American Economic Journal: Applied Economics 2017, 9(3): 223–261
https://doi.org/10.1257/app.20160126

The Big Sort: College Reputation and Labor Market Outcomes†

By W. Bentley MacLeod, Evan Riehl, Juan E. Saavedra, and Miguel Urquiola*

We explore how college reputation affects the "big sort," the process by which students choose colleges and find their first jobs. We incorporate a simple definition of college reputation—graduates' mean admission scores—into a competitive labor market model. This generates a clear prediction: if employers use reputation to set wages, then the introduction of a new measure of individual skill will decrease the return to reputation. Administrative data and a natural experiment from the country of Colombia confirm this. Finally, we show that college reputation is positively correlated with graduates' earnings growth, suggesting that reputation matters beyond signaling individual skill. (*JEL* I23, I26, J24, J31)

Each year, millions of high school graduates choose a college with the hope of a good career. The problem they face is daunting. Each potential college they consider may provide them with a different set of skills. Moreover, the labor market may use the identity of their college as a signal of ability. We call the process by which students choose a college, and find their first job, the "big sort." This paper provides evidence on the role of college identity in the big sort using unique data and a natural experiment from the country of Colombia.

We introduce a simple measure of a college's reputation: the mean admission score of its graduates. Our data allow us to observe which college individuals attended, as well as their subsequent performance in the labor market. We show that their earnings are positively correlated with the reputation of their colleges after controlling for individual characteristics, including their own admission scores. This correlation may arise because high reputation colleges provide more skill, or because college identity signals graduates' ability. To differentiate between these mechanisms, we exploit the staggered introduction of a national college *exit* exam

* MacLeod: Columbia University, 420 West 118th Street, New York, NY 10027 (email: bentley.macleod@columbia.edu); Riehl: Columbia University, 420 West 118th Street, New York, NY 10027 (email: e.riehl@columbia.edu); Saavedra: University of Southern California, 635 Downey Way VPD, Los Angeles, CA 90089 (email: juansaav@usc.edu); Urquiola: Columbia University, 420 West 118th Street, New York, NY 10027 (email: miguel.urquiola@columbia.edu). For useful comments we thank Joseph Altonji, Costas Meghir, Michael Mueller-Smith, Phil Oreopoulos, Kiki Pop-Eleches, and Russell Weinstein. For invaluable help with the data we are grateful to Julian Mariño and Adriana Molina at the Colombian Institute for Educational Evaluation (ICFES), Luz Emilse Rincón at the Ministry of Social Protection, and Luis Omar Herrera at the Ministry of Education. All errors are ours.
† Go to https://doi.org/10.1257/app.20160126 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

that provided employers with a new signal of individual skill. A competitive labor market model predicts that the exit exam should reduce the correlation of earnings with college reputation if reputation serves to signal ability. The empirical evidence is consistent with this prediction, suggesting that college identity plays an informational role.

Finally, we measure the effect of college reputation upon subsequent earnings *growth*. We find that the correlation between reputation and log earnings is not constant, but rather increases with a worker's labor market experience. This descriptive result contrasts with a large literature on the Mincer wage equation, which finds that the correlation of log wages with workers' years of schooling does not vary with experience (Lemieux 2006). Thus, differences in educational attainment are an initial but stable source of inequality (Katz and Murphy 1992, Autor 2014). Our result shows that sorting across different types of colleges may be a further source of inequality—one that grows in importance over the course of workers' careers.

The paper proceeds as follows. In Section I, we develop the model that guides our empirical analysis. The big sort is a complex process, and preferences over colleges may depend upon who else chooses to attend (Rothschild and White 1995). Moreover, there are multiple sorting equilibria depending upon colleges' selection strategies (MacLeod and Urquiola 2015). Rather than attempting to model this complexity, we build on the standard competitive model of wage formation (Burdett 1978, Jovanovic 1979).

We make two assumptions that yield clear predictions regarding the effect of school reputation on wages. First, colleges select students based on their scores on a standardized admission test, as is the case in Colombia. Hence, the most desirable colleges tend to enroll the highest scoring students. This allows us to propose a simple definition of a college's reputation: the mean admission score of its graduates. Second, we assume employers observe graduates' college of graduation, but not their individual admission scores. Thus, in setting wages employers use college reputation to infer graduates' ability as measured by the admission test.

We explore the implications of these assumptions for the introduction of an individual-specific measure of skill that employers do observe—a college exit exam score. The key prediction is that in earnings regressions that include both college reputation and individual admission scores, the availability of the exit exam reduces the return to reputation and increases the return to individual scores.

In Section II, we explore the empirical evidence on this prediction using administrative data that link, for all college graduates: scores on a standardized national admission exam, college of graduation, and labor market outcomes. During the period we study, Colombia introduced national college exit exams, and many students began listing their exit scores on their CVs. These exams were gradually rolled out across 55 fields of study such as accounting, dentistry, economics, and law. This allows us to implement an approach analogous to Card and Krueger (1992) who analyze how time-varying state policies (e.g., class size levels) affect a slope—the relation between years of schooling and wages. In our case, the question is how time-varying college major characteristics (e.g., the existence of an exit exam in a related field) affect two slopes—the earnings return to reputation and the earnings return to admission scores. Consistent with the assumption that employers use college reputation to infer individual ability, we find that the new signal of skill reduced the return to reputation and increased the return to admission scores.

In addition, we find that the exit exams increased average earnings, a result that is consistent with improved employer-employee match quality.[1] The exit exams also prompted student behavioral responses in the form of delayed graduation and preference for colleges and programs with better exit exam performance. In short, these results provide evidence that college identity transmits information on ability, and that the reliance upon reputation fell in the presence of a better performance signal.

In Section III, we ask whether college reputation relates to earnings exclusively through an informational channel. The competitive labor market model predicts that employers update their evaluation of a worker's skill based upon performance on the job (Harris and Holmstrom 1982). Thus, wages should become more correlated with ability as workers gain experience. Farber and Gibbons (1996) and Altonji and Pierret (2001) use this result to show that over workers' careers, observable characteristics like years of schooling become less correlated with wages in regressions that include unobserved measures of ability.[2] In our model, the fact that we define a college's reputation as the mean admission score of its graduates yields a clean prediction for regressions that also include individual scores. If reputation is solely a signal of ability as measured by admission scores, the correlation of earnings and reputation should decrease with experience conditional on individual scores.

The evidence is inconsistent with this prediction. We find that conditional on individual admission scores, the correlation between earnings and reputation *increases* with worker experience. This contrasts with Altonji and Pierret (2001), who find that the correlation of earnings and years of schooling decreases with experience conditional on measures of unobserved ability. This result suggests that college reputation may also affect earnings through channels other than signaling. We cannot disentangle which of several competing hypotheses might explain this finding. Students and parents likewise cannot measure the counterfactual effects of college reputation on earnings. However, they may observe the correlation of reputation and career prospects, which likely increases the demand for the most reputable colleges.

Our findings relate to four distinct literatures on: reputational markets, college choice, the impact of selective schools, and costly signaling.

Reputational Markets.—Nelson (1970) introduced the idea that consumer goods are either inspection or experience goods. The quality of an inspection good can be determined before purchase; that of an experience good can only be determined after. Work in an industrial organization (Melnik and Alm 2002, Hubbard 2002, Jin and Leslie 2003, Cabral and Hortaçsu 2010, and Dranove and Jin 2010) observes

[1] Şahin et al. (2014) suggest that there is a role for policies that improve such matches. They point out that occupational mismatch in the United States has become more severe for college graduates since the Great Recession.

[2] Farber and Gibbons' (1996) and Altonji and Pierret's (2001) results suggest that schooling signals ability, while other factors correlated with schooling have a deterministic effect on wages. In related work, Lange (2007) finds that errors regarding worker skill decline markedly after a few years of employment, although Kahn and Lange (2014) find greater persistence.

that with experience goods the reputation of the seller affects the price; for example, a bottle from a good winery commands a high price even if it ultimately proves to be corked. We show that a similar effect arises in education: employers are sensitive to college reputation, and this sensitivity is reduced when better information becomes available (as recommended by Bishop 2004). Further, consistent with college being a complex, composite good (e.g., Black and Smith 2006), we find that students in turn respond to employers' changing perception of college reputation.

College Choice.—Hoxby (1997, 2009) shows that stratification by ability has increased significantly among US colleges. Thorough sorting may account for the fact that Arcidiacono, Bayer, and Hizmo (2010) find that college identity in the United States fully reveals Armed Forces Qualification Test (AFQT) scores. In contrast, we find that college identity only partially reveals admission scores in Colombia. This discrepancy may reflect that college stratification in Colombia, although increasing, is not as thorough as in the United States.[3] This suggests that college preferences, and hence reputations, are endogenous and may change over time. In particular, the introduction of college exit exams affected the labor market return to college reputation and preferences of college applicants. The endogeneity of preferences is relevant to theoretical work on matching in college and other markets (e.g., Roth and Sotomayor 1989, He 2014).[4] These models assume that students have clear exogenous preferences over the colleges they wish to attend. Future research could explore if peer effects impact optimal market design.

The Effects of Attending a Selective College.—Our work complements studies that estimate the wage effects of attending a selective college. Using US data, Dale and Krueger (2002, 2014) find a positive effect, but one that is concentrated among minorities (see also, Hoekstra 2009). Using Chilean data, Hastings, Neilson, and Zimmerman (2013) find evidence of significant variation in effects across colleges and majors, and less heterogeneity across family background (see also Rodríguez, Urzúa, and Reyes 2015). Our contribution is to explore the mechanisms underlying these effects by explicitly measuring reputation in an entire market. While our results suggest that information-related channels may account for some of the effects in this literature, they do not foreclose other mechanisms like peer effects (Epple, Romano, and Sieg 2006) and network externalities (Zimmerman 2013; Dustmann et al. 2016; and Kaufmann, Messner, and Solis 2013). If such externalities are more important for high level managerial jobs, then this may explain the effect of college reputation upon wage growth.

Costly Signals.—The celebrated Spence (1973) model shows that if schooling is a signal whose cost is declining in ability, a college wage premium can exist even if college has no value added. Our focus is on *which* college students attend

[3] In addition, Hoxby and Avery (2013) show that even controlling for ability, individuals from disadvantaged backgrounds are less likely to apply to reputable colleges. Their results are generally consistent with a role for "brand name" reputations.

[4] See Abdulkadiroğlu and Sönmez (2013) for a recent review of the large literature on this issue.

rather than *whether* they attend. As MacLeod and Urquiola (2015) show, students may care about college identity even if college value added does not vary with college reputation. Informational concerns alone can lead to ability sorting and stratification.

Finally, such informational channels cannot explain our finding of a positive correlation between reputation and earnings growth. Greater traction might arise from models with peer effects (e.g., Epple and Romano 1998) especially if peer interactions provide networks that become more valuable with experience. Alternately, graduates from high reputation colleges might be more likely to obtain positions in firms with higher levels of on-the-job human capital investment. For example, larger firms sometimes set pay bands for positions they wish to fill with applicants of certain characteristics (this is known as the Hay compensation system; see Milkovich, Newman, and Gerhart 2010). Those characteristics might include college identity, leading to a correlation between reputation and investment in human capital that results in a superstar-type effect (Rosen 1981).

I. College Reputation, Signaling, and Wages

This section adds college reputation to the standard Bayesian model of wage formation (Jovanovic 1979). It presents two propositions that we take to the data in Sections II and III. A full derivation of the model and these propositions is in online Appendix A.

A. *Ability*, *Admission Scores*, and *College Reputation*

Let α_i denote the log ability of student i, where by ability we mean the type of aptitude measured by precollege admission tests. We define two measures of ability from our data. First, we observe each student's score on a college admission exam, τ_i, and we assume it provides a noisy measure of ability:

$$\tau_i = \alpha_i + \epsilon_i^{\tau}.$$

The second measure is college reputation. Reputation may incorporate many aspects of college quality, such as peer composition and faculty research output. We define the reputation of a college s to be the mean admission score of its graduates, and denote it by R_s:

$$R_s = E\{\tau_i | i \in s\} = \frac{1}{n_s} \sum_{i \in s} \tau_i,$$

where n_s is the number of graduates from college s. This measure has two analytical advantages. First, in settings where selective schools use test scores to determine admission, R_s will be mechanically related to other attributes that lead students to prefer certain colleges. Second, as we discuss below, this reputation measure delivers clear predictions in regressions that also include individual admission scores.

B. *Employers' Information and Wage Setting Process*

We let θ_i denote the log skill of student i and suppose it is given by

$$\theta_i = \alpha_i + v_{s_i}.$$

Skill includes both precollege ability, α_i, and v_{s_i}, which we will interpret as attributes related to an individual's membership at college s_i. These may include factors that contribute to skill formation at school, such as teaching or peer effects, as well as access to alumni networks. These may also include individual traits (not perfectly correlated with α_i) along which individuals sort into colleges, such as family income or motivation.

We suppose that the market sets log wages, w_{it}, equal to expected skill given available information, I_{it}, regarding worker i in period t:

$$w_{it} = E\{\theta_i | I_{it}\} + h_{it},$$

where h_{it} is time-varying human capital growth due to experience and on-the-job training. We consider Mincer wage equations that net out human capital growth to focus on the time-invariant component of skill that is generated by education and revealed over time to the employer (see Lemieux 2006):

$$\hat{w}_{it} = w_{it} - h_{it} = E\{\theta_i | I_{it}\}.$$

We suppose that employers' information set, I_{it}, includes college reputation, R_{s_i}.[5] While employers likely care about individuals' precollege ability as captured by R_{s_i}, they also care about other attributes related to graduates' postcollege skill. We therefore define a college's *labor market reputation* as the expected skill of its graduates: $\mathscr{R}_s = E\{\theta_i | i \in s\}$. It follows that $\theta_{i \in s} \sim N\left(\mathscr{R}_{s_i}, \frac{1}{\rho^{\mathscr{R}}}\right)$, where $\rho^{\mathscr{R}} = \frac{1}{\sigma^2_{\mathscr{R}}}$ denotes the precision of $\mathscr{R}_s$.[6]

Our data do not contain $\mathscr{R}_s$, and it may differ from R_s if colleges with higher reputation provide more value added or select students based upon dimensions of ability that we do not observe. For instance, if colleges prefer motivated students, and students prefer more value added, R_s and v_s will be positively correlated. To allow for this, we suppose v_s satisfies $E\{v_s | R_s\} = v_0 + v_1 R_s$, where v_1 is the reputation premium, i.e., the return to reputation beyond that captured by admission scores. If this premium is positive ($v_1 > 0$), then a college with a better reputation provides higher value added, broadly understood.

[5] Employers likely observe college identity, but they may not perfectly observe our measure of reputation. Below we discuss how our definition helps to address the possibility that this assumption does not hold.

[6] We assume all variables are mean zero and normally distributed, and we characterize their variability using precisions. The precision, $\rho^{\mathscr{R}}$, could also be indexed by s and hence be school-specific. We did not find robust evidence that the variance has a clear effect on earnings, and so set this aside for further research.

To summarize, employers observe a signal of worker i's skill given by the labor market reputation of her college of origin:

$$\begin{aligned}\mathscr{R}_{s_i} &= E\{\alpha_i + v_{s_i} | R_{s_i}\} \\ &= E\{\alpha_i | R_{s_i}\} + v_0 + v_1 R_{s_i}.\end{aligned}$$

In words, labor market reputation captures employers' expectations of ability, α_i, and attributes related to college membership, v_s, under the assumption that they observe our measure of reputation, R_s.

At the time of hire, employers observe other signals of skill that we do not see (Farber and Gibbons 1996). We denote these by

$$y_i = \alpha_i + v_s + \epsilon_i,$$

with associated precision ρ^y. Importantly, y_i does not include τ_i because we assume that employers do not observe graduates' individual admission test scores. This is consistent with the standard assumption in the employer learning literature that AFQT scores are unobserved, and with anecdotal evidence that in our setting, graduates' CVs rarely feature their college admission exam score (we present evidence supporting this assumption below).

Lastly, employers observe signals related to worker output after employment begins:

$$y_{it} = \alpha_i + v_s + \epsilon_{it},$$

where ϵ_{it} includes human capital growth and other fluctuations in worker output. These are observed *after* setting wages in each period t (where $t = 0$ is the year of graduation). Let $\bar{y}_{it} = \frac{1}{t+1}\sum_{k=0}^{t} y_{ik}$ denote mean worker output, and let $\rho^{\bar{y}}$ be the time-invariant precision of y_{it}.[7]

The market's information set in period t is thus $I_{it} = \{\mathscr{R}_{s_i}, y_i, y_{i0}, \ldots, y_{i,t-1}\}$. Assuming all variables are normally distributed, log wages net of human capital growth are

$$\hat{w}_{it} = \pi_t^{\mathscr{R}} \mathscr{R}_{s_i} + \pi_t^y y_i + \left(1 - \pi_t^{\mathscr{R}} - \pi_t^y\right) \bar{y}_{i,t-1}, \tag{1}$$

where the weights on the signals satisfy $\pi_t^{\mathscr{R}} = \frac{\rho^{\mathscr{R}}}{\rho^{\mathscr{R}} + \rho^y + t\rho^{\bar{y}}}$ and $\pi_t^y = \frac{\rho^y}{\rho^{\mathscr{R}} + \rho^y + t\rho^{\bar{y}}}$. Note that $\pi_t^{\mathscr{R}}, \pi_t^y \to 0$ as wages incorporate new information from worker output.

Equation (1) describes employers' wage setting process given available information, I_{it}. We do not observe I_{it}, and instead derive the implications of the wage equation for regressions on characteristics in our data. Below we estimate regressions that include controls for experience and graduation cohort to capture the time-varying

[7] The assumption that the precision of y_{it} is time stationary also follows Farber and Gibbons (1996). We note that this assumption implies that any human capital growth included in ϵ_{it} is not serially correlated.

effects (recall that $\hat{w}_{it} = w_{it} - h_{it}$). Here we focus upon the implications for the relationship between the signals of individual ability and wages net of human capital growth.

We define the *return to reputation* at time t, r_t, and the *return to ability*, a_t, as the coefficients from the regression:

$$\hat{w}_{it} = r_t R_{s_i} + a_t \tau_i + e_{it}, \tag{2}$$

where e_{it} is the residual. The return to reputation, r_t, is the wage impact of a change in R_s for students with similar admission scores, τ_i. The return to ability, a_t, is the wage impact of a change in τ_i for students from colleges with similar reputations.

C. *Predictions for the Introduction of a College Exit Exam*

While the returns to reputation and ability are not causal, changes in these parameters are informative as to the signaling role of reputation. In Section II, we ask how these returns were affected by the introduction of a new measure of individual skill—a college exit exam. We suppose that the exit exam increases the amount of information contained in y_i; its precision is $\rho^{y,exit} > \rho^y$ when the exit exam is offered. This could arise because students list exit exam scores on their CVs, receive reference letters as a result of their performance, or modify job search behavior after learning their position in the national distribution of exam takers.

The increase in the precision of y_i reduces the weight on reputation in wage setting, $\pi_t^{\mathcal{R}}$. Let $\delta_i = 1$ if and only if a student is exposed to the possibility of writing the exit exam. We can rewrite regression (2) as follows:

$$\begin{aligned} \hat{w}_{it} &= (1 - \delta_i)(r_t R_{s_i} + a_t \tau_i) + \delta_i \left(r_t^{exit} R_{s_i} + a_t^{exit} \tau_i\right) + e_{it}^{exit} \\ &= \left(r_t R_{s_i} + a_t \tau_i\right) + \delta_i \left(\beta_t^r R_{s_i} + \beta_t^a \tau_i\right) + e_{it}^{exit}, \end{aligned} \tag{3}$$

where $\beta_t^r = r_t^{exit} - r_t$ and $\beta_t^a = a_t^{exit} - a_t$. Online Appendix A.D shows that $\beta_t^r < 0$ and $\beta_t^a > 0$.[8] Thus, we have the following proposition.

PROPOSITION 1: *If wages are set to expected skill given the available information* (*equation* (1)), *then the introduction of an exit exam reduces the return to college reputation* ($\beta_t^r < 0$) *and increases the return to ability* ($\beta_t^a > 0$).

Proposition 1 yields a prediction regarding the role of college reputation in transmitting information on ability. If employers do not use reputation to set wages, a new signal of skill should have no effect on the relative weights of reputation and

[8] Since our regressions use log wages, the experience profiles reflect the reduction in uncertainty as information about the worker accumulates. Experience profiles can therefore differ for individuals with $\delta_i = 1$ and $\delta_i = 0$. To account for such effects, our regressions will include controls for experience that vary with individuals' potential access to the exit exams.

admission scores. If instead the exit exam causes employers to rely less on labor market reputation, $\mathscr{R}_s$, and more on other signals of worker skill, y_i, this reduces the effect of R_s (which is a better predictor of $\mathscr{R}_s$) and increases the effect of the admission score (which is a better predictor of y_i).

Though one could measure college reputation in many ways, our definition isolates a signaling mechanism because R_s contains no additional information on α_i given a student's individual score, τ_i. Proposition 1 thus captures how the introduction of new information shifts the weight in wage determination from the group to the individual level measure of ability. In contrast, other measures of reputation may be correlated with α_i even conditional on individual scores.

Our definition also helps distinguish a signaling channel from competing hypotheses such as accountability effects. For example, in our context, there is anecdotal evidence of colleges adding test preparation sessions after the exit exam introduction. The exit exams may also have prompted colleges to change their curricula or students to work harder. Such changes would affect skill formation while at college, included in v_s; they would not affect precollege ability, α_i, the focus of our analysis. Thus, while accountability related responses can explain changes in the return to reputation, they cannot explain a shift in the weight from R_s to individual admission scores. We describe the empirical evidence on signaling and accountability effects in Section II.

It is worth emphasizing that reputation as defined above is an *equilibrium* phenomenon (MacLeod and Urquiola 2015). It depends upon the more desirable colleges selecting individuals based on their observed ability rather than on other factors. Such an equilibrium is self-enforcing in the sense that students have an interest in working hard to get into the best college. However, a consequence of this effect is that it makes it difficult for the market to observe the quality of education. This happens because the market only observes the overall skill of graduates, and thus cannot easily disentangle college value added from selection. Our results will provide some direct evidence of this effect.

D. *Predictions for Wage Growth*

In Section III, we describe how the returns to reputation and ability change with experience, t, thereby comparing college reputation to other signals of ability studied in the literature. Previous research makes a distinction between *conditional* returns, given by equation (2), and *unconditional* returns, given by

$$\hat{w}_{it} = r_t^u R_{s_i} + e_{it}^R; \tag{4}$$

$$\hat{w}_{it} = a_t^u \tau_i + e_{it}^\tau. \tag{5}$$

The unconditional returns to reputation, r_t^u, and to ability, a_t^u, are the coefficients on reputation and the admission exam score in these separate regressions. In online Appendix A.E, we show that the evolution of the regression coefficients from (2), (4), and (5) satisfy Proposition 2.

PROPOSITION 2: *If wages are set equal to expected skill given the available information, then:*

(*i*) *The unconditional return to reputation, r_t^u, does not change with experience.*

(*ii*) *The unconditional return to ability, a_t^u, rises with experience.*

(*iii*) *The conditional return to reputation, r_t, is smaller than the unconditional return, and with experience falls to v_1, the reputation premium.*

(*iv*) *The conditional return to ability, a_t, is smaller than the unconditional return, and rises with experience.*

Parts (i)–(ii) of Proposition 2 mirror Farber and Gibbons' (1996) predictions that observable characteristics are fully incorporated in initial wages, while employers gradually learn about unobservable traits. Reputation, R_s, has a constant effect because it is observed at the time of hire; and signals from worker output, y_{it}, merely confirm employers' expectations. The effect of the admission score, τ_i, grows with experience because it is initially unobservable to employers and correlated with y_{it}.

Parts (iii)–(iv) predict a declining conditional return to reputation, and an increasing conditional return to ability. These match Altonji and Pierret's (2001) predictions for observable and unobservable characteristics, but our measure R_s makes for a clean test of the role of reputation in signaling. Since reputation is mean college admission score, τ_i is a sufficient statistic for ability, α_i, in regression (2). Thus, part (iii) of Proposition 2 holds even if employers imperfectly observe R_s, or if α_i is correlated with human capital growth; all of these effects are captured in the admission score coefficients in equation (2).[9] The return to reputation should decline unless there is a time-varying effect of other college membership attributes, v_s, and these attributes are correlated with reputation $(v_1 > 0)$.

Thus, Proposition 2 allows us to explore whether the return to reputation arises solely because college identity signals ability as measured by admission scores. Rejection by the data would suggest that other college membership attributes lead reputation to be correlated with wage growth. We examine these hypotheses in Section III.

II. The College Exit Exam

Proposition 1 provides predictions for the introduction of an exit exam under a competitive labor market model. This section explores the empirical evidence related to these predictions. We first discuss institutional background and our measure of reputation. We then turn to the exit exam, sample, empirical specifications, and results.

[9] The assumption that R_s contains no information for α_i conditional on τ_i may not hold if α_i affects individuals' choice of college beyond their admission scores. In this case, the optimal predictor of α_i could also include R_s since τ_i is a noisy measure of ability. We discuss this possibility in Section IIIC, where we explore sorting into colleges on other dimensions.

A. *Background and Data Sources*

Colombia's higher education system consists of public and private institutions that award various types of degrees. In this paper, we refer to "colleges" as institutions that award the equivalent of US bachelor's degrees after four or five years of study. Colombia also has institutions that specialize in two-year or three-year degrees. We set these aside to focus on institutional identity within a single schooling level.[10]

To apply to college, students are required to take a standardized exam, the Icfes.[11] The Icfes is generally analogous to the SAT, but it is taken by the vast majority of high school seniors regardless of whether they intend to apply to college.[12] The Icfes plays a major role in college admissions: many schools extend admission offers based solely on students' Icfes performance; others consider additional factors, and a handful administer their own exams.

We use student names, birthdates, and national ID numbers to link individual-level administrative datasets from three sources:

- The Colombian Institute for Educational Evaluation provided scores for all high school seniors who took the Icfes between 1998 and 2012. It also provided college exit exam fields and scores for all exam takers in 2004–2011 (discussed below).
- The Ministry of Education provided enrollment and graduation records for students entering college between 1998 and 2012. These include enrollment date, graduation or dropout date, program of study, college, and aggregate percentile on the Icfes exam. These data cover roughly 90 percent of all college enrollees; the Ministry omits a number of smaller colleges due to poor and inconsistent reporting.
- The Ministry of Social Protection provided monthly earnings records for formal sector workers during 2008–2012. These come from data on contributions to pension and health insurance funds. We calculate average daily earnings by dividing base monthly earnings for pension contributions by the number of formal employment days in each month and averaging across months.[13] This agency also provided four-digit economic activity codes for the first job in which a worker appears in their records.

[10] The Ministry of Education classifies institutions into five types: universities, university institutes, technology schools, technology institutes, and technical/professional institutes; we define the first two as colleges. We also focus on the Ministry's "university-level" majors, which have normative durations of four to five years.

[11] Icfes stands for Institute for the Promotion of Higher Education, the former acronym for the agency that administers the exam. The agency is now the Colombian Institute for Educational Evaluation, and the exam is called Saber 11. We use the name Icfes to match the designation during the period covered by our data.

[12] Angrist, Bettinger, and Kremer (2006) and our personal communications with the Colombian Institute for Educational Evaluation suggest that more than 90 percent of high school seniors take the exam. The test-taking rate is high in part because the government uses Icfes exam results to evaluate high schools.

[13] Our theoretical predictions are for log wages, but our records only allow us to calculate earnings per day, not per hour. Colombian labor market survey data shows that hours are relatively constant early in college graduates' careers, which suggests that our results are not due to the use of daily earnings.

B. *Ability and College Reputation*

We define two measures of ability that correspond to those in the theory (Section I). The first is student i's score on the Icfes admission exam, which we denote by τ_i. Throughout, we express Icfes scores as percentiles relative to all high school seniors who took the exam in the same year. The second is the reputation of a college s, denoted by R_s, defined as the mean Icfes score of its graduates.[14] To avoid capturing any effects from the exit exam rollout on reputation, we calculate R_s using graduates who took the Icfes exam in 2000–2003.

Icfes and reputation are divided by ten so that both measures range from 0–10 and one unit is 10 percentile points. One unit of reputation is about one standard deviation in this measure, and it is roughly sufficient to move from either the seventy-fifth to the one-hundredth percentile, or from the fiftieth to the seventy-fifth percentile. Anecdotally, a student applying to a very top college might also apply to one with one point lower in reputation as a "safety school."

Figure 1 shows that there is substantial variation in ability both across and within colleges. The horizontal axis depicts the reputation of 136 colleges that have at least 10 graduates per cohort. The height of the black dots indicates the median Icfes percentile among graduates from each school, while the vertical bars show twenty-fifth to seventy-fifth percentile ranges. There is a mass of colleges near the middle of the reputation distribution and fewer near the extremes. In addition, graduates from the same college differ significantly in ability. For example, the interquartile range at the median institution is 32 percentile points, which extends beyond the mean Icfes values of more than 80 percent of all colleges.

C. *The Exit Exam*

In 2004, the agency that administers the Icfes test began another major initiative by introducing field-specific college *exit* exams. These exams are standardized and administered in every college that offers a related program. Exam fields range from relatively academic in orientation (e.g., economics and physics) to relatively professional (e.g., nursing and occupational therapy). The stated intent of this effort was to introduce elements of accountability into the college market. School-level aggregate scores were made available and used by news outlets as part of college rankings.

Rather than focus on its accountability dimension, we analyze the exit exam as potentially affecting students' capacity to signal their skill. This is consistent with anecdotal evidence that many students list exit exam scores on their CVs or on online profiles.[15] The exit exam may also affect faculty recommendations or

[14] In Colombia, students apply not just to a college but to a college/major pair. We define reputation at the college level to focus on the signaling component of a student's choice of institution. Major choice may also convey information about a student's ability. Below we show that our main results are similar when we define reputation at the college/major level.

[15] It may be puzzling that, anecdotally, some students list their exit but not their Icfes exam scores on their CVs. One potential explanation is that the Icfes scores are more difficult to interpret. The Icfes exam yields scores on eight or more different subjects, and during the period we analyze, the testing agency did not provide an aggregate score to students. By contrast, during the period of our analysis, the exit exams yielded a single score in a subject related to a student's major.

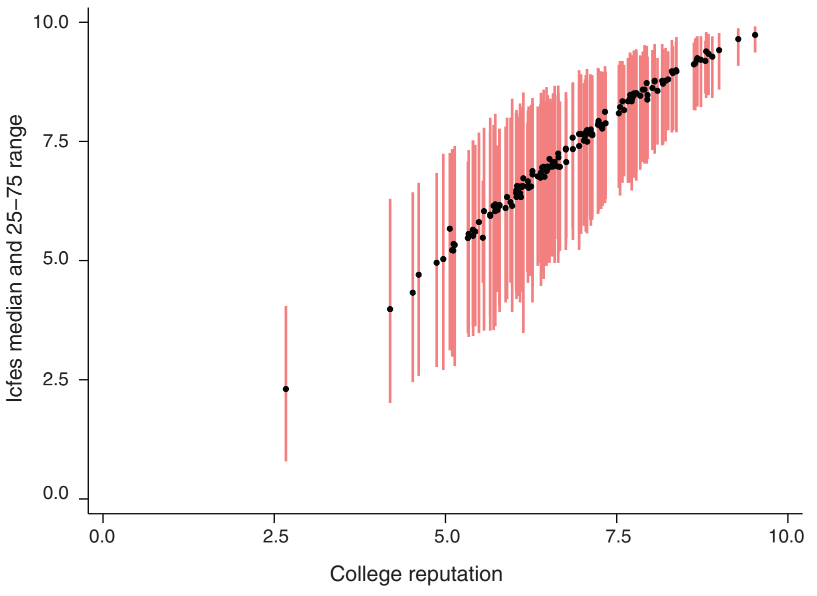

Figure 1. College Reputation and Individual Ability

Notes: The sample for this figure includes all high school seniors who took the Icfes in 2000–2003 and graduated from one of the 136 colleges with 40 or more graduates from the 2000–2003 Icfes cohorts (i.e., not less than ten per cohort). We define Icfes percentiles based on students' performance relative to all 11th grade exam takers in their same year. Percentiles are calculated using the average of eight core component scores: biology, chemistry, geography, history, language, mathematics, philosophy, and physics. College reputation is the mean Icfes percentile among graduates from each of the 136 colleges. Black dots are the median Icfes percentiles among graduates from each school, and vertical lines are the twenty-fifth to seventy-fifth Icfes percentile ranges.

students' search behavior after learning their position in the national distribution of exam takers.

D. *Identification*

To identify the effects of this new signal of skill, we exploit the gradual rollout of the exam fields in an "intent to treat" spirit. Exams were introduced in 55 fields between 2004 and 2007. The initial fields were those related to popular majors such as economics and industrial engineering; fields corresponding to less common degrees were introduced later (online Appendix B.A lists all fields and their introduction year). During this time the exams were not required, although they were taken by the majority of students in related majors. In 2009, the exit exam became mandatory for graduation, and a "generic competency" exam was made available for majors without a corresponding field.

Although the exit exams were field-specific, during the period we study there was no formal system assigning college majors to exam fields. This match is necessary to determine which majors were treated. We therefore perform this assignment ourselves using the Ministry of Education's 54 major groups, which we label

Table 1—Introduction of Exit Exam Fields and Matched College Programs

Exit exam fields	Matched programs	Program area	College programs		
2004 fields	30	Agronomy	Agronomy	Animal husbandry	Veterinary medicine
		Business	Accounting	Administration	Economics
		Education	Education		
		Engineering	Agricultural eng.	Architecture	Chemical eng.
			Civil eng.	Electrical eng.	Electronic eng.
			Environmental eng.	Food eng.	Industrial eng.
			Livestock eng.	Mechanical eng.	Systems eng.
		Health	Bacteriology	Dentistry	Medicine
			Nursing	Nutrition	Optometry
			Physical therapy		
		Social sciences	Communication	Law	Psychology
			Sociology		
2005 fields	5	Natural sciences	Biology	Chemistry	Geology
			Math/statistics	Physics	
2006 fields	1	Health	Surgical tools		
2007 fields	1	Social sciences	Physical education		
2009 generic exam	17	Engineering	Administrative eng.	Biomedical eng.	Mining eng.
			Other eng.		
		Fine arts	Advertising	Design	Music
			Plastic/visual art	Representative art	Other fine arts
		Health	Public health		
		Social sciences	Anthropology	Geography/history	Language/literature
			Library science	Philosophy	Political science
Total	54				

Notes: This table displays the match of college programs to the exit exam field and generic exam years. Programs are the Ministry of Education's 54 core knowledge groups, which are further categorized into the listed eight program "areas." Online Appendix B.A lists the exam fields and details how we match them to programs.

programs.[16] We assign each of the 54 programs to one of the 55 exam fields if the program name appears in the name of the field exam. We assign programs without matching names to the generic competency exam introduced in 2009. Online Appendices B.A and B.B describe this matching procedure and show that our main results are robust to several alternative matching methods.

Table 1 summarizes the resulting match. For each year, it lists the number of matched programs and the program areas they originate in. Programs related to agronomy, business, education, and health received exam fields almost exclusively in 2004, while natural science programs did so in 2005. Programs related to fine arts had no corresponding field until the introduction of the generic exam in 2009. Some programs in engineering and social sciences received fields in 2004, while others had none up to 2009. Most of our identification comes from a comparison of 2004

[16] These programs aggregate approximately 2,000 college major names that vary across and within schools. For instance, the Ministry might combine a major named "Business Administration" at one college with one labeled "Business Management" at another if it considers that these have similar content.

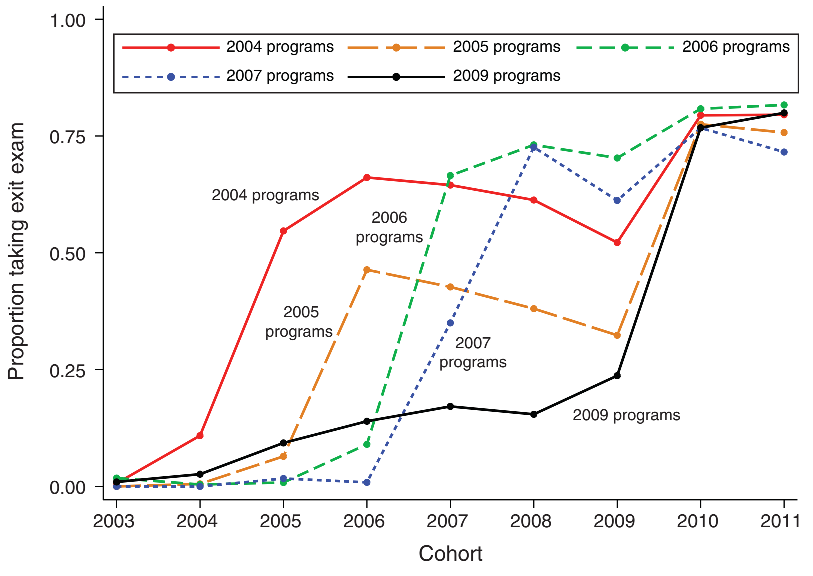

FIGURE 2. PROPORTION OF STUDENTS TAKING EXIT EXAM BY PROGRAM GROUP

Notes: Lines represent program groups defined by the year in which the program's assigned exit exam field was introduced (see Table 1). The figure includes 2003–2011 graduates from all programs in our data, even those excluded from our main analysis sample for reasons described below.

programs and 2009 programs. Engineering and social science programs potentially provide a compelling comparison because they appear in both groups.

We define a binary treatment variable δ_{pc}, which equals one if students in program p and graduation cohort c had an available exit exam in the matched field. Because students typically take the exam one year before graduating, the first treated cohort is that which graduated one year after the introduction of the field assigned to its program.[17] For example, $\delta_{pc} = 1$ for psychology students who graduated in 2005 or later because the psychology field exam was introduced in 2004; $\delta_{pc} = 0$ for all anthropology students who graduated before 2010 because the testing agency did not produce a related exam field.

Figure 2 shows that the introduction of exit exam fields led to sharp increases in the fraction of students taking the test. For example, the test-taking rate in 2004 programs jumped from 10 to 55 percent with the 2005 cohort, the first we define as treated for this program group. Students in 2009 programs rarely took the exam until the cohort following the exit exam mandate in 2009.[18]

To summarize, we define a treatment indicator, δ_{pc}, at the program-cohort rather than at the individual level. Thus, we analyze the introduction of the exams in an "intent to treat" spirit. This reflects that beyond the fact that students were not

[17] Across all cohorts in our sample, approximately 58 percent of test takers took the exam one year before graduation; 20 percent took it in the year of graduation; and 22 percent took it two or more years before.

[18] The existence of exam takers in the 2003–2004 cohorts indicates that a small number of students took the exam in their final year or after graduating. The 75 percent test-taking rate in the 2010–2011 cohorts suggests that compliance with the exam mandate was not universal.

Table 2—Summary Statistics for Exit Exam Sample

	Year program received exit exam				
Variable	2004	2005	2006	2009	All
Number of graduates in 2003–2009	131,962	2,014	1,043	11,033	146,052
Number of earnings obs. in 2008–2012	528,435	7,418	4,516	41,433	581,802
Number of programs	27	1	1	10	39
Number of colleges	94	5	5	21	94
Reputation	7.45	8.50	5.88	8.26	7.52
	(1.21)	(0.66)	(0.42)	(0.96)	(1.21)
Icfes	7.66	9.04	6.36	8.60	7.74
	(2.29)	(1.09)	(2.27)	(1.71)	(2.26)
log average daily earnings	10.87	10.71	10.66	10.84	10.87
	(0.70)	(0.66)	(0.51)	(0.76)	(0.70)
Return to reputation	0.138	0.041	−0.224	0.031	0.133
	(0.019)	(0.040)	(0.063)	(0.049)	(0.020)
Return to ability	0.028	0.009	0.015	0.049	0.029
	(0.003)	(0.010)	(0.014)	(0.013)	(0.003)

Notes: log average daily earnings are for the year 2012. Parentheses contain standard deviations except for the returns to reputation and ability. These rows display coefficients on reputation and Icfes from a regression of log average daily earnings in 2008–2012 on these two variables, program-cohort dummies, and a quadratic in experience (defined as calendar year minus graduation cohort) interacted with program dummies. We run these regressions separately for each program group using only 2003–2004 graduates. The parentheses under these coefficients contain standard errors clustered at the college level.

required to take exit exams during the period we study, they had no obligation to disclose their performance if they did (although not doing so might in itself convey information). Thus, while we can assert that the introduction of the exam into a student's field potentially affected the information available in that individual's labor market, we do not know precisely how it affected what firms observed about her.[19]

E. *Sample*

We analyze the effects of the exit exam using the 2003–2009 graduation cohorts. With these we can focus cleanly on the period in which signals of skill were introduced into a subset of fields.[20] Table 2 presents summary statistics separately for program groups defined by the year each program received its assigned exit exam field. Approximately 90 percent of students graduate from programs that received an exam field in 2004; most of the remaining graduates had no corresponding field until the 2009 generic exam.

We observe earnings for these graduates in 2008–2012. This means that we only observe earnings several years after graduation for cohorts prior to the exit exam introduction (2003–2004), while we observe earnings closer to graduation for cohorts after. The next section describes how we address this data constraint.

[19] The potential endogeneity of exam taking also explains why we do not use the exit exam scores in our main analysis, either to define reputation or as a measure of graduates' skill. It is also possible that employers' perceptions of students from programs without exit exams were altered by information on programs in the same colleges that received exams. However, such spillover effects would bias our results toward finding no effects of the exit exams.

[20] This is no longer clearly the case after the 2009 cohort due to several structural changes in the exit exams.

Our sample includes 39 programs offered at 94 colleges. These numbers are smaller than the total number of programs defined by the Ministry of Education (54) and the number of colleges in our records (136). We exclude programs and colleges that have too few observations to precisely estimate a return to reputation among graduates from the same program—a necessity for our empirical specification below. Online Appendices B.C and B.D provide details on the sample selection and show that our main results are robust to the key restrictions.

All colleges in the sample offer at least one of the 27 programs with a 2004 exam field, while only 25 schools offer one or more of the 12 programs with post-2004 exam fields. The distribution of Icfes scores is right-skewed with mean around the seventy-seventh percentile—or 7.7 points. This reflects the fact that less than half of all high school graduates eventually enroll in college and, of those, about 50 percent graduate. Colleges that offer 2009 programs have reputations that are about 8 percentile points higher on average than colleges that offer 2004 programs, but their graduates have slightly lower average daily earnings.

The last two rows in Table 2 report the returns to reputation and ability (Icfes) within each program group. These are analogous to the r and a coefficients from equation (2) in Section I, except that these are averages across the multiple years of earnings we observe (2008–2012). In Table 2, we use only the two pre-exit exam cohorts (2003–2004) to estimate these returns; this provides a useful benchmark for the results below. 2004 programs have higher returns to reputation than the other program groups; a 10 percentile increase in college reputation is associated with a 14 percent increase in earnings for 2004 programs, but only a 3 percent increase for 2009 programs.[21]

These differences in program characteristics and returns raise questions as to whether delayed exit exam programs are a good counterfactual for early exit exam programs. We adopt several strategies to address these in our empirical analysis below.

F. *Empirical Specifications and Results*

This section estimates a benchmark specification that tests the effects of the exit exam on the returns to reputation and ability. We complement these results with four types of robustness checks. First, we add further controls for labor market experience and graduation cohort to address issues related to the structure of our data and to the years for which we observe earnings. Second, we restrict identification to programs with similar characteristics to address the nonrandom rollout of exam fields. Third, we explore the sensitivity of our results to competing hypotheses and other measures of college reputation. Fourth, we use balance and placebo regressions to test for differential sorting or concurrent macroeconomic trends.

[21] The negative return to reputation for the 2006 program illustrates the empirical challenge of trying to estimate a return to reputation within each program. Not only can these returns be noisy when only a few schools offer a program, but the value of going to a higher ranked school depends on the labor market that students from the program commonly enter (in this case, the program trains surgical instruments technicians). For related issues, see Hastings, Neilson, and Zimmerman (2013) and Rodríguez, Urzúa, and Reyes (2015).

Benchmark Specification.—We follow Card and Krueger (1992), who ask how state-level policies affect the rate of return to education. Note that the return to education is a slope—the impact of years of schooling on earnings. The issue we tackle is analogous—we ask how the exit exams affected the impacts of college reputation and Icfes on earnings. Our benchmark specification relates changes in the returns to reputation and ability to the staggered rollout of the exam fields. Consider the regression

$$w_{ipct} \;=\; d_{pc} + f_p(t) + r_{pc} R_{s_i} + a_{pc} \tau_i + e_{ipct}, \tag{6}$$

where w_{ipct} is the log average daily earnings for student i in program p, graduation cohort c, and with potential labor market experience t, defined as calendar year minus graduation cohort. Variables d_{pc} are dummies for program-cohort cells and $f_p(t)$ is a quadratic in experience interacted with program dummies. This "first-step" specification estimates returns to college reputation, r_{pc}, and to ability, a_{pc}, separately for each program-cohort cell.

A second-step regression relates these returns to our treatment variable δ_{pc}, which equals one for students with exit exam fields assigned to their program and cohort. For example, the second-step specification for the return to reputation is

$$\hat{r}_{pc} \;=\; \mu_p + \mu_c + \beta^r \delta_{pc} + \upsilon_{pc}, \tag{7}$$

where μ_p and μ_c are program and cohort dummies and υ_{pc} is the residual. This is a standard differences-in-differences specification applied to slopes rather than to levels—it controls for average program and cohort differences in the returns to reputation (via the fixed effects μ_p and μ_c) and identifies the effect of the exit exam, β^r, through changes in returns across both programs and cohorts.

Card and Krueger (1992) use a two-step procedure. We opt for a single-step specification to identify changes in the *relative* weights of college reputation and Icfes on earnings. Plugging (7) and a similar equation for $\hat{a}_{pc}$ into (6) yields our benchmark specification:

$$w_{ipct} \;=\; d_{pc} + f_p(t) + (\mu_p + \mu_c + \beta^r \delta_{pc}) R_{s_i} + (\nu_p + \nu_c + \beta^a \delta_{pc}) \tau_i + e_{ipct}. \tag{8}$$

Specification (8) is analogous to equation (3) from Section I, but it uses differences-in-differences variation in treatment. It controls for program-specific experience effects and level differences in daily earnings across program-cohort cells; and it allows each program and cohort to have different returns to reputation and Icfes through the μ and ν dummies. The coefficients of interest, β^r and β^a, are identified off variation in exposure to the exit exam across both programs and cohorts, defined by our treatment variable δ_{pc}.

Proposition 1 predicts $\beta^r < 0$ and $\beta^a > 0$. This comes from the assumption that employers use both labor market reputation, $\mathscr{R}_s$, and other signals of worker skill, y_i, in setting initial wages. We assume that the exit exam increases the precision of y_i, for example, through the appearance of scores on CVs. Our measure of reputation, R_s, is a better predictor of $\mathscr{R}_s$, while Icfes scores, τ_i, are a better predictor

Table 3—Exit Exam Effects on Returns to Reputation and Ability

Dependent variable: log average daily earnings	Benchmark specification (1)	Experience and cohort controls: Within experience (2)	Experience and cohort controls: Linear trends (3)	Restriction to similar programs: S. sciences and engineering (4)	Restriction to similar programs: Within $\hat{r}_p$ quartiles (5)	Restriction to similar programs: Within $\hat{a}_p$ quartiles (6)
Reputation × δ_{pc}	−0.041	−0.033	−0.034	−0.046	−0.018	−0.053
	(0.017)	(0.015)	(0.028)	(0.026)	(0.010)	(0.017)
Icfes × δ_{pc}	0.017	0.018	0.012	0.038	0.018	0.008
	(0.006)	(0.007)	(0.009)	(0.010)	(0.007)	(0.004)
Observations	581,802	267,924	267,924	273,590	581,802	581,802
R^2	0.258	0.224	0.224	0.266	0.258	0.258
Number of programs	39	39	39	22	39	39
Experience levels	0–9	4–7	4–7	0–9	0–9	0–9

Notes: All columns report coefficients on the interactions of reputation and Icfes with the treatment variable δ_{pc}. Regressions in columns 1 and 3–6 include a quadratic in experience interacted with program dummies, dummies for program-cohort cells, and interactions of both reputation and Icfes with program and cohort dummies. Column 2 includes dummies for program-cohort-experience cells and interactions of both reputation and Icfes with program-experience and cohort-experience dummies. The sample for each regression is restricted to the experience levels listed in the bottom row. Parentheses contain standard errors clustered at the program level. Column 3 adds interactions of both linear experience and cohort terms with college reputation and Icfes for each program. Column 4 restricts the sample to social sciences and engineering program areas and adds interactions of dummies for social-science-area-cohort cells with both reputation and Icfes. Column 5 adds interactions of both reputation and Icfes with dummies for cells defined by cohort and each program's quartile of the returns to reputation estimated from 2003–2004 cohorts. Column 6 adds interactions of both reputation and Icfes with dummies for cells defined by cohort and each program's quartile of the returns to Icfes estimated from 2003–2004 cohorts.

of y_i. Thus, as the market relies less on $\mathscr{R}_s$ and more on y_i, the return to reputation falls $(\beta^r < 0)$ and the return to ability rises $(\beta^a > 0)$.[22]

Column 1 of Table 3 estimates benchmark specification (8). Like all other columns in Table 3, it reports only the β^r and β^a coefficients on the interactions of reputation and Icfes with our treatment variable δ_{pc}. The results suggest that relative to students in programs and cohorts without exams, students exposed to the exit exams see their daily earnings become more correlated with incoming collegiate ability and less correlated with college reputation. The reputation effect is slightly lower than one-third of the mean return to reputation in Table 2; the Icfes coefficient is slightly higher than one-half of the mean return to Icfes.[23]

Figure 3 illustrates the benchmark results in column 1 using only 2004 and 2009 programs. Panel A displays the linear relationship between reputation and residuals from a regression of log earnings on Icfes, experience, and program-cohort cells. The lighter colored lines depict programs with 2004 exit exam fields (Table 1) and the black lines contain programs that did not receive a field until 2009. In each case the solid lines describe students who graduated prior to the introduction of all exit exams, and the dashed lines describe students who graduated after the introduction

[22] Although this prediction results from higher precision in employers' initial information set, the changes in the relative returns to reputation and Icfes are also evident (but less pronounced) at periods $t > 0$ because wages continue to reflect initial information. Our data do not allow us to observe early career earnings for pre-exit exam cohorts (2003–2004), so our estimates reflect changes in returns at higher experience levels.

[23] Online Appendix B.E presents the program-cohort level returns to reputation and Icfes from the first-step equation (6). Averaging and differencing these returns yields estimates similar to column 1 of Table 3.

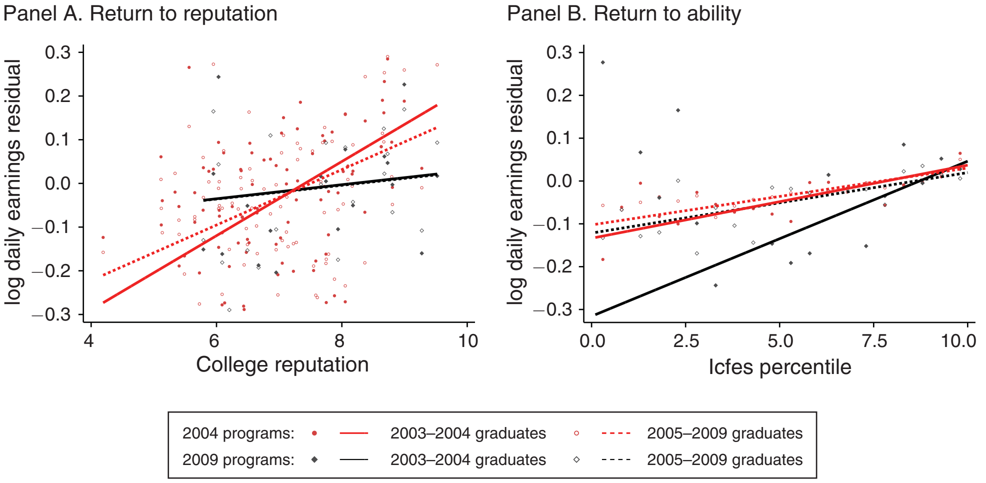

Figure 3. Exit Exam Effects—2004 and 2009 Programs

Notes: In panel A, the dependent variable is the residual from regressing log average daily earnings on Icfes, an experience quadratic interacted with program dummies, and program-cohort cell dummies separately for each program and cohort group. Lines depict the linear relationship between these earnings residuals and college reputation for each program and cohort group. Dots are the mean earnings residual at each college, calculated separately for each program and cohort group. In panel B, the dependent variable is the residual from regressing log average daily earnings on reputation, an experience quadratic interacted with program dummies, and program-cohort cell dummies separately for each program and cohort group. Lines depict the linear relationship between these earnings residuals and Icfes percentiles for each program and cohort group. Dots are the mean earnings residual in each of 20 equally-spaced Icfes percentile bins, calculated separately for each program and cohort group.

of the initial exam fields. In 2004 programs, earnings are less correlated with reputation in cohorts following the exit exam introduction. In 2009 programs, the correlation between reputation and earnings is similar in all cohorts.

Panel B of Figure 3 displays the analogous linear relationship between Icfes and log earnings residuals that control for reputation. The correlation between Icfes and earnings declines across cohorts in both program groups, but the decline is more pronounced in programs without an exam field. This is consistent with a stronger correlation between earnings and ability in early exit exam programs in the presence of an aggregate decline in the return to Icfes.

There are two sources of caution in interpreting the results from equation (8)—one related to data constraints and one related to identification. The first arises because our data cover only seven cohorts with earnings observed over five years; hence, we do not observe pretreatment cohorts at very early experience levels. The second relates to possible violations of the usual assumption of parallel trends implicit in differences-in-differences estimation; evidence that such violations may be important comes from Table 2 and from the different pre-exit exam slopes in Figure 3. We now describe robustness checks that address these two issues.

Experience and Cohort Controls.—Our sample includes 2003–2009 cohorts with earnings measured in 2008–2012. This means we cannot disentangle a first-period effect of the exit exam from an effect that varies with experience because we do not observe initial earnings for pre-exit exam cohorts. As a result, our benchmark results are based on returns to reputation and ability that average across experience levels.

Our data structure raises concerns if there is variation across programs in how college reputation or ability correlate with the returns to experience. For example, suppose that the return to reputation rises more quickly with experience in programs with early exit exam fields. This could mechanically generate a $\beta^r < 0$ estimate since the post-exam cohorts (2005–2009) have lower potential experience than the pre-exam cohorts (2003–2004).

To address this, we add further controls for experience to the benchmark specification. To illustrate, suppose we estimated equation (8) using only earnings at five years of potential experience, thus ensuring that we are comparing exposed and unexposed cohorts at the same seniority. This regression could only include 2003–2007 cohorts because we do not observe earnings five years out for 2008–2009 graduates. We could repeat this estimation for any level of potential experience at which we observe cohorts prior to the introduction of all exit exams, which is between four (using 2004–2008 graduates) and seven (using 2003–2005 graduates) years of experience.[24] This procedure would yield four college reputation treatment effects and four Icfes treatment effects, one for each year of potential experience. We combine these into a single estimate by removing the experience quadratics from equation (8), restricting observations to those between four and seven years of experience, and fully interacting all fixed effects with experience dummies:

$$(9) \qquad w_{ipct} = d_{pct} + (\mu_{pt} + \mu_{ct} + \beta^r \delta_{pc}) R_{s_i} + (\nu_{pt} + \nu_{ct} + \beta^a \delta_{pc}) \tau_i + e_{ipct},$$

where d_{pct} are fixed effects for program-cohort-experience cells, and μ and ν are fixed effects for program-experience and cohort-experience cells. The coefficients β^r and β^a are thus averages of the experience-specific estimates, identified only off variation within experience levels. If unobserved program-level variation in the interaction of reputation and experience mechanically biases our estimate of β^r downward, including these experience controls should move the estimated coefficient toward zero.

The addition of experience controls decreases the magnitude of the reputation effect only slightly (column 2, Table 3). Program differences in the returns to experience do not appear to drive the reduction in the return to reputation. This is also true for the return to Icfes; the estimates in columns 1 and 2 of Table 3 are nearly identical.

A related test is to allow the returns to reputation and ability to follow program-specific linear trends in both experience t and cohort c. For this we add linear trend interactions with reputation ($\mu_p t R_s$ and $\mu_p c R_s$) and with Icfes ($\nu_p t \tau_i$ and $\nu_p c \tau_i$) to the benchmark specification.[25] Including experience trends alone

[24] In principle, we can identify treatment effects using post-2004 cohorts since two programs in our sample received the exit exam in 2005 and 2006. In practice, over 90 percent of our sample is comprised of students from 2004 programs, so regressions that exclude the 2003–2004 cohorts yield noisy estimates.

[25] The full specification with linear trends in experience and cohort is

$$w_{ipct} = d_{pc} + f_p(t) + (\mu_p + \mu_p t + \mu_p c + \mu_c + \beta^r \delta_{pc}) R_{s_i} + (\nu_p + \nu_p t + \nu_p c + \nu_c + \beta^a \delta_{pc}) \tau_i + e_{ipct}.$$

yields similar estimates to those from specification (9) since we limit the sample to earnings between four and seven years of experience. Adding cohort trends is the typical differences-in-differences test of adding linear terms in the "time" dimension. Cohort trends absorb linear program-specific paths in the returns to reputation and ability that predate the exit exam and should have a measurable impact on our point estimates if these paths are important.[26]

The results appear in column 3 of Table 3. The coefficient on the reputation effect is nearly identical to column 2, while the Icfes effect falls only slightly. The consistency of these magnitudes argues against the hypothesis of divergent trends across programs, although the estimates in column 3 are substantially less precise. This loss in precision suggests the effects of exit exam were not immediate but rather materialized over several years—an intuitive result if the market processed the tests gradually.

Restriction to Similar Programs.—Our key identifying assumption is that in the absence of the exit exams, there would have been parallel trends in the returns to reputation and ability among programs exposed and not exposed to the exams. One fact that might cast doubt on this is that programs that received exams early have higher returns to reputation (Table 2). To address this we focus on comparable programs. We do so in three ways: restricting attention to social sciences and engineering, areas that have multiple programs in different exam year groups (see Table 1);[27] stratifying programs by quartiles of the pre-exit exam returns to reputation; and stratifying programs by quartiles of the pre-exam returns to Icfes. In each case, we define program groups G and supplement equation (8) with dummies for group-cohort cells interacted with reputation and Icfes (e.g., $\mu_{Gc} R_s$ and $\nu_{Gc} \tau_i$).[28] Thus, β^r and β^a are only identified by variation in exposure to the exit exam within groups of programs that have common characteristics.

Column 4 in Table 3 uses only programs in social sciences and engineering. The reputation effect is similar in magnitude to those in previous columns, while the Icfes effect is more than double. Both are statistically significant at the 10 percent level despite the fact that the program restriction substantially reduces precision.

In column 5 of Table 3, we define program groups by pre-exit exam returns to reputation. We first estimate a return to reputation for each of the 39 programs in our sample using 2003–2004 graduates (i.e., $\hat{r}_{p,2003-2004}$).[29] We then define program groups G by quartiles of these returns, with nine to ten programs per group. This directly addresses the concern that 2004 programs have higher returns to reputation—in this case, we compare delayed exam programs with low reputation returns

[26] Our ability to control for preexisting cohort trends is limited, however, because we only observe two cohorts prior to the exit exam introduction (2003–2004).

[27] The health program area also includes a single program with a delayed exit exam field (surgical tools). Estimates analogous to column 4 of Table 3 that include health programs yield similar coefficients, but they are not significant because identification in the health program area comes from this single program.

[28] The full specification with program group controls is

$$w_{ipct} = d_{pc} + f_p(t) + (\mu_p + \mu_c + \mu_{Gc} + \beta^r \delta_{pc}) R_{s_i} + (\nu_p + \nu_c + \nu_{Gc} + \beta^a \delta_{pc}) \tau_i + e_{ipct}.$$

[29] For this, we estimate equation (6) using only 2003–2004 graduates and replace the r_{pc} and a_{pc} coefficients with r_p and a_p. Online Appendix B.E presents these program-specific returns to reputation (and returns to ability).

only to the subset of 2004 programs with similarly low returns. The reputation effect in column 5 is smaller than in earlier specifications, consistent with some inflation in our estimates due to pretreatment differences; but it is still significant because the standard error decreases. This suggests that the effects in this specification are identified off more similar programs because there is less noise in estimating treatment effects.

Column 6 of Table 3 is similar to column 5, but we define program groups as quartiles of pre-exit exam returns to Icfes (i.e., $\hat{a}_{p,\,2003-2004}$). This specification tests the influence of pretreatment program differences in returns to ability. The resulting Icfes effect is also smaller than in the benchmark specification but more precisely estimated.

Competing Hypotheses.—The results in Table 3 are consistent with the exit exams transmitting information on ability, but the exams may have had other non-informational consequences. For example, school-mean exit exam scores were publicized, which may have altered employers' perceptions of colleges' labor market reputations. The exit exams may also have prompted colleges to change curricula or add test preparation sessions, or individuals to work harder in preparation for the exams. Such accountability related reforms may also have affected the observed returns to reputation and ability.

In online Appendix A.F, we develop theoretical predictions that help distinguish between the informational and accountability related impacts of the exit exams. The key insight is that accountability responses affect individuals' skill accumulation while in college—which is captured by the v_s term in our model—and not their precollege ability, α_i. This generates different predictions as to how the exam introduction affects the returns to college reputation and to ability.

For example, in regressions that include both R_s and τ_i, an informational mechanism predicts a decrease in the return to reputation and an increase in the return to Icfes scores. However, accountability mechanisms predict no effect on the conditional return to τ_i because R_s is a better measure of college membership attributes, v_s. Thus, while accountability responses could potentially explain the reputation effects in Table 3, they cannot explain the Icfes effects.

Further, signaling and accountability mechanisms generate different predictions regarding the *unconditional* returns to R_s and τ_i —that is, the coefficients in regressions that include only one of these characteristics. Under a signaling hypothesis, the exit exams should have no effect on the unconditional return to reputation if R_s is observed by employers, since information from individuals' exit exam scores merely confirms the labor market's expectations on average. Further, the exit exams should lead to an increase in the unconditional return to τ_i, but this increase should be smaller than when one also includes R_s in the regressions. By contrast, the main prediction from an accountability hypothesis is that the unconditional returns to R_s and to τ_i should have the same sign; this arises because any effects of the exams on v_s that are correlated with R_s will also, by definition, be correlated with τ_i.

Columns 1 and 2 in Table 4 present these unconditional returns to reputation and to ability by replicating our benchmark specification with only reputation terms (column 1) or only Icfes terms (column 2) included. As predicted by the signaling

Table 4—Exit Exam Effects under Other Reputation Measures and Hypotheses

	Unconditional returns		Other reputation measures		
Dependent variable: log average daily earnings	R_s only (1)	Icfes only (2)	Mean Icfes at college-program (3)	1 − admit rate at college (4)	Mean log earnings at college (5)
Reputation × δ_{pc}	−0.029		−0.038	−0.122	−0.044
	(0.018)		(0.022)	(0.066)	(0.080)
Icfes × δ_{pc}		0.005	0.019	0.012	0.012
		(0.007)	(0.007)	(0.006)	(0.006)
Observations	581,802	581,802	581,802	581,802	581,802
R^2	0.253	0.231	0.258	0.236	0.274
Number of programs	39	39	39	39	39
Mean return to reputation	0.161		0.132	0.098	0.700
Mean return to ability		0.069	0.027	0.064	0.035

Notes: All columns report coefficients on the interactions of reputation and Icfes with the treatment variable δ_{pc}. Parentheses contain standard errors clustered at the program level. Regressions in columns 1–2 are identical to column 1 in Table 3, except column 1 excludes Icfes and all its interaction terms; and column 2 excludes reputation and all its interaction terms. Regressions in columns 3–5 are identical to column 1 in Table 3, except they use different reputation measures. Column 3 defines reputation as in our benchmark procedure (i.e., mean Icfes), but at the college-program level rather than the college level. Column 4 defines reputation as one minus the college admission rate (i.e., 1 − admitted/applied) using aggregate admission data from the Ministry of Education. We include only university-level programs with a positive number of applicants and admitted students in a given cohort, and we average across all cohorts for which we have data (2007–2013). Column 5 defines reputation as the mean log daily earnings at each college using 2003–2004 graduates in our sample. We include only earnings at five years of potential experience, the earliest we can observe for both cohorts. The mean returns to reputation and ability are calculated from specifications similar to those in each column, but they use only 2003–2004 graduates and include only a single reputation and Icfes term.

hypothesis, the unconditional return to reputation is not statistically different from zero, and the unconditional return to ability is positive but smaller than the conditional return (column 1, Table 3). Further, the unconditional returns to reputation and ability are oppositely signed. Thus, although the standard errors on these unconditional returns are too large to draw definitive conclusions, the results are more consistent with a signaling mechanism than with an accountability related mechanism.

Taken together, the effects of the exit exams on the conditional and unconditional returns show that the strongest empirical result is the shift in weight from a group-level measure of ability—reputation—to an individual measure—Icfes scores. This is the effect captured by our benchmark specification, and it is harder to explain through channels other than signaling.

Other Reputation Measures.—Our measure of reputation, R_s, captures the expected "admission exam" ability of graduates from a given college. The exit exams may also have provided information to employers on other dimensions of graduates' skill. Table 4 explores some of these. Columns 3–5 present results that use different measures of college reputation but are otherwise identical to our benchmark specification (Table 3, column 1).

Column 3 defines reputation as mean Icfes at the college-program level rather than the college level, which allows schools to have strengths that vary by major. This is relevant because Colombian students apply to college/major pairs. We use a school-level definition for our main analysis to focus on the information conveyed by a student's choice of institution, but major choice may provide additional information on an individual's ability. The magnitudes of the results in column 3 are nearly identical to those for our benchmark results, though the standard errors are larger. This likely reflects the fact that the college-program reputations are calculated from smaller samples.[30]

Column 4 defines a college's reputation as one minus its admission rate (this measure is thus positively correlated with R_s). The results mirror those in our benchmark specification. The similarity of these results reflects the fact that R_s is mechanically correlated with other desirable school attributes when colleges use admission scores to select students.

Column 5 defines reputation as the average log earnings of a college's graduates.[31] This yields our best measure of labor market reputation, $\mathscr{R}_s$, which includes both precollege ability, α_i, and attributes related to college membership, v_s. The exit exams led to an increasing return to Icfes and a lower return to reputation, though the reputation effect is statistically insignificant. Reputation measures like average earnings do not provide a clean test of signaling, however, because they may be correlated with ability even conditional on individual Icfes scores.

A related alternative hypothesis is that the exit exams enhanced the transmission of information on characteristics other than college reputation. This could explain the pattern of results in Table 3 if these characteristics are correlated with college reputation. To explore this hypothesis, in online Appendix B.F we replicate our benchmark regressions including other individual characteristics—gender, mother's education, and family income—instead of reputation. These alternative specifications show that the returns to other characteristics also fell with the exit exams, but none of the results are statistically significant. Further, in a specification that includes all characteristics jointly, only the reputation effects are significant. Although we cannot rule out signaling effects on characteristics not included in our data, this provides evidence that the strongest effects of the exit exams were on the return to college reputation.

Placebo and Balance Tests.—A further placebo test replicates our main analysis using college *dropouts* rather than graduates. Dropouts are a compelling placebo group because they enrolled in the same colleges and programs as graduates but exhibited little change in exam taking. Columns 1 and 2 in Table 5 document this by regressing an indicator for taking the exit exam on program dummies, cohort dummies, and our treatment variable, δ_{pc}. For graduates, exposure to the exit exam is associated with a 50 percentage point increase in the likelihood of taking the exam; for dropouts, it is unrelated.

[30] Online Appendix B.L replicates all the robustness tests in Table 3 using college-program level reputation. The point estimates for both the reputation effects and the Icfes effects are similar to our main results across all specifications. The standard errors are typically larger, however, and thus the reputation effects corresponding to columns 2 and 4 of Table 3 are not statistically significant.

[31] We calculate this using only pre-exit exam cohorts (2003–2004) and earnings measured five years after graduation, the earliest we can observe for these cohorts. Results are similar when we use earnings measured in the year of graduation for cohorts exposed to the exit exams.

Table 5—Placebo Test Using College Dropouts

	Dependent variable: Took the exit exam		Dependent variable: log average daily earnings	
	Graduates (1)	Dropouts (2)	Graduates (3)	Dropouts (4)
Exposed to exit exam (δ_{pc})	0.500	0.025		
	(0.054)	(0.020)		
Reputation × δ_{pc}			−0.041	0.011
			(0.017)	(0.032)
Icfes × δ_{pc}			0.017	−0.002
			(0.006)	(0.011)
Observations	146,052	77,586	581,802	259,258
R^2	0.335	0.026	0.258	0.118
Number of programs	39	39	39	39
Mean exam taking rate	0.070	0.162		
Mean return to reputation			0.133	0.040
Mean return to ability			0.029	0.032

Notes: The sample for columns 1 and 3 includes college graduates and their earnings observations (i.e., the same sample as in Table 2). The sample for columns 2 and 4 includes students from the same colleges and programs who dropped out in 2003–2009, and their earnings observations. The dependent variable in columns 1 and 2 is an indicator for taking the exit exam. The regressions include program dummies and cohort dummies, where cohorts are defined by graduation year for college graduates and drop-out year for college dropouts. We report the coefficient on the treatment variable δ_{pc}, which we define identically for graduation and dropout cohorts. The dependent variable in columns 3 and 4 is log average daily earnings. We report coefficients on the interactions of reputation and Icfes with the treatment variable δ_{pc}. Column 3 is identical to column 1 in Table 3. The specification includes a quadratic in experience interacted with program dummies, dummies for program-cohort cells, and interactions of both reputation and Icfes with program and cohort dummies. Column 4 uses the same specification with cohorts and experience defined by drop-out year. The means at the bottom of the table are calculated using only 2003–2004 graduates. In all regressions, parentheses contain standard errors clustered at the program level.

Column 3 of Table 5 replicates our benchmark result for graduates from specification (8) (Table 3, column 1). Column 4 of Table 5 estimates the same specification using dropouts. There is little evidence that changes in dropouts' returns to reputation and ability are correlated with the introduction of the exit exams. If anything, the return to reputation for dropouts increases with the exam rollout, although the coefficient is noisily estimated. The point estimate on the Icfes effect is close to zero. To the extent that dropouts and graduates are subject to similar enrollment or macroeconomic trends, this finding supports the notion that our main results are attributable to the exit exams.

Dropouts are not a perfect counterfactual for graduates for several reasons. First, as shown in Table 5, dropouts have smaller pretreatment returns to reputation, although these returns are positive on average. Thus, the information conveyed by dropouts' college identities may differ from that of graduates. Further, dropouts spent less time in college, and thus have less exposure to any potential accountability responses induced by the exit exam. However, if our main results were driven by accountability reforms rather than signaling, one would expect to see that dropouts who stayed in college longer have treatment effects more similar to those of graduates. We find no evidence of this.

The dropout placebo test is also consistent with balance regressions reported in online Appendix B.G, which ask whether the exit exam rollout was correlated with changes in graduates' observable characteristics. If the field-specific introduction of the exit exams were correlated with trends in school or program choice, this should appear as changes in average reputation or Icfes scores across programs. There is little evidence of this channel. Changes in reputation and Icfes scores in programs with access to the exit exams are small and statistically insignificant.[32]

Online Appendix B.G also explores the effect of the exit exams on the probability of formal employment—a potential sample selection concern since we do not observe earnings for non-employed or informal workers. The estimated effect is not statistically significant and small relative to the mean formal employment rate.

In sum, the introduction of a new signal of skill—the field-specific college exit exams—reduced the return to reputation and increased the return to ability. These results are most consistent with an informational effect of the exit exams, and they provide evidence that college reputation signals individual ability to the labor market.

G. *Complementary Effects of the Exit Exam*

There is suggestive evidence that the exit exam affected other outcomes. For example, column 1 in Table 6 shows its impact on time to graduation. This estimate is from a standard differences-in-differences regression that includes program dummies, cohort dummies, and our treatment variable, δ_{pc}. The result suggests that individuals in programs with exam fields took about one-quarter of a year longer to graduate. This is consistent with increased student effort, or with colleges taking steps to prepare students for the test. For instance, there is anecdotal evidence of colleges seeking to influence their students' performance, with activities ranging from "boot camp" preparation to more overt "gaming" via exclusion of certain students.[33]

Using a similar specification, column 2 in Table 6 presents evidence that earnings increased by 7 percent more in programs with early exam fields. This could have occurred if the exam improved match quality, raising overall productivity. It could also reflect students with access to the exam getting higher paying jobs at the expense of college dropouts and vocational school students, who are excluded from our sample.

Finally, we ask whether the exit exams altered individuals' school or program choices. This would be consistent with the government's stated intent. Column 3 of

[32] These results likely reflect high costs to switching programs in Colombia and the fact that our sample predominantly includes students who enrolled prior to the existence of any exit exams. Colombian colleges do not make it easy for students to change majors; switching essentially requires applying de novo.

[33] These results suggest that graduation cohort may be endogenous. We address this concern by estimating equation (8) with cohorts defined by *predicted* rather than actual graduation date, where predicted graduation is based on the year of enrollment. Online Appendix B.D shows that the results from this regression are similar to our benchmark specification; this suggests that selective graduation timing is not driving our main results.

Table 6—Complementary Effects of the Exit Exam

	Dependent variable		
	Years in college (1)	log daily earnings (2)	Enrollees' Icfes scores (3)
Exposed to exit exam (δ_{pc})	0.237	0.070	
	(0.110)	(0.019)	
Icfes reputation × $\delta_{p\tilde{c}}$			−0.162
			(0.053)
Exit exam reputation × $\delta_{p\tilde{c}}$			0.147
			(0.063)
Observations	146,052	581,802	485,350
R^2	0.132	0.201	0.277
Number of programs	39	39	39

Notes: The dependent variable in column 1 is graduation year minus enrollment year. The sample includes all students from Table 2. We report the coefficient on our treatment variable, δ_{pc}. The regression also includes program dummies and cohort dummies. The dependent variable in column 2 is log average daily earnings for all observed experience levels (zero to nine years). The sample includes all earnings observations from Table 2. In addition to δ_{pc}, the regression includes program dummies, cohort dummies, and a quadratic in experience interacted with program dummies. The dependent variable in column 3 is individual Icfes percentile. The sample includes all students who enrolled in one of the 94 colleges and 39 problems in Table 2 between 2003 and 2009. We calculate Icfes and exit exam reputation using students who took the Icfes in 2000–2008, took the exit exam in 2009–2011 (when the exam was mandatory), and graduated from one of the school-programs in our sample. We convert Icfes and exit exam scores into percentiles relative to this sample and within exit exam fields and years. We calculate reputation as means at the school-program level and normalize both measures so one unit represents 10 percentile points in this distribution of exam takers. We define the treatment variable $\delta_{p\tilde{c}}$ using enrollment cohorts $\tilde{c}$, with $\delta_{p\tilde{c}} = \delta_{pc}$ for $\tilde{c} = c$. We report coefficients on the interactions of Icfes reputation and exit exam reputation with the treatment variable, $\delta_{p\tilde{c}}$. The regression includes dummies for program-cohort cells and interactions of both reputation measures with program dummies and cohort dummies. In all regressions, parentheses contain standard errors clustered at the program level.

Table 6 explores how the ability of incoming students changed with the exit exam introduction. For this regression, we define two measures of reputation using a population of graduates who took the exit exam in 2009–2011, when it was required of all graduates. We define Icfes reputation as mean Icfes percentile at the *school-program* level. Similarly, exit exam reputation is the school-program mean exit exam percentile. We convert Icfes and exit exam scores to percentiles within this population so that both reputation measures are on the same scale.

Icfes and exit exam reputations are highly correlated but not perfectly so. We suppose that the exit exam reputation contains new information, and that this information gradually became available to students entering college starting with the 2005 enrollment cohort.

Column 3 presents a specification analogous to the benchmark equation (8) with two key differences. First, the sample includes 2003–2009 *enrollees* rather than graduates, and we define students as treated by the exit exam ($\delta_{p\tilde{c}} = 1$) if they began a program p in an enrollment cohort $\tilde{c}$ after the introduction of the assigned field. Second, the dependent variable is the Icfes percentile of entering students, and we replace the independent variables R_s and τ_i with the school-program measures of Icfes and exit exam reputation. The reported coefficients in column 3 reflect how the correlations of Icfes and exit exam reputation with incoming students' Icfes scores changed with the exit exam rollout.[34]

The results show that in programs with exams, the ability of incoming students became more correlated with exit exam reputation, and less correlated with Icfes reputation. In other words, school programs whose exit exam performance exceeded their average Icfes performance saw increases in the ability of their incoming classes. This suggests students selected different programs and/or colleges as new information on their quality became available.

III. College Reputation and Earnings Growth

The previous section showed that college reputation plays a signaling role. This section asks whether college reputation serves *only* to signal ability as measured by admission scores. To do so, it explores the predictions from Proposition 2 (Section I) on how college reputation correlates with initial earnings and with earnings growth.

A. *Sample*

We follow Farber and Gibbons (1996) and Altonji and Pierret (2001) in studying individuals making their initial transition to the labor force. We restrict our sample to individuals who: graduated in 2008 or 2009 (this allows us to observe earnings in the year of graduation and the next three years); and entered the labor market immediately upon graduation and remained during four consecutive years (i.e., they did not attend graduate school or leave the formal labor force).[35] The results are thus not attributable to movements into and out of the labor market.

B. *Empirical Specifications and Results*

Our basic specification is

$$w_{it} = d_{c_i t} + r_0 R_{s_i} + r\left(R_{s_i} \times t\right) + a_0 \tau_i + a\left(\tau_i \times t\right) + e_{it}. \tag{10}$$

The dependent variable, w_{it}, is log daily earnings for student i measured at potential experience t, which as before is employment year minus graduation year; $d_{c_i t}$ are graduation cohort c_i by experience t cell dummies; college reputation, R_{s_i}, and Icfes score, τ_i, are as before; r_0 is the return to reputation in the year of graduation; and r is the average change in the return to reputation from an additional year of potential experience; a_0 is period-zero return to ability; and a is the average yearly change in this return.[36] We report only coefficients on reputation, Icfes, and their interactions with experience, where the latter two are estimated using earnings only up to three

[34] The full specification, of which column 3 reports only the γ^{τ} and γ^{exit} coefficients, is

$$\tau_{ip\tilde{c}} = d_{p\tilde{c}} + (\mu_p + \mu_{\tilde{c}} + \gamma^{\tau}\delta_{p\tilde{c}})[\text{Icfes reputation}]_{sip} + (\nu_p + \nu_{\tilde{c}} + \gamma^{exit}\delta_{p\tilde{c}})[\text{Exit exam reputation}]_{sip} + e_{ip\tilde{c}}.$$

[35] Online Appendix B.H provides further details on the sample.

[36] Formally, we parametrize the experience-specific r_t (and a_t) coefficients in equation (2) as $r_t = r_0 + r \times t$.

Table 7—Returns to Reputation and Ability, and Experience Interactions

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

Notes: The dependent variable is log average daily earnings. The sample includes students in column (D) of Appendix Table B8 and earnings in the four years after graduation. Columns 1–3 estimate equation (10) excluding and including Icfes terms. In addition to the reported variables, both regressions include dummies for cohort-experience cells. Column 4 adds the following controls to column 3: age at graduation, a gender dummy, dummies for eight mother's education categories, dummies for missing age and mother's education values, college program dummies, and dummies for college municipalities. Each control is interacted with a quadratic in experience. Parentheses contain standard errors clustered at the college level.

years after graduation, the maximum we can observe for our sample of 2008–2009 graduates.

In estimating equation (10), our goal is not to identify the causal effect of reputation or admission scores. Our interest is in how their returns change with worker experience—the r and a coefficients—and whether these changes match the predictions from our signaling model.

Table 7 estimates equation (10) both excluding and including Icfes terms, which yields the unconditional return to reputation and the conditional returns to reputation and Icfes. This corresponds to regressions (4) and (2) from Section I and the various subparts of Proposition 2.[37] We discuss results from each of these regressions separately in the subsections below.

Unconditional Return to Reputation.—Column 1 of Table 7 estimates equation (10) including reputation but not Icfes terms, such that the estimates represent the unconditional return to reputation, r^u. The period-zero estimate shows that a one point increase in college reputation is associated with a 10 percent increase in daily earnings in the year of graduation ($r_0^u \approx 0.10$). Proposition 2 predicts that the unconditional return to reputation should not change with experience, implying a

[37] Proposition 2 also contains predictions for regressions that include Icfes but not reputation terms. Online Appendix B.I shows that the results match the predictions: the unconditional return to Icfes increases with experience. This is consistent with findings in Farber and Gibbons (1996) and Altonji and Pierret (2001).

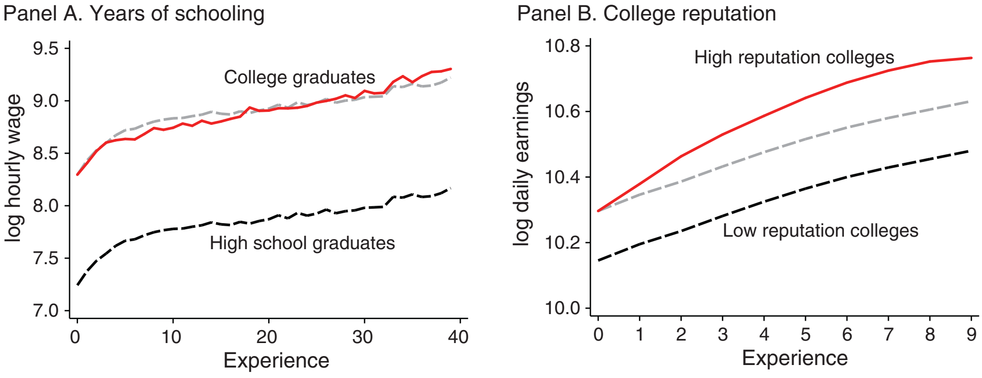

Figure 4. Earnings-Experience Profiles

Notes: Panel A includes high school and college graduates from the 2008–2012 monthly waves of the Colombia Integrated Household Survey (*Gran Encuesta Integrada de Hogares*). Lines depict the mean log hourly real wage (in 2008 pesos) for each schooling group, where we calculate means using survey weights. High school graduates are workers with exactly 11 years of schooling; college graduates have exactly 16 years of schooling. We define experience as min(age − years of schooling − 6, age − 17). The dashed light grey line is parallel to the high school profile starting from the college intercept. Panel B includes 2003–2012 graduates from the 136 colleges represented in Figure 1 with earnings observations in 2008–2012. Lines depict the mean log daily real earnings (in 2008 pesos) for graduates from high and low reputation colleges, which we define by the unweighted median reputation of the 136 colleges. We define experience as age − 16 − 6 and omit levels of experience above nine years because they appear only for workers who took especially long to graduate. The dashed light grey line is parallel to the low reputation profile starting from the high reputation intercept.

zero coefficient on the interaction of reputation and experience. This arises because initial wages fully incorporate information employers observe, including college reputation. Thus, reputation cannot predict innovations in wages; this is identical to wages being a martingale in Farber and Gibbons (1996).

Column 1 strongly rejects this prediction; the return to reputation *increases* with experience. Taken at face value, the coefficient implies that the advantage of having gone to a college with a one point greater reputation increases by about 50 percent within the first four years of employment. This contrasts with the results in Farber and Gibbons (1996) and Altonji and Pierret (2001), who find no evidence of an increasing effect of years of schooling, another educational trait workers might use to signal ability.

The contrast between the reputation and years of schooling results can also be depicted using earnings-experience profiles. Mincer (1974) noted that the wage profiles of workers with different schooling levels are approximately parallel throughout the earnings lifecycle. Panel A of Figure 4 replicates this finding using 2008–2012 household survey data from Colombia.[38] It plots the mean log hourly real wage among workers with two schooling levels—completed high school and completed college—i.e., the gap between the two profiles is the college premium. This gap

[38] In Figure 4, we define potential labor market experience as min(age − years of schooling − 6, age − 17). This definition differs from the one we use elsewhere in the paper (earnings year minus graduation year) because the Colombian household survey does not include school completion dates. However, the age and schooling definition matches those in Mincer's (1974) original analysis and in Altonji and Pierret (2001).

remains roughly constant across 40 years of potential experience, consistent with results in the United States (Lemieux 2006).[39]

Panel B of Figure 4 uses our administrative data to plot earning profiles by college reputation. To match the cross-sectional analysis in panel A, panel B includes 2008–2012 earnings from all 2003–2012 college graduates. We plot mean log daily real earnings separately for graduates from high and low reputation colleges, defined by the median reputation. The earnings gap between the two profiles roughly doubles over the first ten years of experience, as indicated by the divergence of the high reputation profile from the light grey dashed line that is parallel to the low reputation profile.

These results thus suggest that the *slope* of workers' earnings-experience profiles increases with reputation. One potential explanation for this is that reputation may be imperfectly observed. Employers likely observe college identity, but they may not have access to our measure of reputation defined by mean Icfes scores. In this case, employers would further learn about reputation through workers' output, resulting in a return to reputation that rises with experience. To address this possibility, we consider a stronger signaling prediction on regressions that also include individual admission scores.

Conditional Returns to Reputation and Ability.—Columns 2 and 3 of Table 7 add Icfes terms to the regression from column 1. We first add only Icfes scores, and then add an interaction term between Icfes and potential experience. Thus, column 3 estimates equation (10) as written. In these joint specifications, the coefficients reflect the conditional returns to reputation and to ability from equation (2). As Proposition 2 predicts, the period-zero reputation coefficient in column 2 of Table 7 falls relative to its unconditional return in column 1. Consistent with employer learning about ability, column 3 of Table 7 also shows a positive and significant coefficient on the interaction of Icfes and experience.[40]

The main coefficient of interest is on the interaction of reputation with experience. Proposition 2 states that the conditional return to reputation should fall over time. This is similar to the Altonji and Pierret (2001) prediction for observable traits like race or schooling, but our definition of reputation yields a clean test of signaling. Since reputation is a group-level mean of Icfes, Icfes scores are a sufficient statistic for "admission exam" ability, α_i; conditional returns to reputation *mechanically* do not reflect the transmission of information on α_i. The conditional return to reputation should, therefore, decline with experience even if employers do not perfectly observe our measure of reputation; learning about reputation is reflected in

[39] The constant relationship between years of schooling and earnings in Colombia also holds in standard Mincerian regressions reported in online Appendix B.J.

[40] The positive coefficient on the Icfes-experience interaction is similar to the Farber and Gibbons (1996) and Altonji and Pierret (2001) findings using Armed Forces Qualification Test (AFQT) scores as an unobserved characteristic. However, it is in contrast with findings in Arcidiacono, Bayer, and Hizmo (2010), who also study AFQT scores but make a distinction between graduates who enter the labor market after high school and those who do so after college. For college graduates, they show that AFQT is strongly related to wages in the year of graduation, and this relationship changes little over the next ten years. Their conclusion is that AFQT revelation is complete for college graduates, and they suggest that this revelation occurs through college identity. Online Appendix B.I discusses one potential explanation for the difference in findings: sorting by ability in Colombia—although increasing—appears to be less extensive than in the United States.

the Icfes coefficients. Unlike Altonji and Pierret (2001), our model predicts a negative coefficient on Reputation $\times$ t even if there are interactions between ability, α_i, and human capital growth, h_{it}. These effects are also captured by the Icfes $\times$ t term.

In sum, if college reputation serves purely as a signal of ability, Proposition 2 predicts a negative coefficient on the interaction of reputation and experience. Column 3 of Table 7 clearly rejects this. The reputation-experience interaction, although smaller in magnitude than in column 1 of Table 7, is still positive and significant.

The increasing correlation of reputation and earnings is a descriptive result, but it is robust to a wide range of specifications and samples. For example, column 4 of Table 7 adds controls for graduates' gender, age, socioeconomic status, college program, and regional market. All controls are interacted with a quadratic in potential experience to allow earnings trajectories to vary with each characteristic. The coefficient on the reputation-experience interaction decreases slightly, but it is still highly significant and roughly of the same magnitude. Online Appendix B.K shows that this interaction term remains positive with further controls, different definitions of labor market experience, and in alternate samples. Online Appendix B.L also shows that the interaction term is positive when we define reputation at the college-program level rather than at the college level.

C. *Potential Explanations for the Increasing Return to Reputation*

The above results reject a model in which reputation relates to wages only as a signal of ability, α_i, and instead suggests that other attributes related to college membership influence earnings growth. In our model, these attributes are denoted by v_{s_i}, which we define to include both sorting on traits like socioeconomic status, and factors that contribute to skill acquisition at school such as teaching or peer effects. We suppose that employer expectations are given by $E\{v_{s_i} | R_{s_i}\} = v_0 + v_1 R_{s_i}$, where v_1 is the reputation premium. If v_1 is positive, an increasing return to reputation could arise for two reasons. First, if the market does not perfectly observe our measure of reputation, it may become increasingly correlated with wages as employers learn about other college membership attributes. Second, the return to reputation may rise if college membership attributes are related to human capital growth.

Figure 5 provides suggestive evidence that both of these channels may be at work. First, panel A considers one potential component of v_{s_i}: socioeconomic status as measured by whether a student's mother has a college degree.[41] The x-axis contains reputation when observations are colleges, and Icfes when observations are individuals (the scale is the same). The solid line shows that as one moves from the college with the lowest reputation to that with the highest, the mean percentage of students with college-educated mothers increases from below 20 to above 50. The dashed line describes the individual-level relationship between students' Icfes scores and their mother's education, i.e., this is the relationship that would exist if sorting into colleges were by Icfes only. Socioeconomic sorting is less pronounced

[41] Similar patterns emerge for traits related to family income, parents' occupation, and geography.

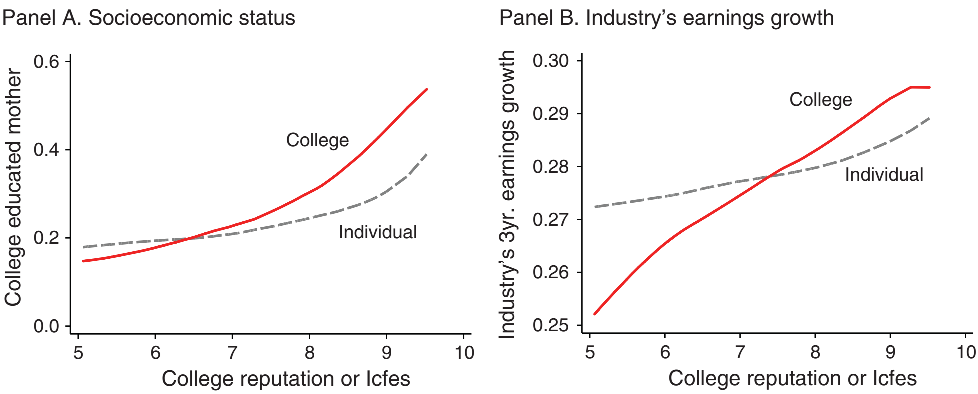

Figure 5. College Membership Attributes and Their Time-Varying Effects

Notes: The sample for panel A is identical to Figure 1. The dependent variable is a dummy equal to one if a student's mother has a college/postgraduate degree. The sample for panel B includes any student in panel A with a four-digit economic activity code from the Ministry of Social Protection. For each four-digit industry, we calculate the mean 2008 log daily earnings for 2005 college graduates and for 2008 college graduates. The dependent variable is the difference between the 2005 and 2008 cohort averages for the industry of each graduate's first job. Dashed lines are local linear regressions of the dependent variable on Icfes percentile. Solid lines are local linear regressions of school means of the dependent variable on college reputation with weights equal to the number of graduates.

in this hypothetical scenario than in the actual one; i.e., there is more sorting across colleges on mother's schooling than is predicted by Icfes scores alone.[42] This is consistent with a positive reputation premium $(v_1 > 0)$; sorting on mother's education is positively correlated with reputation. This could lead to a rising return to reputation if employers imperfectly observe both reputation and mother's education.

Second, panel B of Figure 5 shows that the reputation premium, v_1, may be correlated with human capital investment. The y-axis depicts the average three-year earnings growth in the industry of each graduate's first job. We define industries using four-digit codes, and we calculate earnings growth rates within industry as the mean difference in 2008 log earnings between 2005 and 2008 graduates. The dashed line shows the population-level relationship between industry earnings growth and Icfes scores. Graduates with fiftieth percentile Icfes scores have first jobs in industries where earnings increase by 27 percent within four years, and this growth rate rises by 1.5 percentage points across the Icfes distribution. The solid line shows that the relationship between earnings growth and college reputation is more pronounced. On average, graduates from colleges with reputations at the fiftieth percentile enter industries in which earnings increase by only 25 percent within four years. Mean earnings growth is 4.5 percentage points higher in the industries that employ graduates from top colleges. In short, graduates from higher ranked colleges obtain jobs in industries with greater earnings growth, and this relationship holds even for students with similar ability.

[42] The fact that Colombian financial aid markets are less developed suggests that straightforward ability to pay—beyond the lack of information or ability to take advantage of financial aid opportunities highlighted by Hoxby and Avery (2013) and Hoxby and Turner (2013)—may account for some of the substantial role that socioeconomic status plays in college choice.

Table 8—Industries of First Employment Ranked by Reputation Residuals

Industry	Mean reputation residual	3 year earnings growth
Panel A. Top ten industries		
Securities trading	0.52	0.67
Higher education	0.41	0.27
Oil and gas extraction	0.31	0.50
Specialized animal breeding	0.30	0.15
Chemical manufacturing	0.28	0.52
Petroleum and natural gas extraction	0.28	0.47
Pharmaceutical products wholesale	0.27	0.42
Rubber products manufacturing	0.27	0.20
Cleaning products manufacturing	0.26	0.45
Radio and television	0.26	0.54
Top ten average	0.33	0.36
Panel B. Bottom ten industries		
Public transportation	−0.45	0.20
Financial cooperatives	−0.41	0.17
Social security services	−0.37	0.21
Preschool education	−0.34	0.22
Intermediate products wholesale	−0.33	0.19
Basic primary education	−0.32	0.19
Gambling services	−0.31	0.23
Public services and administration	−0.30	0.08
Elementary education facilities	−0.29	0.17
Soda and mineral water production	−0.27	0.32
Bottom ten average	−0.34	0.19

Notes: This table includes industries with at least 100 graduates from the sample for Figure 5. We define industries and calculate earnings growth as in that figure. Reputation residuals are from a regression of college reputation on Icfes. We display only the top and bottom ten industries ranked by the mean of these residuals. Averages for the top ten and bottom ten are weighted by the number of graduates in each industry.

Table 8 further illustrates this point by displaying examples of these industries. For this table, we regress college reputation on individual Icfes scores and calculate the residuals. We display the top ten and bottom ten industries according to the average value of these residuals. This indicates whether graduates are sorting into industries beyond what their Icfes scores predict. For example, the top ranked industry by this metric—securities trading—has a reputation residual of 0.52. This indicates that graduates whose first job is in securities trading come from colleges with 5.2 percentile points higher reputation than is predicted by their Icfes scores alone. Further, workers in securities trading experience rapid earnings growth, with earnings increasing by 67 percent within the first four years.

Many of the other industries that disproportionately employ graduates from top colleges are related to engineering, and they also tend to have high early career earnings growth. By contrast, the mean earnings growth in the bottom ten industries by reputation residual is 17 percentage points lower than that in the top ten. Many of these low-ranked industries are in the public sector, offering careers in government administration or elementary education.

These results suggest that the increasing return to reputation may reflect a career effect (Topel and Ward 1992) in which better college reputation allows some individuals to be matched to jobs with steeper wage profiles, or to firms that facilitate more on-the-job training. Higher reputation schools might also provide better networks (e.g., Kaufmann, Messner, and Solis 2013; Zimmerman 2013) that ultimately make individuals more productive.[43]

Our setting and data do not reveal whether the correlation between college reputation and earnings growth is due to unobserved dimensions of sorting or due to a causal effect of college identity. In particular, we cannot rule out the hypothesis that college reputation is merely serving as a *signal* for unobservable characteristics that themselves are related to human capital accumulation. Further, even if sorting into colleges occurs only on the ability dimension (α_i in our model), the increasing conditional return to reputation could arise because admission scores are imperfect measures of ability. Nonetheless, the widening of earnings profiles across Colombian colleges is starkly different from the parallel nature of earnings profiles across schooling levels. This may lead students to suspect that their choice of college quality matters for their earnings trajectories in a way that their choice of educational attainment might not.

IV. Conclusion

Debates like those surrounding affirmative action suggest that college plays a key role in determining the distribution of opportunity. As a consequence, a large literature studies the implications of college attendance. Some papers (e.g., Card 1995) ask if college has a causal return, while others (e.g., Goldin and Katz 2008) consider the evolution and determinants of the college wage premium.

Such work does not address the dilemma faced by the millions of students who—having decided to go to college—must choose one. The size of the test preparation industry, for example, suggests that students and parents believe that college choice is important, and that life opportunities are better if one goes to a better college. We call the process by which students are matched to colleges and subsequently to jobs, "the big sort."

This paper has explored the role that college reputation plays in the big sort. Specifically, we have shown that if colleges are selective and more able students choose more "reputable" colleges, then one can produce a single dimensional measure of college *reputation*. We chose a particular measure—the average admission test score of a college's graduates—because it allows a clean examination of signaling mechanisms.

We showed that, consistent with work on other markets, employers use college reputation to make inferences about individual graduates. Specifically, while the cross-sectional data are consistent with this, we exploited a natural experiment in Colombia to show that providing more information about student skill reduces the importance of reputation. Thus, college identity performs a signaling function, and students may be right to worry about *which* college in addition to *whether* college. In other words, we find support for MacLeod and Urquiola's (2015) assumption

[43] Other candidate explanations for the increasing return to reputation arise from violations of the assumptions of the competitive model itself. For example, labor contracts may be such that there is compression in starting wages. In US law firms, for instance, it is not uncommon to observe entering associates being paid the same regardless of their law school of origin. Compensation may later diverge in a way correlated with an LSAT-based reputation measure (Heisz and Oreopoulos 2006).

that labor markets do not immediately observe all individual characteristics (such as Icfes or AFQT scores), and college membership may transmit some of them.

However, we also find that signaling is not the whole story. Even after controlling for admission scores, a graduate's starting earnings and earnings *growth* are positively correlated with her college's reputation. These results are consistent with the hypothesis that colleges add to skill, and that their value added varies systematically with their reputation. Although we cannot establish that this is a causal link, these correlations matter because they are observable—students may notice that individuals from better schools seem to get careers with higher earnings trajectories, which may lead them to prefer more reputable schools.

The purpose of the big sort is to match individuals to jobs. A literature documents significant differences in compensation across firms and occupations that cannot be explained by worker ability (Krueger and Summers 1988; Gibbons and Katz 1991; and Abowd, Kramarz, and Margolis 1999). Our evidence is consistent with the hypothesis that college choice is partially driven by what students *believe* is the consequence of choice. We find evidence that in Colombia, as in the United States, students prefer colleges that are more selective; this in turn leads employers to offer higher wages to the graduates of such colleges. These results illustrate that increasing access to college will not necessarily reduce wage and income inequality. The big sort is a complex system that moves individuals from high school to their first job. Our finding that the exit exams reduced the reputation premium suggests other countries could use analogous policies to address wage inequality.

REFERENCES

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2013. "Matching Markets: Theory and Practice." In *Advances in Economics and Econometrics*, Vol. 1, edited by Daron Acemoglu, Manuel Arellano, and Eddie Dekel, 3–47. Cambridge: Cambridge University Press.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333.

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *Quarterly Journal of Economics* 116 (1): 313–50.

Angrist, Joshua, Eric Bettinger, and Michael Kremer. 2006. "Long-Term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review* 96 (3): 847–62.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2 (4): 76–104.

Autor, David H. 2014. "Skills, education, and the rise of earnings inequality among the 'other 99 percent'." *Science* 344 (6186): 843–51.

Bishop, John H. 2004. "Drinking from the Fountain of Knowledge: Student Incentives to Study and Learn—Externalities, Information Problems, and Peer Pressure." Center for Advanced Human Resource Studies (CAHRS) Working Paper 04-15.

Black, Dan A., and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24 (3): 701–28.

Burdett, Kenneth. 1978. "A Theory of Employee Job Search and Quit Rates." *American Economic Review* 68 (1): 212–20.

Cabral, Luís, and Ali Hortaçsu. 2010. "The Dynamics of Seller Reputation: Evidence From eBay." *Journal of Industrial Economics* 58 (1): 54–78.

Card, David. 1995. "Using geographic variation in college proximity to estimate the return to schooling." In *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*, edited by

Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, 201–22. Toronto: University of Toronto Press.

Card, David, and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100 (1): 1–40.

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 (4): 1491–1527.

Dale, Stacy B., and Alan B. Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *Journal of Human Resources* 49 (2): 323–58.

Dranove, David, and Ginger Zhe Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature* 48 (4): 935–63.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. "Referral-based Job Search Networks." *Review of Economic Studies* 83 (2): 514–46.

Epple, Dennis, and Richard E. Romano. 1998. "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects." *American Economic Review* 88 (1): 33–62.

Epple, Dennis, Richard Romano, and Holger Sieg. 2006. "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education." *Econometrica* 74 (4): 885–928.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (4): 1007–47.

Gibbons, Robert, and Lawrence F. Katz. 1991. "Layoffs and Lemons." *Journal of Labor Economics* 9 (4): 351–80.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race Between Education and Technology.* Cambridge, MA: Belknap Press.

Harris, Milton, and Bengt Holmstrom. 1982. "A Theory of Wage Dynamics." *Review of Economic Studies* 49 (3): 315–33.

Hastings, Justine, Christopher A. Neilson, and Seth D. Zimmerman. 2013. "Are Some Degrees Worth More Than Others? Evidence from College Admission Cutoffs in Chile." National Bureau of Economic Research (NBER) Working Paper 19241.

He, Yingha. 2014. "Gaming the Boston School Choice Mechanism in Beijing." Toulouse School of Economics Working Paper 12-345.

Heisz, Andrew, and Phil Oreopoulos. 2006. "The Importance of Signaling in Job Placement and Promotion." Analytical Studies Branch Research Paper 236.

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *Review of Economics and Statistics* 91 (4): 717–24.

Hoxby, Caroline M. 1997. "How the Changing Market Structure of U.S. Higher Education Explains College Tuition." National Bureau of Economic Research (NBER) Working Paper 6323.

Hoxby, Caroline M. 2009. "The Changing Selectivity of American Colleges." *Journal of Economic Perspectives* 23 (4): 95–118.

Hoxby, Caroline, and Christopher Avery. 2013. "The missing 'one-offs': The hidden supply of high-achieving, low income students." *Brookings Papers on Economic Activity* 43 (1): 1–65.

Hoxby, Caroline, and Sarah Turner. 2013. "Expanding College Opportunities for High-Achieving, Low Income Students." Stanford Institute for Economic Policy Research (SPIEPR) Discussion Paper 12-014.

Hubbard, Thomas N. 2002. "How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market." *Journal of Law and Economics* 45 (2): 437–68.

Jin, Ginger Zhe, and Phillip Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *Quarterly Journal of Economics* 118 (2): 409–51.

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy* 87 (5, Pt. 1): 972–90.

Kahn, Lisa B., and Fabian Lange. 2014. "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures." *Review of Economic Studies* 81 (4): 1575–1613.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107 (1): 35–78.

Kaufmann, Katja Maria, Matthias Messner, and Alex Solis. 2013. "Returns to Elite Higher Education in the Marriage Market: Evidence from Chile." https://www.aeaweb.org/conference/2014/retrieve.php?pdfid=826.

Krueger, Alan B., and Lawrence H. Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica* 56 (2): 259–93.

Lange, Fabian. 2007. "The Speed of Employer Learning." *Journal of Labor Economics* 25 (1): 1–35.

Lemieux, Thomas. 2006. "'The Mincer Equation' Thirty Years after *Schooling, Experience, and Earnings*." In *Jacob Mincer: A Pioneer of Modern Labor Economics*, edited by Shoshana Grossbard, 127–45. New York: Springer.

MacLeod, W. Bentley, Evan Riehl, Juan E. Saavedra, and Miguel Urquiola. 2017. "The Big Sort: College Reputation and Labor Market Outcomes: Dataset." *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20160126.

MacLeod, W. Bentley, and Miguel Urquiola. 2015. "Reputation and School Competition." *American Economic Review* 105 (11): 3471–88.

Melnik, Mikhail I., and James Alm. 2002. "Does a Seller's Ecommerce Reputation Matter? Evidence from eBay Auctions." *Journal of Industrial Economics* 50 (3): 337–49.

Milkovich, George, Jerry Newman, and Barry Gerhart. 2010. *Compensation*, 10th ed. Chicago: McGraw-Hill Education.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York: Columbia University Press.

Nelson, Phillip. 1970. "Information and Consumer Behavior." *Journal of Political Economy* 78 (2): 311–29.

Rodríguez, Jorge, Sergio Urzúa, and Loreto Reyes. 2015. "Heterogenous Economic Returns to Postsecondary Degrees: Evidence from Chile." Unpublished.

Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review* 71 (5): 845–58.

Roth, Alvin E., and Marilda Sotomayor. 1989. "The College Admissions Problem Revisited." *Econometrica* 57 (3): 559–70.

Rothschild, Michael, and Lawrence J. White. 1995. "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs." *Journal of Political Economy* 103 (3): 573–86.

Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. 2014. "Mismatch Unemployment." *American Economic Review* 104 (11): 3529–64.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87 (3): 355–74.

Topel, Robert H., and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics* 107 (2): 439–79.

Zimmerman, Seth D. 2013. "Making Top Managers: The Role of Elite Universities and Elite Peers." http://faculty.chicagobooth.edu/workshops/micro/pdf/ZimmermanJMP.pdf.