then update the truncation point by $\tau_k \equiv t_{j_k^*} + b_k$, and re-compute bandwidth b_{k+1} for the KDE based on the censored sample $\{t_j\}_{j=1}^{j_k^*}$.

(5) Stop once k is found such that j_{k+1}^* does not exist (meaning that for the censored sample $\{t_j\}_{j=1}^{j_k^*}$ a KDE is strictly positive everywhere).

We chose a KDE based on the Epanechnikov kernel, which is known to be marginally more efficient than other kernel functions. This choice, in combination with Silverman's automatic bandwidth selection rule, implies a bandwidth formula of $b_1 = 2.345\widehat{\sigma}_1 J^{-1/5}$ in the first iteration, and $b_k = 2.345\widehat{\sigma}_k j_{k-1}^{*-1/5}$ in the k^{th} iteration ($k \geq 2$), where $\widehat{\sigma}_k$ is the sample standard deviation within the k^{th} iteration. Notice that the algorithm does not actually require computation of a KDE at each iteration, only a bandwidth, though choice of the specific kernel is needed to pin down the leading constant on the bandwidth selection rule.

Executing this process on our data leads to a final truncation point of $\tau_2 = 27.81$ minutes per question (the 99.35^{th} percentile of the un-censored sample), after 2 iterations. Figure 9 displays a histogram of time spent per question, including observations above and below the truncation point. Time units are depicted in logs rather than levels for ease of visualization since the largest and smallest observations differ by several orders of magnitude.

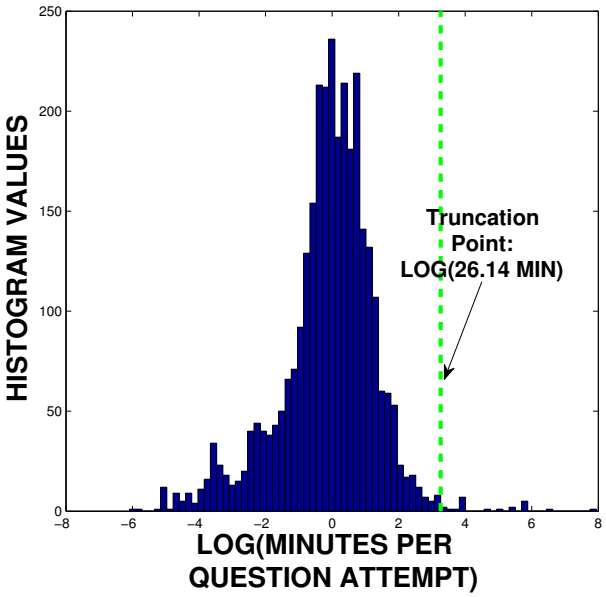
8.2. **ADDITIONAL FIGURES.** Here we present some additional figures depicting the empirical distributions of investment activities by group and treatment status. In interpreting these figures, one caveat should be kept in mind. Proposition 2 only directly applies to the plots in Figures 4 – 5, since these depict CDFs of exam scores, the variable being directly incentivized within the experimental study. Thus, theory predicts that those plots should qualitatively resemble the patterns in Figure 2. It has nothing directly to say about other intermediate variables such as time spent on the website, or number of questions attempted, as these may combine in different ways for different agents to produce exam scores. However, for illustrative purposes, we present additional CDF plots in Figures 10 – 11 here.

8.3. **ROBUSTNESS CHECKS.**

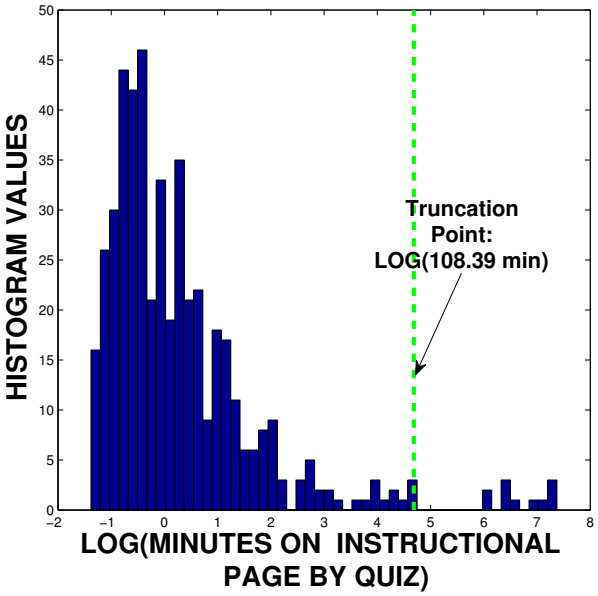
8.3.1. *Testing Average Differences.* Table 4 above tested for average treatment differences

8.3.2. *Selective Attrition.* Figures 12 – 15 illustrate a robustness check on our quantile function estimator, when we attempt to adjust for selective attrition. The upper panels in Figures 12 and 13 plot the empirical CDFs of pre-test scores, restricted to the subsample of students who took the final exam as well. The bottom panels re-produce the CDFs of final exam scores for comparison. From these figures it appears that selective attrition may be working slightly against finding our results in general. Figures 14 and 15 are an attempt at adjusting our quantile function estimator for the possible influence of selective attrition. To do so, we once again, restrict ourselves only to the sample of test subjects who took the final exam, and we convert all pre-test and final exam scores into standardized units by dividing by the within-exam standard deviation. Then, for point estimates and all bootstrapped estimates, we compute the quantile difference function using both

FIGURE 9. TIME TRUNCATION RULE

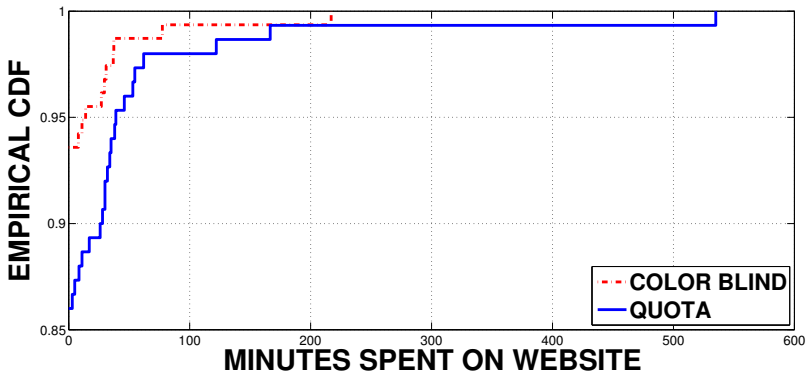


(A) This panel displays a histogram of observed time spent on each question. Each datum in the histogram is a student-question-attempt observation.

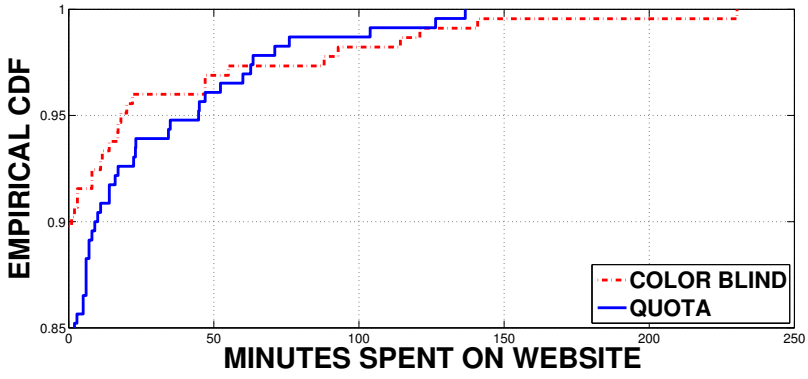


(B) This panel displays a histogram of time per instructional page view. Each datum in the histogram is a student-quiz-attempt observation.

FIGURE 10. TIME SPENT:
PRO vs. *RQ*



(A) Seventh Graders



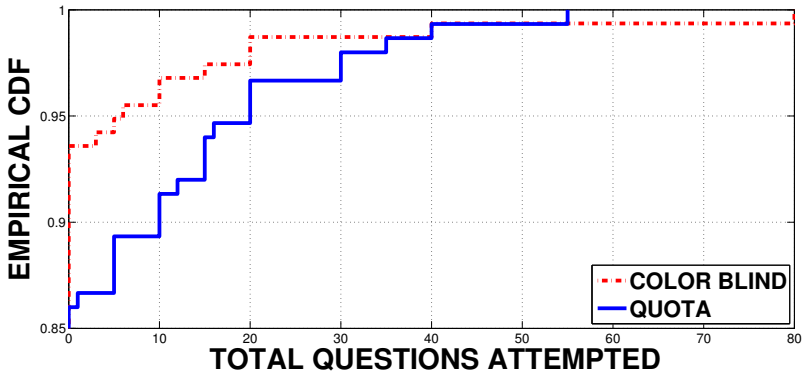
(B) Eighth Graders

the pre-test sample, $\hat{\Delta}_j^{pre}(q),\ j = \mathcal{A}, \mathcal{D}$, and final exam sample, $\hat{\Delta}_j^{final}(q),\ j = \mathcal{A}, \mathcal{D}$. Then, we compute an adjusted quantile difference function estimate by taking the difference

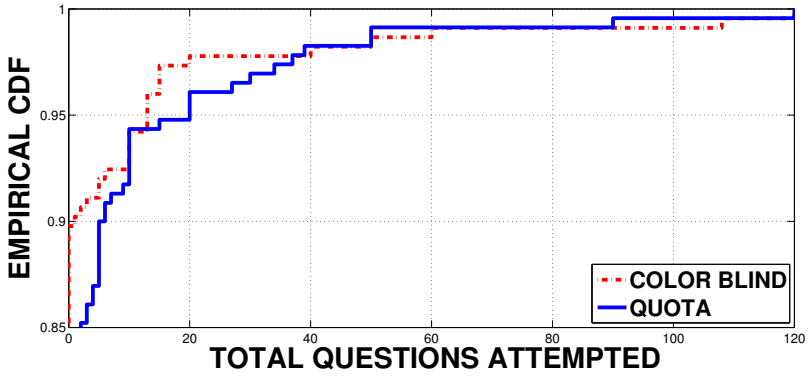
$$\hat{\Delta}_j^{final}(q) - \hat{\Delta}_j^{pre}(q), \tag{B.1}$$

in order to remove any possible pre-existing difference there may have been within the sample of non-attriters. The results produce plots that look very similar to those in the body of the paper, from which we conclude that selective attrition does not appear to be driving our results.

FIGURE 11. QUESTION ATTEMPTS:
PRO vs. *RQ*



(A) Seventh Graders



(B) Eighth Graders

FIGURE 12. 7^{th} GRADE BY TREATMENT:
Pre-Test vs. Final Exam

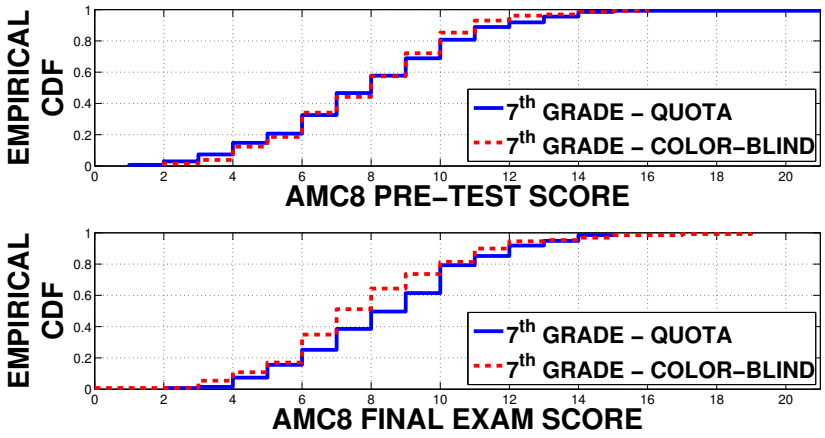


TABLE 6. (RE-)TESTING DIFFS BY DEMOGRAPHICS AND TREATMENT

	Investment				Performance	
	Used Website	# Subjects Attempted	Total Time	# Questions Attempted	Final Exam Score	Exam Score Change
Constant ($\widehat{\beta}_0$)	0.065***	0.131**	2.749	1.154	7.848***	0.096
Std. Err.	(0.024)	(0.059)	(1.966)	(0.736)	(0.229)	(0.239)
Quota ($\widehat{\beta}_1$)	0.089***	0.149*	5.584**	1.350	0.654**	0.591*
Std. Err.	(0.033)	(0.083)	(2.756)	(1.031)	(0.316)	(0.330)
P-Value:	*[0.008]*	*[0.072]*	*[0.043]*	*[0.191]*	*[0.039]*	*[0.074]*
Advantaged $*$ *Quota* ($\widehat{\beta}_2$)	-0.050	-0.116	-5.153	-0.933	-0.874**	-0.546
Std. Err.	(0.042)	(0.105)	(3.505)	(1.311)	(0.397)	(0.415)
Advantaged ($\widehat{\beta}_3$)	0.042	0.104	2.096	1.151	1.145***	-0.324
Std. Err.	(0.030)	(0.075)	(2.509)	(0.939)	(0.288)	0.300
School Fixed Effects	yes	yes	yes	yes	yes	yes
N	992	992	992	992	895	895
Additional Test: Effect of Quota on Advantaged Group						
$\widehat{\beta}_1 + \widehat{\beta}_2$	0.038	0.033	0.431	0.417	-0.221	0.045
P-Value:	*[0.139]*	*[0.611]*	*[0.842]*	*[0.607]*	*[0.359]*	*[0.857]*

Notes: Each column is a separate regression. Advantaged is an indicator variable for whether the student is a 6th or 8th grader (the older group in each school type). Standard errors are in parentheses. 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.

FIGURE 13. 8th GRADE BY TREATMENT:
Pre-Test vs. Final Exam

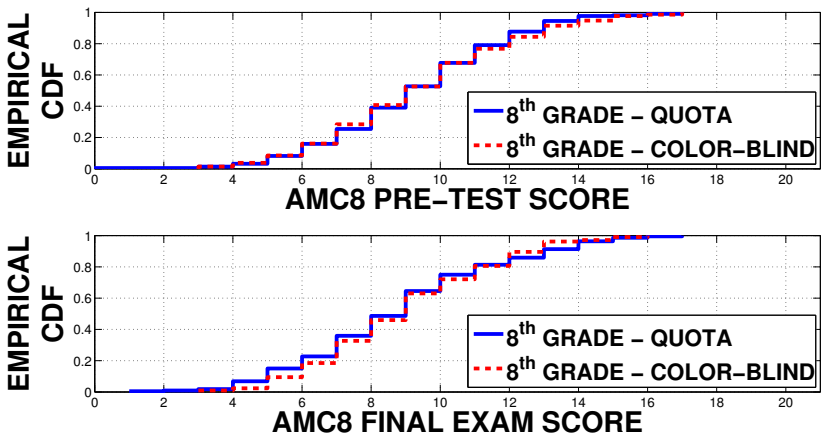


FIGURE 14. 7^{th} GRADE QUANTILE FUNCTION DIFFERENCE, ADJUSTED FOR ATTRITION

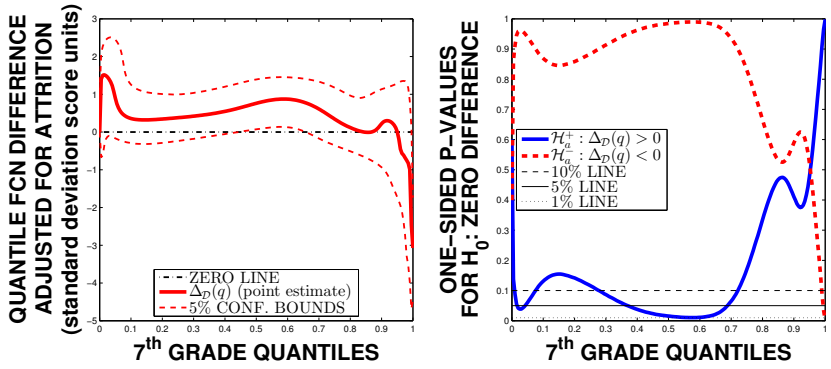


FIGURE 15. 8^{th} GRADE QUANTILE FUNCTION DIFFERENCE, ADJUSTED FOR ATTRITION