Labor Economics
Education

Dynarski et al. (2018) – College Intervention

Dynarski et al. (2018) **NBER**
Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students

In an earlier paper, Hoxby and Avery (2013) showed a striking empirical puzzle: the vast majority of low-income high achievers do not apply to any selective college, even though selective colleges would typically cost these students less because of financial aid packages. Given that attending a high-quality college can increases earnings, this is an important problem to solve as it may play a role in perpetuating income inequality.

This paper implements an RCT to test an inexpensive, targeted, personalized outreach campaign to try to alter the college decisions of low-income students. The intervention aimed to address three issues that high-achieving, low-income students faced:

1. Uncertainty about their suitability for an elite school
2. Over-estimates of the (net) cost of college
3. Procedural barriers, e.g. financial aid forms

The RCT was called the "HAIL (High Achieving Involved Leader) scholarship" and run at the University of Michigan (UM). The aim was to incentive low-income, high-achieving students across the state of Michigan to apply to UM. The intervention consisted of the following:

- Students received personally-addressed packets at their homes in the first week of September of their senior year
- The packet contained a letter from the UM president encouraging them to apply and promised four years of free tuition and fees if admitted, with no requirement to complete financial aid forms. The packet also contained a flyer with application process, UM brochures, and fee waivers for application materials. The materials clearly stated that they did not have to apply for financial aid (e.g. FAFSA) but were encouraged to do so to obtain even more aid
- The offer was also mailed to the parents and emailed to the principal of the student. The parents received letter two weeks after the student. The principals received emails in late August
- Students received a personalized web address to learn more about the scholarship and UM. Since these links were personalized, this allows the researchers to track whether and when a student logs onto a website

The control group received only a postcard listing UM application deadlines in September. This was a low-cost intervention: each packet cost less than $10 to produce and deliver.

The target population for HAIL are rising seniors in public high schools that are low-income (i.e. receiving subsidized-lunch) and high-achieving (met GPA and standardized test score cutoffs). This resulted in 2,000 students in 500 schools that met these criteria over 2015 and 2016 when the experiment was run. Note two important things. First, all public school students in Michigan had to take the ACT (2007-2015) or SAT (since 2015) in their junior year. Second, all the students in the study are eligible for HAIL (and the free tuition) regardless of their treatment status. Therefore, the treatment is simply informing them of their eligibility. To reduce issues that the the control group might hear about the scholarship from their treated classmates, the randomization was done at the school-level (i.e. it was schools and not individuals that randomly received treatment; for treated schools, all eligible students in the school received the intervention).

The results of this were striking. Students in the treated schools were more likely to apply to, gain admission to, and enroll at UM (see Table 1). The treatment increased the application rate by 41 percentage points, and the (unconditional) admission rate increasing by 17 percentage points. Put another way, the application rate for the treatment group (67.5%) was more than double the control group (26%).

Table 1: Effect of HAIL Scholarship on UM

Outcome	Treatment effect		Control mean
Applied	0.416	0.413	0.259
	(0.021)	(0.019)	
Admitted	0.174	0.163	0.149
	(0.019)	(0.017)	
Enrolled	0.149	0.141	0.117
	(0.018)	(0.016)	
Strata dummies	X	X	
Covariates		X	
Number of schools	1,026		
Number of students	3,910		

Source: Dynarski et al. (2018), Table 3

Since this intervention was focused on UM it raises an interesting question about college choice: did HAIL increase the share of low-income students attending highly selective colleges or did it only increase it at UM at the expense of its peer schools? The authors use a national database on college attendance to test these ideas, as shown in Table 2. They find no diversion from colleges that are at least as selective as UM. Instead, HAIL increased the share of students enrolling at any four-year college and decreased the share attending two-year colleges. Moreover, it increased the likelihood of enrolling any college by 3.9 percentage points. In other words, one quarter of the increase in enrollment at UM is due to students who would not have attending *any* college in the absence of the treatment.

Table 2: Effect of HAIL Scholarship on College Choice

College attended	Treatment effect	Control mean
Highly competitive or above	0.146 (0.018)	0.135
UM	0.146 (0.016)	0.107
Highly competitive or above other than UM	0.000 (0.007)	0.028
Four-year	0.074 (0.020)	0.675
Two-year	-0.035 (0.013)	0.116
Any	0.039 (0.018)	0.791
In Michigan	0.045 (0.020)	0.727
Public in Michigan	0.062 (0.021)	0.645
Outside Michigan	-0.006 (0.010)	0.064
Number of schools	1,026	
Number of students	3,910	

Source: Dynarski et al. (2018), Table 7

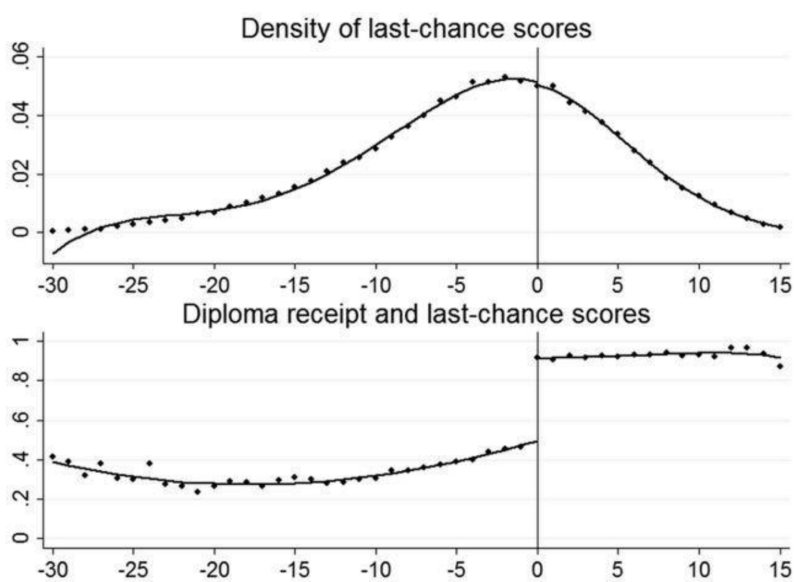
Clark and Martorell (2014) – High School Diploma

Clark and Martorell (2014) **JPE**
The Signaling Value of a High School Diploma

This paper studies the signaling value of a high school diploma. A high school diploma is interesting for two reasons. First, it is most commonly held credential in the US. Second, it is a measure of schooling that does not affect productivity. A diploma is simply a piece of paper that does not confer any extra knowledge onto a student. Therefore, it should have purely a signaling effect on students. An ideal experiment would randomly assign diploma to students once they finish high school coursework and then observe wage differences. Given that this is infeasible and definitely unethical, the authors look for a natural experiment.

The paper looks at the case of Texas, where students must take an high school exit exam. Students must pass the exam in order to graduate high school and receive a diploma. Typically, students take the exam in the 10 or 11th grade. They are also allowed to retake the exam (retests are administered in the fall, spring, and summer of each year). The authors focus on students who retake the exam at the end of their 12th grade, because this is typically their last chance to take the exam (they can re-take in the summer after their class graduates, but this is very rare). They focus on students who barely pass and barely fail this last-chance exam. Barely passers and barely failers should be very similar in many dimensions, with the key difference being that the passers formally graduated high school. This regression discontinuity design closely mimics the ideal the experiment we would like to run. They could have looked at retakes in earlier grades, but this could cause more complications (e.g. those who just fail in the 11th grade retake might then drop out of school, which results in different years of schooling between the barely passers and barely failers).

Figure 1: Last-Chance Exam Scores and Diploma Receipt

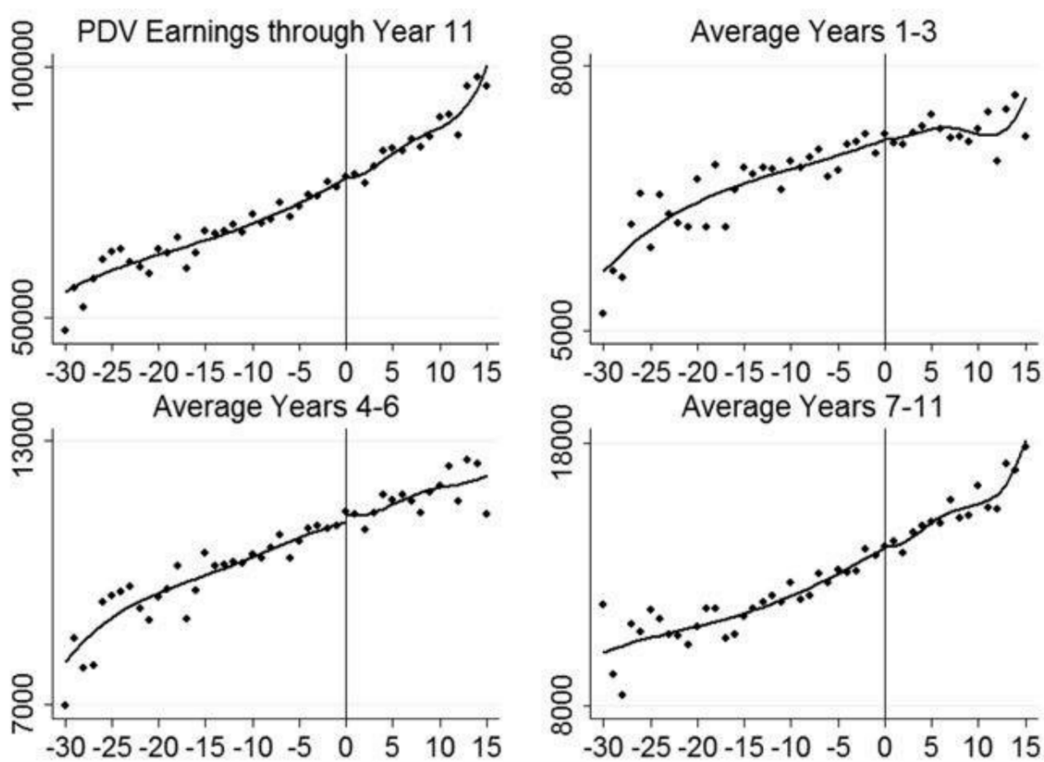


Source: Clark and Martorell (2014), Figure 1

Figure 1 shows the research design. The lower panel shows that those who pass (0 indicates the minimum passing score) have an almost probability 1 of receiving a panel. However at 0, the likelihood of getting a diploma falls to around 40-50%. Notice that this is not a perfect assignment (i.e. it is a “fuzzy” regression discontinuity). So we should interpret the passing of the last-chance test as an *instrument* for high school graduation. Also note that the upper panel supports the design. One worry could be that since is the last-chance exam, teachers and schools could be manipulating the scores in order to ensure their students pass. However, if that were the case, we might observe evidence of "bunching" at the cut-off, which we don't. This suggests that this should be close to random assignment of treatment status.

The authors link the education data (high school records) to earnings data. They observe students who were in the 10th grade in 1991-1995 and who took the last-chance exam in 1993-1997. The earnings data goes until 2004, so they can observe 7-11 years of post-high school earnings for each person. Since there is a discontinuity in earning a diploma, they can then check if there is a discontinuity in earnings. If there is, given the identification strategy, we should be able to assign all of this difference to the signaling value of the diploma. However, as we can see in Figure 2, there is no discontinuity in earnings (whether you look at present discounted value, or different years after the exam). We can see that there is an indeed a relationship between scores and earnings, but that actually passing the test provided no extra signaling benefit.

Figure 2: Earnings by Last-Chance Exam Scores



Source: Clark and Martorell (2014), Figure 2

Arteaga (2018) – Colombia: Course Reform

Arteaga (2018) **JPubE**
The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia's Top University

This paper aims to look at the college premium and separate the effects of human capital accumulation from signaling. The author looks at the case study of Universidad de Los Andes, the top university in Colombia. In 2006, the curriculum for business and economics degrees changes such that the number of courses required dropped by 6 and 12, respectively (and the instruction time dropped by one semester). Given the lower course-load, this reform should have decreased human capital accumulation. However, it did not have an effect of the selection of students into the programs.

The rationale for this is motivated by a simple model. Colleges have a clear rank from top to bottom, and each school has different levels of human capital accumulation (lets call this coursework). For an individual, it is always more costly to go to a school with harder coursework (more human capital development). However, individuals have different abilities. So for the same school, high ability students find it easier (less costly) than low ability students. Firms set wages based on a combination of their beliefs about the average ability of students at each school and the level of coursework. Given this, to understand signaling versus human capital, we need to change the level of coursework, but hold fixed the average ability of students in the school. This is not easy to find. Typically, if a school lowers its coursework, that means it becomes less costly to attend for lower ability students, and therefore the average ability of the student body would fall. This means that we usually are not able to separately explain how much of an effect on wages is caused because of the change in coursework (human capital accumulation) or the change in the quality of the school (signaling).

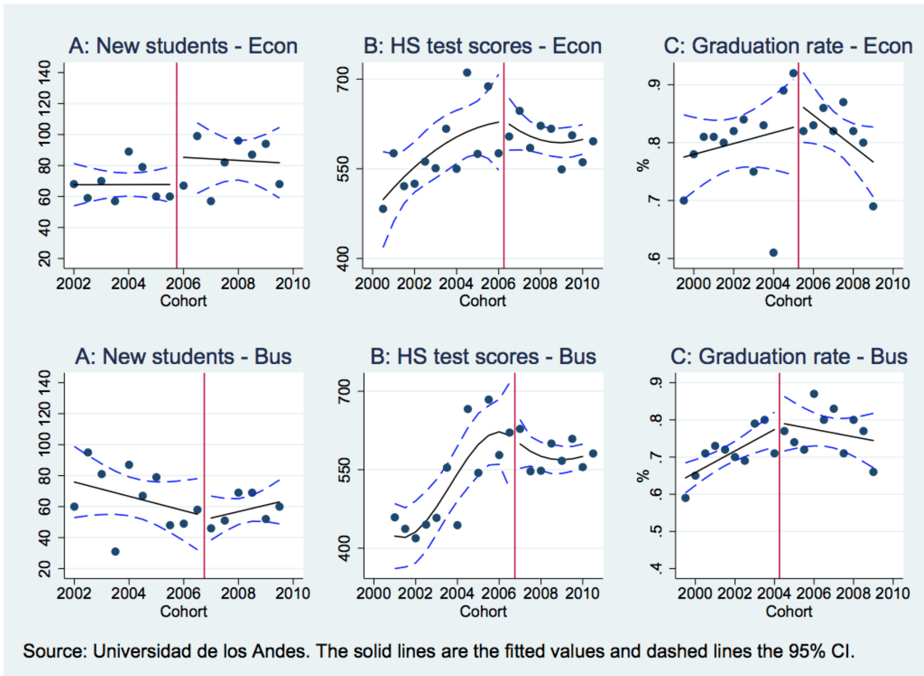
The author finds a setting where this exactly happens. In 2006, Los Andes decided to reduce the coursework for most of its majors. It was the only school to do this at the time and aimed to make degrees shorter and more in-line with international standards. The paper focuses on business and economics because they only changed the number of credits, while other departments also did a complete overhaul of the curriculum. The reform only applied to new students, 2nd-year economics, and 3rd-year business students. To fit the story of the model, the author needs to make sure that the newly admitted students were not of a lower quality than past students. This should not be an issue here because it is the top ranked school (contingent on the reform not changing), there are a fixed number of slots available, and the admissions process is based solely on standardized test scores. Indeed, the author does not find an effect on student quality after the reform (Figure 3).

To estimate the effect of the reform, the author uses a difference-in-difference strategy by comparing Los Andes to the next 10 schools in the national rankings ("Top 10"). The equation to estimate is for a person i at time t:

$$\ln W_{it} = \beta_0 + \beta_1 Andes_i \times Post_t + \beta_2 Andes_i + \beta_3 Post_t + \beta_4 X_{it} + \varepsilon_{it}$$

where W_{it} is average monthly earnings, $Andes_i$ is an indicator for whether they attended Los Andes, and $Post_t$ is an indicator if they were in school for the reform. First, we want to see the graphs to see if the parallel trends assumption holds. This is shown in 4, where we seem to have parallel trends. There also appears to be an immediately large effect for economics and much smaller effect on business graduates that eventually grows over time. Note that given the data, the author only observes people on average 3 years after graduation. This is an issue since wage profiles can still be rapidly changing early in people's

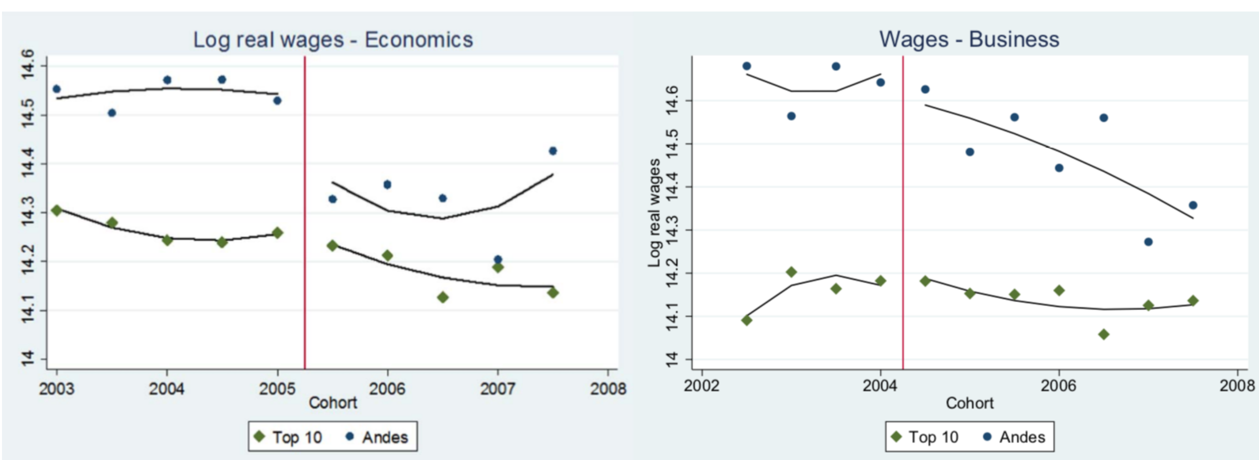
Figure 3: Effects of the reform on Los Andes class selection



Source: Arteaga (2018), Figure 3

careers, though at the same time, the role of signaling is likely to play the biggest role just after college graduation.

Figure 4: Effects of the reform on Los Andes graduate wages



Source: Arteaga (2018), Figure 4

For the actual estimation, we are particularly interested in the estimate of the β_1 coefficient. This captures the effect of the reform. The author finds that $\beta_1 = -0.163$ for economics and -0.136 for business. Given that this is a log-linear equation, we interpret this as saying the reform led to a 16.3% and 13.6% reduction in wages for economics and business graduates, respectively. This rejects the signaling model as such a reform arguably only changed the human capital of graduates. One institutional feature that supports this is story is that recruitment for Colombia college graduates involves not only interviews but also tests of specific knowledge. This should give firms an objective way to test the overall abilities of students. Indeed, the author interviewed some employers and this was one of the reasons they thought the coursework was important (this is also a good example of how researchers should talk to people in the field – even if it is just anecdotal – rather than assume how people behave).

MacLeod et al. (2017) – Colombia: The Big Sort

MacLeod et al. (2017) **AEJ:AE**
The Big Sort: College Reputation and Labor Market Outcomes

This paper studies how college reputation affects the "big sort", the process where students choose colleges and find their first jobs. The earnings for college graduates is correlated with their college's reputation, even after controlling individual characteristics. But why does this correlation exist? Is it because schools with better reputation teach more skills or is it because they simply send a signal about the graduate's ability? The authors come up with a fairly intuitive model to test this. The "big sort" involves two parts. First, students have to choose and attend college (matriculation). Second, students finish college and look for a job (graduation).

Let's first think about matriculation. Suppose that a student's admissions test score (e.g. SAT) is a *noisy* measure of their ability (as in, it is correlated with ability, but it doesn't allow you to perfectly observe it). Each college has a reputation. We need some way to measure this, so the authors assume that the reputation is as being the average admissions test score of its graduates. Note that these scores are taken before the students even attend the school. However, more selective schools will pick students with the highest scores, which will give them a higher reputation.

Now, let's think about graduation. When a student graduates, they have (hopefully!) learned something at college. This means that their skill at graduation is their baseline pre-college ability plus the skills the school teaches. Firms want to hire new graduates but they cannot observe the student's skill. What they can observe is the school's reputation.[1] From this, they have to predict the skill of the student, given the college's reputation. For example, they ask themselves: "given that Columbia has an average SAT score of 1540, how smart is the average person that Columbia admits?". Once they have a prediction, they will set the wage according to the graduate's expected skill. After hiring someone, an employer can observe them working and over time, they will eventually figure out the true ability of the worker. They do this by weighing up factors such as their skill at work and the expected skills of graduates (e.g. "This person went to Columbia so I expect them to be smart... but at the same time, they don't seem to know how to use Excel..."). Over time though, your school's reputation doesn't matter any more and the wage just depends on your output as a worker. Putting this together, we can think that your wage at any particular time depends on the reputation of your college and your ability (which will be measured using the test score).

This model seems fairly reasonable, especially for their setting. But the usefulness of the model is that it provides some very intuitive testable predictions:

1. The introduction of an exit exam allows employers to rely less on the reputation of the school and more on the employer's skills. Therefore, this should reduces the return to college reputation and increases the return to ability.

2. At the time of hiring, employers have already incorporated everything that is observable (i.e. college reputation). This means that the return to reputation should fall with experience (not necessarily to zero, because there should still be some reputation premium). But over time, as they learn more about unobservables (i.e. skills), they should become more important. This means that the return to ability should increase with experience.

[1] Note, in this setting of Colombia, graduates rarely report their admissions test on their CVs. This is quite different to the US, where some employers will ask for your test scores (WSJ)

Now with this model, we can take it to the data. The authors study colleges in Colombia. Students apply to colleges by taking a standardized exam called the Icfes (very analogous to SAT, but it plays an even bigger role in admissions).[2] In 2004, the agency that runs Icfes (like the College Board) begin running field-specific college *exit* exams. These were standardized exams that graduates took that tested their knowledge in their field of study. These exit exams provided a new signal of a student's skill. One useful feature here is that these exit exams were rolled out gradually. It wasn't until 2009 that the exit exams because mandatory for everyone.

The authors have data on college enrollees from 1998-2012, including graduation/drop out date and their program of study. They can link this up to monthly earnings data for 2008-2012. For test scores, they have data on students who took the Icfes between 1998-2012 and the exit exam scores for 2004-2011. They focus on the cohorts graduating from 2003-2009 during the gradual rollout of the exit exams (there were some other changes after 2009 that make it hard to study). This rollout was not random - more popular programs such as economics and engineering got the exams first. Therefore, we can't just compare the programs who got the early exams to the ones who got the later exams because there are likely a lot of unobserved differences. Instead, the authors compare what happens before and after the exams. Effectively, this is going to be a difference-in-differences specification, but with a twist!

The authors run the following regression for a student i in program p at school s, graduation cohort c, at time t years after graduation:

$$\begin{aligned} w_{ipct} = &\beta_1 Rep_s + \beta_1^e Rep_s \times Exit_{pc} \\ &+ \beta_2 Test_i + \beta_2^e Test_i \times Exit_{pc} \\ &+ \beta_3 Exp_{pt} + \phi_{pc} + \varepsilon_{ipct} \end{aligned}$$

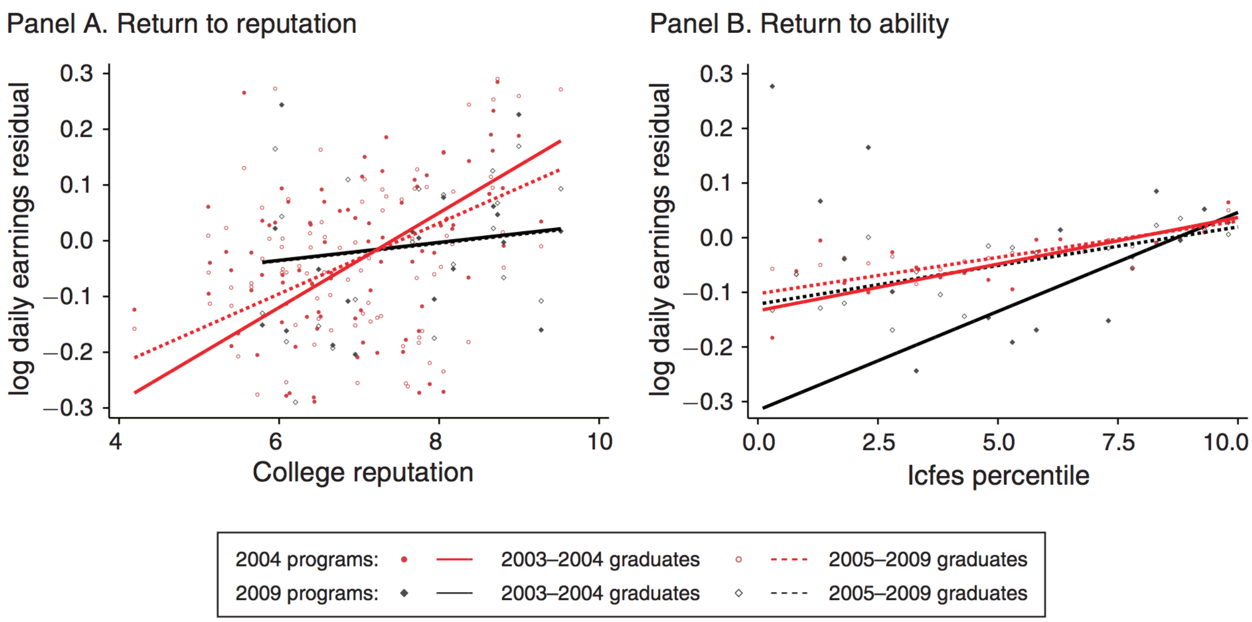
where w is the log earnings, Rep_s is the school's reputation, $Test_i$ is the student's admission test score, and $Exit_{pc}$ indicates whether the program-cohort had an exit exam prior to 2009. They also include Exp_{pt}, a measure of experience that also depends on program), and program-cohort fixed effects ϕ_{pc}. This specification says that there is a return to reputation (β_1) and ability (β_2), but that the exit exam changes these returns. For example, if you were in a program-cohort that did not have an exam, then the return to reputation is β_1. But if you were in one that did have an exam, then the return to reputation is $\beta_1 + \beta_1^e$. Therefore, β_1^e and β_2^e represent the effect of the exam on the returns of reputation and ability. If their model is correct, then we should find that $\beta_1^e < 0$ and $\beta_2^e > 0$.

Intuitively, this is just like a difference-in-difference model. The treatment is getting an exit exam, which some program got earlier in prior to 2009. The control group did not receive the exams until 2009. The usual DID is to look at how the treatment affected the average outcome. This DID is looking at how the treatment affected the *slope* of a variable. So we are comparing how the *difference* in slopes between the treatment and control group changes before and after the introduction of exit exams. This idea is captured in Figure 5. The 2004 programs received exit exams in 2004 and so are in the treated group (red lines). The 2009 programs received exit exams in 2009 after the gradual rollout, so they are the control group (black lines). The 2003-2004 cohort graduated before the exit exams were in place (for any program), so they represent the pre-period in the DID (solid lines). The 2005-2009 cohorts had the exit exams, but only for the 2004 programs, so they represent the post-period in the DID (dashed lines). Panel A shows how the exit exams affected the return to reputation. We see that in the pre-period, the slopes of the treated and control groups are quite different: the treated group tended to have a stronger relationship between

[2] To make interpretation easier, the author report Icfes scores as percentiles (0=lowest scores, 10=highest scores), and similarly college reputation is the mean Icfes percentile score

reputation and earnings. In the post period, the slope of the treated group falls, indicating that there is less of a relationship between reputation and earnings. Additionally, the control group's slope barely changes in the post-period, so we can attribute the change in slope for the treated groups as being driven by the introduction of exit exams. In Panel B, we see the same graph but now for return to ability. Now, the relationship between test scores and earnings falls for both groups over time. However, the decline is much stronger for the control group. Therefore, if we do a DID, this means that the exit exam must have been off-setting this decline for the treated group. So, consistent with the model, the authors find that the exit exams reduced the returns to reputation but increased the returns to ability. Therefore, this suggests that college reputation plays a role in signaling ability on the labor market.

Figure 5: Exit Exam Effects



Source: MacLeod et al. (2017), Figure 3

This result shows that reputation plays a role in signaling and fits with the first prediction of the model. The second prediction says that the role of reputation should be decreasing over time. For this, they estimate the following regression:

$$\begin{aligned} w_{ict} = &\beta_1 Rep_s + \beta_1^t Rep_s \times t \\ &+ \beta_2 Test_i + \beta_2^t Test_i \times t \\ &+ \phi_{ct} + \varepsilon_{ict} \end{aligned}$$

Now, they include an interaction of reputation and time ($Rep_s \times t$) and an interaction of test scores and time ($Test_i \times t$). This is not a causal relationship, but rather we are looking to see whether the sign is consistent with the model, i.e. we should expect $\beta_1^t < 0$ and $\beta_2^t > 0$. The results are shown in Table 3. While they do find that the returns to test score increase over time, they also find that the returns to reputation are also increasing over time. This is inconsistent with the model, which is quite surprising and as other papers in the literature did not find such an effect for years of schooling (i.e. that the returns to years of schooling increase over time).

Figure 6(a) gives a graphical representation of the idea. We expect earnings to be increasing in experience.

Table 3: Returns to Reputation and Ability with Experience

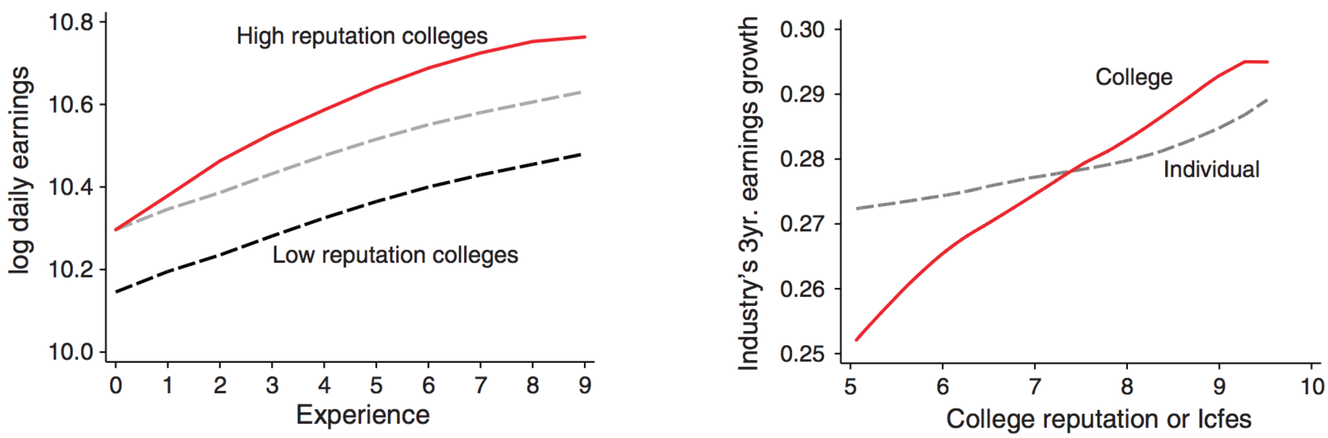
Dependent variable: log average daily earnings	(1)	(2)	(3)	(4)
Reputation	0.101	0.071	0.079	0.055
	(0.017)	(0.017)	(0.017)	(0.016)
Reputation × t	0.017	0.017	0.012	0.008
	(0.003)	(0.003)	(0.003)	(0.002)
Icfes		0.033	0.024	0.017
		(0.003)	(0.002)	(0.002)
Icfes × t			0.006	0.002
			(0.001)	(0.001)
Observations	83,492	83,492	83,492	83,492
R^2	0.179	0.189	0.190	0.306
Number of colleges	130	130	130	130
Extra controls				Yes

Source: MacLeod et al. (2017), Table 7

However, we might have also suspected that going to a high reputation college causes a parallel shift in the earnings-experience profile (i.e. going to high reputation school is associated with a constant increase in earnings at each experience level). This would be represented by the grey line. The data actually shows the red line - that the gap in earnings between those who went to a high and low reputation college increases over time. The authors cannot pin down an exact reason for this. They suggest two possibilities. One is that reputation is in fact not perfectly observed by firms. Therefore, over time they not only learn about the ability of a worker, but also the school's reputation. With a better understanding of the school's reputation, it becomes more useful and important in determining the ability of a worker. Another possibility is that reputation may help in human capital accumulation (i.e. not only do colleges add to skill, but that higher ranked schools add more value). Figure 6(b) shows the correlation between test scores and earnings growth for the industry of each graduate's first job (growth over the first four year). The graph shows two lines. The grey line represents the correlation for an individual; for example, someone with a median test score (i.e. Icfes score of 5) on average have a first job in an industry where earnings increase by 27% in four years. The red line represents the correlation for a college; for example, a school with a median reputation (i.e. an average Icfes score of 5) has graduates who on average end up in an industry with 25% earnings growth. Notice that for median individuals, college placements are lower than we would expect for an individual. However, eventually the lines cross. This tells us that graduates from high reputation colleges tend to obtain jobs in industries with higher earnings growth than you would expect from an individual with a similarly high score. The authors list some possible explanations for this finding, e.g. colleges provide better networks or help students in finding better job matches.

So this paper has mixed findings for their model. It seems that reputation plays a role as the return to reputation did decrease after the introduction of the exit exams (suggesting that employers could now use these exams to predict skill rather than relying so much on reputation). However, they also find that return to reputation increases with experience, which cannot be explained by their model. While they cannot find a causal relationship, this is still an important finding. It further shows the importance of the "big sort": students may be able to observe that graduates of more reputable schools have higher earnings trajectories, which would then influence their college choice.

Figure 6: Reputation Over Time



(a) Earnings-Experience Profile (b) Industry Growth vs Test Scores

Source: MacLeod et al. (2017), Figures 4, 5

References

Arteaga, Carolina. 2018. “The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia’s Top University.” *Journal of Public Economics*, 157: 212–225.

Clark, Damon, and Paco Martorell. 2014. “The Signaling Value of a High School Diploma.” *Journal of Political Economy*, 122(2): 282–318.

Dynarski, Susan, C J Libassi, Katherine Michelmore, and Stephanie Owen. 2018. “Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students.” *NBER Working Paper*.

Hoxby, Caroline, and Christopher Avery. 2013. “The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students.” *Brookings Papers on Economic Activity*, 44(1 (Spring)): 1–65.

MacLeod, W Bentley, Evan Riehl, Juan E Saavedra, and Miguel Urquiola. 2017. “The Big Sort: College Reputation and Labor Market Outcomes.” *American Economic Journal: Applied Economics*, 9(3): 223–261.