%)	29	39	78	47
Non-native English speaker (%)	37	41	75	40
Prior CST well below threshold	31	37	63	
Prior CST near threshold	24	25	19	
Prior CST well above threshold	45	38	18	
Academic Progress Index (API)	906	860	741*	752
N	2,604	2,362	16,492	1,979,678

* Weighted mean of Sudden Valley Unified's 6 other regular high schools.

However, the most notable differences in card placement rates occur not between the two schools, but within them. Across the two schools, Asians are nearly 5 times more likely to receive platinum cards than their Hispanic classmates. Just 5 percent of Hispanic students and 3 percent of African American students receive platinum cards, while 24 percent of Asians receive platinum cards. White students are also under-represented among platinum card holders in these schools, with 9 percent of white students receiving platinum cards. Although less pronounced, racial gaps also exist in gold card receipt. As a result, 75 percent of Hispanic students and 78 percent of black students have low-status white ID cards, compared to 41 percent of Asian students. These large discrepancies in ID card receipt may reinforce academic stereotypes and exacerbate stereotype threat and lift effects. It is also possible that the ID card program may draw on existing racial hierarchies at Live Oak and Mann to legitimate the status that the gold and platinum ID cards convey to their recipients.

The remaining inequalities in ID card placement reported in Table 2 are relatively small. The proportion of students who receive platinum cards peaks at 19 percent in the tenth grade and declines to 9 percent by the twelfth grade. Males are slightly more likely to receive platinum cards than females, and free/reduced lunch recipients are less likely to receive platinum or gold cards than their relatively affluent peers. Non-native English speakers – most of whom are Asian – are over-represented among platinum and gold card holders. In an online supplement, we report the results of a more detailed multivariate examination of these ID card inequalities. While these analyses indicate that Asians and boys are over-represented

Table 2: Card attainment rates by race, gender, free/reduced lunch, language status and grade, Live Oak and Mann High Schools, 2012

	White (below proficient)	Gold (proficient on all)	Platinum (advanced on all)
All students	59	29	13
Mann	60	31	10
Live Oak	58	26	16
Hispanic	75	20	5
Asian	41	36	24
White	62	29	9
African-American/Other race	78	19	3
9 th grade	66	22	12
10 th grade	51	30	19
11 th grade	57	32	11
12 th grade	60	31	9
Male	57	27	14
Female	60	28	12
Free/reduced price lunch	66	25	9
Non-native English speaker	49	32	19
N	2,910	1,414	642

among high-status ID card holders even after controlling for their prior achievement, these gaps appear to be due to student achievement growth patterns, rather than inconsistencies in ID card assignment.

Findings

Figure 1 provides a descriptive portrait of mean ELA CST score trends for tenth grade students in Live Oak and Mann High Schools as well as their peers in other schools across the Sudden Valley district and the state of California between 2004-05 and 2011-12. Students at both Live Oak and Mann consistently outperformed their peers elsewhere in Sudden Valley and in the State of California, both before and after the introduction of the ID card program. While mean CST scores improved for students statewide and throughout Sudden Valley during this time period, scores for Live Oak and Mann students grew at a particularly rapid pace. Mean CST scores for Live Oak and Mann students grew by approximately 3 percent annually in the years prior to the ID program implementation; compared to less than 2 percent for students in other Sudden Valley high schools. As Figure 1 demonstrates, however, the rate of mean test score growth for Live Oak and Mann students further improved with the introduction of the ID card program in 2010-11 and 2011-12. Mean test scores at Live Oak and Mann grew by nearly 7 percent annually during

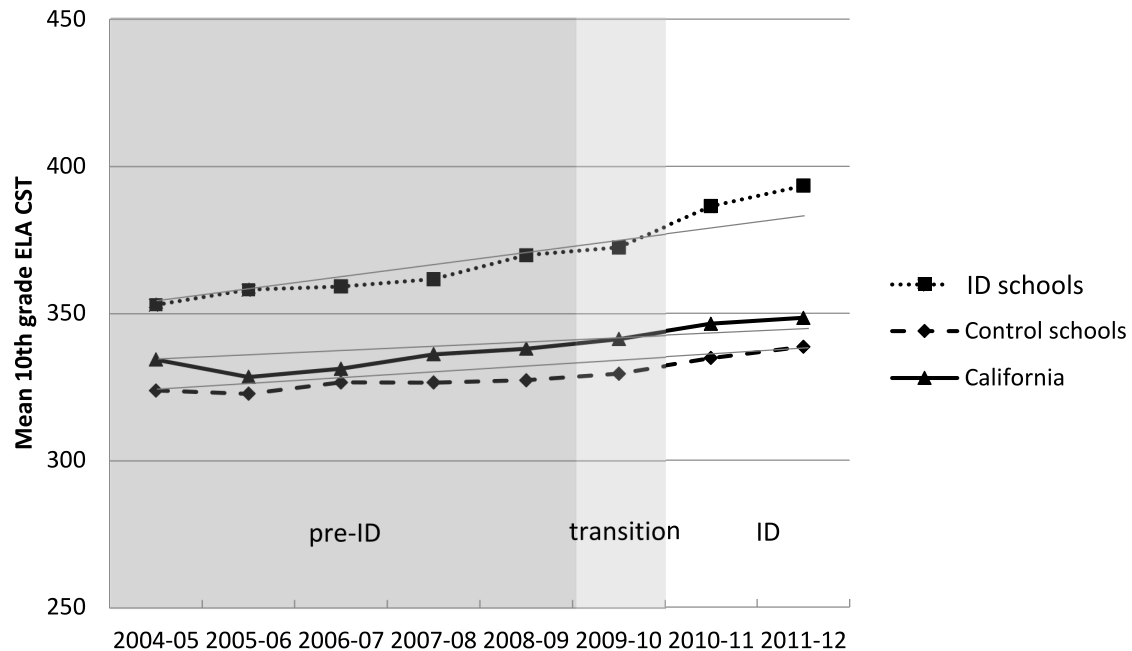


Figure 1: Mean ELA CST score trends for students at Live Oak and Mann (ID schools), other Sudden Valley Unified High Schools (Control schools), and California. Source: California Department of Education

the years in which the ID program was in place – a rate of growth that is twice the rate of pre-program growth in these schools and twice the rate of growth posted by other Sudden Valley and California high schools during the same years. While this evidence is by no means conclusive, it is consistent with the prediction that exposure to the ID program corresponds with rapid improvements in student performance on the CSTs upon which ID card assignment is based.

Table 3 provides a more thorough set of descriptive statistics for Live Oak, Mann, and other Sudden Valley high schools in the pre- and post- ID program periods using student level data. Unfortunately, we only have student-level data from the last four years of the time-series summarized in Figure 1 (2008-09 to 2011-12). However, since Figure 1 suggests that achievement trends were fairly consistent prior to 2008-09, this shorter time-series likely provides an accurate view of ID program implementation and its relation to student achievement.⁵ Consistent with Figure 1, the table indicates that mean achievement levels were higher at Live Oak and Mann on each of the metrics for which we have data, including CST scores as well as scores on California’s high school exit exam (CAHSEE) and teacher-assigned grades. Further, this figure indicates that mean achievement levels on each of these metrics with the exception of math CAHSEE scores improved substantially in Live Oak and Mann during the years in which the ID program was implemented. While other Sudden Valley high schools also posted gains on many of the measures of student achievement, these gains tended to be relatively small.

Table 3: Mean standardized achievement levels for students in Live Oak and Mann (ID schools) and other Sudden Valley Unified high schools (Control schools), with standard deviations in parentheses

	ID schools		Control schools	
	Live Oak & Mann HS		Other SVUSD HS	
	Pre-ID period	ID period	Pre-ID period	ID period
ELA CST	0.43 (0.85)	0.64 (0.93)	-0.23 (0.92)	-0.19 (0.90)
Math CST	0.26 (0.98)	0.74 (1.08)	-0.33 (0.82)	-0.16 (0.89)
ELA CAHSEE	0.40 (0.81)	0.57 (0.89)	-0.20 (0.90)	-0.13 (0.93)
Math CAHSEE	0.58 (0.86)	0.57 (0.97)	-0.19 (0.88)	-0.15 (0.92)
ELA Grades	0.31 (0.93)	0.38 (0.98)	0.01 (0.96)	-0.12 (0.98)
Math Grades	0.21 (0.90)	0.32 (0.95)	-0.15 (0.98)	-0.04 (0.99)
N	1,060	2,528	3,203	7,927

Notes: The pre-ID period is the 2008-09 school year; the ID period is the 2010-11 and 2011-12 school years. Although Live Oak and Mann did not implement the ID card program until the 2010-11 school year, ID assignments in the first program year were based on spring 2010 CST scores. Therefore, we consider the 2009-10 school year a transitional period in the implementation process.

Prize systems and student achievement

In the analyses that follow, we build on the logic of Figure 1 and Table 3, using a multivariate difference-in-difference approach to understand how student achievement changed when the ID card program was implemented. These analyses track changes in educational outcomes for tenth grade students at Live Oak and Mann before and after the implementation of the ID card program, and compare them to changes in other Sudden Valley high schools over the same period. To conduct these analyses, we pool data for all tenth grade students enrolled in any Sudden Valley high school in the 2008-09, 2009-10, 2010-11, and 2011-12 school years. We focus on tenth grade students because we have exit exam scores for these students in addition to CST scores, allowing us to examine the extent to which program effects transfer from the directly incentivized CSTs to the exit exam, which was already high-stakes for students, and was not incentivized by the ID card program. By comparing cross-cohort changes for Live Oak and Mann students with the changes that occurred in other Sudden Valley high schools during the same time period, we are able to account for the confounding effects of time-invariant school characteristics

and broad secular trends. These analyses take the general form:

$$Y = \beta_0 + \beta_1(LM) + \Sigma\beta_{2-4}(cohort) + \Sigma\beta_{5-7}(LM * cohort) + \beta_k(controls) \quad (1)$$

where Y is a measure of student achievement from the spring of the tenth grade year, LM is an indicator variable for Live Oak and Mann tenth grade students, $cohort$ is a series of indicator variables comparing students who were tenth grade students in each of the two ID card program years and the transition year with their peers in the 2008–09 reference cohort,⁶ $LM*cohort$ is a series of interaction terms tracking the difference between cross-cohort changes in the ID card schools with cross-cohort changes in the control schools, and $controls$ include measures of student race, gender, free/reduced lunch status (which is an imperfect proxy for poverty), and English language proficiency status. β_1 therefore estimates the conditional difference in test scores between the ID card schools and other schools in the baseline school year. β_{2-4} measure cross-cohort changes in the control schools, net of demographic controls. β_5 measures the unique cross-cohort change that occurred at Live Oak and Mann in the transition year before the ID card program was fully implemented. β_{6-7} , the parameters of interest in these models, provide estimates of the unique cross-cohort changes that occurred at Live Oak and Mann in the ID card program years, net of district-wide changes and within-school demographic shifts.

Table 4 summarizes the results of our difference-in-difference analysis of the ID card program on ELA CST scores as well as several other academic outcomes (see Table 3 in the online supplement for full model results). The analyses reported in the first column of Figure 3 indicate that prior to the introduction of the ID card program, students at Live Oak and Mann outperformed their peers at other Sudden Valley high schools on the ELA CST by 0.34 standard deviations. ELA CST scores in control schools were 0.07 standard deviations higher in 2012 than in 2010. However, Table 4 shows that scores improved particularly rapidly in Live Oak and Mann, so that the introduction of the ID card program was associated with significant positive changes in Live Oak and Mann student achievement on ELA CSTs, above and beyond the contemporaneous changes in other Sudden Valley schools (0.18 SD in 2011 and 0.19 SD in 2012). Figure 2 provides a visual representation of these results.

Table 4 further suggests that the ID card program's introduction was associated with even larger gains in mathematics CST scores (0.31 SD in 2011 and 0.33 SD in 2012).⁷ The remaining analyses summarized in Table 4 explore the extent to which the ID card program influenced student performance on California's High School Exit Exam (CAHSEE) and grades. Both of these outcomes were incentivized in the absence of the ID card prize, since students must pass the exit exam in order to earn a diploma and classroom grades play an important role in college admissions. Since the ID card program does not change these incentives, hypothesis 1 posits that exposure to the ID card program will be associated with smaller gains on these outcomes than on the CSTs. Consistent with this hypothesis, we find no evidence of ID card program effects on mathematics exit exam scores or mathematics grades. However, ELA grades are higher at Live Oak and Mann in both years the ID card program was in place, as were ELA CAHSEE scores in the first year of the program.

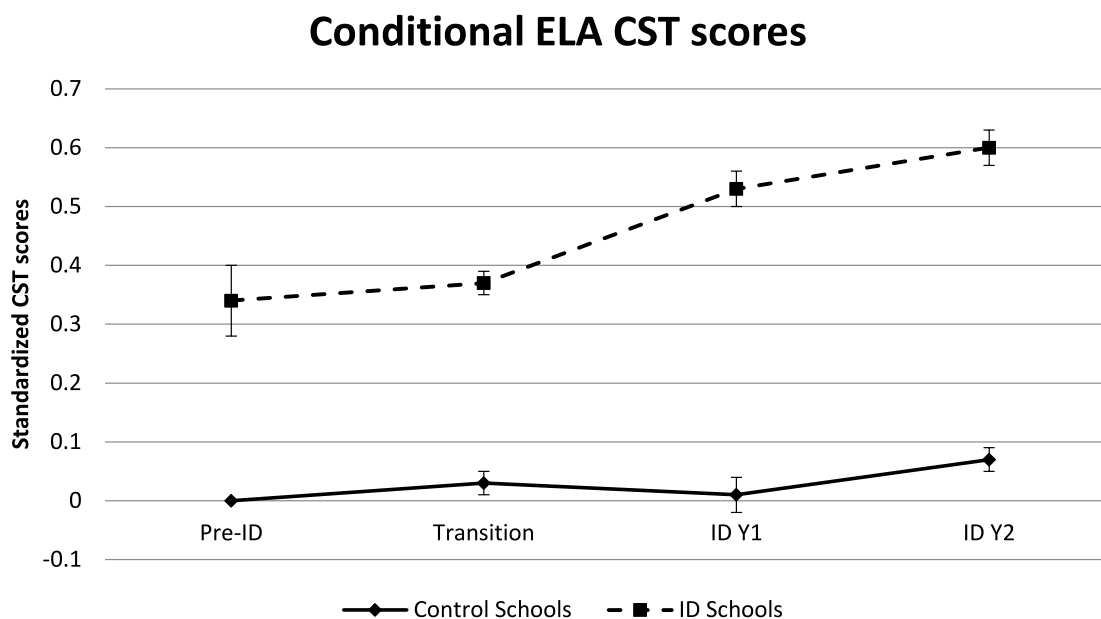
Table 4: Difference-in-difference estimates of the effects of ID card program implementation on CST scores, CAHSEE scores, and grades (selected results)

	ELA CST (std)	Math CST (std)	ELA exam (std)	Math exit exam (std)	ELA grades (std)	Math grades (std)
ID school	0.34 [†] (0.06)	0.27 (0.22)	0.25 [†] (0.03)	0.42 [†] (0.08)	0.01 (0.22)	0.08 (0.07)
2010	0.03 (0.02)	0.08* (0.02)	0.04* (0.01)	0.02 (0.04)	-0.20 [†] (0.06)	0.02 (0.04)
2011	0.01 (0.03)	0.17 [†] (0.04)	0.12 [†] (0.03)	0.02 (0.05)	-0.16* (0.07)	0.10 (0.09)
2012	0.07* (0.02)	0.21 [†] (0.02)	0.07 [†] (0.03)	0.10* (0.04)	-0.09 (0.07)	0.15* (0.07)
ID·2010	0.00 (0.02)	0.08 (0.07)	0.00 (0.04)	-0.16* (0.06)	0.11 (0.06)	(0.12) -0.07
ID·2011	0.18 [†] (0.03)	0.31* (0.10)	0.14 [†] (0.03)	0.01 (0.05)	0.21 [†] (0.07)	0.01 (0.10)
ID·2012	0.19 [†] (0.03)	0.33 [†] (0.02)	0.06 (0.03)	-0.09 (0.04)	0.21 [†] (0.07)	-0.02 (0.11)
N	19,567	18,763	19,504	19,466	19,686	19,722

Note: See supplement Table 3 for full model results. ELA and Math grade analyses use tobit estimation to correct for floor effects, and all models use school-clustered standard errors to account for school clustering. * $p < 0.05$; † $p < 0.01$.

The analyses reported in Table 4 thus align with hypothesis 1, insofar as they suggest that the ID card program's status incentives boosted student performance on the CSTs that the program aimed to reward. It is less clear whether the gains associated with exposure to the ID program spilled over to other aspects of student academic behavior. While there is evidence of comparable increases in ELA grades, it is difficult to know how to interpret this, given that the teachers were encouraged to consider CST scores as they calculated student grades.⁸ By contrast, our standardized estimates of ID card program effects on math CST scores are significantly larger than our estimates of ID card program effects on CAHSEE scores ($p < 0.01$) and grades ($p < 0.001$). On balance, we thus conclude that although the ID card program's effects may have transferred to non-incentivized outcomes, its effects were largest and most consistent on the tests it directly incentivized.⁹

We further expect that students will calibrate their response to prize incentives based on their perceived likelihood of winning. Hypothesis 2, therefore, holds that the relationship between exposure to the ID card program and student outcomes will vary by students' prior CST achievement. According to this hypothesis, students who scored close to the threshold for a platinum or gold ID card in the



Note: Confidence intervals represent standard errors from models in supplement Table 3.

Figure 2: ELA CST score trends, Live Oak and Mann (ID schools) and other Sudden Valley Unified High Schools (Control schools), controlling for race, gender, special education status, and English-language learner status.

year before will post larger achievement gains under the program than students whose prior CST achievement was far below the threshold for a gold ID card and who are thus likely to doubt their chances of earning a gold ID card. To test this hypothesis, we add to equation 1 a series of three-way interaction effects allowing the relationship between ID card exposure and student outcomes to vary based on whether students had previously scored well below the gold card threshold, close to the gold card threshold, or well above the gold card threshold (in which case they were potentially motivated by the prospect of earning a platinum card). In these analyses, the well below threshold group includes any student who scored 25 points or more below the gold card threshold on their lowest prior-year CST, the near-threshold group includes students whose lowest CST score was between 25 points (approximately 0.4 standard deviations) below the gold card placement threshold and 25 points above the gold card threshold, and the well-above threshold group includes any student whose lowest score was scored at least 25 points above the gold card placement threshold on all CSTs.

The analyses reported in the first two columns of Table 5 test hypothesis 2. The ID*2011 and ID*2012 coefficients in model 1 show that exposure to the ID card program was not associated with statistically significant changes in ELA CST scores for students who had previously scored at least 25 points above the gold card threshold on both CSTs. Model 2 reports similar findings for mathematics CST scores. However, the three-way interactions in these models suggest that the ID card program was associated with significantly larger gains in both ELA and mathematics for students who had previously scores close to the gold card threshold.

We see a similar pattern for students who scored well below the threshold, though in this case only the interactions for math are statistically significant. These results confirm that there is variation in the test score gains associated with the ID card program, and suggest that the ID card program is associated with the largest gains for students who scored near the threshold.¹⁰ Supplemental analyses confirm that students scoring near the threshold experienced a test score increase in the years that the ID card program was in place (see Table 5 in the online supplement for details).

Stereotype threat and stereotype lift

Our third hypothesis suggests that the relationship between exposure to the ID card program's incentives and achievement may also vary by student race. If the prize program activates stereotyped social identities, we hypothesize that exposure to the program will have significantly larger achievement gains for Asian students (who benefit from positive academic stereotypes) than white students, and smaller achievement gains for Hispanic students (who are subject to negative academic stereotypes). To test this hypothesis, we estimate a similar three-way interaction effect allowing the association between ID card exposure and student outcomes to vary by student race.¹¹

Models 3 and 4 in Table 5 report the results of these analyses. Consistent with hypothesis 3, these models suggest that the relationships between program exposure and CST scores are somewhat smaller for Hispanic students than for white students, and the program is associated with slightly larger gains for Asian students than white students. The race-specific increases associated with exposure to the ID card program are uneven across subject domains. In math we find that students from all racial backgrounds benefit from exposure to the ID card program (see supplement Table 5 for details). Further, there are no statistically significant differences in the degree to which different groups benefit. By contrast, we find that only the ELA scores of Asian and white students increase under the ID card program, so that Hispanic students receive a smaller ELA score from exposure to the ID card program than their Asian and white counterparts (though in many cases these race differences are only marginally significant). In sum, the pattern of ELA results in Table 5 (but not the pattern of math results) is broadly in keeping with predictions from the stereotype threat literature. It is important to note, however, that we find no evidence to suggest that Hispanic students do worse under the ID card program than in its absence. That is, while Hispanic students may not experience the gains experienced by white and Asian students in ELA, they do not appear to score significantly worse than Hispanic students in non-ID card program schools and years.¹²

While conceptually distinct, hypotheses 2 and 3 are difficult to distinguish empirically since student achievement is associated with race. Models 5 and 6 in Table 5 report the results of models that include all interaction terms to test the extent to which the heterogeneous effects consistent with hypothesis 2 are independent of the heterogeneous effects consistent with hypothesis 3. The relevant interaction effect coefficients are similar in both direction and magnitude to the

Table 5: Difference-in-difference estimates of effects of ID card program implementation on CST scores, by students' prior achievement and race (selected results)

	Model 1 ELA CST (std)	Model 2 Math CST (std)	Model 3 ELA CST (std)	Model 4 Math CST (std)	Model 5 ELA CST (std)	Model 6 Math CST (std)
ID school	0.71 [†] (0.07)	0.76 [†] (0.18)	0.41 [†] (0.06)	0.41 (0.20)	0.70 [†] (0.06)	0.74 [†] (0.17)
2011	0.34 [†] (0.06)	0.59 [†] (0.03)	0.09 [†] (0.02)	0.24 [†] (0.04)	0.33 [†] (0.08)	0.58 [†] (0.04)
2012	0.50 [†] (0.07)	0.60 [†] (0.04)	0.17 [†] (0.03)	0.31 [†] (0.02)	0.48 [†] (0.07)	0.64 [†] (0.06)
ID·2011	-0.02 (0.07)	0.02 (0.14)	0.13 [†] (0.03)	0.26* (0.13)	0.09 (0.07)	0.03 (0.13)
ID·2012	-0.09 (0.10)	0.00 (0.11)	0.14 [†] (0.04)	0.27 [†] (0.04)	-0.06 (0.12)	-0.02 (0.10)
Under-ID·2011	0.09 (0.11)	0.21* (0.08)			0.09 (0.07)	0.25* (0.13)
Under-ID·2012	0.17 (0.14)	0.25* (0.10)			0.18 (0.16)	0.26* (0.09)
Close-ID·2011	0.27* (0.08)	0.26* (0.09)			0.26* (0.08)	0.27* (0.08)
Close-ID·2012	0.34* (0.05)	0.36 [†] (0.10)			0.34 [†] (0.06)	0.36 [†] (0.09)
Hispanic-ID·2011			-0.20 (0.10)	0.02 (0.04)	-0.16* (0.05)	-0.08 (0.04)
Hispanic-ID·2012			-0.21 (0.11)	0.02 (0.05)	-0.08 (0.05)	0.02 (0.05)
Asian-ID·2011			0.10 (0.10)	0.23 (0.20)	-0.04 (0.07)	0.19 (0.10)
Asian-ID·2012			0.07 (0.07)	0.08 (0.12)	-0.04 (0.04)	0.09 (0.09)
N	19,567	19,283	19,567	19,283	19,567	19,283

Notes: See supplement Table 4 for full model results and supplement Table 5 for supplemental subgroup significance tests. ELA and Math grade analyses use tobit estimation to correct for floor effects and all models use school-clustered standard errors to account for school clustering. Although Live Oak and Mann did not implement the ID card program until the 2010-11 school year, ID assignments in the first program year were based on spring 2010 CST scores. Therefore, we consider the 2009-10 school year a transitional period in the implementation process.

* $p < 0.05$; † $p < 0.01$.

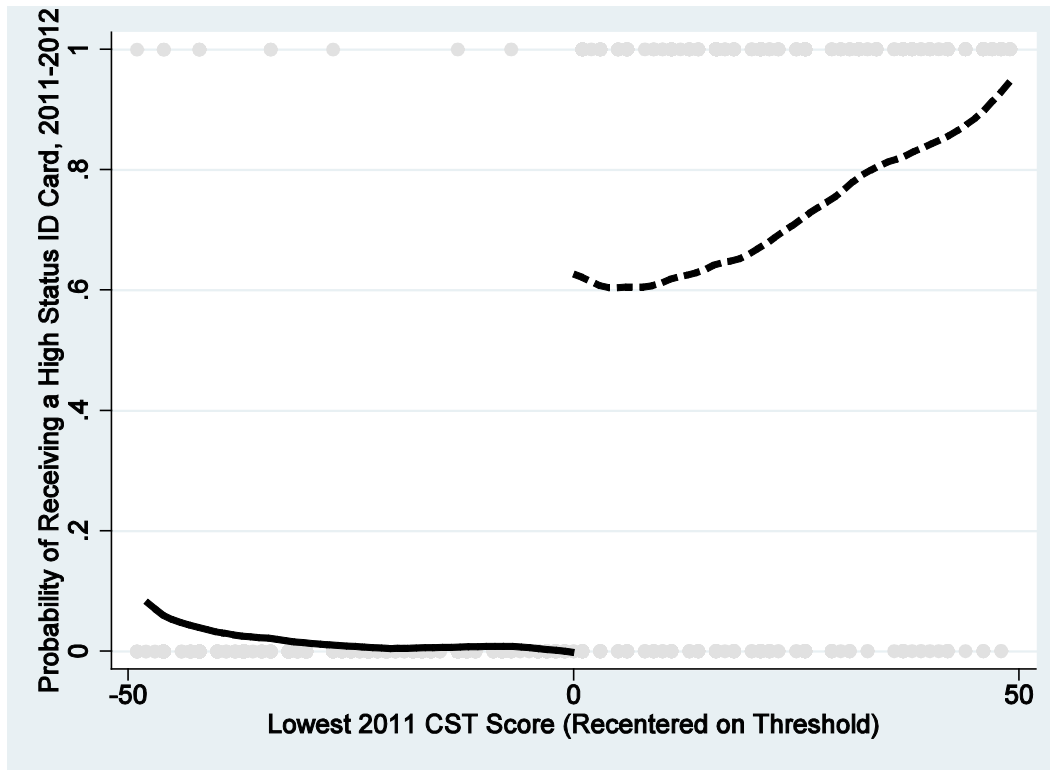
coefficients in the earlier models in Table 5, suggesting that heterogeneous effects by prior achievement and by race are largely independent.

The construction of categorical inequality

In the process of incentivizing academic achievement, the prize program also creates social categories based on student test scores. Building on theories of categorical inequality (cf. Tilly 1998), our fourth hypothesis suggests that this program may influence students' academic outcomes by creating new institutionally sanctioned, status-laden identity categories for students. If, for example, ID card assignment influences students' identities or teachers' views of students, we might expect new discontinuities in student academic achievement to emerge along the ID card assignment thresholds. In particular, we hypothesize that receiving a low-status white ID card has a negative effect on students' subsequent academic achievement.

We test this hypothesis using a series of binding score regression discontinuity analyses. Under the ID card program assignment system utilized by both Live Oak and Mann, any student who scores below the proficiency threshold on any CST is likely to earn a white ID card, while students who score above the proficiency threshold on all CSTs earn gold or platinum ID cards. Figure 3 demonstrates that this assignment procedure creates a pronounced discontinuity in student ID card assignment odds at the proficiency threshold. Using data from all 2012 Live Oak and Mann students, Figure 3 graphs student ID card assignment rates against an index defined as the lowest of students' CST scores (after recentering scores around the proficiency threshold).¹³ It indicates that this index closely tracks card assignment. A small number of students below the threshold earn gold cards (presumably via the growth categories in the gold card assignment system — see the online supplement for more details). However this rate jumps up at the placement threshold from less than 0.05 to 0.60.¹⁴ This figure thus suggests that approximately 40 percent of students scored just above the gold card threshold on both the ELA and mathematics CSTs, but failed to receive a gold card, presumably because they scored below the threshold on science or social studies CSTs, for which we have limited data (see the online supplement for additional discussion). Nonetheless, Figure 3 clearly demonstrates that a pronounced discontinuity exists in student ID card assignment at the proficiency threshold. Consistent with the assumption that students' locations just above and below the proficiency threshold are random, supplementary analyses indicate that there are no threats to validity due to discontinuities in other student characteristics such as race, race, prior grades, or school attrition.¹⁵

Our analyses exploit this discontinuity in the ID card assignment process to estimate the effect of receiving a low-status white ID card on student achievement. Intuitively, our analysis assumes that in the absence of the ID card program, there would be a continuous (if not necessarily linear) relationship between prior CST scores and later achievement for students across the prior-CST distribution. Discontinuities in that relationship at the assignment threshold indicate that ID card placement has an effect on student achievement (Imbens and Lemieux 2008; Reardon and Robinson 2012). In settings like this, an RD design can provide causal inferences that are “as good as random assignment” (Lee and Lemieux 2010)



Note: Proportion of students with high status (gold or platinum) cards, by students' lowest prior year CST (recentered around the proficiency threshold).

Figure 3: Discontinuity in ID card placement, 2011-12 Live Oak and Mann High School students.

regarding the effects of being on either side of an arbitrary test score threshold that sorts students into categories imbued with social meaning. The RD design is implemented by estimating equations of the following general form:

$$Y_{ist} = \alpha + \beta I(CST_{ist-1} < 0) + f(CST_{ist-1}) + \varepsilon_{ist} \quad (2)$$

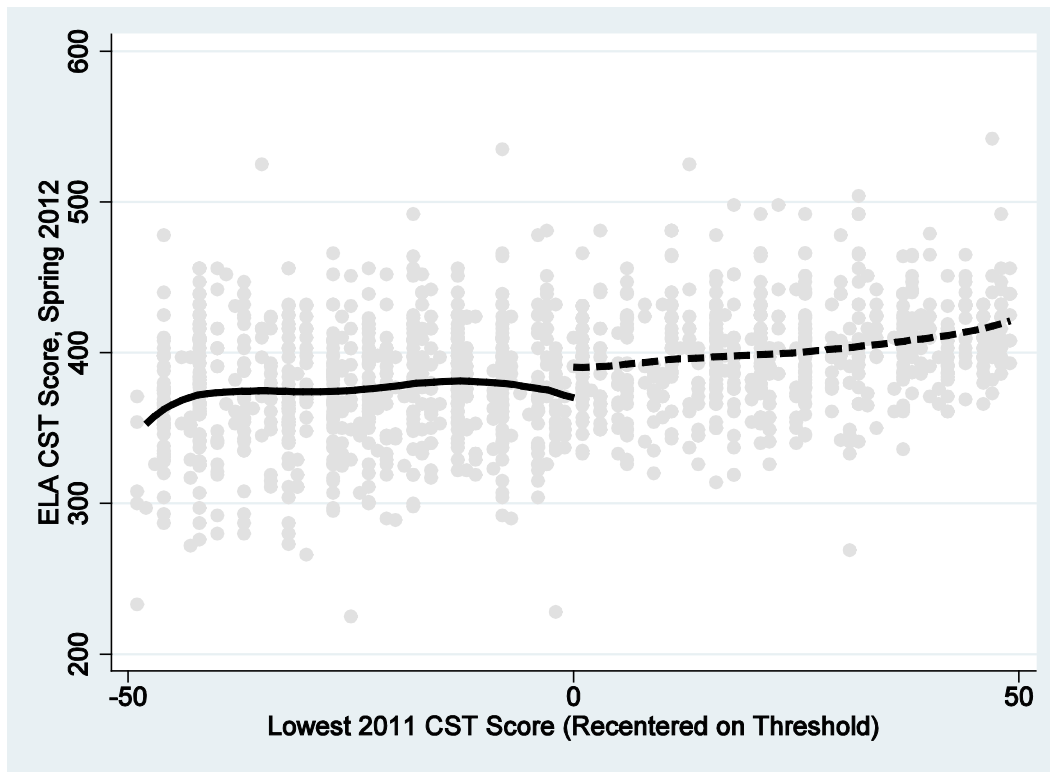
where Y_{ist} is a student outcome (e.g., 2012 CST) for student i in school s in year t . The variable, CST_{ist-1} , is the "assignment variable" in this RD design: the lowest 2011 CST centered on the proficiency threshold. The parameter of interest, β , identifies the jump in outcomes when 2011 CST is below the proficiency threshold, conditional on $f(CST_{ist-1})$, a function of the assignment variable (which we estimate using local linear regressions). We use a binding score regression discontinuity (Reardon and Robinson 2012) because the 2011 CST test on which students receive the *lowest* score determines their eligibility for a white versus gold card.

Given that an RD design leverages comparisons between students that are relatively near to the assignment cutoff, we exclude from these analyses students whose lowest score is greater than 50 points (approximately 0.8 standard deviations) above or below the white ID card assignment threshold.

In addition, the analyses exclude approximately 100 students who fell under a little-used growth clause in the program, which allowed students to earn a gold card by improving their performance by at least one band on two or more CSTs, without declining on any CSTs (see online supplement for details). These students are excluded from the analyses since they were not receiving ID cards based on whether their lowest CST score was above or below the proficiency threshold. Within these constraints, the analyses include all Live Oak and Mann students who have data on the relevant outcomes. As a result, CST analyses include ninth through eleventh grade students; exit exam analyses include tenth grade students only; and grade and suspension analyses include students from grades 9-12. The online supplement provides additional details regarding the regression discontinuity analyses.

Figure 4 provides a visual representation of the regression discontinuity analyses that we use to estimate the effects of receiving a low-status white ID card (rather than a higher-status gold ID card) for students' spring 2012 ELA CST scores. The solid line in this figure represents the relationship between students' lowest 2011 CST score and their 2012 ELA CST score, for students whose lowest 2011 score is below the threshold for a gold ID card. The dashed line represents the local linear regression estimate of the relationship between students' lowest 2011 CST score and their 2012 ELA CST score for students whose lowest 2011 score is above the gold ID card threshold. The gap between these two lines represents the effect of receiving a white ID card on students' 2012 ELA scores. This gap, which is equal to approximately 20 CST points (0.35 standard deviations) is statistically significant, indicating that receiving a white ID card has a substantial negative effect on student achievement for students just below the cutoff.

Table 6 summarizes the results of our regression discontinuity analyses of the effects of receiving a white ID card on Live Oak and Mann students' CST scores, exit exam scores, and grades. Like the estimate of the white ID card effects on ELA CST, our estimate of the effect on ELA CAHSEE scores are also negative and statistically significant, with a standardized effect of -0.67. We also find evidence suggesting that the white ID card had large negative effects on Live Oak and Mann students' grades. The difference in ELA grades between students who received a white ID card and students who received a gold ID card is statistically significant and equivalent to the difference between a C+ average and a D average. The findings on math grades are somewhat smaller, but still statistically significant and large, representing roughly the difference between a C- and a D. Supplemental analyses using alternative bandwidths and estimation strategies reported in the online supplement, suggest that the estimates of the effects of ID card assignment on student grades are sensitive to model specification, and that the effects for ELA and math grades are perhaps closer to 0.8 and 0.6, respectively. We suspect that white ID card receipt might exert a larger negative effect on grades than test scores because it alters both students' behavior and potentially changes the way teachers view and thus grade students. However, the fact that the negative effects of white ID cards occur on ELA CST and CAHSEE scores as well as grades suggests that receiving a white ID card discourages students and changes their learning behaviors, and that the grade effects are not occurring because of teacher bias alone.



Note: 2012 ELA CST scores, by students' lowest prior year CST (recentered around the proficiency threshold).

Figure 4: Regression discontinuity analysis of the effect of white card placement on 2012 ELA CST scores, 2011-12 Live Oak and Mann students.

While not statistically significant, we also see a small change in student suspension rates at the ID card assignment threshold. Students whose scores are just short of a higher-status gold ID card spend approximately 3 more course periods per year in suspension (including in-school and out-of-school) than students whose scores are just over that threshold. This coefficient is very imprecisely estimated, in part because the distribution of suspensions is highly non-normal. More than half of Live Oak and Mann students receive no suspensions in a given year, while nearly 10 percent spend multiple school days in suspension.¹⁶ Nonetheless, this analysis provides suggestive evidence regarding the effects of a low-status white ID card in non-academic realms.

One possible confounding explanation for these findings is that they reflect the effects of receiving a performance band identification of basic (which results in a white card) rather than receiving a white ID card to publicly mark this score. To examine this concern, we estimated the same RD specifications in Sudden Valley schools that did *not* participate in the ID card program. Importantly, we see no effects of being above and below this cutoff in schools that do not assign color-coded ID cards based on this threshold. While all students in California receive information about their scores, as well as their score category (i.e. "basic" vs. "proficient"), simply knowing their score and proficiency band is not sufficient to

Table 6: Regression discontinuity estimates of the effects of white vs. gold ID card eligibility on student achievement and behavior, 2011-12 Live Oak and Mann students

	Reduced form point estimate	Standardized effect	First Stage	N
ELA CST (Bandwidth=13.3)	-20.42 [†] (-2.67)	-0.35	0.60 [†]	1,328
Math CST (Bandwidth=12.3)	-7.57 (-0.58)	-0.10	0.61 [†]	1,321
ELA exit exam (Bandwidth=9.3)	-22.17*	-0.67	0.68 [†]	395
Math exit exam (Bandwidth=10.0)	5.20 (0.53)	0.14	0.70 [†]	398
ELA grades (Bandwidth=2.4)	-1.56 [†] (-2.82)	-1.46	0.55*	1,803
Math grades (Bandwidth=3.5)	-0.84* (-2.09)	-0.80	0.52 [†] *	1,654
Suspensions (Bandwidth=5.9)	2.96 (1.43)	0.26	0.58 [†]	1,805

Notes: Regression discontinuity models estimated via local linear regression using a triangle kernel and optimal bandwidths (Imbens and Kalyanaraman, 2011). Standardized effects report the reduced form point estimate in terms of the 2012 Sudden Valley Unified standard deviation for the outcome. CST data are available for students in grades 9-11; exit exam data are available only for tenth grade students; grades and suspensions are available for all ninth through twelfth grade students. Z-statistics from standard errors clustered at the school-level are reported in parentheses.

* $p < 0.05$; † $p < 0.01$.

create the effects that we find in Table 6. Rather, these results are consistent with the hypothesis that these categories need to be infused with social status and made public to have the impact that we observe.¹⁷

Discussion

Prize systems are an increasingly important organizing force in many social realms, including contemporary American schools. As students progress through elementary and high school, they compete in a wide range of competitions ranging from the relatively trivial (spelling bees, athletic field days, homecoming court elections) to the highly consequential (academic course placements). At the peak of this interlocking system of competitions sits college admissions, a system in which panels of judges confer extremely desirable and highly stratified status rewards upon select students (Stevens 2009; Weis, Cipollone, and Jenkins 2014). Contemporary educational accountability policies construct similar prizes for schools and teachers, conferring honor and shame on educators based on their students' performance. These status incentives are powerful. One principal describes the humiliation asso-

ciated with leading a “failing” school to “having leprosy in the Bible” (Aviv 2014: 6); in another highly publicized case, a teacher committed suicide after being identified as a “least effective teacher” (Felch, Song, and Smith 2010; Lovett 2010).

The Live Oak and Mann color-coded ID card program’s attempt to better align student incentives with the accountability pressure facing educators provides a powerful case study for understanding how prize systems influence behavior and create inequality. We argue that prize systems like the Live Oak and Mann High School color-coded ID card program operate through two mechanisms. First, prizes act as incentives. By articulating criteria and allocating rewards accordingly, prizes attempt to define excellence in their field and encourage actors to pursue it. Second, prizes convey distinction upon winners. In doing so, they create new social categories and opportunities for the construction of categorical inequalities.

Our analyses indicate that both mechanisms are important to understanding the color-coded ID card program and its consequences. Exposure to the Live Oak and Mann color-coded ID card program’s incentives corresponded with modest, concentrated improvements in student achievement. Specifically, we show that the ID card program boosted average student performance on end-of-course CSTs by attaching incentives to this achievement measure. These gains represent a success for the creators of the ID card policy, whose primary interest was increasing average student achievement on the CST. However, we find less evidence to suggest that these gains spilled over to other measures of student achievement. Furthermore, although we do not find consistently statistically significant differences in the effects of ID program exposure based on students’ prior test scores, the pattern of results in these analyses suggests that the policy’s impact may have been driven by students who judged themselves most likely to gain from the prize system.

In sum, these findings are broadly consistent with prior research in the rational choice tradition that documents the ways in which students’ educational behaviors respond to incentives. While several prior studies have examined student reactions to pecuniary and other tangible incentives, ours is among the first to document the powerful effects that status-based incentives have on student behavior. Our findings indicate that students respond to status incentives in much the same way as other incentives as they develop their educational aspirations and allocate time and energy. As in other incentive structures, students who were most likely to reap rewards from the ID card prize system seemed to respond most strongly (Bénabou and Tirole 2003).

However, our analyses also reveal that the program exacerbated inequality. While we find no evidence to suggest that the program *depresses* Hispanic student achievement, we do find evidence suggesting that some differences between ethnic groups widened. To the degree that Hispanic students benefited less than Asian students, the ID card program may have exacerbated existing inequalities between racial groups at Live Oak and Mann.

But most strikingly, we find that the ID card program created powerful new categorical inequalities among students at Live Oak and Mann. Our regression discontinuity analyses provide strong evidence that receiving a low-status white ID card has detrimental effects on a wide range of student outcomes, including student scores on the tests that are central to the ID card assignment process as

well as student grades. Further, we find some evidence suggesting that white card assignment might increase disciplinary suspensions, although this estimate is imprecise and not statistically significant.

Bourdieu's notion of a "rite of institution" helps to explain why this prize structure might create new inequalities. Much like a dubbing ceremony for the initiation of a knight, the color-coded ID card program is "an act of communication, but of a particular kind: it signifies to someone what his identity is, but in a way that both expresses it to him and imposes it on him by expressing it in front of everyone" (Bourdieu 1991: 121). Consistent with Bourdieu's description of a rite of institution, we find that the color-coded ID cards communicated identities both to ID recipients themselves and to the broader school communities. Receiving a low-status ID card negatively influences students' achievement (as is apparent in the negative effects of white card receipt on standardized test scores) as well as their teachers' assessments (as is apparent in the negative effects of white card receipt on grades.)

It is tempting to describe the categorical inequalities that the ID card program created as unfortunate unintended consequences. Indeed, administrators at Live Oak and Mann utilize the language of incentives exclusively. It certainly was not their intent to activate existing social categories or create new ones. But we argue that it is impossible to separate the program's incentive effects from its categorical inequality effects. Prizes like the ID card program work to motivate students only to the extent to which they manage to create privileged social categories. Put differently, if students ignored the distinctions between white, gold, and platinum ID cards – or if they prefer low-status white ID cards over higher status gold and platinum ID cards – the program would not have its intended effects on student behavior. The ID card program attempted to establish a new, purely meritocratic status hierarchy in these schools. However, given ethnic and socio-economic inequalities in gold and platinum ID card receipt, it seems likely that the program was both legitimated by and served to further legitimate existing educational inequalities.

Although the ID card program is distinctive in many ways, we suspect that similar links between incentives and categorical inequality hold in a wide array of social settings. Like the ID card program, all prize structures operate by naming winners and conferring status advantages upon them. Since these status advantages are positional – prize winners gain status at the cost of losers – mutual dependence between incentives and categorical inequality seems inherent. The link between status rewards and inequality production is particularly clear in educational settings, where formal status-based competition – for high-status track placements, selective university admissions, and competitive scholarships – is a central organizing principle. But this link likely holds in any setting in which there are status-related competitions, including labor markets, honorary societies, and neighborhoods.

Notes

- 1 Interestingly, some research indicates that incentives that reward academically desirable behaviors (such as reading books and attending school) are more effective than awards that are tied directly to test scores (Fryer 2011), a finding that is congruent with the

notion that students may not fully understand the link between effort and academic achievement (cf. Rosenbaum 2001).

- 2 In agreeing to provide student records for an evaluation of this program, “Sudden Valley Unified” district officials requested that we withhold any information that would make it possible to identify the district or its schools, and that we not contact Sudden Valley High School students or employees about the program. We thus draw extensively on several published sources, including local and student newspapers and school administrator memoirs to provide important background information regarding the ID card program and its operation. This quotation and others from Sudden Valley educators, students, and parents are paraphrased from them. Since references to these public sources would compromise our agreement with the district, we do not cite them.
- 3 A parallel literature suggests that women may also encounter stereotype threat, particularly in mathematics. We find no evidence of gender differences in the increases associated with exposure to the ID card program in our analyses. This is perhaps not surprising as there are no clear gender achievement gaps at Live Oak and Mann.
- 4 Students typically take math and ELA tests in each of these years and take social studies and sciences tests in some, but not necessarily all of these years depending on their course sequence (California Department of Education, 2014).
- 5 We would be particularly concerned, for example, if we saw that Live Oak and Mann scored anomalously low in the 2008-2009 school year (or if the control schools scored anomalously high in that year).
- 6 We consider the 2009-2010 year as a transitional year and the 2010-2011 and 2011-2012 school years as the years in which the program was fully implemented. Student end-of-course test scores in the 2009-2010 year determined student ID cards in the following year, even though the ID card program was not yet in existence when students took that test. We do not know the extent to which students knew about or understood the program at the time of testing.
- 7 To interpret β_6 and β_7 as unbiased causal estimates of the ID card program’s effect on educational outcomes, one must assume that the implementation of the ID card program is the only important change that occurred uniquely in Live Oak and Mann during the treatment years, conditional on controls. This assumption is restrictive. While we are able to control for some major demographic changes, we have limited data on students’ socioeconomic backgrounds and less observable data on changes in school administration. We know that Live Oak and Mann had no administrative turnover during the period under study. Supplemental analyses further indicate that teacher turnover rates varied little across Sudden Valley schools and over time. Furthermore, most decisions regarding textbooks, teacher training, and course placement practices were made at the district level and we are thus confident that all Sudden Valley schools followed very similar curricular policy trajectories during the study period. However, we lack comprehensive data on administrator turnover or teacher assignments at control schools. Given these threats to validity, we avoid strong causal language in discussing the findings of these difference-in-difference analyses. However, we note that any potentially confounding factor must (a) co-occur with program implementation, (b) occur uniquely in the program schools, and (c) be imperfectly correlated with shifts on the measured control variables.
- 8 Supplementary analyses indicate that ELA CST scores correlate with course grades in the 0.5 to 0.6 range. However, we find no evidence to suggest that this correlation changed over time, either in the schools that implemented the ID card program or in other Sudden Valley Unified high schools. Thus, we do not think that the differences observed in grades are driven solely by teachers using CST scores to determine grades.

- 9 While all tenth grade students take the same ELA and math CAHSEE tests, their CSTs and grades are specific to the courses that they are taking. This is particularly important in math, where students are exposed to a strongly tracked sequence of courses. It is thus noteworthy that we see large differences in math CST scores, but not CAHSEE scores, as we might expect the CST scores to be somewhat noisier than the CAHSEE scores.
- 10 The interaction effects reported in Table 5 show that students near the threshold experienced larger test score increases than students well above the threshold. However, when we compare students near the threshold to those well below the threshold, we find statistically significant differences only for the 2011 ELA CST scores; in other cases the differences between the interaction effects are smaller and not statistically significant. See additional analyses reported in the supplement, Table 5.
- 11 The logic of hypothesis 3 also suggests that the effects of the ID card program may vary with student class. However, we lack a reliable indicator of family class or socio-economic status.
- 12 Supplement Table 5 indicates the sum of the ID*program year and the ID*Hispanic*program year coefficients is not significantly different from zero.
- 13 We computed this index variable in two steps: (1) We recentered students' 2011 math and ELA test scores around their relevant proficiency threshold. (For example, we subtracted 350 from the ELA scale scores for students who took the ninth grade ELA CST to create this recentered ELA CST score variable.) (2) We then computed the index variable as the lower of students' recentered 2011 ELA and mathematics scores.
- 14 These first-stage discontinuities in card assignment probabilities vary across the analyses presented in Table 6 because the sample of students contributing data on each of the outcomes varies. This is primarily a function of students taking a different number of CSTs (of varying difficulties) in different grades (see the online supplement for additional discussion).
- 15 The sole exception is gender. However, as noted in the supplement, additional analyses find that our results are robust to the inclusion of gender as a control variable, and we find no evidence that the effects of receiving a white card vary by gender.
- 16 Results from models using an inverse hyperbolic sine transformation to account for this distribution yield similar results.
- 17 Further robustness checks reveal that there are no regression discontinuity effects at other theoretically inconsequential thresholds (e.g. 10 points above or below the proficiency threshold). While we do find some evidence of similar treatment effects at the platinum ID card threshold, these effects are imprecisely estimated due to the relatively small number of students at risk for this ID card.

References

- Alba, R. (2005). Bright vs. Blurred Boundaries: Second-Generation Assimilation and Exclusion in France, Germany, and the United States. *Ethnic and Racial Studies*, 28(1), 20-49. <http://dx.doi.org/10.1080/0141987042000280003>
- Allan, B. M., and Fryer, R. G. (2011). *The Power and Pitfalls of Education Incentives*. Brookings Institution, Hamilton Project.
- Altonji, J. G. (1993). "The Demand for and Return to Education When Education Outcomes Are Uncertain." *Journal of Labor Economics*, 11(1), 48-83. <http://dx.doi.org/10.1086/298317>

- Anand, N. and Mary R. Watson. 2004. "Tournament Rituals in the Evolution of Fields: The Case of the Grammy Awards." *Academy of Management Journal* 47:59–80. <http://dx.doi.org/10.2307/20159560>
- Angrist, Joshua and Victor Lavy. 2009. "The Effects of High-Stakes High School Achievement Awards: Evidence from a randomized trial." *The American Economic Review*, 1384-1414. <http://dx.doi.org/10.1257/aer.99.4.1384>
- Aronson, Joshua, Michael J. Lustina, Catherine Good, Kelli Keough, Claude M. Steele, and Joseph Brown. 1999. "When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat." *Journal of Experimental Social Psychology*, 35(1), 29-46. <http://dx.doi.org/10.1006/jesp.1998.1371>
- Ashforth, Blake E., and Fred Mael. 1989. "Social Identity Theory and the Organization." *Academy of Management Review*, 14(1), 20-39.
- Aviv, Rachel. 2014. "Wrong Answer: In an Era of High-Stakes Testing, a Struggling School Made a Shocking Choice." *The New Yorker*. June 21, 2014.
- Ball, Sheryl, Catherine Eckel, Philip J. Grossman, and William Zame. 2001. "Status In Markets." *The Quarterly Journal of Economics*, 116(1), 161-188. <http://dx.doi.org/10.1162/003355301556374>
- Beattie, Irene R. 2002. "Are all 'Adolescent Econometricians' Created Equal? Racial, Class, and Gender Differences in College Enrollment." *Sociology of Education*, 19-43. <http://dx.doi.org/10.2307/3090252>
- Ben-Zeev, Talia, Steven Fein, and Michael Inzlicht. 2005. "Arousal and Stereotype Threat." *Journal of Experimental Social Psychology*, 41(2), 174-181. <http://dx.doi.org/10.1016/j.jesp.2003.11.007>
- Bénabou, Roland and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *The Review of Economic Studies*, 70(3), 489–520. <http://dx.doi.org/10.1111/1467-937X.00253>
- Behrman, Jere R., Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin. 2015. "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools." *Journal of Political Economy*, 123(2), 325-364. <http://dx.doi.org/10.1086/675910>
- Best, Joel. 2008. "Prize Proliferation." *Sociological Forum*, 23:1–27 <http://dx.doi.org/10.1111/j.1573-7861.2007.00056.x>
- Bishop, John. 2006. "Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn—Externalities, Information Problems and Peer Pressure." *Handbook of the Economics of Education*, 2, 909–944.
- Blume, Howard. 2012. "Posted Images of California School Tests Raise Cheating Concerns." April 28, 2012 *Los Angeles Times*. <http://articles.latimes.com/2012/apr/28/local/la-me-0428-test-cheating-20120428>
- Booher-Jennings, Jennifer. 2005. "Below the Bubble: 'Educational Triage' and the Texas Accountability System." *American Educational Research Journal*, 42(2), 231-268. <http://dx.doi.org/10.3102/00028312042002231>
- Boudreau, Kevin J., Nicola Lacetera and Karim R. Lakhani. 2011. "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis." *Management Science*, 57(5), 843-863. <http://dx.doi.org/10.1287/mnsc.1110.1322>
- Bourdieu, Pierre. 1991. *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.
- Braun, Henry, Irwin Kirsch and Kentaro Yamamoto. 2011. "An Experimental Study of the Effects of Monetary Incentives on Performance on the 12th-Grade NAEP Reading Assessment." *Teachers College Record*, 113(11), 2309-2344.

- Breen, Richard and Meirr Yaish. 2006. "Testing the Breen–Goldthorpe Model of Educational Decision Making." *Mobility and inequality*, 232-258.
- Brown, Jennifer. 2011. "Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars." *Journal of Political Economy*, 119(5), 982-1013. <http://dx.doi.org/10.1086/663306>
- California Department of Education. (2014). STAR CST Blueprints — Standardized Testing and Reporting (STAR) (CA Dept of Education). Retrieved March 6, 2015, from <http://www.cde.ca.gov/ta/tg/sr/blueprints.asp>
- Croizet, Jean-Claude, and Theresa Claire. 1998. "Extending the Concept of Stereotype Threat to Social Class: The Intellectual Underperformance of Students from Low Socioeconomic Backgrounds." *Personality and Social Psychology Bulletin*, 24(6), 588-594. <http://dx.doi.org/10.1177/0146167298246003>
- Crosnoe, Robert. (2011). *Fitting in, Standing out: Navigating the Social Challenges of High School to Get an Education*. Cambridge University Press. <http://dx.doi.org/10.1017/cbo9780511793264>
- Davies, Paul G., Steven J. Spencer, and Claude M. Steele. 2005. Clearing the Air: Identity Safety Moderates the Effects of Stereotype Threat on Women's Leadership Aspirations. *Journal of Personality and Social Psychology*, 88(2), 276. <http://dx.doi.org/10.1037/0022-3514.88.2.276>
- Davies, Paul G., Steven J. Spencer, Diane M. Quinn, and Rebecca Gerhardtstein. 2002. "Consuming Images: How Television Commercials that Elicit Stereotype Threat can Restrain Women Academically and Professionally." *Personality and Social Psychology Bulletin*, 28(12), 1615-1628. <http://dx.doi.org/10.1177/014616702237644>
- Deci, Edward L., Richard Koestner, and Richard M. Ryan. 1999. A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125(6), 627. <http://dx.doi.org/10.1037/0033-2909.125.6.627>
- Dee, Thomas S. and James Wyckoff. 2015. Incentives, Selection, and Teacher Performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*. <http://onlinelibrary.wiley.com/doi/10.1002/pam.21818/pdf>
- Eckert, Penelope. 1989. *Jocks and Burnouts: Social Categories and Identity in the High School*. Teachers College Press.
- Eden, Dov, and Gad Ravid. 1982. Pygmalion versus Self-Expectancy: Effects of Instructor-and Self-Expectancy on Trainee Performance. *Organizational Behavior and Human Performance*, 30(3), 351-364. [http://dx.doi.org/10.1016/0030-5073\(82\)90225-2](http://dx.doi.org/10.1016/0030-5073(82)90225-2)
- Ehrenberg, Ronald G. and Michael L. Bognanno. 1990. The Incentive Effects of Tournaments Revisited: Evidence from the European PGA tour. *Industrial and Labor Relations Review*, 74S-88S. <http://dx.doi.org/10.1177/001979399004300305>
- Entwisle, Doris R. and Murray Webster, M. 1978. Raising Expectations Indirectly. *Social Forces*, 57(1), 257-264. <http://dx.doi.org/10.1093/sf/57.1.257>
- Felch, Jason, Jason Song, and Doug Smith. 2010. "Who's Teaching L.A.'s Kids?" *Los Angeles Times*. August 14, 2010. Retrieved August 21, 2014. <http://www.latimes.com/news/local/la-me-teachers-value-20100815,0,2695044.story>
- Frey, Bruno S. 2006. "Giving and Receiving Awards." *Perspectives on Psychological Science* 1:377-88. <http://dx.doi.org/10.1111/j.1745-6916.2006.00022.x>
- Goette, Lorenz, David Huffman, and Stephan Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to

- Real Social Groups." *The American Economic Review*, 212-216. <http://dx.doi.org/10.1257/000282806777211658>
- Goode, William J. 1978. *The Celebration of Heroes: Prestige as a Social Control System*. Berkeley: University of California Press.
- Hallett, Tim. 2010. "The Myth Incarnate: Recoupling Processes, Turmoil, and Inhabited Institutions in an Urban Elementary School." *American Sociological Review*, 75(1), 52-74. <http://dx.doi.org/10.1177/0003122409357044>
- Hamilton, Laura S., Brian M. Stecher, Julie A. Marsh, Jennifer Sloan McCombs, and Abby Robyn. 2007. *Standards-Based Accountability Under No Child Left Behind: Experiences of Teachers and Administrators in Three States*. Rand Corporation.
- Harvey, Oliver J., B. Jack White, William R. Hood, and Caroline W. Sherif. 1961. *Intergroup Conflict and Cooperation: The Robbers Cave Experiment* (Vol. 10). Norman, OK: University Book Exchange.
- Imbens, Guido W., and Kalyanaraman, K. (2011). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, rdr043.
- Imbens, Guido W., and Thomas Lemieux. (2008). "The Regression Discontinuity Design: Theory and Applications." *f*, 142 (2).
- Inzlicht, Michael and Talia Ben-Zeev. 2000. "A Threatening Environment: Why Females are Susceptible to Experiencing Problem-Solving Deficits in the Presence of Males." *Psychological Science* 11:365-371. <http://dx.doi.org/10.1111/1467-9280.00272>
- Jackson, C. Kirabo. 2010. "A Little Now for a Lot Later: A Look at a Texas Advanced Placement Incentive Program." *The Journal of Human Resources*, 45(3), 591-639. <http://dx.doi.org/10.3368/jhr.45.3.591>
- Jackson, C. Kirabo, Jonah E. Rockoff, and Douglas O. Staiger. 2014. "Teacher Effects and Teacher-Related Policies." *Annual Review of Economics*, 6, 801-825. <http://dx.doi.org/10.1146/annurev-economics-080213-040845>
- Jaeger, Mads Meier. 2007. Economic and Social Returns to Educational Choices Extending the Utility Function. *Rationality and Society*, 19(4), 451-483. <http://dx.doi.org/10.1177/1043463107083739>
- Jiménez, Tomás R., and Adam L. Horowitz. 2013. "When White Is Just Alright: How Immigrants Redefine Achievement and Reconfigure the Ethnoracial Hierarchy." *American Sociological Review* 78: 849-871. <http://dx.doi.org/10.1177/0003122413497012>
- Kalkhoff, Will, and Shane R. Thye. 2006. "Expectation States Theory and Research: New Observations from Meta-Analysis." *Sociological Methods & Research*, 35(2), 219-249. <http://dx.doi.org/10.1177/0049124106290311>
- Kerckhoff, Alan C. 1995. Institutional Arrangements and Stratification Processes in Industrial Societies. *Annual Review of Sociology*, 323-347. <http://dx.doi.org/10.1146/annurev.so.21.080195.001543>
- Khan, Shamus R. 2011. *Privilege: The Making of an Adolescent Elite at St. Paul's School*. Princeton, NJ: Princeton University Press.
- Kohn, Alfie. 1999. *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. New York: Houghton Mifflin Harcourt.
- Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48, 281-355. <http://dx.doi.org/10.1257/jel.48.2.281>
- Lovaglia, Michael J., Jeffrey W. Lucas, Shane R. Thye, and Barry Markovsky. 1998. "Status Processes and Mental Ability Test Scores." *American Journal of Sociology* 104:195-228. <http://dx.doi.org/10.1086/210006>

- Lovett, Ian. 2010. "Teacher's Death Exposes Tensions in Los Angeles." *New York Times* November 9, 2010: A23.
- Maehr, Martin L. and Carol Midgley. 1996. *Transforming School Cultures*. Westview Press.
- Manski, Charles F. 1993. Adolescent Econometricians: How do Youth Infer the Returns to Schooling?. Pp. 43-60 in *Studies of Supply and Demand in Higher Education*. University of Chicago Press.
- McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2), 698-714. <http://dx.doi.org/10.1016/j.jeconom.2007.05.005>
- McGuinn, Patrick J. 2006. *No Child Left Behind and the Transformation of Federal Education Policy, 1965-2005*. Lawrence: University Press of Kansas.
- Merton, Robert K. 1968. "The Matthew Effect in Science." *Science*. 159:56-63 <http://dx.doi.org/10.1126/science.159.3810.56>
- Meyer, John W., and Brian Rowan. 1975. *Notes on the Structure of Educational Organizations*. Stanford Center for Research and Development in Teaching.
- Morgan, Steven L. 2005. *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford University Press.
- O'Brien, Laurie T. and Christian S. Crandall. 2003. Stereotype Threat and Arousal: Effects on Women's Math Performance. *Personality and Social Psychology Bulletin*, 29(6), 782-789. <http://dx.doi.org/10.1177/0146167203029006010>
- Papay, J.P., Murnane, R.J., and Willett, J.B. (forthcoming). How Performance Information Affects Human-Capital Investment Decisions: The Impact of Test-Score Labels on Educational Outcomes. *Journal of Human Resources*.
- Peters, William. 1971. *A Class Divided*. Doubleday, Garden City, NY.
- Reardon, Sean F. and Joseph P. Robinson. 2012. Regression Discontinuity Designs with Multiple Rating-Score Variables. *Journal of Research on Educational Effectiveness*, 5(1), 83-104. <http://dx.doi.org/10.1080/19345747.2011.609583>
- Ridgeway, Cecilia L. 2014. Why Status Matters for Inequality. *American Sociological Review*, 79(1), 1-16. <http://dx.doi.org/10.1177/0003122413515997>
- Ridgeway, Cecilia L., and Shelley J. Correll. 2006. Consensus and the Creation of Status Beliefs. *Social Forces*, 85(1), 431-453. <http://dx.doi.org/10.1353/sof.2006.0139>
- Rhodes, Jessie. 2012. *An Education in Politics: The Origins and Evolution of No Child Left Behind*. Cornell University Press.
- Rosenbaum, James E. 2001. *Beyond College for All: Career Paths for the Forgotten Half*. Russell Sage Foundation.
- Rossman, Gabriel and Oliver Schilke. 2014. Close, But No Cigar: The Bimodal Rewards to Prize-Seeking. *American Sociological Review*, 79(1), 86-108. <http://dx.doi.org/10.1177/0003122413516342>
- Spencer, Steven J., Claude M. Steele, and Diane M. Quinn. 1999. Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, 35(1), 4-28. <http://dx.doi.org/10.1006/jesp.1998.1373>
- Springer, Matthew G. (Ed.). 2009. *Performance Incentives: Their Growing Impact on American K-12 Education*. Brookings Institution Press.
- Steele, Claude M., and Joshua A. Aronson. 1995. Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797. <http://dx.doi.org/10.1037/0022-3514.69.5.797>

- Stevens, Mitchell L. 2009. *Creating a Class: College Admissions and the Education of Elites*. Harvard University Press.
- Stoet, Gijsbert and David C. Geary. 2012. Can Stereotype Threat Explain the Gender Gap in Mathematics Performance and Achievement? *Review of General Psychology*, 16(1), 93. <http://dx.doi.org/10.1037/a0026617>
- Taylor, Kate. 2015. "At Success Academy Schools, High Scores and Polarizing Tactics." *New York Times*. April 7, 2015. A1.
- Thistlethwaite, D. L., and Campbell, D. T. (1960). Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment. *Journal of Educational Psychology*, 51(6), 309. <http://dx.doi.org/10.1037/h0044319>
- Tilly, C. (1999). *Durable Inequality*. Berkeley, CA: University of California Press.
- Tran, A. and Zeckhauser, R. (2012). Rank as an Inherent Incentive: Evidence from a Field Experiment. *Journal of Public Economics*, 96, 645-650. <http://dx.doi.org/10.1016/j.jpubeco.2012.05.004>
- Turner, R. H. (1960). Sponsored and Contest Mobility and the School System. *American Sociological Review*, 855-867. <http://dx.doi.org/10.2307/2089982>
- Walen, Alanna. 2012. "How do Students Describe the Effects of the Jostens Renaissance Rewards and Recognition Program on their Academic Achievement?" Masters Thesis, Hamline University, St. Paul, MN.
- Walton, G. M., and Cohen, G. L. (2003). Stereotype Lift. *Journal of Experimental Social Psychology*, 39(5), 456-467. [http://dx.doi.org/10.1016/S0022-1031\(03\)00019-2](http://dx.doi.org/10.1016/S0022-1031(03)00019-2)
- Wang, M. T., and Eccles, J. S. (2013). School Context, Achievement Motivation, and Academic Engagement: A Longitudinal Study of School Engagement Using a Multidimensional Perspective. *Learning and Instruction*, 28, 12-23. <http://dx.doi.org/10.1016/j.learninstruc.2013.04.002>
- Weick, K. E. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly*, 1-19. <http://dx.doi.org/10.2307/2391875>
- Weis, Lois, Kristin Cipollone, and Heather Jenkins. 2014. *Class Warfare: Class, Race, and College Admissions in Top-tier Secondary Schools*. Chicago: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226135083.001.0001>
- Wise, S. L., and DeMars, C. E. (2005). Low Examinee Effort in Low-Stakes Assessment: Problems and Potential Solutions. *Educational Assessment*, 10(1), 1-17. http://dx.doi.org/10.1207/s15326977ea1001_1
- Xie, Y. (2014). Undemocracy: Inequalities in Science. *Science*, 344 (6186), 809-10. <http://dx.doi.org/10.1126/science.1252743>

Acknowledgements: Research reported in this article was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under Award Numbers P01HD065704 and K01HD073319, by the Institute for Education Sciences under Award Number R305B130017, and by the Spencer Foundation under Award Number 201400180. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Institute for Education Sciences, or the Spencer Foundation. The authors are grateful to Marianne Bitler, Thomas S. Dee, Ken Dodge, Greg Duncan, David Frank, Eric Grodsky, Andrew McEachin, Evan Schofer, Jeff Smith, and participants in the colloquia at Brown University and University of Wisconsin for useful comments and discussions.

Thurston Domina: School of Education, University of North Carolina, Chapel Hill.
E-mail: tdomina@email.unc.edu.

Andrew M. Penner: Department of Sociology, University of California, Irvine.
E-mail: penner@uci.edu.

Emily K. Penner: Center for Education Policy Analysis, Stanford University.
E-mail: epenner@stanford.edu.