INCENTIVE PROVISION IN INVESTMENT CONTESTS: THEORY AND EVIDENCE*

CHRISTOPHER S. COTTON, BRENT R. HICKMAN†, AND JOSEPH P. PRICE

Abstract. We build on recent models of large-scale contests–in which many heterogenous agents compete for heterogenous prizes–to develop novel theoretical predictions concerning incentive provision in scenarios where competitive human capital investment determines rank-order allocations of non-divisible resources such as college seats. We then conduct a field experiment in which hundreds of subjects compete for an array of prizes, and find empirical patterns in competitive learning behaviors that are highly-consistent with the theory. Our findings deliver two key insights. First, when human capital serves a dual role as an intrinsic asset and a rank-order index, investment behavior varies significantly with the strength of one's competition, holding own ability and the set of prizes fixed. Second, if an observable noisy signal of ability is available, we find evidence of a counterintuitive role for preferential treatment in promoting overall investment.

Date: **Original Version: October 2013; Current version: August 2017**.

Key words and phrases. Large Contest, All-Pay Auction, Affirmative Action, College Admissions, Field Experiment, Human Capital.

JEL subject classification: J15, J24, C93, D82, D44.

Cotton: Department of Economics, Queen's University, Dunning Hall 230, Kingston, Ontario K7L 3N6; cotton@econ.queensu.ca; phone: (613) 533-2251.

†Hickman (corresponding author): Department of Economics, Queen's University, Dunning Hall 345, Kingston, Ontario K7L 3N6; hickmanbr@gmail.com; phone: (613) 533-2252.

Price: Department of Economics, Brigham Young University, 162 FOB, Provo, UT 84602; joe_price@byu.edu; phone: (801) 422-5296.

We greatly appreciate comments and helpful input from, Andrew Sweeting, Aaron Bodoh-Creed, John-Eric Humphreys, John List, Lieutenant Colonel Mike Linder (USAF), and participants of seminars at the University of Chicago, the 2013 meeting of the American Economic Association, the 2013 North American summer meetings of the Econometric Society, 2014 Canadian Public Economic Group conference, the 2015 Econometric Society World Congress, the 2015 European Meetings of the Economic Science Association, Brigham Young University, University of Missouri-Columbia, Queen's University, Arizona State Univeristy, Ryerson University, SUNY Binghampton, UBC-Okanagan, University of Guelph, and London School of Economics. We also wish to acknowledge the outstanding research assistance of Joe Patten, who played a crucial role in executing this project. Cotton is grateful for financial support provided by his position as the Jarislowsky-Deutsch Chair in Economic and Financial Policy at Queen's University. Cotton and Price gratefully acknowledge partial funding from the Spencer Foundation which supported an initial pilot study that lead to this research.

***NOTE:* Previous versions of this paper have circulated under the title "*Affirmative Action and Human Capital Investment: Theory and Evidence from a Randomized Field Experiment*."**

1. INTRODUCTION

The economics literature applies contest theory to model a variety of environments involving allocation of fixed, non-divisible resources, from competition across firms for contracts, to competition within a firm for promotion, to competition within a classroom for grades. Most of the literature focuses on settings where agents are homogeneous, or where there are a small number of heterogeneous agents or prizes. However, real-world settings often involve many-to-many rank-order matching between heterogeneous agents and heterogeneous prizes. There is also a vast literature depicting single-agent models of human capital investment. Once again though, many important real-world settings include competitive investment, where agents' expected rewards depend on relative performance, and therefore the returns to one's own investment hinge on choices of one's competitors.

Economically important examples of markets with many-to-many matching and competitive investment abound. The academic arms race among high-school students for university admissions and scholarships is an annual competition for roughly 2 million freshman seats at a diverse set of post-secondary institutions. Throughout high school, students exert effort in their studies and extra curricular activities to acquire abilities and experiences which increase their relative standing during the application process. Similarly, more than 500,000 people take the GRE (Graduate Record Examination) each year, and approximately 100,000 take the LSAT (Law School Admission Test) as they compete for admissions to graduate programs and law schools of varying quality. These exam scores and their GPAs serve as both a reflection of real human capital acquired during college, and as a rank-order index for post-graduate admissions.

Another prominent example is career track assignment in the US armed forces. Each year thousands of people enroll in officer training programs which are designed to teach academic, leadership, and military skills, and one's performance in acquiring and demonstrating these skills contributes to an explicit cumulative score. For example, the US Air Force Academy feeds an annual class of 1000 graduates into dozens of career tracks, the most desirable (and best paid) of which are pilot, intelligence, JAG (law), and medical. For each cohort there are only a limited number of slots open for each track, and a cadet's cumulative score rank determines the order in which he or she may choose. US Air Force Flight School trains a large fraction of the pilots who eventually end up in American commercial aviation, and is run similarly. For the roughly 1000 junior officers who go on to flight school each year, training takes place in two stages where performance on every learning task, both in and out of the classroom, is measured and added to a cumulative score. In the first stage one's score rank determines order of choice for limited slots in different flight tracks, with fighter/bomber being the most highly paid and sought-after track. Scores are then re-set at the beginning of the second stage, and one's final score rank determines order of choice for limited slots pertaining to specific planes which pilots will go on to fly.

Large-scale investment competitions are not limited to academic pursuits and job markets either. Authors choose how much time and effort to invest in their writing as they compete with each other for publication and readership, and performing artists compete similarly for audiences. Even the marriage market may be seen as competition amongst many suitors for spouses, where real-effort investment occurs on traits such as physical condition, education, and/or income. Each of the above examples share three salient features: (1) agent heterogeneity—a wide array of competitors differ in their investment costs; (2) prize heterogeneity—schools, programs, jobs, firms, or match partners differ in value, holding agent characteristics fixed; and (3) a dual role for investment—producing both direct, intrinsic returns to the investor, and indirect returns by controlling his or her desirability to the other side of the matching market.

We combine these three features into a theoretical framework of a rank-order competitive investment game between many heterogenous agents for many heterogeneous prizes. The model is adapted from a more general theoretical framework developed in Aaron Bodoh-Creed and Brent R. Hickman [2017], with our focus being on a simplified version that allows for comparative static predictions. Our goal is to explore how incentives for human capital investment vary with shifts in competition, and how the impact may depend on individual characteristics. In the model, a continuum of agents have a privately-known cost of human capital production θ; there is a fixed continuum of vertically heterogeneous "prizes" P, which we can think of as match partners, jobs, admissions slots, or other outcomes. There is a rank-order mechanism which allocates prizes based on relative, observed human capital production after investment decisions are made.

Our model delivers two novel insights concerning how behavior reacts to changes in the intensity of competition. First, holding fixed an agent's own costs and the set of prizes, a stochastic dominance shift toward lower investment costs among the agent's competitors leads a low-cost agent to invest more aggressively, and a medium- or high-cost agent to reduce investment. In other words, the strongest agents invest more, while weaker agents invest less, as the strength of their competitors increases. We refer to this latter shift in middle- and high-cost agent investments as the *discouragement effect*. Moreover, for any cost type θ in the interior of the support there exists some stochastic shift toward lower cost competitors such that θ will become discouraged and reduce investment. In that sense, the distinction between "low-cost" and "high-cost" agents is relative as well, being a function of one's quantile rank, rather than one's absolute ability level.

This comparative static is interesting because it involves no change to the direct marginal costs or benefits of improving one's human capital. There are two canonical economic models which have been used to explain educational choices: the Becker model [Gary Becker, 1973] which interprets schooling as investment in productive human capital, and the Spence model [Michael Spence, 1973] which interprets schooling as a signaling game to separate agents by their unobservable types in an incentive-compatible way. While consensus among labor economists has generally favored the Becker interpretation, our first main result illustrates how Spence-like competition can still play an economically important role, even when the signaling variable is intrinsically valued by the agent.

Our second theoretical result establishes a somewhat surprising role for preferential treatment in provision of investment incentives. Suppose that there is some observable binary characteristic R that is correlated with unobservable types, so that increasing R from 0 to 1 induces a stochastic dominance shift in the distribution of human capital production costs. For ease of discussion we will refer to the group of agents for whom $R = 1$ as the *disadvantaged group*.[1] Within this context, we consider the effects of a policy that provides preferential treatment to the disadvantaged group: a representative quota, which reserves a representative set of prizes for the disadvantaged group, before the competition begins. This means that disadvantaged agents compete only among themselves, effectively increasing the cost distribution of their competition compared to a situation in which all agents compete against all other agents. A representative quota leads to the majority of disadvantaged agents *increasing* their investment, and to a small share of top agents reducing their investment. Intuitively, preferential treatment can improve average investment of disadvantaged agents by shifting the distribution of competitors and thereby mitigating discouragement effects. Predictions for the $R = 0$ group are in opposite directions, but with theoretical ambiguity about how the average agent in that group will react.

These two theoretical predictions have several policy-relevant implications. First, it has been widely speculated that today's admissions process for elite colleges represents a burdensome academic arms race that imposes unnecessary costs on students and their families. Our first result shows that a shift in the distribution of competitors without any change to market-wide school quality leads elite students to further escalate the arms race. Examples of phenomena which might lead to a shift in the equilibrium distribution of human capital include proliferation of new investment technologies—*e.g.*, advanced-placement courses and extra-curricular academic programs—and universities increasingly marketing their services to top students from abroad as a means of improving revenues.

Second, preferential treatment schemes have become commonplace in many settings, including race-based affirmative action (widespread among most American universities), and income-based preferential admissions rules like the Texas Top 10% program. Traditional wisdom has held that these programs will generally erode incentives by lowering admissions standards. Our second theoretical result largely refutes this idea: a preferential treatment policy which targets disadvantaged students tends to *increase* their effort by placing within reach outcomes which would otherwise be unattainable. This mitigates discouragement effects that arise when disadvantaged agents find themselves too far behind the competitive curve. It is even possible, within the model, to mitigate allocative inequality through a preferential admission policy while at the same time narrowing demographic achievement gaps and increasing average human capital investment, market-wide.

Finally, our second theoretical result tells us something about the potential sources of achievement gaps in primary and secondary education. In the US, there is a vast disparity in public education funding between schools located in poor communities versus schools located in wealthy ones. If we think of the idiosyncratic production technology θ as encapsulating factors both internal (*e.g.*, personal drive) and external (*e.g.*, K-12 education quality) to the individual, then one would expect students in under-funded schools to have systematically higher learning costs. Conventional wisdom has held that (holding internal factors fixed) asymmetric school funding creates achievement gaps in a mechanical way, simply by creating gaps in human capital production technology. However, our

[1] Throughout we assume that cost types exist on a common support but relative masses of high and low costs differ across groups.

second theoretical result illustrates how asymmetric funding and rank-order incentives can have a combined synergistic impact. The comparative static scenario in our theory and experimental design holds gross payoffs and individual production technology fixed, but produces meaningful changes in learning behaviors, simply by altering the distribution of one's competitors. In a full-information world where college admissions decisions could be resolved prior to investment (*e.g.*, at birth), Becker-style investment would serve to equate direct marginal costs and benefits of human capital, and any observed academic gap between rich and poor students would be directly attributable to production costs alone. However, in the actual world with private information on ability, human capital serves a second role as a rank-order index, and *relative differences* in production technology contribute to investment choices as well. In short, our model suggests that academic achievement gaps are endogenous equilibrium objects, and should not be thought of as exogenous or mechanical.

Following the theoretical analysis, several questions remain as to the empirical relevance of these insights. In order for relative investment incentives to be a first-order consideration for social scientists and policy-makers, real-world participants must make complex, strategic labor-leisure decisions consistent with Bayes-Nash equilibrium. Moreover, this must be true at all ability levels, including for middle- and low-ability participants, and not just among the best and brightest. After presenting our theoretical results, we develop an academic field experiment in partnership with real schools to investigate these concerns empirically. Our experiment involves paying large groups of middle-school-aged students based on their relative performance on the American Mathematics Competition 8 (AMC8), a national mathematics exam. Our experiment includes students from two adjacent grades competing in math achievement for fixed monetary payoffs of varying value, where students in the lower grade (having one year less math education) serve as the disadvantaged group. Students are individually randomized into two treatments for our math competition: a control treatment in which both grades compete head-to-head, and a quota treatment, where a proportional set of prizes is reserved for the students in the disadvantaged group. The representative quota entails two primary effects: first, it reserves a more favorable set of prizes for students in the lower grade, relative to what they would receive in the control treatment; second, it directly alters one's set of relevant competitors, since it implies that competition occurs only within one's own group. Therefore, our treatment and control groups can be interpreted as creating a counterfactual scenario which simultaneously tests both of our comparative static results on shifts in competition and preferential treatment.

Because we are interested in how competition shapes incentives to invest real effort building human capital, a typical experiment in which subjects are assigned to a treatment immediately before completing some task is inadequate. Instead, we worked with teachers to incorporate the experiment into students' schooling over a course of two weeks. We assessed two in-class AMC8 exams—a pre-exam to measure baseline proficiency and a post-exam to allocate prizes and measure progress—and we provided the students with an interim investment period in which they could study math problems ahead of the second AMC8 exam. During this period we gave students access to a website we designed with math learning materials. We then tracked their use of this site, including how many people in each group used it, how much time they spent, how many practice questions they attempted, and how they divided their study among different math subjects.

Within the resulting data we find strong evidence in favor of the qualitative predictions of the theory. First, a shift toward higher-cost competitors increases observed effort and measured proficiency progress. Although the highest-performing lower-grade students did decrease performance as predicted by theory, the drop was modest, with the majority of lower-grade students increasing both learning time and test performance. We find that lower-grade students in the treatment group are more than twice as likely to use the practice website, spend almost triple the amount of total time studying, and attempt more practice questions relative to their counterparts in the control group. Moreover, in so doing they measurably improve their test scores as well: on average, the treatment increased test scores among lower-grade students by $1/5^{\text{th}}$ of a standard deviation over the course of the study. One might worry that these gains within the beneficiary group targeted by the treatment come at the cost of commensurate weakening in incentives for other students, but we find no such evidence. Higher-grade students maintain roughly the same average level of investment activity and exam performance across control and treatment, which possibility is also allowed for by the theory. In other words, the strongest behavioral impacts are predominantly positive and concentrated among the more disadvantaged students.

Finally, the theory also makes fine predictions on how the signs and magnitudes of the incentive effects differ by quantiles and demographic groups. We perform a test for these finer predictions, based on flexible, semi-nonparametric quantile functions. We find that the interquartile range of students within the disadvantaged group displays large and statistically significant improvement in achievement under our treatment, whereas there is a small but positive mass of the highest ability types within the same group who reduce their achievement in a statistically detectable way. We also find modest evidence for effects in the opposite direction within the non-beneficiary group—the highest-performing students in the upper grade tend to perform better under the quota treatment, while the lowest-performing students tend to do worse.

It is remarkable the degree to which observed behavioral responses to our exogenous variation conform to the theory, both in their signs and relative magnitudes across the two groups. Disadvantaged students react strongly to the preferential treatment and in a largely positive way because, on average, they begin with high costs of competition and it therefore produces a substantial improvement in their standing within the rank-order investment competition. On the other hand, students from the advantaged group respond weakly because, despite seeing their expected gross payoffs decline, their learning costs were on average low enough to maintain roughly the same relative performance under more anemic incentives. Our treatment had the effect of narrowing achievement gaps by incentivizing disadvantaged individuals to "catch up" to others with lower overall costs.

Our experimental analysis makes several novel contributions. Ours is the first controlled experiment of a large-scale, many-to-many contest involving hundreds of individuals and a diverse set of prizes. Our experimental evidence provides strong support for theoretical predictions on strategic forces that shape human capital choices: our subjects, who range from fifth through eighth grade, respond to changes in relative incentives in a way that is surprisingly consistent with the sophisticated behavior depicted in our game-theoretic model. Additionally, we are the first experimental study to identify the impact of a preferential treatment policy on real learning effort by students. Our experiment took place in a natural classroom and home learning environment, where the free-time/ study-time trade-offs were familiar to test subjects, and mimic scenarios they face when preparing for college in their everyday lives. Although it is not possible to directly measure intrinsic value tied to human capital accumulation, we are still able to establish a link to real investment by directly monitoring time use decisions and changes in math proficiency. In this way, we are able to show that relative incentives like those tied to many real-world examples can have economically meaningful impacts on investment in productive human capital.

The remainder of this paper has the following structure. Section 2 gives an overview of the previous literature and explains how our contribution relates to what has been done before. Section 3 presents a theoretical model of HC investment and university admissions to motivate our experimental design. We also explore model predictions to be tested in our study. Section 4 describes the structure of our field experiment in more detail. Section 5 presents and discusses our experimental results. Section 6 concludes with a discussion of our findings, the inherent limitations within our research design, and directions for future research. An appendix contains mathematical proofs and relevant extensions of the theoretical model. An accompanying online appendix presents additional technical details of our empirical analysis, as well as additional tables and graphs.

2. RELATED LITERATURE

Aaron Bodoh-Creed [2013], Wojciech Olszewski and Ron Siegel [2016], and Bodoh-Creed and Hickman [2017] also develop models of large-scale contests with many heterogeneous agents and prizes. The first two papers have substantially different focus from our analysis. Bodoh-Creed [2013] considers a uniform-price auction with incomplete information, and Olszewski and Siegel [2016] consider when a contest with many heterogeneous agents can be approximated by a model with a single agent with a continuum of possible types.

The third paper, Bodoh-Creed and Hickman [2017], provides the general theoretical framework on which our model is based. The virtue of this framework is that it is rich enough to encompass a large set of possibilities, but the drawback is that with so many moving parts it is difficult to prove sharp comparative statics predictions for behavioral responses to policy change. Therefore, we present a simplified version of the model focusing on aspects of the environment that are most readily testable using experimental methods. Bodoh-Creed and Hickman focus on foundational theory which establishes existence of equilibrium and provides justification for the simple continuum model we adopt.Bodoh-Creed and Hickman also derive general results concerning the properties and design of market mechanisms for college placement. In this paper, our focus is more narrow in one way and more broad in another. First, we restrict attention to experimental comparisons between two simple canonical mechanisms: a pure rank-order rule that assigns more valuable prizes to those who perform best, and a quota that reserves a representative set of prizes for a disadvantaged subgroup. This allows us to assess the real-world applicability of the theoretical predictions on performance, and measure their associated magnitudes. Second, we use a novel experimental design with online monitoring of study effort to explore background phenomena—*e.g.*, labor-leisure trade-offs—on which the theory is silent. This allows us to confirm whether observed changes in performance induced by our treatment are associated with actual learning activities, or something less socially valuable such as exam-day effort or wasteful signaling.

Our paper also contributes to the literature on effort and contest design. Francis Galton [1902] first considered how to divide a fixed pot of prize money between two prizes in a contest, and how the result depends on the number of competitors. More recently, papers have considered the optimal allocation of prizes in game theoretic models of contests [e.g. Benny Moldovanu and Aner Sela, 2001, 2006].[2] Our paper focuses on contests with an exogenous distribution of heterogenous prizes, and considers how the distribution of competitors within a contest affects effort provision and performance. The reason that the distribution of competitor ability (given a set of prizes) affects performance in our environment is similar to the reason that the distribution of prizes (given a set of competitors) affects performance in the other environment. Richard L. Fullerton and R. Preston McAfee [1999] and Yeon-Koo Che and Ian Gale [2003] explore how a contest designer may be able to influence the distribution of competitor ability, but in very different ways than us. None of the earlier work involves large-scale contests, or an environment with many heterogeneous prizes.

Additionally, our paper contributes to the literatures on affirmative action and incentives for learning. In a theoretical model, Stephen Coate and Glenn Loury [1993] show how affirmative action can make high value jobs more-easily attainable and can therefore reduce the incentives for minorities to undertake costly investment. The key difference between our model—with many-to-many matching—and theirs—with one-to-one matching—is that there is no scarcity of high value positions in Coate and Loury's framework. In our large-contest framework, agents compete with one another for a fixed set of heterogeneous prizes. We show that, although a preferential treatment policy will decrease incentives for the highest ability beneficiaries to invest, the opposite effect will predominate for most of them. Although our results are consistent with Coate and Loury [1993] for the highest ability beneficiaries of preferential treatment, our predictions are quite different for the overall beneficiary population. We also test our predictions with an experiment, and find evidence consistent with our theoretical model.

Previous experimental work has also tested the link between affirmative action policies and effort. Examples in the laboratory include Andrew Schotter and Keith Weigelt [1992], where asymmetry was exogenously imposed by researchers assigning cost functions to subjects, and Anat Bracha, Alma Cohen and Lynn Conell-Price [2015] which focused on gender-based asymmetry in quantitative problem solving. Caterina Calsamiglia, Jorg Franke and Pedro Rey-Biel [2013] conducted a related field experiment in which 10-13 year old children compete in Sudoku puzzles in a pair-wise (one-to-one) tournament, with asymmetry stemming from previous exposure to Sudoku. Each of these studies found that affirmative action can increase average performance by disadvantaged players.

[2] Some experimental work explores related concerns in a laboratory [e.g. Timothy N. Cason, William A. Masters and Roman M. Sheremeta, 2010; Roman M. Sheremeta, 2010].

Our analysis offers a number of advantages over previous experimental studies. First, our experiments involve competition between hundreds of heterogenous participants competing for a range of heterogenous prizes. Such a setting more closely mirrors important aspects of the large scale contests implicit in university admissions, hiring, or other settings. Like other papers in the literature, we consider how preferential treatment policies affect the average performance of different groups. Unlike other papers, however, the size of our contests allow us to also consider how preferential treatment policies change the distribution of performance within groups, which is important given that in large scale contests, the theory predicts that some group members will increase, while others decrease, their effort and performance under preferential treatment policies. Second, we are able to observe investment in human capital, rather than just effort during an experimental task. We include an investment period between assignment to a treatment group and our final exam. During this period we monitor student time usage at home in a non-invasive way. Ours is the first paper to do so, and it provides us with a window into individuals' labor-leisure trade-offs in actual learning, rather than focusing solely on in-class effort during a task. Our experimental design allows us to assess the impact of affirmative action on both interim investment *and* final outcomes. Math learning in preparation for the AMC8 exam is close to learning that enables higher performance on college entrance exams, and therefore provides a relevant measure of human capital investment. By working with test subjects' regular teachers, using materials that were already being used as teaching tools in their schools, and allowing for study choices at home, we are the first experiment to create a natural setting in which test subjects are making decisions similar to those which will lead to their ultimate college placement outcomes. Third, we intentionally chose an experimental design to eliminate concerns about stereotype threat—our test subjects were fairly homogeneous culturally and socioeconomically—so as to be able to measure magnitudes of incentive effects in a clean environment.

There is a substantial empirical literature studying preferential treatment regimes in college admissions, especially race-based affirmative action. An early study by William G. Bowen and Derek Bok [1998] quantified the preference given to minority students by admissions officers at elite schools. A lengthy debate in the literature focuses on the mismatch hypothesis, including papers such as Linda Datcher Loury and David Garman [1995], Richard H. Sander [2004], Mark C. Long [2008], Jesse Rothstein and Albert H. Yoon [2008], David L. Chambers, Timothy T. Clydesdale, William C. Kidder and Richard O. Lempert [2005], Peter Arcidiacono, Esteban Aucejo and Joseph Hotz [2016], and Elanor Wiske Dillon and Jeffrey Andrew Smith [2017]. Throughout this literature SAT—"Scholastic Aptitude Test"—scores are used as a proxy for student ability, and assumed to be fixed. However, students' incentives to invest in human capital during middle and high school depend on admission policies they expect to face when applying to college. Test scores are therefore a function of student ability *and* market incentives induced by affirmative action. Our results suggest that more attention to relative incentives is needed in empirical studies of observational datasets concerning college admissions.

This paper contributes to a small but growing literature recognizing that admissions policies shape incentives for pre-college human capital accumulation. Bruno Ferman and Juliano Assuncao

[2011] found evidence that test scores among black Brazilian high-school students decreased in response to an admissions quota at elite universities in Rio de Janeiro.[3] Hickman and Bodoh-Creed [2017] estimated a structural empirical model of the U.S. college admissions market based on the theoretical foundation of Bodoh-Creed and Hickman [2017]. His counterfactual analysis of admissions, investment, and welfare under alternative affirmative action policies found evidence that affirmative action increases the stock of minority human capital overall, and provided further comparisons between color-blind admissions, representative quotas, and American-style preference-based affirmative action (score bonuses).

Finally, our study is also related to others concerning performance pay in primary and secondary schools. Recent studies including Eric Bettinger [2012], Edwin Leuven, Hessel Oosterbeek and Bas van der Klaauw [2010] and Roland Fryer [2011] found mixed results about whether paying students can improve their effort and outcomes. Our experimental incentives are different from these in that we paid participants based on their *relative performance* rather than offering a fixed wage contract. Michael Kremer, Edward Miguel and Rebecca Thornton [2009] also studied a program awarding merit scholarships based on relative performance, and found evidence that students respond to changing financial incentives in a competitive environment. Our experimental design has two main differences from these studies: first, we were able to monitor outcomes *and* interim time inputs; and second, our focus is not on the effectiveness of pay-for-performance *per se*, but rather on *differences* between common relative performance incentive schemes. Finally, in a competitive setting some of a subject's behavior may be driven by an intrinsic desire to win, as in Christopher Cotton, Frank McIntyre and Joseph Price [2013] who presented evidence that simply framing a task as a contest can lead to better performance by some participants. In our study, we frame a common task as *two different contests*, which allows us to pick up on differences across alternative allocation mechanisms.

3. Theory

3.1. **MODEL.** We present a model of a large-scale investment contest which will form the basis of our experimental design. The setup is a special case of Bodoh-Creed and Hickman [2017] who developed the technical foundations of the framework presented here. We present a Bayesian game with a continuum of heterogeneous agents of mass 1. The mass of agents exert effort to improve their achievement on a proficiency examination which measures accumulated human capital (HC) and also establishes a rank ordering among them as they compete for prizes. There is a mass 1 of prizes, which differ in their value, and which are awarded to the agents based on relative achievement. Agents differ by how costly it is to improve their achievement.

[3] While seemingly contradictory to our findings, our model predicts that a quota which affects all points in the quality spectrum of universities will lead to a widening of the demographic achievement gap among the most talented students. Ferman and Assuncao [2011] estimated a negative overall effect of a policy change which applied only to top universities in Rio De Janeiro. Thus, the policy they study would have predominantly impacted high-ability black students who were at less of a disadvantage to begin with. Likewise, Brent R. Hickman and Aaron Bodoh-Creed [2017] estimated that affirmative action negatively affects investment among a small mass of the highest ability minorities in the US, even though on average the effect is positive.

Each agent i simultaneously chooses a level of achievement h_i in the contest. We assume there is a minimum level of achievement necessary to receive a prize, denoted $\underline{h} > 0$.[4] Acquiring HC requires time and effort, and is therefore thought to be costly. Agents differ by their background, abilities, access to help, and other resources that affect the rate at which time is converted into new HC. At the individual level, these factors are summarized by a parameter θ_i, which determines the marginal costs of agent i's HC accumulation. Investment entails a utility cost $c(h_i;\theta_i)$, which is strictly increasing and convex in the amount of output: $c'(h_i;\theta_i) > 0$ and $c''(h_i;\theta_i) > 0$ for all $h_i \geq \underline{h}$. We also assume $\frac{\partial c}{\partial \theta} > 0$ and $\frac{\partial^2 c}{\partial h \partial \theta} > 0$ so that smaller θ implies a more productive agent having lower costs for a given level h, and also a lower marginal cost of increasing output from h to $h + \Delta$. Agents know their own type, θ_i, and the distribution of all types, denoted by random variable Θ.

There is a continuum of prizes of mass 1. We denote an individual prize by quality level p, and model the relative masses of different prizes as following distribution F_P. Throughout the paper, we shall simplify discussion by assuming $F_P = Uniform[\underline{p}, \overline{p}] = [0,1]$ so that prize quality index and quantile rank are the same, with higher p indicating a more valuable prize. An agent matching with prize p_i and having produced h_i units of human capital experiences gross benefit $u(p_i, h_i) = p_i^{\alpha} h_i^{\beta}$, with net utility given by[5]

$$U(p_i, h_i; \theta_i) = p_i^{\alpha} h_i^{\beta} - c(h_i;\theta_i). \tag{1}$$

For tractability we simplify Bodoh-Creed and Hickman [2017]'s general framework by assuming an agent's gross match utility $u(\cdot,\cdot)$ takes the familiar Cobb-Douglas form. This will facilitate derivation of comparative static results pertaining to the research questions we wish to address with our experimental design.

3.1.1. ***Incorporating Demographics***. Each agent observably belongs to one of two mutually exclusive demographic subgroups, $\mathcal{A}$ and $\mathcal{D}$, with $\delta \in (0,1)$ being the mass of the latter. Since costs are privately known to each individual, agents view their competitors' types in group $j = \mathcal{A}, \mathcal{D}$ as realizations of a random variable Θ_j with distribution $F_j(\theta)$ and density $f_j(\theta)$ which is strictly positive on a common support $[\underline{\theta}, \overline{\theta}]$. As convenient shorthand, we define the overall distribution and density as $F(\theta) \equiv \delta F_{\mathcal{D}}(\theta) + (1-\delta) F_{\mathcal{A}}(\theta)$ and $f(\theta) \equiv \delta f_{\mathcal{D}}(\theta) + (1-\delta) f_{\mathcal{A}}(\theta)$.

We assume group $\mathcal{A}$ is "advantaged" and group $\mathcal{D}$ is "disadvantaged" in that a typical agent in $\mathcal{D}$ finds HC more costly than a typical agent in $\mathcal{A}$. Formally, we assume the distributions of $\Theta_{\mathcal{A}}$ and $\Theta_{\mathcal{D}}$ are ordered by *likelihood ratio (LR) dominance*, where

$$\frac{\partial \left(f_{\mathcal{D}}(\theta) / f_{\mathcal{A}}(\theta) \right)}{\partial \theta} > 0 \quad \text{for all} \quad \theta \in [\underline{\theta}, \overline{\theta}].$$

[4]In a university admissions competition, for example, quantity $\underline{h}$ may represent a minimum pre-requisite for market participation, as when university applicants must first earn a high school diploma in order to be considered.

[5]We restrict attention to the case where the agent with the highest cost type, $\overline{\theta}$, chooses achievement $\underline{h}$. In other words, we assume $U(\underline{p}, \underline{h}; \overline{\theta}) = U(\emptyset, 0; \overline{\theta})$, which guarantees that all agents in our game weakly prefer the minimum HC to not participating in the contest. Those with $h_i \geq \underline{h}$ are allocated a prize based upon their relative HC accumulation. This means that the model is one of decisions on the intensive margin, being conditional on participation. The question of how preferential treatment may affect participation decisions on the extensive margin is left for future research.

Intuitively, LR dominance implies, among other things, that cost quantiles within $\mathcal{A}$ are all lower than the corresponding quantiles within $\mathcal{D}$.[6] Although some "disadvantaged" agents have low costs of acquiring HC, and some "advantaged" agents have high costs, the average and median agents in $\mathcal{D}$ have higher costs of producing new HC than their counterparts in $\mathcal{A}$, *etc.* This is how we operationalize the idea that, on average, disadvantaged agents must invest more time and effort to overcome obstacles which are (imperfectly) correlated with their demographic status.[7]

Note that LR dominance implies there exists a unique point, $\widetilde{\theta} \in (\underline{\theta}, \overline{\theta})$, at which the densities $f_{\mathcal{A}}$ and $f_{\mathcal{D}}$ cross. This fact will become useful later when we explore comparative statics under different allocation mechanisms.

3.1.2. ***Prize Allocation Rules.*** We consider two distinct prize allocation rules. Let $H \sim G(H)$ denote the HC output of a randomly selected agent, and let $H_j \sim G_j(H_j)$ be similarly defined for an agent from group $j = \mathcal{A}, \mathcal{D}$. The baseline prize allocation rule is a *pure rank order* (PRO) mechanism, which ignores demographics when allocating prizes. The mechanism determines the quantile rank of i's achievement h_i within the overall HC distribution and then matches her to a prize at the corresponding quantile rank. For example, the 75^{th} percentile agent matches with the 75^{th} percentile prize, *etc.* Formally, under the PRO rule, agent i from group j receives prize

$$P_j^{PRO}(h_i) = P^{PRO}(h_i) \equiv G(h_i), \quad j = \mathcal{A}, \mathcal{D}. \tag{2}$$

The second prize allocation rule we consider is a *representative quota* (RQ) preferential treatment policy, which reserves a similar distribution of prizes for each group, $\mathcal{A}$ and $\mathcal{D}$, and then allocates the prizes *within* each group by rank ordering. By "similar distribution" we mean that fraction δ of all prizes at each point in the quality spectrum are earmarked *ex ante* for group $\mathcal{D}$, thus splitting prizes into two subsets having mass δ and $(1-\delta)$, but with both subsets still following the original quality distribution F_P.[8] Formally, under the RQ rule, agent i receives prize

$$P_j^{RQ}(h_i) \equiv G_j(h_i), \quad j = \mathcal{A}, \mathcal{D}. \tag{3}$$

An agent's prize assignment under the RQ rule depends only on her achievement relative to other members of her own demographic group. It splits the prize pools across two separate contests, with identical distributions of prizes, but different distributions of competitor ability. From an agent's perspective, the distinguishing characteristic of the RQ rule is that it alters the distribution of one's competitors, while leaving all other aspects of the contest—one's ability and the set of all prizes under competition—the same as under a PRO rule.

[6] More concretely, LR dominance implies that for any measurable subset $T \subseteq [\underline{\theta}, \overline{\theta}]$ the distributions of $\Theta_{\mathcal{A}}$ and $\Theta_{\mathcal{D}}$, conditional on the event T follow first-order stochastic dominance, or $F_{\mathcal{D}}(\theta|T) \leq F_{\mathcal{A}}(\theta|T)$, for all $\theta \in T$.

[7] In the context of education and race, for example, it is well-known that Black and Hispanic children in the United States tend to be less affluent and have less access to crucial childhood inputs like health care and high-quality public education; however, some still grow up in affluent environments which are more advantageous to childhood learning.

[8] A subtle but important detail to note here is that our RQ policy calibrates δ to the *fraction of group $\mathcal{D}$ participants*, which may not be the same as the mass of that group within the population at large. For example, South Africa mandates employment quotas for skilled professions, where quotas are pegged to the fraction of blacks in the overall population. One problem in the implementation of this law has been that not enough blacks exist within the skilled labor market (*e.g.*, individuals with prerequisite post-secondary degrees) to fill mandated quotas. Such a rule would be a more extreme version of the preferential treatment policy studied in this paper, and may therefore entail different consequences.

3.1.3. ***Solution Concept:*** The strategic environment presented above forms the basis for our experimental design in the next section. Bodoh-Creed and Hickman [2017] develop a general model of a large scale contest that nests the framework presented here as a special case. They begin with a finite set of agents and prizes, and demonstrate existence of a unique symmetric Bayes-Nash equilibrium. They then establish general conditions under which equilibria of the unwieldy finite model can be well approximated by the solution to a decision problem where the finite sets of agents and prizes are replaced with continua.[9] They refer to this solution as an *approximate equilibrium*, and it lends analytic tractability to an otherwise complicated model. For the understanding we wish to develop in this paper, it suffices to limit discussion to the continuum representation of the model. Our main goal in this section is to illustrate qualitative model predictions which are testable through experimental methods.

3.2. **EQUILIBRIUM ANALYSIS.** Let $h^*(\theta)$ denote the common equilibrium investment function under the benchmark PRO rule, and let $h^*_{\mathcal{A}}(\theta)$ and $h^*_{\mathcal{D}}(\theta)$ denote the group-specific investment functions under the alternative RQ rule. Moreover, let θ^*, $\theta^*_{\mathcal{A}}$, and $\theta^*_{\mathcal{D}}$ denote the relevant inverses, so that $\theta^* \equiv h^{*-1}$ and $\theta^*_j \equiv h^{*-1}_j$, $j = \mathcal{A}, \mathcal{D}$. We now turn attention to our main theoretical results concerning incentive provision and shifts in relative competition. In the next subsection we begin by developing results for a simplified version of the model where we shut down the direct marginal benefit of human capital in order to isolate the implications of strategic investment incentives. In the following two subsections, we show that the strategic aspect of the model plays largely the same role in a more realistic setting where human capital is intrinsically valued by agents.

3.2.1. ***Pure Strategic Incentives***. In this section we consider a special case of the Cobb-Douglas match utility function where $(\alpha, \beta) = (1, 0)$. Here, PRO competition net utility takes the form

$$U(s_i, h_i, \theta_i) = P^{PRO}(h_i) - c(h_i; \theta_i).$$

Agent i chooses HC to maximize her net payoff, given that other agents play according to h^*. Equilibrium HC investment is strictly decreasing in θ, so equation (2) can be re-written as $P^{PRO}(h) = 1 - F\left[\theta^*(h)\right]$, and agent i's objective as

$$\max_{h_i \geq \underline{h}} \left\{ \left(1 - F\left[\theta^*(h_i)\right]\right) - c(h_i; \theta_i) \right\}. \tag{4}$$

Taking a first-order condition (FOC), we get $-f(\theta_i)\theta^{*\prime}(h_i) = c'(h_i; \theta_i)$. In equilibrium, $\theta^*(h_i) = \theta_i$ for all i, and since h^* is the inverse of θ^*, it follows that $h^{*\prime}(\theta_i) = 1/\theta^{*\prime}(h_i)$. Therefore, through a change of variables we can rearrange the FOC to get

$$h^{*\prime}(\theta_i) = \frac{-f(\theta_i)}{c'\left[h^*(\theta_i); \theta_i\right]}, \text{ with boundary condition } h^*(\overline{\theta}) = \underline{h}. \tag{5}$$

Given the assumptions on f and c, it is easy to see that h^* is strictly decreasing in θ. Equation (5) allows us to compare investment under two alternative cost distributions.

[9]Alternatively, one may also interpret the continuum simplification as the basis of a behavioral strategy employed by a cognitively constrained agent. Rather than tracking probabilities over all the order statistics of large, complicated sets of competitors and seats, the agent may drastically reduce computational burden at little cost by pretending as if she was operating within a continuous world instead of a discrete one.

Theorem 1. *Consider two PRO contests, 1 and 2, which differ only by their cost distributions, and assume competition is more fierce under contest 2 in the sense that F_1 LR dominates F_2. Let $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$ denote the unique crossing point of the density functions where $f_1(\tilde{\theta}) = f_2(\tilde{\theta})$. There exists a unique interior crossing point $\ddot{\theta} \in (\underline{\theta}, \tilde{\theta})$, such that $h_1^*(\theta) < h_2^*(\theta)$ for all $\theta < \ddot{\theta}$ and $h_1^*(\theta) > h_2^*(\theta)$ for all $\theta > \ddot{\theta}$.*

A formal proof is in the Appendix. The theorem provides useful insight into how competition shapes incentives. It says that, holding fixed the set of all prizes, an increase in the degree of competition will cause the most able agents to invest more aggressively, and less talented agents to withdraw somewhat, decreasing their HC accumulation. This second shift is a manifestation of a phenomenon in contests known as the *discouragement effect*. To our knowledge, this paper is the first to characterize discouragement effects in a many-to-many rank-order setting.

The intuition embodied in the theorem is that since investment costs must be sunk before prizes are assigned, then, holding one's own cost type θ fixed, if the distribution of competitors shifts so that one's quantile rank falls low enough, investment incentives fall. Moreover, *any* cost type but the lowest possible type can conceivably become subject to discouragement: note that it directly follows from the result that for any interior cost type $\theta' > \underline{\theta}$, there is some stochastic dominance shift toward lower cost competitors that is extreme enough so that the crossing point of the densities to belongs to the interval $\tilde{\theta} \in (\underline{\theta}, \theta')$, and type θ' will reduce investment.

There is much insight to be had from Theorem 1 when comparing alternative allocation rules as well. Under an *RQ* mechanism granting preferential treatment to group $\mathcal{D}$, equation (3) can be re-written as $P_j^{RQ}(h) = 1 - F_j\left[\theta_j^*(h)\right]$, $j = \mathcal{A}, \mathcal{D}$, and the objective for agent i from group j is now

$$\max_{h_i \geq \underline{h}} \left\{ \left(1 - F_j\left[\theta_j^*(h_i)\right]\right) - c(h_i; \theta_i) \right\}. \tag{6}$$

Note the main difference here is that the group-specific distribution enters decision making rather than the unconditional one. That leads to the following FOC

$$h_j^{*\prime}(\theta_i) = \frac{-f_j(\theta_i)}{c'\left[h_j^*(\theta_i); \theta_i\right]}, \text{ with boundary condition } h_j^*(\overline{\theta}) = \underline{h}, \;\; j = \mathcal{D}, \mathcal{A}. \tag{7}$$

Recall our assumption that the random variable $\Theta_{\mathcal{D}}$ LR dominates $\Theta_{\mathcal{A}}$, or in other words, the ratio $f_{\mathcal{D}}(\theta)/f_{\mathcal{A}}(\theta)$ is strictly increasing in θ. This implies that Θ LR dominates $\Theta_{\mathcal{A}}$, and that $\Theta_{\mathcal{D}}$ LR dominates Θ as well. To see why, note that

$$\frac{f(\theta)}{f_{\mathcal{A}}(\theta)} = \frac{\delta f_{\mathcal{D}}(\theta) + (1-\delta) f_{\mathcal{A}}(\theta)}{f_{\mathcal{A}}(\theta)} = \delta \frac{f_{\mathcal{D}}(\theta)}{f_{\mathcal{A}}(\theta)} + (1 - \delta),$$

from which it follows that $f(\theta)/f_{\mathcal{A}}(\theta)$ is strictly increasing in θ. Likewise,

$$\left(\frac{f_{\mathcal{D}}(\theta)}{f(\theta)}\right)^{-1} = \frac{\delta f_{\mathcal{D}}(\theta) + (1-\delta) f_{\mathcal{A}}(\theta)}{f_{\mathcal{D}}(\theta)} = \delta + (1 - \delta)\frac{f_{\mathcal{A}}(\theta)}{f_{\mathcal{D}}(\theta)},$$

so $f_{\mathcal{D}}(\theta)/f(\theta)$ is strictly increasing in θ as well. Therefore, switching between allocation rule *PRO*—where the competition group is all agents—and the *RQ* rule—where competition occurs only within one's own group—entails an effective LR dominance shift in the distribution of competitors while holding the distribution of prizes fixed. Thus, the above argument in conjunction with Theorem 1 leads to the following proposition on HC accumulation under allocative preferential treatment.

Proposition 2. *Assume $F_{\mathcal{D}}$ LR dominates $F_{\mathcal{A}}$ and let $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$ denote the unique crossing point of the cost densities where $f(\tilde{\theta}) = f_{\mathcal{A}}(\tilde{\theta}) = f_{\mathcal{D}}(\tilde{\theta})$. Then there exist crossing points $\ddot{\theta}_{\mathcal{A}}, \ddot{\theta}_{\mathcal{D}} \in (\underline{\theta}, \tilde{\theta})$, such that*

(i) $h^*_{\mathcal{D}}(\theta) < h^*(\theta)$ for all $\theta < \ddot{\theta}_{\mathcal{D}}$ and $h^*_{\mathcal{D}}(\theta) > h^*(\theta)$ for all $\theta > \ddot{\theta}_{\mathcal{D}}$, and
(ii) $h^*_{\mathcal{A}}(\theta) > h^*(\theta)$ for all $\theta < \ddot{\theta}_{\mathcal{A}}$ and $h^*_{\mathcal{A}}(\theta) < h^*(\theta)$ for all $\theta > \ddot{\theta}_{\mathcal{A}}$.

In words, the *ex-ante* strongest agents within group $\mathcal{D}$ *decrease* HC investment, as competition for the top prizes becomes less intense. At the same time, higher-cost individuals exert greater effort and *increase* HC accumulation, as *RQ* mitigates discouragement effects by placing them in a competition group where they are less far behind the curve. The same reasoning implies the opposite effects for group $\mathcal{A}$. The best and brightest group $\mathcal{A}$ agents *increase* HC investment under the *RQ* rule, as competition for the top prizes becomes more intense, and others become discouraged as they find themselves further behind the curve in their new competition group.

Since theory predicts behavioral responses in opposite directions for agents of different abilities within each group, the result above begs the question of which effect will dominate. Once again, Theorem 1 can shed some light on the answer. Using the density crossing point $\tilde{\theta}$, we can partition the cost support into two subsets,

$$T_{\mathcal{A}} \equiv (\underline{\theta}, \tilde{\theta}) \quad \text{and} \quad T_{\mathcal{D}} \equiv (\tilde{\theta}, \overline{\theta}),$$

which we refer to as the *typical cost sets* for each group. Intuitively, T_j is the region of the support where group j's density is strictly higher, meaning it is overrepresented relative to its share in the overall population. LR dominance implies some interesting properties for these sets. Since the densities have a unique crossing and since both must integrate to 1, it follows that

$$\int_{T_{\mathcal{A}}} [f_{\mathcal{A}}(\theta) - f_{\mathcal{D}}(\theta)] = \int_{T_{\mathcal{D}}} [f_{\mathcal{D}}(\theta) - f_{\mathcal{A}}(\theta)] \, .$$

In words, the degree of over-representation of group $\mathcal{D}$ in the high-cost set $T_{\mathcal{D}}$ is the same as the degree of over-representation of group $\mathcal{A}$ within the low-cost set $T_{\mathcal{A}}$.

Proposition 3. *Assume the same conditions as in Proposition 2. Then under the RQ rule (relative to the PRO rule), typical disadvantaged agents—that is, group $\mathcal{D}$ agents with costs $\theta \in \left\{(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}] \cup T_{\mathcal{D}}\right\}$—exert higher effort and accumulate more HC. Moreover, if we define $\Delta : (\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}] \cup T_{\mathcal{D}} \to \mathbb{R}$ as the difference on this set between group $\mathcal{D}$ investment under RQ versus PRO, or*

$$\Delta(\theta) \equiv (h^*_{\mathcal{D}}(\theta) - h^*(\theta)) \, ,$$

then $\Delta(\theta)$ is strictly positive and attains a maximum on the interval $(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}]$. Moreover, if investment costs are strictly convex in h, then $\Delta(\theta)$ attains its maximum on the open interval $(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta})$.

Proposition 3 implies that, relative to their disadvantage in the overall population, a majority of group $\mathcal{D}$ agents actually *increase* HC investment under the preferential treatment scheme. To gain an appreciation for the strength of this result, the proposition also shows that not only is there a positive effect for all group $\mathcal{D}$ cost types in their typical set, but the improved incentives extend well beyond $T_{\mathcal{D}}$ as well. In fact, the largest improvement of investment incentives by type (*i.e.*, where $\Delta(\cdot)$ attains its maximum) actually occurs on the interval $(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}]$. Thus, the result implies an increase of investment activity for a large fraction of the disadvantaged group.

On the other hand, the situation is less clear for the advantaged group. Although an analogous statement can be made—that all group $\mathcal{A}$ cost types $\theta \in \left\{(\ddot{\theta}_{\mathcal{A}}, \tilde{\theta}] \cup T_{\mathcal{D}}\right\}$ will reduce HC output under RQ—the statement is less informative, because investment also increases on part of the typical set $T_{\mathcal{A}}$ as well.

3.2.2. ***Competitive Investment with Intrinsically Valued Human Capital***. Having characterized the workings of strategic forces which produce discouragement effects in our many-to-many matching contest, we now return to the more interesting version of the model to demonstrate that the above results carry over to the case where agents not only value their prize, but also value HC accumulation itself. Here, investments have both direct (through the intrinsic value of HC) and indirect (through the prize allocation contest) effects on the agents' utility.

If $\alpha = \beta = 1$ and costs are linear so that $U(s_i, h_i, \theta_i) = s_i h_i - \theta(h - \underline{h})$, and $[\underline{\theta}, \overline{\theta}] \subset (1, \infty)$, then proving statements analogous to Theorem 1, Proposition 2, and Proposition 3 is relatively straightforward.[10] In this scenario, under PRO student i's objective is now

$$\max_{h_i \geq \underline{h}} \left\{ h_i \left(1 - F\left[\theta^*(h_i)\right]\right) - \theta_i (h_i - \underline{h}) \right\}. \tag{8}$$

Taking a FOC, we get $-h_i f\left[\theta_i^*(h_i)\right] \theta^{*\prime}(h_i) + \left(1 - F\left[\theta^*(h_i)\right]\right) = \theta_i$. In this world, student i can be thought to choose HC production in two parts. First, she raises investment to the level where she optimally benefits from complementarity of her own input h_i and the input p_i awarded to her based on relative performance. This calculation, which Bodoh-Creed and Hickman [2017] refer to as the *productive channel* of incentives, involves the direct marginal benefit of a unit of HC, or $\left(1 - F\left[\theta^*(h_i)\right]\right)$. Above that base level of investment, competitive incentives play an additional role: in a monotone equilibrium i invests further so as to make her higher-cost competitors prefer not to try and jump her. This calculation involves the indirect placement benefit of more HC, or $-h_i f\left[\theta_i^*(h_i)\right] \theta^{*\prime}(h_i)$, and it produces a tendency for over-investment which is referred to as the *competitive channel* of incentives.

If we now define $\eta(\theta) \equiv \log\left(h^*(\theta)\right)$ then through a similar change of variables as before we can rearrange the FOC to get a differential equation

$$\eta'(\theta_i) = -\frac{f(\theta_i)}{\theta_i + F(\theta_i) - 1}, \text{ with boundary condition } \eta(\overline{\theta}) = \log(\underline{h}). \tag{9}$$

[10]The assumption of $[\underline{\theta}, \overline{\theta}] \subset (1, \infty)$ is required because the gross utility and cost function are both linear in h. Since the direct marginal benefit of more human capital is $u_2(s, h) = s$ and the direct marginal cost is θ, we must have $\underline{\theta} > \overline{s}$, in order to rationalize all students choosing finite HC production.

Once again, η is strictly decreasing in θ and we can now prove the analog of Theorem 1 above:

Theorem 4. *Assume HC production costs and gross utility are linear in h, so that $c(h;\theta) = \theta(h - \underline{h})$, and $u(s,h) = sh$. Moreover, consider two pure rank order contests with cost distributions, $F_1(\theta)$ and $F_2(\theta)$, where competition is more intense under F_2 in the sense that F_1 LR dominates F_2 (i.e., the ratio $\frac{f_1(\theta)}{f_2(\theta)}$ is strictly increasing). Then, letting $\tilde{\theta}$ denote the unique crossing point of f_1 and f_2, there exists a unique interior crossing point $\ddot{\theta} \in (\underline{\theta}, \widetilde{\theta})$ such that $h_1^*(\theta) < h_2^*(\theta)$ for $\theta < \ddot{\theta}$ and $h_1^*(\theta) > h_2^*(\theta)$ for $\theta > \ddot{\theta}$.*

A proof of this result is relegated to the appendix, but we can build on it to prove comparative statics for different prize allocation rules just as before. Proposition 5 and Proposition 6 below build on Theorem 4 in precisely the same way that Proposition 2 and Proposition 3 build on Theorem 1 above. In particular, as above, Proposition 6 follows naturally from Theorem 4 and from the properties of LR dominance: we now know that $\eta'_{\mathcal{D}}(\theta) < \eta'(\theta) < \eta'_{\mathcal{A}}(\theta)$ for each $\theta \in [\tilde{\theta}, \overline{\theta}]$, and since the log transformation preserves ordering, it follows that $h^{*\prime}_{\mathcal{D}}(\theta) < h^{*\prime}(\theta) < h^{*\prime}_{\mathcal{A}}(\theta)$ on that same interval as well. This also produces a slight strengthening of the result since the derivatives are strictly ordered at the density crossing $\tilde{\theta}$. Together, the final two results demonstrate that the model still predicts a large fraction of the disadvantaged group increasing investment under RQ with intrinsically valued HC:

Proposition 5. *Let HC production costs and gross utility be linear in h, so that $c(h;\theta) = \theta(h - \underline{h})$, and $u(s,h) = sh$. Moreover, assume $F_{\mathcal{D}}$ LR dominates $F_{\mathcal{A}}$ and let $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$ denote the unique crossing of the cost densities where $f(\tilde{\theta}) = f_{\mathcal{A}}(\tilde{\theta}) = f_{\mathcal{D}}(\tilde{\theta})$. Then there exist crossing points $\ddot{\theta}_{\mathcal{A}}, \ddot{\theta}_{\mathcal{D}} \in (\underline{\theta}, \tilde{\theta})$, such that*

(i) $h^*_{\mathcal{D}}(\theta) < h^*(\theta)$ for all $\theta < \ddot{\theta}_{\mathcal{D}}$ and $h^*_{\mathcal{D}}(\theta) > h^*(\theta)$ for all $\theta > \ddot{\theta}_{\mathcal{D}}$, and
(ii) $h^*_{\mathcal{A}}(\theta) > h^*(\theta)$ for all $\theta < \ddot{\theta}_{\mathcal{A}}$ and $h^*_{\mathcal{A}}(\theta) < h^*(\theta)$ for all $\theta > \ddot{\theta}_{\mathcal{A}}$.

Proposition 6. *Assume the same conditions as in Proposition 5. Then under the RQ policy (relative to the PRO policy), typical disadvantaged students—that is, group $\mathcal{D}$ students with costs $\theta \in \left\{(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}] \cup T_{\mathcal{D}}\right\}$—exert higher effort and accumulate more HC. Moreover, if we define $\Delta : (\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}] \cup T_{\mathcal{D}} \to \mathbb{R}$ as the difference on this set between group $\mathcal{D}$ investment under RQ versus PRO, or $\Delta(\theta) \equiv (h^*_{\mathcal{D}}(\theta) - h^*(\theta))$, then $\Delta(\theta)$ is strictly positive and attains a maximum on the interval $(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta})$.*

3.2.3. ***Numerical Examples: Strictly Concave Match Utility.*** In this section we present some simple numerical examples to further illustrate the model and demonstrate robustness of the qualitative patterns depicted above. In the example, $\delta = 0.5$ and both $\Theta_{\mathcal{D}}$ and $\Theta_{\mathcal{A}}$ follow normal distributions truncated to a common support $[\underline{\theta}, \overline{\theta}] = [1, 2]$ with variance parameter $\sigma_{\mathcal{A}} = \sigma_{\mathcal{D}} = 0.25$. The mean parameters differ, with $\mu_{\mathcal{D}} = 1.5$ and $\mu_{\mathcal{A}} = 1.1$, which ensures that the distributions are ordered by stochastic dominance (see Figure 1). We specify costs as a linear function $c(h;\theta) = \theta h$, so that the maximum distance between investment under *RQ* and *PRO* occurs at the boundary

FIGURE 1. NUMERICAL EXAMPLES:
Cost Densities

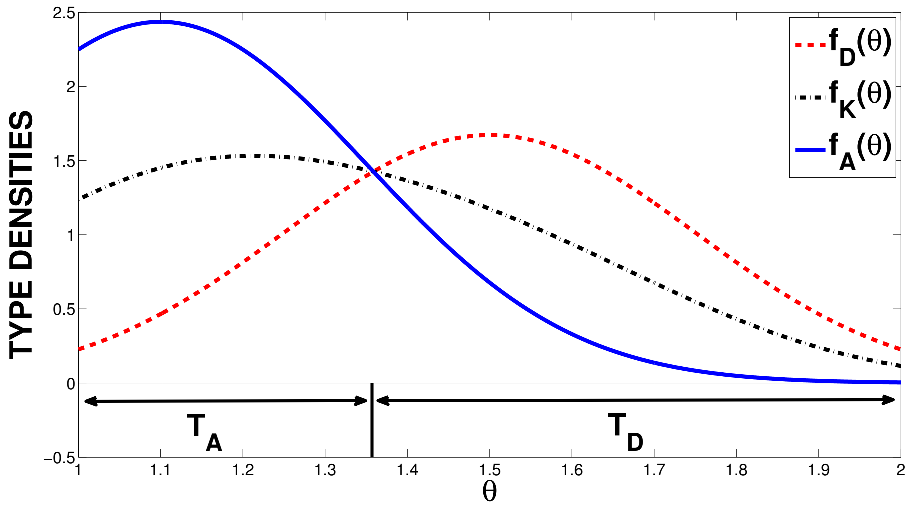
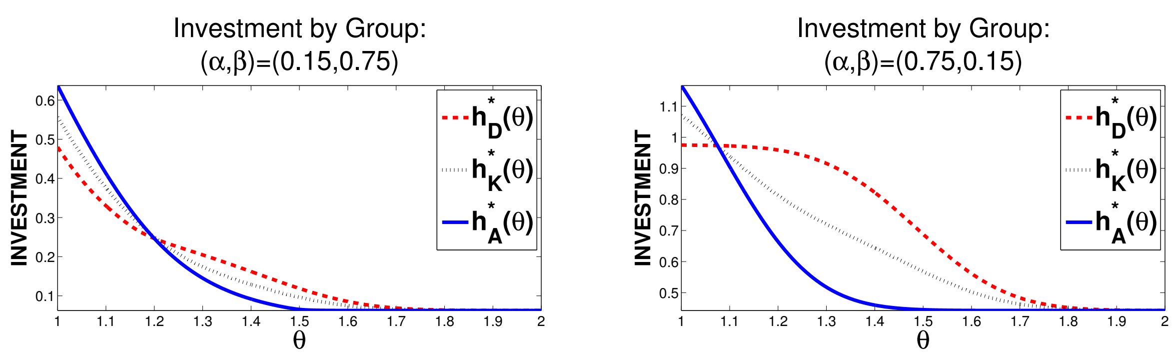


FIGURE 2. NUMERICAL EXAMPLES:
Investment, PRO vs RQ



(A) Example 1 (B) Example 2

between $T_{\mathcal{A}}$ and $T_{\mathcal{D}}$. Finally, we set $\underline{h} = 0.5$; this makes type $\overline{\theta}$ indifferent to market participation when the outside option is normalized to a value of zero.

We numerically solve for equilibria in two examples. In the first (Example 1), HC factors relatively heavily into match utility, with $\alpha = 0.15$ and $\beta = 0.75$. In the second (Example 2), students care less about their own HC and more about the quality of the institution they attend, with $\alpha = 0.75$ and $\beta = 0.15$. Equilibrium outcomes are summarized in figure 2.

The patterns arising from these examples are similar as before. There continues to exist a single crossing point between the investment functions of the two groups under RQ vs. PRO. Under RQ, the highest ability disadvantaged students and the lower ability advantaged students decrease HC investment, while the lower ability disadvantaged students and the higher ability advantaged students increase HC investment. In both examples average investment in group $\mathcal{D}$ rises; for group $\mathcal{A}$ it rises in Example 1 and falls in example 2.

Figure 2 depicts a comparison of the RQ investment functions for each group (solid and dashed lines), as well as the common investment function for both groups under a PRO allocation rule (dotted line). For each group there is crossing point of the investment functions, with the upper bounds of the HC distributions being different by seat allocation rule. For the disadvantaged group, a positive mass of the top agents reduce investment, while middle- and high-cost agents increase it. Intuitively, the policy aids the top agents from $\mathcal{D}$, but since they were already placing close to the upper bound their outcomes cannot be commensurately improved and they rationally reduce effort. For other agents in $\mathcal{D}$, the policy alleviates discouragement effects by placing them in a competition group where their own type is not as far behind the curve, making them more competitive for higher quality outcomes. In turn, they respond with an increased willingness to engage in costly investment. Similar logic holds for group $\mathcal{A}$, but in reverse. A key aspect of our experiment will be to engineer this sort of counterfactual scenario, where similar ability agents find themselves facing different levels of competition for the same set of outcomes. Within this experimental counterfactual we test whether the representative quota policy leads to changes in achievement distributions. We also investigate intermediate inputs behind investment costs depicted in the model: time and effort.

4. EXPERIMENTAL DESIGN

We built our incentives and learning exercise around the American Mathematics Competition 8 (AMC8) exam, sponsored by the Mathematical Association of America for students in 8$^{\text{th}}$ grade and below. It consists of 25 multiple choice questions (five choices each) in 40 minutes, and the questions become progressively more difficult from start to finish. The AMC8 is an early precursor to math competitions such as the International Mathematical Olympiad and the Putnam. The AMC8 website explains that it "provides an opportunity to apply concepts [to] high level of problems which... are designed to challenge and offer problem solving experiences beyond those provided in... junior high school."

4.1. **SAMPLE POPULATION.**

Our total sample includes 992 middle school and junior high students from 10 schools in Utah County, Utah, including both charter schools and regular public schools. Administrators at these schools partnered with us for the study, so participation in our study was at the classroom level on an opt-out basis. Academically and socioeconomically, our sample population was fairly average within the US.[11] Most of our partner schools had previously participated in the AMC8 before partnering with us for this study. One difference between our

[11] Our test subjects are drawn from a population which is somewhat more affluent than the rest of the country, but academically comparable. Based on figures from the National Center for Education Statistics, we estimate a median household income of $59,800 for the sample population, compared to a nationwide median of $53,046.[12] In 2012, approximately 33% of test subjects were eligible for free or reduced-price lunch, compared to a national average of 48%. That year our partner schools housing 5$^{\text{th}}$ and 6$^{\text{th}}$ grade subjects (20% of our sample pool) performed significantly better than other Utah schools in terms of meeting state math standards (approximately 91% vs 76%), while schools housing our 7$^{\text{th}}$ and 8$^{\text{th}}$ graders (80% of our sample pool) performed slightly worse than other Utah schools (81% vs 83% meeting state standards). Utah is a state which typically ranks at or near the median for nationally measured academic outcomes such as NAEP scores and enrollment rates in Advanced Placement programs.

sample population and the national population is that our partner schools exhibited a certain degree of racial and cultural homogeneity: less than 9% of subjects were racial minorities (black or Hispanic), compared to a nationwide average of 37.45%. All schools in this study serve suburban populations.

4.2. **TREATMENT GROUPS AND INCENTIVES.** Participants in our study first took a practice AMC8 test from a previous year. We used this as a baseline measure of each subject's ability. Individuals were randomized into either a control group—with a pure rank order (PRO) competition—or a preferential treatment group—with a representative quota (RQ) competition. Following the notational convention from the previous section, we will henceforth refer to these treatments as PRO and RQ, respectively. For treatment PRO, we ran competitions involving subjects in two adjacent grades; that is, 7^{th} and 8^{th} graders competing together, and 5^{th} and 6^{th} graders competing together. For each of the two age cohort pairings, subjects in the lower grades (5^{th} or 7^{th}) are henceforth referred to as the "disadvantaged" group $\mathcal{D}$, and subjects in the higher grades (6^{th} or 8^{th}) are referred to as the "advantaged" group $\mathcal{A}$, since the latter are one year older and have received one more year of mathematics education on average.[13] For treatment RQ, subjects competed only within their own grade level, but for a proportionally equivalent set of prizes (relative to PRO), as described below. We ran separate competitions for 5^{th}, 6^{th}, 7^{th}, and 8^{th} graders.

The top 30 percent of subjects within each competition group received cash prizes, which were uniformly distributed between \$4 and \$34 in \$2 increments. Prizes were awarded within competition groups, according to final exam scores. For example, 7^{th} grade subjects in treatment PRO needed to score within the top 30 percent of all 7^{th} and 8^{th} grade subjects in their treatment to receive a prize. In treatment RQ, subjects competed against others in their own grade only, but for a representative set of prizes. More specifically, we began with the same aggregate prize distribution as for treatment PRO, and then earmarked prizes at each different level in proportion to the mass of lower-grade subjects in each age cohort pairing (note that the mass of groups $\mathcal{A}$ and $\mathcal{D}$ were also identical across treatments PRO and RQ). This ensured that the moments of the prize distribution (including the 70% mass of zeros) were the same across all competition groups, with each one vying for the same number and variety of prizes on a per capita basis. For example, 7^{th} grade subjects in treatment RQ only had to score within the top 30 percent of 7^{th} graders in their treatment to receive a prize. Moreover, their distribution of prizes, conditional on winning something, was the same as for 8^{th} graders in treatment RQ, and also the same as for all $7^{\text{th}}/8^{\text{th}}$ grade subjects in treatment PRO. Thus, for an advantaged or disadvantaged subject of a given ability level, the only difference across the two treatments is the distribution of one's competitors.

Each subject received an information sheet describing their assigned group, how many subjects from which grade(s) they would be competing against, and the score distribution within their group based on the practice test. Subjects received their own practice score back at the same time so they

[13]As we discuss in Section 6, the difference in average preparation between our disadvantaged and advantaged groups likely understates difference between black and white students at the same age in the US. Using national figures from the NAEP exam, we estimate that the 7^{th}–8^{th} grade gap in our sample pool is roughly 60% of the gap between black and white 8^{th} graders in terms of math proficiency.

could see where they fit within their competition group. The sheet also contained a table describing the prize structure. We printed information relative to each competition group on a different color of paper so that subjects could visually see in their classroom that roughly half of the subjects were assigned to each treatment. Altogether, there were six different groups: four groups for the quota treatment (one for each grade) and two groups for the neutral treatment (one for $5^{th}/6^{th}$ grade and one for $7^{th}/8^{th}$ grade). In a web appendix, we provide an example of the information sheet given to each group.[14]

4.3. **MATH LEARNING WEBSITE.** At the bottom of the information sheet was the url of a website we set up with practice problems drawn from five past AMC8 exams. At 25 questions each, this made for 125 total practice problems covering six different math topics: Arithmetic, Algebra, Combinatorics, Geometry, Logic, and Probability. Problems were divided into a set of 31 total quizzes. Each year, the 25 AMC8 exam questions are numbered in increasing order of difficulty. For each of the previous five year's exams, the website included one quiz covering problems 1-10, a second quiz covering problems 11-20, and a third covering problems 21-25. Test subjects were notified that each grouping of 3 same-year quizzes were ordered by their difficulty level. We also arranged this same battery of math problems into an additional set of 16 quizzes, each containing 5 topic-specific math problems. These topic quizzes were also ordered by their difficulty level.

Subjects could attempt each quiz as many times as they liked, or move on to additional materials they had not yet tried. After completing each quiz, our software displayed an instructional page which reported to each subject her score, the correct answers for each problem, and step-by-step solutions published by the developers of the AMC8. Subjects were provided with a web page that contained links for all of the quizzes we offered, but in order to access the quizzes, they had to input their name, grade, and school on the first page of the web form. This allowed us to track online activities for each quiz session, including which subjects visited the website, how many different topics they tried, how much time they spent, how many questions they attempted, what they answered on each attempt at each question, and how much time they spent viewing the instructional page.

Within each quiz, questions were separated on different web pages in blocks of 3, 4, or 5 questions per page, and the instructional page at the end displayed feedback for all questions on a single page.[15] Time on our website was measured at the page level, meaning that we got a time measure for blocks of either 3, 4, or 5 questions. In order to convert this information into a time spent per question measure, we divided each block-level time observation by the number of questions within that block. Instructional page times for 10-question quizzes were split into two observations a piece by dividing by two in order to make them comparable to 5-question instructional page view times.

4.3.1. ***Time Measurement.*** One difficulty arose in that there were clear instances where subjects left the website in the middle of a quiz for several hours or more. To adjust for this problem we

[14] Copies of the information sheets given to test subjects are available for download at http://home.uchicago.edu/~hickmanbr/uploads/CHP2014_WEB_APPENDIX.zip

[15] Each 10-question quiz was broken into three pages with 3 questions on page one, 3 questions on page two and 4 questions on page three. Each 5-question quiz displayed all 5 problems on a single page.

chose truncation points on the domains of time per question and instructional page view time, and we replaced each observation above that point with the appropriate subject-specific censored mean.[16] In selecting our truncation point we looked for occurrences of "holes" in the support of the distribution of times per question.[17] For our time per question data, this leads to a truncation point of 26.14 min/question (the 99.35th percentile), and for instructional page views, 108.39 min/page view (*i.e.*, 21.68 min/solution, or the 98th percentile). In the Appendix we display a histogram of (uncensored) time per question and instructional page view times.

At the end of the day, the time monitoring capability on our website is not perfect, and it is impossible to directly observe work stoppages in the middle of a quiz question. In particular, it may still be the case that smaller work stoppages occur below the truncation points. Therefore, in terms of time per question we are effectively interpreting work stoppages of less than 27 minutes as time which comes at a positive cost to the child. We argue that 27 minutes is a reasonable truncation point for several reasons. First, work stoppages for our uncensored time observations (most of which were less than 10 minutes) would serve as a poor substitute for longer, unbroken leisure spells. Second, since this potential problem is the same across both treatment groups, there is no reason to believe that our results are being aided by it.

Third, the AMC8 contains fairly challenging material that may require significant time inputs for some subjects. Table 1 in the next section displays the mean and variance of time spent per question attempt, using the censored sample of times. The most difficult topic appears to be combinatorics, with a mean time of 2.839 minutes and a standard deviation of 3.532. Given that the censored distribution of time per question is right skewed, and 10 minutes (the 98th percentile of the un-censored sample) is roughly two standard deviations above the mean for combinatorics, it is plausible that roughly 1.5% of our sample could exist on the interval between 10 minutes and 26 minutes.

4.4. **TESTING.** Subjects took the actual AMC8 test in their regular classrooms, under all of the normal conditions in which students around the country take the AMC8. Most of the subjects in our study attended schools where participation in the AMC8 was already being offered to students by their teachers. The schools that cooperated in this study administered the test to all students within each participating classroom on an opt-out basis, so that all subjects participated in the study, except those whose parents returned an opt-out form. The study involved two in-class exam sessions: the practice test was the AMC8 exam for the previous year, and the final exam was the AMC8 for the current year. The cash prizes were delivered to each school shortly after the final

[16]To illustrate this rule, suppose that Tommy attempted three 5-question quizzes for a total of 15 questions. Suppose further that we observed times of 5 minutes each for seven questions, 15 minutes each for another seven questions, and 2000 minutes for the last one. Then if the truncation point were, say 30 minutes, the last observation of 2000 is replaced by Tommy's idiosyncratic censored mean time of 10 minutes (for all other questions he attempted). As a robustness check, we also ran our analysis by simply dropping truncated observations instead, and results are very similar to those we present below.

[17]More specifically, a hole in the distribution support was defined as the minimum point at which a full-support condition fails, which we estimated as a point where a kernel-smoothed density estimate hit zero. The idea behind this rule is that if the type distribution has full support, then the distribution of times per question should have full support as well since the choice of how much time to spend is continuous. For a more complete description of our truncation point selection rule, see the Appendix.

exam, and handed out to each subject in an envelope. The outcome measures that we use in the next section include both the effort-based measures with website data, as well as a performance-based measure using subjects' scores on the AMC8.

5. EXPERIMENTAL RESULTS

The theoretical model predicts that some subjects in each demographic group will increase their HC investment activities under a representative quota preferential treatment rule, while others in the same group will decrease their efforts. Our empirical investigation takes the next step by testing for the average effects of preferential treatment on effort and performance by demographic group and estimating magnitudes. We also use a non-parametric analysis to consider the effects of preferential treatment on the performance distributions and achievement gaps.

5.1. DESCRIPTIVE STATISTICS.

5.1.1. ***Pre- and Post-Exams.*** Table 1 contains descriptive statistics on AMC8 exam scores. Roughly three quarters of our sample were $7^{\text{th}}/8^{\text{th}}$ graders. The difference between these and our $5^{\text{th}}/6^{\text{th}}$ grade test subjects is that the latter all came from accelerated classes, whereas the former are representative of the overall student body within their schools. This difference is born out in the data: while 8^{th} grade subjects did best on the pre-test with an average score of 9.04 (out of 25 possible), 6^{th} graders as a group came in second at 8.12 on average. 7^{th} and 5^{th} grade average pre-test scores are close, at 7.58 and 7.19, respectively.[18]

We have also broken down test scores by two groups that we refer to as ***investors***—subjects who logged on to our website at least once during the study period—and ***non-investors***—those who did not. Subjects who did better on the pre-test were more likely to be investors, although some subjects who did not do as well also chose to be investors, and many subjects who did quite well on the pre-test chose to be non-investors. For the group of investors, we also present summary statistics concerning their activities on the website. Investors' times ranged between a few minutes and 8.92 hours, or an average of about 53 minutes per day over the study period. If we divide total cash payments by total hours worked by all investors, we get an expected hourly wage of $10.73. Number of questions attempted ranged between 1 and 120, with mean and standard deviation of roughly 19 and 23, respectively. *Topics* represents the number of different topic categories a subject attempted, using the topic-specific quizzes, being about two on average.

5.2. EMPIRICAL ANALYSIS.

5.2.1. ***Testing Overall Differences by Treatment.*** Tables 2 and 3 investigate the effect of a quota on the overall population, including both advantaged *and* disadvantaged groups. The first column of Table 2 displays the fraction of students from each treatment group who logged on to our website at least once to practice math. As for the other investment variables, the reader should keep in mind that Tables 2 - 4 aim to measure a treatment effect of a policy on an entire group,

[18]The national AMC8 population in 2013 (see https://amc-reg.maa.org/reports/generalreports.aspx) had a mean of 10.69 with standard deviation 4.44. The AMC8 is predominantly administered through opt-in participation, whereas our experiment was on an opt-out basis. This accounts for the lower mean among our sample pool.

Table 1. SUBJECT DESCRIPTIVE STATISTICS

	Mean	Median	Std. Dev.	N
Pre-Exam Scores				
All	8.45	8	2.90	992
5^{th} Grade	7.19	7	2.39	48
6^{th} Grade	8.12	8	2.47	155
7^{th} Grade	7.58	7	2.84	275
8^{th} Grade	9.04	9	2.82	396
Investors	9.46	10	3.19	118
Non-Investors	8.32	8	2.83	874
Final Exam Scores				
All	8.64	8	2.88	895
5^{th} Grade	7.40	7	2.22	42
6^{th} Grade	9.17	9	2.82	133
7^{th} Grade	8.12	8	2.90	233
8^{th} Grade	8.75	9	2.80	374
Investors	9.20	9	3.06	113
Non-Investors	8.56	8	2.84	782
Human Capital Investment (Investors Only)				
Total Time	43.65	26.85	64.65	118
Problem Solving Time	32.99	19.31	41.43	118
Instructional Time	10.66	3.37	38.85	118
Questions	18.89	10.00	22.53	118
Topics	1.94	1.00	1.43	118

Notes: All time figures are post-censoring as described in Section 4.3 and quoted in minute units. *Investors* are defined as subjects who logged on to the math learning website at least once during the investment period. *Non-Investors* are those who did not.

including both the intensive and extensive margins of investment. This is why the effort numbers in Table 2 and afterward appear small: they are averaged over both investors *and* non-investors. The results indicate that subjects in the quota treatment, including subjects from all age groups, were 75% more likely to have visited the website than subjects in the color-blind treatment. They also tried out more topics, spent more time on the website and answered more questions. Table 2

Table 2. EFFORT AND PERFORMANCE BY TREATMENT

	Investment				**Performance**
	Used Website	# Topics Attempted	Total Time	# Questions Attempted	Final Exam Score
Quota	0.154	0.284	6.634	2.729	8.680
Std. Err.	(0.015)	(0.037)	(1.216)	(0.456)	(0.139)
Neutral	0.088	0.189	3.932	1.817	8.604
Std. Err.	(0.014)	(0.035)	(1.149)	(0.431)	(0.133)
N	992	992	992	992	895

Notes: Each cell provides the mean of the measure listed in each column. Standard errors are provided in brackets. Estimates under each of the four effort variables are intended to capture the effect of a treatment on human capital investment for the total study population, and are therefore averaged over both investors and non-investors.

also indicates that subjects in both treatments scored roughly the same on the final exam. This is allowed for by the theory, where predictions for the overall population are qualitatively ambiguous, but later on we will see a different story when we separate these measures by demographic group.

Table 3 provides statistical tests for the raw differences displayed in table 2. In the first row we run a simple regression using a dummy for the quota treatment, meaning it represents the experimental difference between an RQ rule and PRO allocations at the population level (*i.e.*, including both demographic groups). Each cell in the table represents a separate regression with the outcome variable labeled in the column header. We report the point estimate and p-value for a test of the hypothesis that there is no difference by treatment group. From the table we see strong evidence that the representative quota increases the fraction of subjects willing to invest at least some time. We also see evidence that it induces them to experiment with more topics, as well as increase the total time invested and number of questions attempted. Although these last two differences are only marginally significant, the estimated magnitudes are large, with RQ subjects logging an estimated 57% and 70% more inputs of time and question attempts, respectively. Additional controls (pre-test score and/or school fixed effects) are added in the bottom two rows as a check on the effectiveness of our randomization. It seems to have worked well, as adding these additional 11 variables caused no significant shifts in point estimates.

5.2.2. ***Testing Differences by Treatment Within Demographic Groups***. Recall that the theory allows for preferential treatment to have differential effects by ability (*i.e.*, cost of competition) and demographic group. In Table 4, we add a demographic dummy to investigate this claim. The first five columns each present estimates for a regression equation of the form

$$Outcome = \beta_0 + \beta_1 Quota + \beta_2 Advantaged * Quota + \beta_3 Advantaged + \beta_4 Pre\text{-}Test + \varepsilon,$$

where $Quota$ is a dummy for treatment status, $Advantaged$ is a demographic dummy, $Pre\text{-}Test$ is a subject's pre-test score in standard deviation units, and the specific $Outcome$ variable is labeled in

Table 3. TESTING DIFFERENCES BY TREATMENT

	Investment				**Performance**
	Used Website	# Topics Attempted	Total Time	# Questions Attempted	Final Exam Score
Quota − Neutral	0.066***	0.095*	2.701	0.912	0.076
P-Value:	*[0.001]*	*[0.061]*	*[0.107]*	*[0.146]*	*[0.693]*
(*Controls: none*)					
Quota − Neutral	0.065***	0.093*	2.650	0.884	0.097
P-Value:	*[0.002]*	*[0.067]*	*[0.113]*	*[0.158]*	*[0.576]*
(*Controls: pre-test scores*)					
Quota − Neutral	0.058***	0.078	2.404	0.773	0.164
P-Value:	*[0.005]*	*[0.130]*	*[0.158]*	*[0.224]*	*[0.346]*
(*Controls: pre-test scores, school FEs*)					
N	992	992	992	992	895

Notes: Each cell represents a separate regression. The number reported is the coefficient for the quota treatment. Row 1 includes no controls and provides a statistical test of the differences in Table 1. Row 2 includes control for practice test score. Row 3 includes school fixed effects. P-values for a two-sided test of the null hypothesis of zero difference are italicized and in brackets. Estimates under each of the four effort variables are intended to capture the effect of a treatment on human capital investment for the total study population, and are therefore averaged over both investors and non-investors.

the column header.[19] With the inclusion of the interaction term $Advantaged*Quota$, the coefficient β_1 represents the average effect of the representative quota specifically on the disadvantaged group. The effect of the policy on the advantaged group is represented by the sum $\beta_1+\beta_2$. The last column in the table moves pre-test score to the left-hand side, to estimate an alternative measure of the treatment on math improvement:

$$FinalExamScore - \textit{Pre-TestScore} = \beta_0 + \beta_1 Quota + \beta_2 Advantaged * Quota + \beta_3 Advantaged + \varepsilon.$$

For completeness, all regressions include controls for school-level fixed effects, and for the primary effects of interest we report p-values in brackets.

For disadvantaged group subjects we find evidence of large and positive effects across all four investment measures. First, we see a highly significant 8.7 percentage point increase in disadvantaged subjects' willingness to spend at least some time on the website when they receive preferential treatment. To put this in perspective, we can compute a within-demographic percent change for

[19]One potential concern with including pre-test score is that it provides a noisy measure of initial human capital, and may therefore introduce an attenuation bias. As a robustness check we re-estimated the regressions in Table 4, omitting *Pre-Test*, and nearly all coefficient estimates and standard errors—except for the ones connected to β_3, the multiplier on higher-grade cohort status—remained virtually unchanged. See Table 6 in the online appendix for coefficient estimates with *Pre-Test* omitted as a control.

Table 4. TESTING DIFFERENCES BY DEMOGRAPHICS AND TREATMENT

	Investment				Performance	
	Used Website	# Topics Attempted	Total Time	# Questions Attempted	Final Exam Score	Exam Score Change
Constant ($\widehat{\beta}_0$)	0.073***	0.144**	3.033	1.316*	8.147***	0.096
Std. Err.	(0.024)	(0.059)	(1.982)	(0.741)	(0.209)	(0.240)
Quota ($\widehat{\beta}_1$)	0.087***	0.146*	5.517**	1.312	0.624**	0.591*
Std. Err.	(0.033)	(0.083)	(2.757)	(1.030)	(0.287)	(0.330)
P-Value:	*[0.009]*	*[0.077]*	*[0.046]*	*[0.203]*	*[0.030]*	*[0.074]*
Advantaged $*$ *Quota* ($\widehat{\beta}_2$)	-0.047	-0.111	-5.034	-0.866	-0.712**	-0.546
Std. Err.	(0.042)	(0.105)	(3.506)	(1.310)	(0.360)	(0.415)
Advantaged ($\widehat{\beta}_3$)	0.028	0.083	1.613	0.877	0.488*	-0.324
Std. Err.	(0.030)	(0.076)	(2.545)	(0.951)	(0.264)	(0.300)
Pre-Test (standardized) ($\widehat{\beta}_4$)	0.029***	0.045*	1.009	0.572*	1.280***	*N/A*
Std. Err.	(0.011)	(0.027)	(0.893)	(0.334)	(0.092)	*N/A*
School Fixed Effects	yes	yes	yes	yes	yes	yes
N	992	992	992	992	895	895
Additional Test: Effect of Quota on Advantaged Group						
$\widehat{\beta}_1 + \widehat{\beta}_2$	0.040	0.035	0.483	0.446	-0.089	0.045**
P-Value:	*[0.123]*	*[0.586]*	*[0.823]*	*[0.581]*	*[0.684]*	*[0.030]*

Notes: Each column is a separate regression. Advantaged is an indicator variable for whether the subject is a 6^{th} or 8^{th} grader (the older group in each school type). We also include each subject's standardized pre-test score, where standardization is based on the mean and variance within each school type (*i.e.*, $5^{th}/6^{th}$ or $7^{th}/8^{th}$), in order to control for differences in starting human capital. Standard errors are in parentheses; p-values for a two-sided test of the null hypothesis of zero effect are italicized and in brackets. Estimates under each of the four effort variables are intended to capture the effect of a treatment on human capital investment for the total study population, and are therefore averaged over both investors and non-investors.

the disadvantaged group by $100 \times (\beta_1/\beta_0)\%$, which amounts to an increase of 119% on the extensive margin, relative to their disadvantaged counterparts under the color-blind treatment. We also see a significant and even larger increase in terms of time investment: disadvantaged subjects under treatment RQ increased investment by 181%. The other two measures capture specific tasks done during time spent on the website: number of topics attempted and number of questions attempted. Although the latter is only marginally significant, both render large point estimates for increases of 101% and 100%, respectively.

Another striking feature of the table are the performance measures. We find a large and significant difference in final exam scores and exam score changes for disadvantaged subjects with preferential treatment: both measures estimate that they lifted their scores by roughly a fifth of a standard deviation, relative to their disadvantaged counterparts in the control group. Although some portion of this effect may also be due to increased effort and concentration on the day of the final exam, we interpret this result and the other columns in Table 4 as evidence that treatment RQ altered labor-leisure trade-offs to induce additional study effort at home for $\mathcal{D}$ subjects, which in turn lead to more learning and increased math proficiency.

One concern is that this strengthening of incentives for disadvantaged subjects may come at the cost of weakening incentives for advantaged subjects. However, Table 4 also shows that this concern is not supported by the evidence. For 4 out of 5 outcome measures, point estimates for the effect of RQ on group $\mathcal{A}$, given by $\beta_1 + \beta_2$, was actually positive, but insignificant. In column 5, the sum of the two coefficients is slightly negative (representing about 3% of a standard deviation) but with a large p-value. The two outcome measures under which $\beta_1 + \beta_2$ is most significant are the binary measure of investment (column 1), with a p-value of 0.123, and the change in exam score (column 6), with a p-value of 0.03. This implies an estimated percent change of $100 \times (\beta_1 + \beta_2)/(\beta_0 + \beta_3) = +39.6\%$ on the extensive margin for advantaged subjects studying under a quota, with an estimated gain of about 1.5% of a standard deviation in their test scores. Thus, we do not find evidence that there is a trade-off between average human capital investment across demographic groups; if anything the data seem to slightly favor a small increase of investment for the advantaged demographic as well.

5.2.3. ***Selective Attrition***. One potential source of bias in our results concerning the performance measure (final exam score) is that 97 of the subjects who took the practice test and were randomly assigned to a competition group (9.8%) did not show up on the day of the final test.[20] We find that among the disadvantaged subjects, those assigned to the quota group were less likely to miss the final exam (10.6% vs. 16.7%). We also find that among the subjects who did not show up for the final test, the disadvantaged subjects assigned to the quota group had higher practice scores than the disadvantaged subjects not assigned to the quota group (7.16 vs 6.35). However, the practice scores among the subjects who did show up for the final test were nearly the same across these two groups (7.91 vs 7.79). These comparisons all point in a direction opposite of our main results and suggest that the effect of the quota on final performance for disadvantaged subjects may have been greater in the absence of this selective attrition.

5.2.4. ***Differing Policy Responses by Ability Level***. The theoretical model predicts that if the underlying cost types for the disadvantaged group stochastically dominate those in the advantaged group, then qualitative patterns like those displayed in Figure 2 should appear. Namely, for group $\mathcal{D}$ the test score distributions under a quota and color-blind mechanism should have a unique interior crossing point, with the quota CDF strictly above the color-blind CDF to the right of the crossing point, and strictly below to the left. In other words, there should be a positive mass of the best subjects in group $\mathcal{D}$ who decrease output, while subjects of medium and low ability from that group

[20]This problem does not arise with the four investment measures, which did not require observing a final score for us to estimate them. In Table 4 the sample size for the first four columns represent the full sample of test subjects.

increase output. Moreover, the upper bound on the output distribution may also be higher for $\mathcal{D}$ under a quota. The theory makes opposite predictions for group $\mathcal{A}$.[21]

While it is impossible to directly observe the distributions of cost types, we can take queues from the distribution of pre-test scores by demographic group, since they reflect initial math proficiency, as well as how much progress each subject will need to increase her payout. We can then examine the distributions of final exam scores within demographic groups under different treatments to see whether our experimental data seem to be consistent with the theory of incentive effects under preferential treatment. Figures 3 – 5 depict these comparisons in three plots of empirical cumulative distribution functions (CDFs) for pre-test and final exam scores for grades 7 and 8. For the sake of comparability, we have limited our sample in these figures to include only subjects for whom we have both test scores. Therefore, Figure 3 plots empirical pre-test CDFs only for 7^{th} and 8^{th} graders who took the final exam.

Figure 3 strongly supports stochastic dominance of initial math proficiency levels across demographic groups. A two-sample Kolmogorov-Smirnov (KS) test rejects the null hypothesis that the 7^{th} and 8^{th} grade distributions are the same, against a one-sided alternative that the latter stochastically dominates with a p-value of 1.03×10^{-5}. This means group $\mathcal{D}$ subjects, on average, had to achieve more progress in order to be competitive for a prize. This is not the same as observing costs, but the two are certainly related and the hypothesis of stochastic dominance in cost types appears plausible.

We find evidence in Figure 4 that by the end of the study period the score distribution within the 7^{th} grade subsample had diverged by treatment status. As a first attempt at investigating the difference, we perform two-sided KS test for group $\mathcal{D}$ final exams across treatments, which results in a p-value of 0.1047, providing marginally significant evidence that the two distributions were not the same.[22] There appears to be a hint of divergence by treatment group among 8^{th} graders in Figure 5 as well, though the effect is weak and we lack sufficient power for a KS test to distinguish between the two distributions (the p-value for a two-sided test is 0.889).

For our application, the KS test has important limitations. Aside from being a relatively low-power test, it also does not provide a readily interpretable measure for the significance and magnitude of effects at a given point in the distribution. In order to evaluate theoretical predictions concerning how the representative quota shapes incentives differently for subjects who vary by ability levels, we wish to make within-group, cross-treatment comparisons which allow for the magnitude, and even the sign, of the effect to vary across different quantiles. One challenge in making such comparisons

[21] Although the figures we present in this paper explore distributions of both inputs and outputs, it is important to remember that the predictions of the theory only directly apply to exam score, as this is the variable on which prize allocations are based. The mapping from inputs to outputs may vary by subject if each one differs by raw math talent and leisure preference.

[22] Figures 12 and Figure 13 in the Appendix contain additional plots comparing pre-test scores by treatment within demographic groups. The differences in the pre-test distributions are due to selective attrition after we omitted subjects for whom we have no final exam. The figures suggest that in general selective attrition is working against our results presented here. For group $\mathcal{D}$ the pre-test color-blind distribution is below the pre-test quota distribution for values at or below the median, and the upper bound of the pre-test distribution for 7^{th} grade quota subjects is highest. Both characteristics of these distributions are substantially reversed by the final exam. Figure 13 suggests that selective attrition within the 8^{th} grade is also working slightly against finding a significant difference in the direction predicted by theory.

FIGURE 3. PRE-TEST SCORES: 7^{th} GRADE VS 8^{th} GRADE

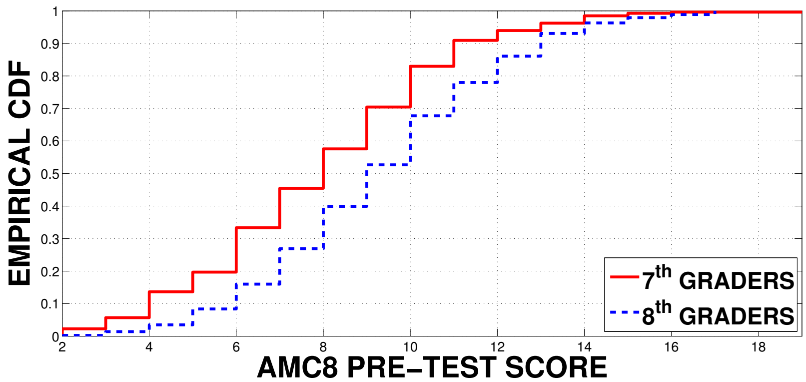


FIGURE 4. SEVENTH GRADE FINAL EXAM SCORES

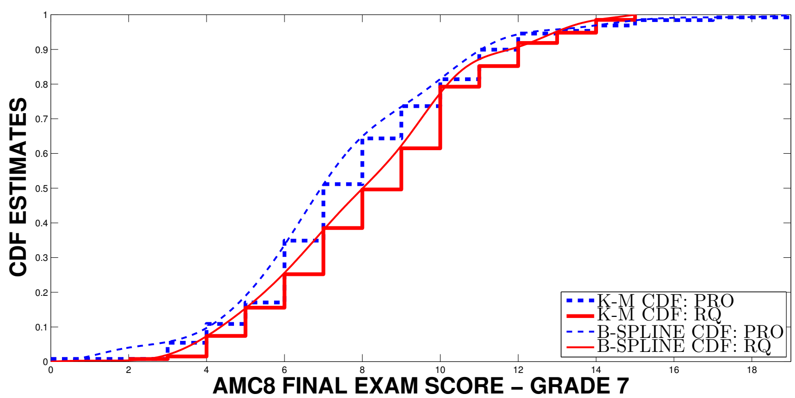
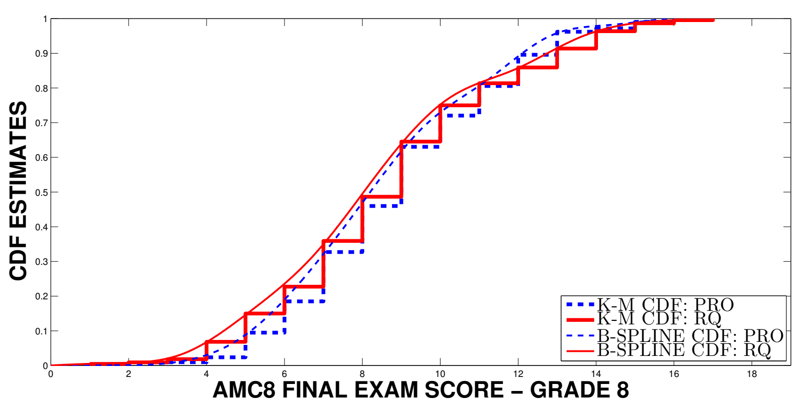


FIGURE 5. EIGHTH GRADE FINAL EXAM SCORES



Notes: For the sake of comparability, Figures 3 – 5 above use only data for 7^{th} and 8^{th} graders who took both the pre-test *and* final exam. See Figures 12 – 13 in the appendix for a discussion on the role of selective attrition.

is that the set of quantile ranks $q \in [0, 1]$ attained by the empirical CDFs for two different samples will rarely overlap. To overcome this problem, we construct smoothed quantile functions based on a flexible B-spline CDF estimator.[23] Our smoothed quantile function estimator allows for comparisons at arbitrary quantile ranks q, with differences that are easily interpretable, being in the original test score units.

Specifically, within each treatment $t \in \{RQ, PRO\}$ and group $j \in \{\mathcal{A}, \mathcal{D}\}$ we partition the support of final exam scores $[0, \overline{h}_j^t]$ into K uniform subintervals and we specify the breakpoints between these intervals as a vector of $K + 1$ knots (including endpoints). These knots define a set of $K + 3$ B-spline basis functions, denoted $\mathcal{B}_{j,k}^t : [0, \overline{h}_j^t] \to \mathbb{R},\ k = 1, \dots, K + 3$, with which we parameterize the CDF of final exam scores as

$$\hat{G}_j^t(h; \boldsymbol{\alpha}_j^t) = \sum_{k=1}^{K+3} \alpha_{j,k}^t \mathcal{B}_{j,k}^t(h).$$

Finally, we estimate the parameter vector $\boldsymbol{\alpha}_j^t$ to achieve a constrained, least-squares, best fit to the empirical CDF, where constraints ensure that $\hat{G}_j^t$ is monotone and attains a value of 0 at the lower bound, and 1 at $\overline{h}_j^t$, or the maximum observed score for group j in treatment t. Our method belongs to the class of *sieve estimators* since, if we allow the number of knots $(K + 1)$ to grow with the sample size, a B-spline can accommodate arbitrary shapes for any well-behaved CDF. We chose a partition of $K = 10$ uniform subintervals for each support, which results in a flexible CDF estimator with 11 free parameters once the terminal conditions are enforced. Figures 4 and 5 display a comparison of the Kaplan-Meier empirical CDFs (thick lines) and the smoothed, B-spline CDFs (thin lines) for each treatment and age group.

With these smoothed CDF estimates in hand, we can construct quantile functions, which we denote by $\hat{h}_j^t(q) = (\hat{G}_j^t)^{-1}(q),\ q \in (0, 1)$. Following the notation of Proposition 2 in Section 3, from this we can compute the quantile difference function,

$$\hat{\Delta}_j(q) \equiv \hat{h}_j^{RQ}(q) - \hat{h}_j^{PRO}(q),\ q \in [0, 1],\ j = \mathcal{A}, \mathcal{D}.$$

This is the primary object of interest, as it allows us to make behavioral comparisons across policies at specific quantiles. Moreover, in order to investigate the role of sampling variability we execute a bootstrap routine, wherein for each age group $j = \mathcal{A}, \mathcal{D}$, we re-sampled from the quota and color-blind samples (with replacement) and re-estimated $\hat{\Delta}_j(\cdot)$ 50,000 times. The left panels of Figures 6 and 7 display point estimates and 95% confidence bounds from this exercise; for each, quantile ranks $q \in [0, 1]$ are on the horizontal axis, and differences displayed in final exam score units are on the vertical axis. Point estimates for the functions $\Delta_{\mathcal{A}}(\cdot)$ and $\Delta_{\mathcal{D}}(\cdot)$ conform well to the patterns predicted in the theory: both have a unique interior point where they cross the zero line. For group

[23]B-splines constitute a parametric family of functions which combines flexibility, stability, and computational convenience. They behave similarly as piece-wise, local polynomials, but afford greater numerical convenience, being made up of globally defined basis functions, like orthogonal polynomials. Moreover, incorporating shape restrictions (*e.g.*, monotonicity and terminal conditions for a CDF function), is quite simple within the B-spline family. A standard text on B-splines is Carl de Boor [2001]. For a brief primer on B-splines and their advantages for empirical economics, see Brent R. Hickman, Timothy P. Hubbard and Harry J. Paarsch [2017, Appendix].

FIGURE 6. SEVENTH GRADE FINAL EXAM SCORES

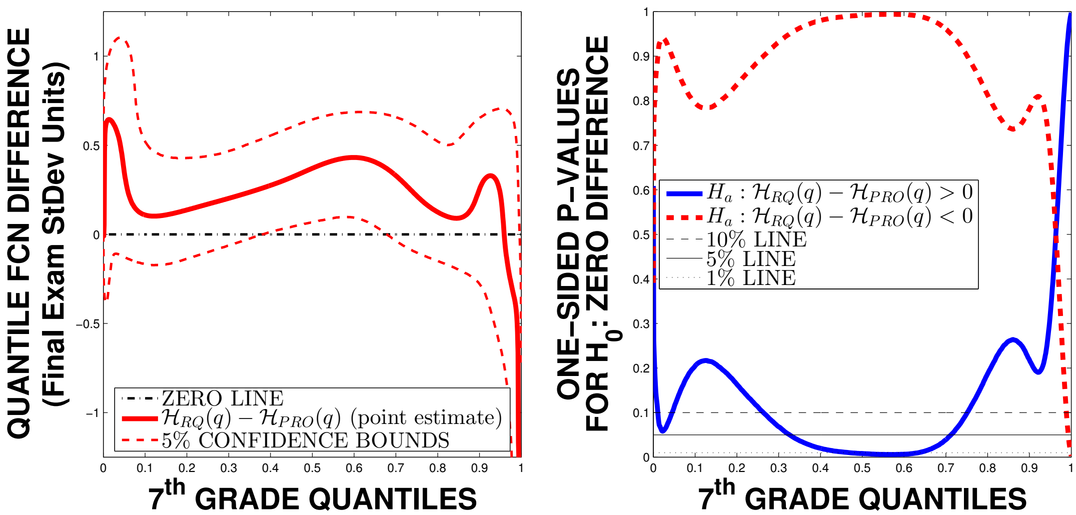
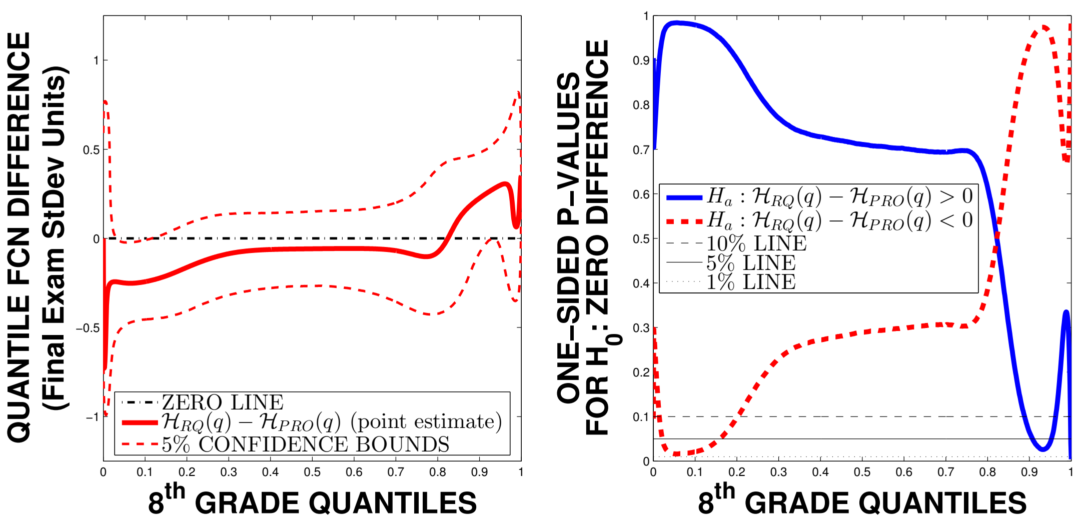


FIGURE 7. EIGHTH GRADE FINAL EXAM SCORES



$\mathcal{D}$, achievement under a quota within the interquartile range improves substantially, by a margin of 17%–43% of a standard deviation. There is a difference in the opposite direction for $\mathcal{A}$, but it is much smaller, with a point estimate from 6%–12% of a standard deviation within the interquartile range.

The right panels in the two figures illustrate the strength of the statistical evidence for the quantile-specific differences by evaluating the null hypothesis $\mathcal{H}_0 : \Delta_{\mathcal{D}}(q) = 0$ at each point, against the two alternative hypotheses where the sign is negative, denoted by $\mathcal{H}_a^-$, or positive, denoted by $\mathcal{H}_a^+$. At each point $q \in [0, 1]$ we computed point-wise one-sided P-values as the fraction of bootstrapped samples resulting in a positive difference (solid lines), or a negative difference (dashed lines) at q. Each panel also displays thin horizontal lines for the 10%, 5%, and 1% thresholds for

Table 5. NARROWING GAPS

	Mean	Median	Std. Dev.	N
Achievement Gaps for All Treatments				
Standardized Pre-Score (GRADE 7)	-0.295	-0.267	0.996	264
Standardized Pre-Score (GRADE 8)	0.181	0.071	0.960	431
Standardized Final Score (GRADE 7)	-0.138	-0.205	1.005	264
Standardized Final Score (GRADE 8)	0.085	0.142	0.987	431
Achievement Gaps for Quota Treatment				
Standardized Pre-Score (GRADE 7)	-0.211	-0.199	1.061	135
Standardized Pre-Score (GRADE 8)	0.231	0.135	0.920	220
Standardized Final Score (GRADE 7)	-0.030	0.142	0.961	135
Standardized Final Score (GRADE 8)	0.055	0.142	1.050	220
Achievement Gaps for Color-Blind Treatment				
Standardized Pre-Score (GRADE 7)	-0.243	-0.199	0.904	129
Standardized Pre-Score (GRADE 8)	0.259	0.135	0.982	211
Standardized Final Score (GRADE 7)	-0.251	-0.552	1.041	129
Standardized Final Score (GRADE 8)	0.115	0.142	0.921	211

Notes: There are three separate panels in the table, each containing standardized scores on the pre-test and post-test. Standardization was performed within each panel-test grouping, excluding scores for subjects who missed the final exam. For example, pre-test scores for the quota treatment were standardized using the mean and standard deviation of pre-test scores for all 7^{th} and 8^{th} graders in the quota treatment who took both the pre-test and post-test.

comparison. The figures show strong evidence for a positive interquartile range difference in group $\mathcal{D}$, and they also show some evidence of the opposite effect near the upper bound of the support. The evidence for policy-induced shifts in group $\mathcal{A}$ is much weaker, but there is some evidence of statistically significant changes near the extremes of the sample, and once again, in directions consistent with those predicted by theory.

5.2.5. ***Narrowing Achievement Gaps.*** We now conclude analysis of our experimental data with a look at the tendency for a preferential treatment policy to narrow achievement gaps across demographic groups. Table 5 displays summary statistics on standardized test scores for the pre-test and final exam, for grades 7 and 8. In the top panel of the table scores were standardized within each exam by subtracting the mean and dividing by the standard deviation for all grade 7 and

8 subjects.[24] Therefore, the means indicate distance between the population average and grade cohort average, in standard deviation units. Without accounting for treatment status we see that 7^{th} grade subjects were roughly half of a standard deviation behind their 8^{th} grade counterparts on average—or $-0.295 - 0.181 = -0.476$ standard deviations to be exact—but by the final exam, the gap between 7^{th} and 8^{th} graders had narrowed by about half—to $-0.138 - 0.085 = -0.223$ standard deviations.

In the lower two panels of Table 5 we break out this effect by treatment group, displaying the same numbers within treatments, but where score standardization now happens within each exam-treatment cell. Part of the test score convergence had to do with differences in the conditions of the pre-test and final exams (likely due to incentives or slightly different content): within the color-blind treatment, about a quarter of the gap disappeared but remained relatively high at 0.367 standard deviations on the final exam. However, the achievement gap under the representative quota closed substantially more, by about 80%, beginning at 0.442 standard deviations, and ending at only 0.085 standard deviations on the final exam.

The medians tell a slightly stronger story, with the median gap beginning about the same within both treatments, closing virtually to zero under the representative quota, and closing only slightly otherwise. Finally, observing that the within-treatment-demographic standard deviations are all close to one suggests that the narrowing of gaps within the two treatments was due predominantly to mean/median shifts in test scores. We interpret these findings as evidence that affirmative action can actually help to narrow achievement gaps by catching up its target demographic to the rest of the population.

6. DISCUSSION AND CONCLUSION

Our theory and field experiment together provide compelling evidence for the role of relative incentives in rank-order contests of human capital investment. From the perspective of the lower-grade group, we create a counterfactual scenario which can be seen to represent either an exogenous shift in outside competition, or to represent the workings of a rule that grants preferential treatment based on demographic status. These phenomena arise in a host of real-world applications, including the increasingly competitive college admissions market, labor markets, military career track assignments, and many others. The purpose of this paper is not to make claims about specific magnitudes of these effects in any particular context; rather our aim is to provide evidence for the empirical relevance of relative incentives in shaping investment behavior, an idea which has been largely overlooked in the human capital literature. Indeed, one important question is whether agents can be thought to display the sophistication needed to react to subtle changes in their competitive environment, even when their own production technology and the set of potential match partners remain fixed. As we have seen above, even middle school children seem to possess a remarkable ability to produce behavioral patterns that conform to Bayes-Nash equilibrium play under shifting

[24]Once again, in order to make the pre-test and final exam figures comparable, we excluded from the analysis any subjects whose final scores were missing due to attrition.

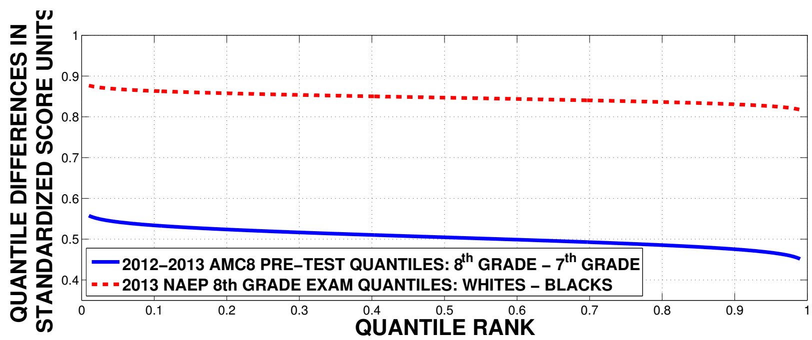
relative incentives. We now conclude by discussing potential differences between our experimental design and some prominent real-world examples.

6.1. **AFFIRMATIVE ACTION.** Affirmative action often denotes the practice of granting preferential treatment based on demographic status in allocating educational opportunities. In the United States, affirmative action has taken race-based forms—most common, favoring African American and Hispanic college applicants—and income-based forms—most notably with the Texas Top 10% program which guarantees admittance to a Texas flagship school for anyone in the top 10% of their high school class. Often the latter is used as a substitute when legal obstacles prevent the former. The rationale behind this policy is based on providing a partial remedy for systemic disadvantages which are highly correlated with demographic status—racial minorities and/or poor students who often receive inadequate funding for their primary/secondary education. Our model speaks to the qualitative impact on incentive provision one might expect from affirmative action which targets disadvantaged groups whose human capital investment costs are systematically higher than the rest of the population. Our experiment shows that middle-school aged children collectively respond to affirmative action-like incentives in a way that is highly-consistent with a theoretical model of a large-scale human capital investment contest.

While interpreting the results from this study, however, it is important to keep in mind what it can and cannot say. The ultimate question of interest is the effect that real world affirmative action has on academic incentives in practice. There are several reasons why this study cannot fully address this policy question. First, in executing field experiments with actual schools, a significant hurdle involves forming partnerships with school administrators whose primary job it is to educate, rather than to produce research. Our school partners in this experiment serve a fairly homogeneous population. On one hand, using grade level as an observable delimiter of learning costs produces a clean exploration of theoretical predictions. On the other hand, at best it serves as an imperfect proxy for race or childhood family income. Although studies on income-based achievement gaps are less common, Figure 8 provides a comparison between our sample pool and the broader US population of pre-college human capital investors. The figure shows that our grade-level delimiter between the advantaged and disadvantaged groups likely understates the competitive difference in math proficiency between blacks and whites vying for actual college seats. The graph compares the difference between black and white 8^{th} grade NAEP quantiles—expressed in standardized NAEP score units (dashed line)—to the difference in 7^{th} and 8^{th} grade quantiles from our sample—expressed in standardized AMC8 score units (solid line). At most NAEP quantiles, black 8^{th} graders trail their white counterparts by about 0.85 standard deviations, whereas the gap among our test subjects is closer to 0.5 standard deviations. Another way to interpret the figure is that wherever the dashed line lay above the solid line indicates that the corresponding quantiles of American Blacks are *more than* one full grade level behind their white, 8th Grade counterparts.

This picture suggests real world discouragement effects minorities face in the competitive college market could be larger than those found in our experimental study. However, other factors may play a role as well, including differences in home environments or cultural factors which may push

FIGURE 8. GRADE-LEVEL DIFFERENCES VS BLACK/WHITE DIFFERENCES



(A) This figure first standardizes AMC8 and NAEP exam scores by subtracting the mean and dividing by the standard deviation for the relevant sample population. Then, the distribution of standardized test scores for each sub-sample–*i.e.*, AMC8 7^{th} grade, AMC8 8^{th} grade, NAEP 8^{th} grade whites, and NAEP 8^{th} grade blacks–is approximated by a normal distribution, which appears to fit the data quite well. The resulting quantile functions are subtracted from each other in order to illustrate relative disparities between groups.

the results in other, less predictable directions. Certainly though, in order to understand how and why a policy works, it is important to cleanly understand the incentive dimension in addition to other sociocultural factors.

6.2. **SHORT-RUN VS. LONG-RUN INCENTIVES.** The second major limitation in the current analysis is that, although it creates a set of incentives which mirror human capital competition in key ways, due to feasibility constraints said incentives are only engineered on a small scale and measured over a short-run horizon. We tracked our students over a period of 10-days during which they were given the opportunity to invest leisure time into their math proficiency. For students in our sample who logged positive amounts of study time, their average expected wage was $10.73/hour, measured as total combined earnings divided by total observed time spent studying on the website, across all students. Junior high and high school students preparing for college optimize labor-leisure division over a much longer horizon (4-5 years until college applications are due), with much larger payoffs that come far in the future.[25] The question of quantifying incentives running up to the US college market is complex and involves many factors, including time preferences, cost of a college education, and many other things that vary widely across individuals.

6.3. **Conclusion.** We present a model of a large contest, in which many heterogeneous agents compete for many heterogenous prizes, which better captures the competitive incentives faced by students competing for university admissions, or workers competing for career advancement. We use to model to predict how a policy that provides preferential treatment (by way of reserving prizes) for a disadvantaged group of agents changes the distribution of effort with in both the targeted and

[25]A large literature has focused on the wage return to higher college quality, including Stacy Berg Dale and Alan B. Krueger [2002], Dan A. Black and Jeffrey A. Smith [2006], and Long [2008]. Although there is some disagreement as to the magnitude for all students, there is broad consensus that the return to attending higher quality colleges is economically significant for poor students.

non-targeted populations. The results give insight into how effort and performance shift under a representative quota affirmative action policy.

To test the theoretical predictions, we design a field experiment in which hundreds of students compete for heterogeneous cash prizes, where we paid students based on their performance relative to other experiment participants on a nationwide math exam. Our experimental design makes a number of contributions. Ours is the first controlled experiment of a large contest. Our focus is on how incentives change the distribution of effort and performance, rather than simply the average level of investment and performance. We observe student study effort, and can therefore have a direct measure of how the competitive incentives change student investment in developing human capital.

Our experimental results show that students respond to preferential treatment policies in a manner that is highly consistent with the theoretical model. Although the highest-ability students in the group targeted by the representative quota decrease their effort, the majority of the targeted population increases its effort and performance. Students who are not targeted by the policy react in largely the opposite way, but to a smaller magnitude. We show that a preferential treatment policy can narrow the achievement gaps between advantaged and disadvantaged students, without having a substantial negative impact on the average effort or performance of those not targeted by the policy.

References

Arcidiacono, Peter, Esteban Aucejo, and Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." *American Economic Review*, 106(3): 525–562.

Becker, Gary. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Bettinger, Eric. 2012. "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores." *Review of Economics and Statistics*, 94: 686–698.

Black, Dan A., and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality Using Multiple Proxies for Quality." *Journal of Labor Economics*, 24(3): 701–728.

Bodoh-Creed, Aaron. 2013. "Efficiency and Information Aggregation in Large Uniform-Price Auctions." *Journal of Economic Theory*, 148: 2436–2466.

Bodoh-Creed, Aaron, and Brent R. Hickman. 2017. "College Assignment as a Large Contest." *working paper, University of Chicago.*

Bowen, William G., and Derek Bok. 1998. *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions.* Princeton, NJ: Princeton University Press.

Bracha, Anat, Alma Cohen, and Lynn Conell-Price. 2015. "Affirmative Action and Stereotype Threat." working paper.

Calsamiglia, Caterina, Jorg Franke, and Pedro Rey-Biel. 2013. "The incentive effects of affirmative action in a real-effort tournament." *Journal of Public Economics*, 98: 15–31.

Cason, Timothy N., William A. Masters, and Roman M. Sheremeta. 2010. "Entry into winner-take-all and proportional-prize contests: An experimental study." *Journal of Public Economics*, 94: 604–611.

Chambers, David L., Timothy T. Clydesdale, William C. Kidder, and Richard O. Lempert. 2005. "The Real Impact of Eliminating Affirmative Action in American Law Schools: An Empirical Critique of Richard Sander's Study." *Stanford Law Review*, 57: 1855–1897.

Che, Yeon-Koo, and Ian Gale. 2003. "Optimal Design of Research Contests." *American Economic Review*, 93(3): 646–671.

Coate, Stephen, and Glenn Loury. 1993. "Anti-Discrimination Enforcement and the Problem of Patronization." *American Economic Review*, 83(2): 92–98.

Cotton, Christopher, Frank McIntyre, and Joseph Price. 2013. "Gender differences in repeated competition: Evidence from school math contests." *Journal of Economic Behavior and Organization*, 86: 52–66.

Dale, Stacy Berg, and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics*, 117: 1491–1528.

de Boor, Carl, ed. 2001. *A Practical Guide to B-Splines, Revised Edition.* New York: Springer-Verlag.

Dillon, Elanor Wiske, and Jeffrey Andrew Smith. 2017. "The Consequences of Academic Match Between Students and Colleges." *CES ifo Working Paper Series*, 6344.

Ferman, Bruno, and Juliano Assuncao. 2011. "Does Affirmative Action Enhance or Undercut Investment Incentives? Evidence from Quotas in Brazilian Public Universities." *Typescript, Massachusetts Institute of Technology Department of Economics.*

Fryer, Roland. 2011. "Financial Incentives and Student Achievement: Evidence From Randomized Trials." *Quarterly Journal of Economics*, 126: 1755–1798.

Fullerton, Richard L., and R. Preston McAfee. 1999. "Auctioning Entry into Tournaments." *Journal of Political Economy*, 107(2): 573–605.

Galton, Francis. 1902. "The Most Suitable Proportion Between the Values of First and Second Prizes." *Biometrika*, 1(4): 380–390.

Hickman, Brent R., and Aaron Bodoh-Creed. 2017. "Human Capital Investment and Affirmative Action: A Structural Policy Analysis of US College Admissions." *Typescript, University of Chicago Department of Economics.*

Hickman, Brent R., Timothy P. Hubbard, and Harry J. Paarsch. 2017. "Identification and Estimation of a Bidding Model for Electronic Auctions." *Quantitative Economics*, 8(2): 505–551.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn." *Review of Economics and Statistics*, 91: 437–456.

Leuven, Edwin, Hessel Oosterbeek, and Bas van der Klaauw. 2010. "The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment." *Journal of the European Economic Association*, 8: 1243–1265.

Long, Mark C. 2008. "College Quality and Early Adult Outcomes." *Economics of Education*

Review, 27: 588–602.

Loury, Linda Datcher, and David Garman. 1995. "Selectivity and Earnings." *Journal of Labor Economics*, 13(2): 289–308.

Moldovanu, Benny, and Aner Sela. 2001. "The Optimal Allocation of Prizes in Contests." *American Economic Review*, 91(3): 542–558.

Moldovanu, Benny, and Aner Sela. 2006. "Contest Architecture." *Journal of Economic Theory*, 126(1): 70–96.

Olszewski, Wojciech, and Ron Siegel. 2016. "Large Contests." *Econometrica*, 84: 835–854.

Rothstein, Jesse, and Albert H. Yoon. 2008. "Affirmative Action in Law School Admissions: What do Racial Preferences Do?" *University of Chicago Law Review*, 75(2): 649–714.

Sander, Richard H. 2004. "A Systemic Analysis of Affirmative Action in American Law Schools." *Stanford Law Review*, 57: 367–483.

Schotter, Andrew, and Keith Weigelt. 1992. "Asymmetric Tournaments, Equal Opportunity Laws and Affirmative Action: Some Experimental Results." *Quarterly Journal of Economics*, 107(2): 511–539.

Sheremeta, Roman M. 2010. "Experimental comparison of multi-stage and one-stage contests." *Games and Economic Behavior*, 68: 731–747.

Spence, Michael. 1973. "Job Market Signalling." *The Quarterly Journal of Economics*, 87(3): 355 – 374.

7. APPENDIX

7.1. THEORETICAL PROOFS APPENDIX.

7.1.1. ***Proof of Theorem 1.*** For the reader's convenience Theorem 1 is restated below:

Theorem 1: *Consider two pure rank order contests,* 1 *and* 2*, which differ only by their cost distributions, and assume competition is more fierce under contest* 2 *in the sense that* F_1 *LR dominates* F_2*. Let* $\tilde{\theta} \in (\underline{\theta}, \overline{\theta})$ *denote the unique crossing point of the density functions where* $f_1(\tilde{\theta}) = f_2(\tilde{\theta})$*. There exists a unique interior crossing point* $\ddot{\theta} \in (\underline{\theta}, \tilde{\theta})$*, such that* $h_1^*(\theta) < h_2^*(\theta)$ *for all* $\theta < \ddot{\theta}$ *and* $h_1^*(\theta) > h_2^*(\theta)$ *for all* $\theta > \ddot{\theta}$*.*

Proof: As mentioned above, the LR dominance property implies first-order stochastic dominance. Therefore, not only do the densities have a unique crossing point, but it must also be true that $f_1(\theta) > f_2(\theta)$ for $\theta > \tilde{\theta}$, and $f_1(\theta) < f_2(\theta)$ for $\theta < \tilde{\theta}$. Since the same boundary condition applies to both contests, $h_1^*(\overline{\theta}) = h_2^*(\overline{\theta}) = \underline{h}$, then equation (5) implies the initial trajectories at the boundary point are ordered in the following way:

$$h_1^{*\prime}(\overline{\theta}) = \frac{-f_1(\overline{\theta})}{c'(\underline{h};\overline{\theta})} < \frac{-f_2(\overline{\theta})}{c'(\underline{h};\overline{\theta})} = h_2^{*\prime}(\overline{\theta}). \tag{A.1}$$

This in turn means that $h_1^*(\theta) > h_2^*(\theta)$ within a neighborhood of $\overline{\theta}$ since the investment functions are continuous and everywhere differentiable. Note that because slopes are negative $h_1^{*\prime}(\theta) < h_2^{*\prime}(\theta)$ means h_1^*, rises in the leftward direction and is *more steep* at θ.

Now suppose there exists at least one point where h_1^* and h_2^* cross, and let $\ddot{\theta} \in (\underline{\theta}, \overline{\theta})$ denote the maximum of all such possible points, with $\ddot{h} \equiv h_1^*(\ddot{\theta}) = h_2^*(\ddot{\theta})$, if any exist. Since h_1^* crosses h_2^* from above at $\ddot{\theta}$, it must be that $h_1^{*\prime}(\ddot{\theta}) \geq h_2^{*\prime}(\ddot{\theta})$ (*i.e.*, h_1^* is *less steep* at the crossing point). However, since $f_1(\theta) > f_2(\theta)$ on $(\widetilde{\theta}, \overline{\theta}]$ by LR dominance, and since $h < h'$ implies $c'(h'; \theta) \geq c'(h; \theta)$ by convexity, the following must be true for any $\theta \in (\tilde{\theta}, \overline{\theta})$:

$$h_1^{*\prime}(\theta) = \frac{-f_1(\theta)}{c'\left[h_1^*(\theta); \theta\right]} < \frac{-f_2(\theta)}{c'\left[h_2^*(\theta); \theta\right]} = h_2^{*\prime}(\theta) \quad \Rightarrow \quad h_1^*(\theta) > h_2^*(\theta). \tag{A.2}$$

Therefore, $\ddot{\theta} < \tilde{\theta}$, if such a point exists. Similarly, since $f_1(\theta) < f_2(\theta)$ on $[\underline{\theta}, \widetilde{\theta})$, then any crossing point would have to obey $h_1^{*\prime}(\ddot{\theta}) = -f_1(\ddot{\theta})/c'(\ddot{h}; \ddot{\theta}) > -f_2(\ddot{\theta})/c'(\ddot{h}; \ddot{\theta}) = h_2^{*\prime}(\ddot{\theta})$. This means that h_1^* can only cross h_2^* from above (*i.e.*, at points where it is *less steep*) and so there can be at most one such crossing.

Finally, to see why a crossing point must exist, suppose for a contradiction that for all $\theta \in (\underline{\theta}, \overline{\theta})$ we have $h_1^*(\theta) > h_2^*(\theta)$. In that case, it follows that

$$\begin{aligned} h_1^*(\underline{\theta}) &= \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{c'\left[h_1^*(u); u\right]} f_1(u) du + \underline{h} \\ &< \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{c'\left[h_1^*(u); u\right]} f_2(u) du + \underline{h} \\ &\leq \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{c'\left[h_2^*(u); u\right]} f_2(u) du + \underline{h} = h_2^*(\underline{\theta}), \quad \rightarrow\leftarrow . \end{aligned}$$

The strict inequality follows because the first and second line depict expectations over the decreasing function $1/c'\left[h_1^*(u); u\right]$ and f_2 places more weight on strictly higher values of u (or lower values of the function). The weak inequality follows from the supposition and from c being convex and having a positive cross-partial derivative. Thus we have a contradiction, so a unique crossing $\ddot{\theta}$ exists on the open interval $(\overline{\theta}, \tilde{\theta})$ and the theorem is proved. ■

7.1.2. ***Proof of Proposition 3.*** For the reader's convenience we re-state the proposition here:
Proposition 3: *Assume the same conditions as in Proposition 2. Then under the RQ rule (relative to the PRO rule), typical disadvantaged agents—that is, group $\mathcal{D}$ agents with costs $\theta \in \left\{(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}] \cup T_{\mathcal{D}}\right\}$—exert higher effort and accumulate more HC. Moreover, if we define $\Delta : (\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}] \cup T_{\mathcal{D}} \to \mathbb{R}$ as the difference on this set between group $\mathcal{D}$ investment under RQ versus PRO, or*

$$\Delta(\theta) \equiv \left(h_{\mathcal{D}}^*(\theta) - h^*(\theta)\right),$$

then $\Delta(\theta)$ is strictly positive and attains a maximum on the interval $(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta}]$. Moreover, if investment costs are strictly convex in h, then $\Delta(\theta)$ attains its maximum on the open interval $(\ddot{\theta}_{\mathcal{D}}, \tilde{\theta})$.

Proof: Corollary 2 directly implies the first part of the results, and that $\Delta(\cdot)$ tends toward zero at its endpoints but is strictly positive everywhere else. Moreover, equation (A.2) in the proof of Theorem 1 (see appendix) establishes that $\Delta'(\theta) < 0$ for all $\theta \in T_{\mathcal{D}}$, meaning the difference between the two investment functions becomes steadily wider as one moves toward the density crossing $\tilde{\theta}$

from the left. If costs are strictly convex, then (A.2) shows that $\Delta'(\theta) < 0$ for all $\theta \in T_{\mathcal{D}} \cup \tilde{\theta}$, so the maximum cannot occur at $\tilde{\theta}$. ■

7.1.3. ***Proof of Theorem 4.*** Once again, we first re-state the result for the reader's convenience: ***Theorem 4:*** *Assume HC production costs and gross utility are linear in h, so that* $c(h;\theta) = \theta(h-\underline{h})$, *and* $u(s,h) = sh$. *Moreover, consider two cost distributions,* $F_1(\theta)$ *and* $F_2(\theta)$, *where competition is more intense under* F_2 *in the sense that* F_1 *LR dominates* F_2 *(i.e., the ratio* $\frac{f_1(\theta)}{f_2(\theta)}$ *is strictly increasing). Then, letting* $\tilde{\theta}$ *denote the unique crossing point of* f_1 *and* f_2, *there exists a unique interior crossing point* $\ddot{\theta} \in (\underline{\theta}, \widetilde{\theta})$ *such that* $h_1^*(\theta) < h_2^*(\theta)$ *for* $\theta < \ddot{\theta}$ *and* $h_1^*(\theta) > h_2^*(\theta)$ *for* $\theta > \ddot{\theta}$.

Proof: Recall that strict LR dominance implies $f_1(\theta) \geq f_2(\theta)$ and $F_1(\theta) < F_2(\theta)$, for $\theta \in [\widetilde{\theta}, \overline{\theta})$. This with $\eta_1(\overline{\theta}) = \eta_2(\overline{\theta})$ and equation (9) together mean that $\eta_1'(\theta) < \eta_2'(\theta)$ and $\eta_1(\theta) > \eta_2(\theta)$, for each $\theta \in [\widetilde{\theta}, \overline{\theta})$. Thus, if η_1 and η_2 cross, the crossing must be on the interval $[\underline{\theta}, \widetilde{\theta})$.

Now, equation (9) can be expressed in integral form by $\eta_j(\theta) = \int_{\theta}^{\overline{\theta}} \frac{f_j(x)}{x+F_j(x)-1} dx + \log(\underline{h})$, $j = 1, 2$. Moreover, if we impose a change of variables $y = F_j(\theta)$ within the integral, we get

$$\eta_1(\underline{\theta}) = \int_0^1 \frac{1}{F_1^{-1}(y) + y - 1} dy + \log(\underline{h}) < \int_0^1 \frac{1}{F_2^{-1}(y) + y - 1} dy + \log(\underline{h}) = \eta_2(\underline{\theta}),$$

where the inequality follows from LR dominance. Therefore, by continuity at least one crossing point exists on the open interval $(\underline{\theta}, \tilde{\theta})$. Let $\ddot{\theta}$ denote the maximum point at which $\eta_1(\ddot{\theta}) = \eta_2(\ddot{\theta})$, and note that since the two functions are negatively sloped and η_1 crosses η_2 from above (moving in the leftward direction) at $\ddot{\theta}$, the following must be true when $\theta = \ddot{\theta}$:

$$\begin{aligned} \eta_1'(\theta) = -\frac{f_1(\theta)}{\theta + F_1(\theta) - 1} &> -\frac{f_2(\theta)}{\theta + F_2(\theta) - 1} = \eta_2'(\theta) \\ \Leftrightarrow \quad f_1(\theta)\,[\theta + F_2(\theta) - 1] &< f_2(\theta)\,[\theta + F_1(\theta) - 1] \end{aligned} \tag{A.3}$$

Now consider approaching $\underline{\theta}$ from above, beginning at $\ddot{\theta}$ and once again moving leftward. Since $\ddot{\theta}$ is to the left of the density crossing $\tilde{\theta}$, then by the LR dominance property, we know that if we begin at $\ddot{\theta}$ and approach $\underline{\theta}$ from above, then $f_2(\theta)$ becomes steadily larger relative to $f_1(\theta)$ as we move leftward. This also implies that $[\theta + F_1(\theta) - 1]$ becomes steadily larger relative to $[\theta + F_2(\theta) - 1]$ in the leftward direction (beginning from $\ddot{\theta}$) as well.[26] Therefore, the ordering between the right-hand and left-hand sides of inequality A.3 only becomes more pronounced as we move leftward from $\ddot{\theta}$. From this fact it follows that $\eta_1'(\theta) > \eta_2'(\theta)$ for each $\theta \in [\underline{\theta}, \ddot{\theta}]$, and the crossing point $\ddot{\theta}$ is therefore unique. ■

[26]To see why, recall that $F_1(\underline{\theta}) = F_2(\underline{\theta}) = 0$ and $f_1(\theta) < f_2(\theta)$ for each $\theta < \tilde{\theta}$. Thus, F_2 becomes steadily larger relative to F_1 when moving in the rightward direction from $\underline{\theta}$ to $\ddot{\theta}$, which is the same as saying that F_1 becomes steadily larger relative to F_2 when moving in the leftward direction from $\ddot{\theta}$.

8. ONLINE APPENDIX

For Incentive Provision in Investment Contests: Theory and Evidence,
by CHRISTOPHER COTTON, BRENT R. HICKMAN, and JOSEPH P. PRICE

8.1. **TIME TRUNCATION RULE.** Time on our website was measured at the page level for each attempt of a quiz by each student. Pages contain blocks of either 3, 4, or 5 questions, so we divided each block-level time observation by the number of questions in order to get a measure of time spent per question. One difficulty arose in that there were a small number of clear instances where students left the website in the middle of a quiz for several hours or more. For example, the largest recorded time spent on a single question was 2,801 minutes, or roughly 47 hours. In order to correct this problem, a small number of implausibly large time observations needed to be corrected. After selecting a truncation point on the time-per-question domain, we replaced each observation above that point with the student-specific censored mean of time per question. For example, suppose that Tommy attempted 11 questions with observed times of 5 minutes for the first five, 15 minutes for the next five, and 300 minutes for the last, and suppose that the truncation point were 30 minutes per question. Then the eleventh observation of 300 minutes is replaced by Tommy's idiosyncratic censored mean of 10 minutes.

In order to select an appropriate truncation point we looked for occurrences of "holes" in the support of the distribution of times per question, or in other words, points at which a full support condition fails. We began with a natural assumption on the student type distribution that there are no interval subsets of the support where the type density assigns zero mass to the entire interval. If this condition holds, then since time spent on a question is a continuous choice related to one's type, that distribution should also have full support too. That is, unless some observations reflect a different data generating process, say time elapsed outside of learning activity due to work stoppages. Thus, a straightforward way to search for spurious time observations is to sort the data and look for points at which a kernel smoothed density estimate (KDE) equals zero for some interval of positive length. This idea gives rise to the following data-driven algorithm for selecting a truncation point:

(1) Sort all time observations from least to greatest, so that the j^{th} and $(j+1)^{\text{st}}$ observations are ordered by $t_j \leq t_{j+1}$ for all $j = 1, \ldots, J$.
(2) Using the sample $\{t_j\}_{j=1}^{J}$, compute an appropriately chosen bandwidth b_1 for a KDE based on a kernel function with support on $[-1, 1]$.[27] Then find the smallest $j_1^* < J$ such that $t_{j_1^*+1} - t_{j_1^*} > 2b_1$. If no such j_1^* exists, then stop; no truncation is needed.
(3) Define initial truncation point $\tau_1 \equiv t_{j_1^*} + b_1$, and compute bandwidth b_2 for the KDE based on the censored sample $\{t_j\}_{j=1}^{j_1^*}$.
(4) In each subsequent iteration $k = 2, 3, \ldots$, if there exists j_k^* defined by

$$j_k^* \equiv \min\{j : t_{j+1} - t_j > 2b_k; j < j_{k-1}^*\},$$

[27] Actually, the only crucial condition here is that the kernel function have bounded support. For example, in this context a Gaussian kernel would not do, as it places positive mass on the entire real line for any dataset. This would be equivalent to assuming full support *ex ante*.