Employer Learning, Statistical Discrimination and University Prestige*

Paola Bordón[†] and Breno Braga[‡]

November, 2015

Abstract

We investigate whether employers use university prestige to statistically discriminate among college graduates. The test is based on the employer learning literature which suggests that if employers use a characteristic for statistical discrimination, this variable should become less important for earnings as a worker gains labor market experience. In this framework, we use a regression discontinuity design which estimates a 19% wage premium for recent graduates of the most selective universities in Chile. However, we find that this premium decreases by 3 percentage points per year of labor market experience. These results suggest that employers use college selectivity as a signal of workers' unobservable productivity when they graduate from college. Nevertheless, as workers reveal their quality throughout their careers, they are rewarded based on their productivity rather than the prestige of their college.

JEL Classification: I21, J31, J71

Keywords: Employer Learning, Statistical Discrimination, College Selectivity, Regression Discontinuity.

---

*We would like to thank John Bound, Charlie Brown, Brad Hershbein, John Kennan, David Lam, Fabian Lange, Brian McCall, Jeff Smith, Christopher Taber, and participants at EconCon, LACEA, SOLE-EALE World Conference, University of Michigan, University of Wisconsin, the Urban Institute, and SAIS Johns Hopkins University seminars for their valuable comments. The authors can be contacted by email at: pbordon@fen.uchile.com and bbraga@urban.org. Paola Bordón gratefully acknowledges support from CONICYT/FONDAP/15130009.

[†]Centre for Social Conflict and Cohesion Studies, Department of Economics, Universidad de Chile

[‡]Urban Institute

1 Introduction

Labor markets are characterized by incomplete information on the productivity of workers (Spence, 1973). There are some characteristics of workers, such as labor market ability, that are important for performance on the job but are not easily observable by employers. In this context, employers often have to make judgments about unobservables on the basis of the available information. Within this framework, statistical discrimination is defined as employers using a group identity of workers to infer their unobservable productivity. The most traditional group identity studied in the statistical discrimination context is race (Phelps, 1972 and Aigner and Cain, 1977). In this literature, the racial wage gap is justified not because employers are prejudiced against a particular race but because they use race identity to predict the unobservable quality of workers. More recently, evidence was found that firms use years of education to statistically discriminate workers (Altonji and Pierret, 2001 and Lange, 2007).

In this paper we study a new dimension of statistical discrimination: we investigate whether employers use the prestige of the university attended by a worker to predict his or her unobservable productivity. We believe that college prestige satisfies the criteria of a group identity that might be used for statistical discrimination for two main reasons. First, this information is easily observable by employers: college graduates present university names in their resumes and prestigious universities are widely recognized in the labor market. Second, there is evidence that more talented individuals attend more prestigious universities (Hoxby, 1998 and Dale and Krueger, 2002). Overall, elite universities have a very competitive application process and tend to select higher quality candidates.[1] Within this framework, it is natural to believe that employers use university prestige to infer the unobservable labor market quality of workers.

In order to test whether employers use university prestige as a signal of unobservable productivity of workers, we rely on the employer learning and statistical discrimination (EL-SD) literature (Altonji and Pierret, 2001).[2] The underlying assumption is that at the early stages, employers as-

[1] As will become clear later, the underlying assumption is that universities are better at screening candidates than employers.

[2] Other important papers in this literature include Lange (2007), Schönberg (2007), Arcidiacono et al. (2010), Mansour (2012) and Bordon (2014).

sess workers on the basis of easily observable variables that are correlated with their unobservable productivity. As a worker gains experience in the labor market, employers weigh these characteristics with other information that becomes available, such as references and on-the-job performance. If employers use a characteristic to statistically discriminate a worker in the early stage of his career, this characteristic should become less important for earnings as a worker reveals his productivity over time.

We use data from the *Futuro Laboral Project* of the Chilean Ministry of Education to test whether employers use college selectivity as a signal of worker's unobservable quality. This data satisfies the purpose of the paper for several reasons. First, it follows different cohorts of Chilean college graduates during their first years in the labor market, the period in which most employer learning happens (Lange, 2007). Second, the data presents information on labor market outcomes such as earnings, and we can identify workers who graduate from elite universities. Finally, the data contains information on the Chilean centralized test for admission to universities. This information will be used to provide both a measure of a worker's ability not easily observed by employers and as a running variable in the regression discontinuity test we propose.[3]

We perform the EL-SD test in two different ways. We first follow the EL-SD test proposed by Altonji and Pierret (2001) and estimate an earnings equation where both returns to graduating from a prestigious university and hard-to-observe ability measures can change with experience. If employers statistically discriminate among workers on the basis of college prestige, then as employers learn about a worker's productivity over time, the coefficients on university prestige should fall, and the coefficients on hard-to-observe ability measures should rise with experience. We use the math and reading scores in the university admission test as the measure of a worker's ability not easily observed by employers. We present further evidence that math and reading scores are good measures of hard-to-observe correlates of productivity.

[3]Kaufmann et al. (2012) , Hastings et al. (2013) and Zimmerman (2013) are recent papers that have also explored the discontinuities generated by the centralized admission process for universities in Chile. Kaufmann et al. (2012) looks at effect of graduating from an elite university on marriage outcomes. Hastings et al. (2013) studies labor market returns to college admission. Zimmerman (2013) shows that students admitted to prestigious universities are more likely to reach managerial positions, but this achievement is concentrated among students who graduated from private high schools. None of these papers explore how the selective university wage premium changes throughout a worker's career, which is the main object of interest of this paper.

In addition to the traditional EL-SD test, we take advantage of Chile's centralized university admission process to propose a statistical discrimination test based on a regression discontinuity design (RD). Using information from the admission test scores, we are able to identify workers who were just above or just below the admission thresholds to the two most prestigious universities in Chile. We propose a RD statistical discrimination test that compares the earnings' dynamics between these two groups of workers as they gain experience in the labor market. The test assumption is that workers just above and just below the admission thresholds are similar in terms of their pre-college unobservable characteristics.

The RD statistical discrimination test predicts that if employers use university prestige to statistically discriminate workers: i) individuals barely admitted to the most selective universities in Chile should be paid substantially more than those barely rejected when they graduate from college; ii) the wage differential between these two groups of workers should shrink as individuals progress in their careers. In contrast to the traditional EL-SD test, the estimation of the regression discontinuity test can be interpreted as the causal effect of graduating from a prestigious university on earnings for those near the admission cutoff. We show that, different from the traditional EL-SD test, our RD approach is not biased by other individual's characteristics that might be used by employers for statistical discrimination, such as family socioeconomic status.[4]

We find evidence that employers statistically discriminate using both the traditional EL-SD and the regression discontinuity test. Following the traditional EL-SD test we estimate that college graduates from the two most selective universities in Chile earn on average 26% higher wages after graduation. However, we find this wage premium decreases by 1.7 percentage points per year of experience. We also estimate that the Math component of the admission test, one of our measures of ability not observed by employers, increases in importance for wages as workers accumulate experience. These results are in accordance with the predictions of EL-SD test proposed by Altonji and Pierret (2001) if employers use prestige of universities to statistically discriminate workers.

The regression discontinuity design approach provides further evidence for statistical discrimina-

[4]As it will become clear later, the traditional statistical discrimination test presented in Altonji and Pierret (2001) is biased on the basis of the treatment variable. However, as a test of employer learning itself it is unbiased.

tion. We estimate a 19% wage premium for recent graduates of the two most prestigious universities in Chile, but this wage premium decreases by 3 percentage points per year of experience. The difference between the traditional EL-SD test and the RD test can be interpreted as evidence that there are individual characteristics used for statistical discrimination by employers that are correlated with university selectivity, but not available in the data.

Finally, under some stronger assumptions, we are able to distinguish the contribution of the human capital effect from the signaling effect on the prestigious university wage premium. Following the estimation procedure proposed by Lange (2007), we use the RD coefficients estimated at each experience level to recover the parameters of the employer learning model. Based on this procedure, we find that the human capital effect represents at most 14% of the college prestigious premium for recent university graduates, with the remaining 86% being driven by signaling effects.

Based on these findings, this paper contributes to different dimensions of the existing literature. First, this paper is a contribution to the EL-SD literature because we study statistical discrimination on the basis of a different group identity. While there is extensive literature analyzing the use of race, gender, and schooling, ours is one of the first papers to study whether employers use university prestige to statistically discriminate workers. To the best of our knowledge, Lang and Siniver (2011), Hershbein (2013), and Araki et al. (2015) are the only other papers that have addressed this issue. While Lang and Siniver have a similar approach to estimating how returns to attending an elite university in Israel change with labor market experience, the authors are unable to properly exploit the regression discontinuity in the college admission process.

Furthermore, as far as we know, our paper is the first to propose an employer learning-statistical discrimination test based on a regression discontinuity design. We show that the traditional test proposed by Altonji and Pierret (2001) is biased for the treatment variable if employers also statistically discriminate workers on the basis of characteristics that are not present in the data and are correlated with graduating from a prestigious university.

Second, we contribute to the literature that studies the effect of graduating from an elite university on labor market outcomes. There is an extensive series of papers that estimate the returns to graduating from a selective university on earnings. Brewer et al. (1999) uses US survey data and Heckman selection corrections to find significant returns to attending an elite private institution. Dale and Krueger (2002) compare the earnings of individuals attending more selective colleges to those of individuals who were accepted at similarly selective colleges but who chose to attend less selective institutions. Dale and Kruger find negligible effects from attending a selective college, especially for students from middle and high-income families.[5] Black and Smith (2006) discuss several approaches for identifying the university selectivity wage premium. Using the NLSY79 and their preferred GMM method, the authors find a selectivity premium that is smaller than the earlier studies, but still statistically significant.

There are also more recent papers that have used a regression discontinuity design to estimate the effect of attending a prestigious university on earnings. Hoekstra (2009) uses a test score admissions cutoff to estimate the returns to attending a flagship state university in the U.S. The main finding of the paper is that male students above the admission cutoff earn 20% more five to ten years after graduation relative to those below the cutoff. Saavedra (2008) estimates a 20% wage premium one year after students graduate from a very selective university in Colombia. Saavedra also finds that students from a selective university perform better at a college exit test. None of these papers have explored how the university selectivity wage premium changes throughout worker's career and if that translates into differences in earnings in the long run.

Most importantly, while there is a significant effort in all of the papers cited above to overcome the selection bias associated with attending a prestigious university, little attention has been paid to the mechanisms that generate the college selectivity wage premium. In contrast to past work, in this paper we shed some light on the reasons why workers from prestigious universities receive higher wages after graduation. On one hand, attending a selective university could be associated with receiving better instruction and having more accomplished peers. In this context, prestigious universities have an advantage of increasing a worker's productivity in comparison to less prestigious universities. On the other hand, the main effect of attending a selective university might be to signal

[5] It is interesting to note that the authors estimates the zero wage premium approximately 15-19 years after a worker's graduation from college. The zero effects for individuals with similar pre-college characteristics later in their careers does not contradict the empirical findings of this paper.

to employers the unobservable inherent ability of a worker. In this context, the value added from a selective college education might not be significantly higher than that from a less prestigious university.

Consistent with the employer learning literature, we interpret our finding of rapid decreases in the elite college premium for workers with similar pre-university characteristics as evidence that strong signaling mechanisms drive the prestigious university wage premium.[6] In addition, under some stronger assumptions, we are able to fully distinguish the human capital effects from signaling effects on earnings when a worker graduates from a prestigious university. To the best of our knowledge, this is the first paper that can provide such insight.

2 Institutional Framework

Higher education in Chile comprises three types of institutions: universities, professional institutes, and technical formation centers. Universities provide the highest degree of learning, combining teaching, research and outreach activities; they teach accredited degree programs (2.5 to 4 years) and award academic degrees (5 to 7 years). Professional institutes grant professional degrees other than those awarded by universities. Technical formation centers are intended to equip higher level technicians with the competencies and skills needed for responding to industry needs in the public and private sectors.

Universities in Chile can be divided into two main categories: traditional and non-traditional institutions. Traditional institutions comprise the oldest and most prestigious universities created before 1981, and those institutions created after 1980 that were derived from the older universities. Traditional establishments consist of 25 fully autonomous universities coordinated by the Council of Chancellors of Chilean Universities and are eligible to obtain partial funding from the state. They employed a single admission process: the Chile's university selection test, the *Prueba de Aptitud Académica* (PAA) until 2003[7]. This test is made up of three compulsory sub-tests including

[6] Although, in the conclusion of this paper we discuss other possible theories that could explain this pattern.
[7] In 2004, Chile's university selection test was modified and it is now called *Prueba de Selección Universitaria* (PSU).

language, mathematics, and the history and geography of Chile. Additionally, depending on which programs a student is applying to, he or she may be required to take the following major specific PAA tests: advanced mathematics, physics, chemistry, biology, and history.

The time-line of the admission process for traditional universities is described in figure 1. First, students take the PAA test. After receiving their scores, they make their application decisions. Students apply to a major and university (or program) simultaneously. They may only apply to 8 programs, ranking them by preference. The final admission scores consist of a weighted average of the compulsory and major specific tests and applicant's high school GPA, with each program setting its specific PAA weights.[8] The number of vacancies for each program is announced before the application process and programs fill their vacancies solely based on the final weighted scores. The admission score cutoff is defined by the score of the last student admitted into a program. Since the cutoffs are not known before the application decisions are made, students cannot manipulate the side of the cutoff on which their scores fall.[9] Non-traditional universities, those created after 1981, may not necessarily use PAA scores to select their incoming students. Nevertheless, the anecdotal evidence is that the majority of students who wish to attend higher education in Chile take the PAA at the end of high school, independent of which university they are planning to attend. The test is relatively inexpensive and administrated throughout the country.

All higher education institutions charge tuition and fees. As of 2001, the Chilean higher education system consisted of 60 universities (25 traditional universities and 35 new private universities operating without direct public subsidy), 42 professional institutes, and 117 private technical formation centers.

[8]For example, the engineering programs in a prestigious university uses a weight of 20% mathematics, 10% language, 10% history, 20% high school GPA, 30% advanced mathematics, and 10% physics to calculate applicant's weighted scores to this program. Different universities might use different weights to calculate the weighted scores for the same majors.

[9]Students might use the admission score cutoff of previous years as a reference. Given the variation of the admission cutoff overtime and the possibility to applying to 8 different programs, we believe that students with scores near the margin for admission to prestigious universities tend to apply to these competitive programs.

3 Data

The data used in the study comes from *Futuro Laboral*, a project of the Chilean Ministry of Education that follows individuals over the first years after graduation from higher education programs. The panel dataset matches tax returns with transcripts of students' majors and the institutions they graduated from. The unit of analysis concerns only those who graduated from traditional or non-traditional universities; those who haven't obtained a higher education degree or graduated from professional institutes or private technical formation centers are not in the sample. Income information is available between the years 1996 and 2005. We have data for the 1995, 1998, 2000 and 2001 graduating classes. Given that the income data is only available until 2005, each cohort is observed for a different length of time. For example, while we observe 10 years of labor market experience for the 1995 graduation class, and only 4 years of labor market experience for the class of 2001.[10] [11]

The information provided by Chile's Internal Revenue Service comprises age, sex, name of the institution that individuals graduated from, major, year of graduation, annual income reported on tax returns, city or cities of employment, number of employers and economic sector. The raw data contains individuals in Chile that had positive earnings between 1996 and 2005, even those who were exempt from tax.[12] [13] From a random sample, the Ministry of Education gathers information on the PAA scores, high school grades and the name of the high schools where students graduated from. As the PAA scores have an important role in both the traditional EL-SD test and regression discontinuity analysis, we restrict our study to this sample.

The wage measured in the sample is the annual income received from jobs and services provided by the individual and it does not include self-employment income.[14] We use the consumer price

[10]Unfortunately, the project was discontinued and the income data for more recent years was not collected.

[11]We present robustness checks using only the 1995 graduation class in the appendix of the paper.

[12]Note that in Chile, married couples always file taxes separately.

[13]A concern is that a portion of the individuals from prestigious universities might go to graduate school after finishing their baccalaureate studies and would therefore be omitted in the earnings sample. However, the fraction of workers who go to graduate school in Chile is very low. Using data from the National Socioeconomic Characterization Survey in the year 2000, we find that only 0.65% of 25-34 year old with a bachelor degree were enrolled in graduate school or had obtained a graduate degree.

[14]We do not have information on weeks or hours worked in the sample and for this reason we cannot explore how much of the annual income increase is due to changes in hours or work weeks. Nevertheless, workers with a bachelor

index as a deflator to compute real wages. The experience variable is computed as the number of years in which an individual has earned income and has paid taxes after graduation. The final sample consists of 58,179 individuals and 313,077 observations.

We divide universities into two groups: selective and non-selective. The selective universities comprise two of the oldest and most prestigious universities in the country and non-selective compromise all other universities. These two schools attract students with the highest PAA scores and therefore are the most selective schools in the country. The programs of these two universities have also been consistently ranked among the highest in Chile and their prestige is well recognized nationwide.[15] Throughout the paper we will use the terms selective, prestigious and elite universities interchangeably to refer to these two universities.[16] Due to a confidentiality agreement with the Ministry of Education, we cannot name these institutions. Data on program admission cutoffs by major and year were collected at the universities' websites (for later application years) and newspapers (for earlier application years). [17]

Table 1 shows descriptive statistics regarding these two groups. As expected, selective universities have higher average scores in the math and language components of the PAA tests, and their students have higher high school grades. We observe that 11% of students at selective universities have graduated from a private high school, compared to 7% from non-selective universities. We plot the distribution of language and math PAA scores for college graduates from selective and non-selective universities on Figures 2 and 3 respectively. One can see from these figures that the language and math scores of graduates from selective universities are concentrated at the higher end of the distribution. Finally, Table 2 we show that workers from the two selective universities have higher average earnings than those from the less prestigious schools.

degree in Chile present both a high employment attachment and the majority work full time. Using the National Socioeconomic Characterization Survey in the year 2000, we find that 86.7% of 25-34 years old with a bachelor's degree work are employed, and of those, 88% work more than 35 hours per week.

[15] In the appendix of this paper we provide further evidence that these two universities are the most selective of the country.

[16] Note that according to this definition, non-selective universities comprise all other universities in the country and most of them do not offer a slot to every applicant.

[17] We find that 4% of individuals in our restricted sample with a prestigious university degree have weighted scores lower than the admission cutoffs. This could be justified by measurement errors in the admission cutoffs and the weights used in the paper, transfers from less prestigious universities or athletes' admissions. We drop these individuals from the sample used in the RD analysis.

4 Employer Learning Statistical Discrimination Model

The standard employer learning model specifies the log-productivity of a college graduate worker i with experience level t:

$$y_{it} = rs_i + \alpha_1 q_i + \lambda z_i + \eta_i + H(t) \tag{1}$$

where s_i captures information that is available to both employers and researchers. In this paper, s_i is defined as an indicator of whether a worker graduated from a prestigious university or not. The variable q_i describes information available to employers but not present in the data, such as family socioeconomic background, z_i is a characteristic present in the data but not available to employers, and η_i is a measure of a worker's inherent ability that is not available in the data or to employers. Finally, $H(t)$ describes the relation between log-productivity and experience, which is assumed to be independent of the other variables in the model.

In the absence of information on z_i and η_i, employers form expectations based on other observed characteristics of workers. Altonji and Pierret (2001) assume that these conditional expectations are linear on s and q:[18]

$$z = \mathbb{E}[z|s,q] + v = \gamma_1 q + \gamma_2 s + v$$

$$\eta = \mathbb{E}[\eta|s,q] + e = \alpha_2 s + e$$

where v and e have mean zero and not correlated with s and q by construction. We suppress the i subscript for most of the rest of the analysis. Under this assumption, one can characterize the expected value of y given information on s and q:

$$\mathbb{E}[y|s,q] = (r + \lambda\gamma_2 + \alpha_2)s + (\alpha_1 + \lambda\gamma_1)q + H(t)$$

In the traditional EL- SD model, employers do not observe y but have access to a noisy measure of a worker's productivity after each period that an individual spends in the labor market:

[18] A normalization allows a suppression of q in the second expectation.

$$\tilde{y}_\tau = y_t + \varepsilon_\tau$$

where the noise ε_t is independent of all the variables in the model. As in Altonji and Pierret (2001), employers share equal information about workers, labor markets are competitive and there is a spot market for labor services. As a consequence, wages are equal to the expected productivity of a worker, given the information available to employers at each period.

$$W_t = \mathbb{E}[\mathrm{exp(y_t)|s, q, \tilde{y}_0, ..., \tilde{y}_{t-1}}]$$

Lange (2007) assumes that ε_t is independently, identically and normally distributed with a finite variance. Under this assumption, the process of updating the expectations of employers has a very simple structure and the log-wage process can be represented by:

$$w_t = (1 - \theta_t)\mathbb{E}[y|s, q] + \theta_t \frac{1}{t} \sum_{\tau=0}^{t-1} \tilde{y}_\tau + \tilde{H}(t) \tag{2}$$

where $\tilde{H}(t)$ is a linear transformation of $H(t)$ and θ_t is a function of the variances of and $\varepsilon_{i\tau}$, s and q. Furthermore $\theta_0 = 0$ and θ_t strictly increases with t converging to 1 as t goes to infinity.[19] This expression states that as a worker progresses in his or her career, employers weight less of their initial beliefs on the worker's productivity based on s and q, and weight more on the new information that becomes available during the worker's career.

4.1 Traditional EL-SD Test

The object of interest in the traditional employer learning model is the linear projection of the log-wage w_{it} on s, z and t.

$$\mathbb{E}^*[w_t|s, z, x] = b_{sx}s + b_{zx}z + \widetilde{H}(t)$$

Without loss of generality, one can define the projections of unobservable variables (q, η) on observable variables (s, z):

[19] See Lange (2007) for the formal derivations of these parameters.

$$q = \gamma_3 s + \gamma_4 z + u_1$$

$$\eta = \gamma_5 s + \gamma_6 z + u_2$$

Using the independence of ε_τ with respect to all the variables of the model, Lange (2007) shows that the coefficients of the projections can be expressed as:

$$b_{st} = (1 - \theta_t) b_{s0} + \theta_t b_{s\infty} \tag{3}$$

$$b_{zt} = (1 - \theta_t) b_{z0} + \theta_t b_{z\infty} \tag{4}$$

where, as discussed before, $\theta_0 = 0$ and $\lim_{t \to \infty} \theta_t = 1$. The traditional EL-SD test consists of estimating how b_{st} and b_{zt} change with experience level t. Indeed, Altonji and Pierret (2001) propose that if employers statistically discriminate workers on the basis of s and if z is positively related to s, one should observe that b_{st} falls with t and b_{zt} rises with t. This is the approach we follow in subsection (5.1), where we verify whether the coefficient on prestigious university rises with experience and if the coefficient on Math and Language components of the PAA test decreases.

Furthermore, under the assumptions above, Lange (2007) shows that:

$$b_{s0} = \underbrace{r}_{\text{A}} + \underbrace{\alpha_1 \gamma_3}_{\text{B}} + \underbrace{\alpha_2 + \lambda(\gamma_2 + \gamma_1 \gamma_3)}_{\text{C}} \tag{5}$$

$$b_{z0} = \underbrace{(\alpha_1 + \lambda \gamma_1) \gamma_4}_{\text{D}} \tag{6}$$

where the coefficient b_{s0} represents the relation between graduating from a prestigious university and wages at the beginning of a worker's career. The term A captures the direct effect of graduating from a prestigious university on productivity. The term B represents the direct impact of q on wages and the fact that q is not present in the data but it is correlated to s. This can be interpreted as the traditional omitted variable problem associated with estimating the returns to graduating from a prestigious university (Dale and Krueger, 2002). It captures the wage impact of any variable correlated to graduating from a prestigious university that is not present in the data, but it is observed by employers (e.g. family socioeconomic status).

Finally, the term C reflects the fact that employers do not observe η and z in the beginning of a worker's career, but are aware of their relation with s. Therefore, employers use s as a signal of unobservable components of a worker's productivity. In the same way, the relation between z and the log wages of a worker at the beginning of his career is given by the coefficient b_{z0}. As employers do not observe z, this coefficient only captures the fact that we are omitting q from the linear prediction, and that z and q are correlated.

$$b_{s\infty} = \underbrace{r}_{\text{E}} + \underbrace{\alpha_1\gamma_3 + \gamma_5}_{\text{F}} \tag{7}$$

$$b_{z\infty} = \underbrace{\lambda}_{\text{G}} + \underbrace{\alpha_1\gamma_4 + \gamma_6}_{H} \tag{8}$$

The coefficients $b_{s\infty}$and $b_{z\infty}$ represent the relation between s and z respectively with wages as $t \to \infty$ and $\theta_t \to 1$. As before, E represents the direct effect of graduating from a prestigious university on wages. The coefficient F captures the fact that η and q have an impact on long-term wages and are related to s but they are omitted in the linear prediction because they are not observed in the data. Note that F is different from the term B because employers only learn η with time. In the same way, the term G captures the direct impact of z on productivity and H captures the correlation of z to the omitted variables η and q.

One important issue omitted from the employer learning literature (Altonji and Pierret, 2001 and Lange, 2007) is how the correlation between s and the unobservable factor q can affect the conclusions of the statistical discrimination test. This issue arises if employers statistically discriminate workers on the basis of variables that are not observed in the data, such as family socioeconomic background, that are correlated to graduating from prestigious university. In this case, the traditional employer learning test might suggest that employers statistically discriminate a worker on the basis of university prestige, when in fact firms might be using family socioeconomic status as a signal of a worker's unobservable characteristics. In other words, while the EL-SD test presented in Altonji and Pierret (2001) is unbiased as a test of employer learning itself, it could be biased as a statistical discrimination test on the basis of the treatment variable s.

For instance, we analyze the extreme case where s is not correlated to η and z ($\alpha_2 = 0$, $\gamma_2 = 0$ and $\gamma_5 = 0$). In this situation, employers do not use s as a signal of a worker's unobservable characteristics, and therefore, workers are not statically discriminated on the basis of university prestige. Furthermore, assuming that q is correlated with η and z ($\gamma_4 \neq 0$), and therefore q is used by employers to statistically discriminate workers. Under this assumption, the traditional employer learning test would suggest that firms statistically discriminate workers on the basis of university prestige because $b_{s\infty} < b_{s0}$ and $b_{z\infty} > b_{z0}$. Note, however, that this conclusion is being driven by the correlation of s and q, and the fact that employers use q to predict z.

4.2 Regression Discontinuity EL-SD Test

The object of interest in the regression discontinuity test we propose is how the difference between average log-wages of individuals just above and just below the admission cutoff to a prestigious university changes with experience. Precisely, we define $Dist.Cutoff_i$ as the distance between a student's test score and the admission threshold of a prestigious university. For simplicity, we assume that all students admitted to a prestigious university enroll and graduate from this university, such that $s_i = 1$ if $Dist.Cutoff_i \geq 0$ and $s_i = 0$ otherwise.[20]

The RD parameter of interest is:

$$\tau_t = \lim_{\text{Dist.Cutoff}\downarrow 0} \mathbb{E}[w_{it}|Dist.Cutoff_i] - \lim_{\text{Dist.Cutoff}\uparrow 0} \mathbb{E}[w_{it}|Dist.Cutoff_i] \tag{9}$$

which represents local average difference of log-wages by experience levels at the admission cutoff. The employer learning statistical discrimination RD test consists of testing whether τ_t decreases

[20] In reality, there is the case that some students admitted to a prestigious university decide to attend a less prestigious university (fuzzy regression discontinuity). For the sake of exposition, we ignore this possibility here. In addition, we asume the same dropout rates for all schools.

with t.

Note that by definition, we have that:

$$\lim_{\text{Dist.Cutoff}\downarrow 0} \mathbb{E}[s|Dist.Cutoff_i] = 1 \text{ and } \lim_{\text{Dist.Cutoff}\uparrow 0} \mathbb{E}[s|Dist.Cutoff_i] = 0$$

Furthermore, the distribution of the other variables in the model $\{z, q, \eta, \varepsilon_\tau\}$ is continuous around the admission cutoffs. In this case, the expected values of these variables just above and just below the admission cutoff are the same:

$$\lim_{\text{Dist.Cutoff}\downarrow 0} \mathbb{E}[X|Dist.Cutoff_i] = \lim_{\text{Dist.Cutoff}\uparrow 0} \mathbb{E}[X|Dist.Cutoff_i]$$

for $X = q, z, \eta$. Using these two conditions, the assumption that employers do not have access to $Dist.Cutoff_i$, and the log-wage process derived in (2), one can easily show that:

$$\tau_t = (1 - \theta_t)(r + \lambda\gamma_2 + \alpha_2) + \theta_t r$$

$$= \underbrace{r}_{I} + (1 - \theta_t)(\underbrace{\alpha_2 + \lambda\gamma_2}_{L}) \tag{10}$$

where θ_t is defined as before. The regression discontinuity effect of graduating from a prestigious university on wages at experience level t is composed of two terms. The first term, I, represents the direct effect of s on the worker's productivity. The second term L represents the fact that employers do not observe η and z, and use s as a signal for these two variables. In other words, if employers statistically discriminate among workers on the basis of university prestige, we have that $L > 0$. However, the signaling term L becomes less important for earnings as firms learn about a workers true productivity, τ_t decreases with t and converges to r as θ_t goes to 1.

There is an important difference between the regression discontinuity test we propose and the traditional employer learning test: the parameter τ_t does not depend on the relation between s and q. Specifically, the regression discontinuity test is robust to the existence of characteristics that could be used for statistical discrimination that are related to graduating from a prestigious university and that are not present in the data. This difference is important because, as discussed above, the traditional EL-SD test might confound statistical discrimination based on family socioeconomic status and statistical discrimination based on college prestige since these factors are intrinsically related and we do not observe family socioeconomic status in the data. In addition, the coefficient τ_t converges to the r as t increases, which represents the causal effect of graduating from a prestigious university on a worker's earnings.

5 Results

5.1 Traditional EL-SD test

We first investigate statistical discrimination on the basis of university selectivity by following the employer learning statistical discrimination (EL-SD) test suggested by Altonji and Pierret (2001). An important innovation of this paper is that we use the math and reading components of the PAA as a measure of ability correlates not easily observed by employers. We have a number of reasons to justify our choice. First, these are the components of the PAA test formulated to measure inherent abilities of applicants. Their purpose is to give opportunities for those who didn't have adequate formal education to demonstrate their ability in the admission processes of traditional universities. Second, there is evidence that employers do not have access to PAA scores at the time of determining wages. According to an interview with Juan Swett, the CEO of "www.trabajando.com" which is the biggest job search web portal in the country, most Chilean employers do not ask for PAA scores in the resumes of prospective workers. A justification for this statement is that, in contrast to universities, employers do not have access to the full distribution of PAA scores. The absolute value of the PAA score for a single worker would not be very informative to a firm.

We present the estimate of the traditional EL-SD model in table 3. All of the standard errors are clustered at the individual level and we use White-Huber corrections for possible heteroscedasticity. Experience is modeled with a cubic polynomial and we control for gender, majors, private high school indicator and year dummies. In order to facilitate the interpretation of the coefficients, we standardize the PAA score by test year. Column 1 of table 3 reports the results of the estimation when the interaction of experience with ability and the selective university dummy are excluded. Similar to past studies (Brewer et al., 1999 and Black and Smith 2006) we estimate that graduating from a very selective university is associated with higher earnings. In exact terms, graduating from a prestigious university increases log-wage by 0.196. Our proxy for innate ability, PAA has a positive and statistically significant effect on earnings. An increase of one standard deviation in the language PAA test increases wages 3%, whereas an increase of one standard deviation in the math test increases wages 7%.

In column 2 we introduce the interaction between the selective university dummy and experience. If university selectivity provides a signal of worker's ability to employers, we should expect the earnings of recent college graduates from prestigious universities to be higher relative to those from less selective institutions. However, we should not observe an increase in the importance of college selectivity on earnings as a worker gains experience if employers learn about a worker's true ability over time. This is in fact the finding of the equation presented in column 2, where we estimate a coefficient of -0.009 for the interaction between the selective university and experience.[21]

The important result of table 3, shown in column 3, where we add PAA scores interacted with experience to capture the idea that employers should increase the reward to unobservable ability with time. The coefficient on selective universities is 0.26, which is large and statistically significant. We estimate a coefficient of -0.017 for the interaction between the selective university and experience. Therefore, the effect of graduating from the most selective and oldest universities on earnings decreases by 1.7% by year, supporting the theory of the EL-SD model. Finally, the coefficient of 0.007 on standard math PAA interacted with experience suggests that the effect of a shift in standard math PAA score changes significantly as workers accumulate experience, which is consistent with the employer learning thesis that wages increasingly reflect productivity, augmenting the correlation between wages and ability. The positive but insignificant effect of the interaction between language PAA and experience can be justified by the fact that language and communication skills are more easily observed in interviews during the hiring process at the beginning of a worker's

[21]Both Farber and Gibbons (1996) and Altonji and Pierret (2001) estimate insignificant effects of interaction between schooling and experience using the same specification. Our estimates are small but marginally significant.

career. Therefore, there is not much learning from experience regarding this attribute.

Figure 4 presents additional empirical evidence of the decreasing returns to graduating from a selective university in Chile. Each circle of the graph represents an estimation of the coefficient of the selective university dummy controlling for math and language scores within each experience group. Note that the impact of graduating from a selective university is greater for recent graduates (21-26%) but tends to slowly decrease in the first years of a worker's career. Note that the wage premium finally stabilizes at around 9% after the 8-th year of labor market experience. Note that, as presented in equation 7, this coefficient represents both the direct effect of graduating from a prestigious university and the indirect effect of η and q on wages.

We present a robustness check for the evidence that employers statistically discriminate workers on the bases of university selectivity. In Table 4, rather than using an indicator that a worker graduated from one of the two most prestigious universities in Chile, we use a continuous measure of prestige for all colleges. Precisely, we assign to each university its quality score as presented in the 2011 "*Que Pasa*" college ranking, one of the most widely recognized in the country. In this framework, we estimate how earnings vary with this score, defined as a university quality index in the table, its interactions with experience and the remaining controls. We lose 5,447 observations of those individuals whose colleges do not present an available "*Que Pasa*" score in 2011. The results from this estimation are very similar to the ones presented in table 3: there are gains from attending a more prestigious university for recent college graduates but these returns tend to decrease with work experience. We also estimate that returns to math PAA scores increase with experience in this specification.

5.2 Regression Discontinuity Test

In order to provide further evidence for statistical discrimination based on college prestige, we use a regression discontinuity (RD) design. The test consists of comparing how earnings change as workers accumulate experience in the labor market for those just above and just below the admission cutoff to the most selective universities in Chile. The identification assumption is that other factors that could affect earnings are continuous at the admission cutoff and students have limited power to manipulate on which side of the admission cutoffs they might fall.[22] As discussed in section 4, different from the traditional EL-SD test presented in the past section, the regression discontinuity test can be interpreted as the causal effect for those individuals around the admission cutoff and is not biased by other factors observed by employers that can be used for statistical discrimination.

5.2.1 The Admission Process and the RD Design

Our data contains information on the year a student took the PAA test, his or her scores on each component of the test, his or her high school grade, the college he or she graduated from and his or her major. We do not observe application decisions and must therefore have additional assumptions and sample restrictions to perform the regression discontinuity design.[23] Precisely, we restrict the data to individuals who graduated with engineering, business, medical and law degrees (competitive majors) and assume that these workers would prefer to graduate with these majors in a less prestigious university rather than study a different major in a prestigious university. Under this assumption, we can interpret that workers just above the admission cutoff (competitive major at prestigious universities) are those who were accepted to the highest program of their preference and those below the threshold (competitive major in less prestigious college) are those who were accepted to the second highest program of their preference.

We find evidence that this is a plausible assumption. First, these are the programs with the highest admission cutoffs and therefore should be the top choices of applicants. Second, there is a positive wage differential between workers with competitive majors in less prestigious universities and workers with less competitive majors in prestigious universities. We interpret this as evidence that students have incentives to pursue engineering, business, medicine or law degrees at a less prestigious universities rather than pursuing a different major at a prestigious university. Finally, in the Appendix we use application data for 2004 to show that individuals that choose engineering, business, medical and law major as their first choice in a prestigious university are very likely to

[22] Students can retake the test in the following year, but cannot retake the test in the same year after receiving results, diminishing the possibility of manipulation.

[23] We only have access to application data starting in 2004.

choose the same majors as their second and third options in the application. After this restriction, the sample consists of 9,376 individuals, with 2,899 of them graduating from a prestigious university.

Using additional data on the PAA weights used by these programs in the two prestigious universities we are able to reconstruct the final weighted score for all individuals in the restricted sample.[24] As a result, we derive $Univ1.Score_i$ and $Univ2.Score_i$ that represents the PAA weighted score of individual i at prestigious university 1 and 2 respectively.

Given the possibility that a student can be accepted to one or neither of the prestigious universities, we define the running variable used in the RD as follows:

$$Dist.Cutoff_i = max\{Univ1.Score_i - Univ1.Cutoff_i, Univ2.Score_i - Univ2.Cutoff_i\}$$

where $Univ1.Cutoff_i$ and $Univ2.Cutoff_j$ are the admission score cutoffs used by universities 1 and 2 for individual i's major in the year of application to college.[25] Note that individuals with $Dist.Cutoff_i$ slightly greater than zero were barely admitted to at least one of the two prestigious universities and individuals with slightly lower than zero were barely rejected by both schools.

In the RD design we will be interested in the following object:

$$\tau_t = \frac{\lim_{\text{Dist.Cutoff}\downarrow 0} \mathbb{E}[w_{it}|Dist.Cutoff_i] - \lim_{\text{Dist.Cutoff}\uparrow 0} \mathbb{E}[w_{it}|Dist.Cutoff_i]}{\lim_{\text{Dist.Cutoff}\downarrow 0} \mathbb{E}[g_i|Dist.Cutoff_i] - \lim_{\text{Dist.Cutoff}\uparrow 0} \mathbb{E}[g_i|Dist.Cutoff_i]}$$

where g_i is an indicator of whether worker i graduated from an elite university, t measures years of experience in the labor market, and w_{it} is the log(wages) after t years of experience. Note that the parameter τ_t represents the local average treatment effect on earnings after t years of experience for workers around the admission cutoffs who would enroll in a prestigious university if they were admitted (intent-to-treat effect). [26]

The employer learning-statistical discrimination RD test we propose consists of estimating if τ_t decreases with t. The test is based on the assumption that the unobserved ability (η_i) is positively

[24]We were only able to obtain PAA weights for years starting in the year 2000. In order to construct final scores for individuals that took the PAA prior to 2000, we assume that programs used the same weights for previous years. The evidence is that programs did not change weights over the period of time considered in this paper.

[25]Data on admission cutoffs were collected at the universities' websites (for later application years) and newspapers (for earlier application years). Therefore they are not affected by potential dropouts in the bottom of the application ranking.

[26]For a discussion of the relationship between regression discontinuity design and treatment effects, see Lee and Lemieux (2010).

correlated to graduating from a selective university but is continuous around the admission cutoff. In this framework, assuming that employers do not observe $Dist.Cutoff_i$, they will use information on college prestige to predict that workers just above the admission cutoff have a higher η_i.[27] However, the wage differential between those above and below the cutoff should decline if employers learn the true distribution of η_i as workers gain experience and therefore should rely less on college prestige to set wages.

5.3 Results

We first address the empirical question of whether the probability of graduating from one of the two prestigious universities in Chile is discontinuous at the admission cutoff. Note that it is possible that individuals with a higher score than the admission cutoffs decided to attend a less prestigious university, which implies that we have a fuzzy regression discontinuity design. Figure 5 shows graphically the discontinuity in the probability of graduating from a prestigious university at the cutoff. From the figure, we find that the discontinuity in graduating from a prestigious university is approximately 60 percent. This means that around 60% of the individuals with PAA scores just sufficiently high enough for admission choose to attend an elite university. Consequently, being just above the admission cutoff causes a large increase in the probability of graduating from a prestigious university in Chile, a necessary condition for the validity of the RD design.

In Figures 6 and 7 we present evidence for the validity of the RD design. In Figure 6 we test for the presence of a density discontinuity at admission cutoff by running kernel local linear regressions of the log of the density separately on both sides of the cutoff, as proposed by McCrary (2008). We do not find evidence of any discontinuity in the density, suggesting that students cannot precisely manipulate their scores around the cutoff. Next, in Figure 7 we search for a jump at the discontinuity for per-treatment variables that should not be affected by the treatment. Precisely, if being above or below the cutoff is random, we should observe a zero treatment effect on the

[27] Note that in section 4 we also assume that employers cannot observe $Dist.Cutoff_i$. Furthermore, screening workers is expensive and employers learn fast (Lange, 2007), therefore it is not economically attractive to perform ability tests on recent college graduates.

probability of being female or graduating from a private high school (Imbens and Lemieux 2008). The figure suggests that there is no discontinuity of these variables around the cutoff. In fact, from a formal test using the same specification in columns (1) to (3) of Table 5 but using female or private high school indicator as a dependent variable, we cannot reject at reasonable levels of significance that there are zero effects of being above the cutoff on these pre-treatment outcomes.[28] We also present evidence in the appendix of the paper that individuals below and above the cutoff are not different in terms of their age when they enter college.

In order to present evidence of the effects of admission to a selective university on earnings, in Figure 8 we plot in the figure the unconditional means of log annual earnings on the vertical axis and the distance from the admission cutoff on the horizontal axis for the first years of labor market experience. The open circles represent 10 points local average and the lines represent linear fits of the data below and above the admission cutoff. The figure shows that there is a jump in earnings in the first year of labor market experience for workers who are just above the cutoff. This discontinuity is consistent with previous literature that finds a significant effect on earnings for being just above the admission cutoff of recent college graduates (Saavedra, 2008). However, as workers gain labor market experience, the discontinuity in earnings tends to decrease to the point that there is no apparent difference in terms of earnings between workers just above and just below the cutoffs five years after graduation. In addition, we observe that workers tend to be paid more in accordance with their weighted score as they accumulate experience in the market.

Table 5 presents further statistical evidence for discontinuity in earnings at the admission cutoff. In columns (1) to (3) of panel A of the table, we show that workers above the admission cutoff have on average 6-8% higher earnings than just below the admission cutoff in their first 10 years of labor market experience (varying little with bandwidth). In columns (4) to (7) we present specifications that allow the return from being admitted to a selective university to change over a worker's career. Under this specification, we estimate a 10%-14% wage premium for those above the cutoff in their first year of labor market experience, but this differential decreases by 1.5 to 2.7 percentage points

[28] Due to space constraints we omit the tests here, but they are available upon request.

per year of experience.

In Panel B of Table 5 we present the earnings discontinuity estimates taking into consideration that not all applicants with sufficiently high scores enroll in the top universities. For this purpose, we estimate an earnings equation using a two-stage least square method, where both graduating from a prestigious university and its interaction with experience are instrumented with an indicator for PAA scores above the admission cutoff and its interaction with experience. We estimate a 16-22% effect of graduating from a selective university on earnings of recent college graduates. However, this gap decreases by 2.1-3.7 percentage points per year of experience in the labor market. Note that these estimates should be interpreted as the causal effect only for those applicants that would enroll in a prestigious university and graduate in the event of achieving a sufficiently high score (intent-to-treat effect).

In order to provide robustness checks for the main RD findings, in Table 6 we present estimates for the earnings discontinuity at the admission cutoff and its interaction with experience for different model specifications. Precisely, we show in row (1) that our estimates are not sensitive to the exclusion of controls, which is expected if treatment is random around the admission cutoff. In rows (2) and (3) we test how our estimates change with different specifications for the distance from the admission cutoff. Finally we estimate our preferred model for males and females separately. While we estimate similar coefficients for these two groups, we do not find a significant change in the returns to being approved by a prestigious university with experience for women. We notice however that this result is due to large standard errors that might be explained by the fact that we have a smaller fraction of women in the restricted sample.

6 Job Market Signaling and Human Capital Effects[29]

In the previous section we tested whether employers use university prestige to statistically discriminate among college graduates using a regression discontinuity test. The reminder of this paper examines the importance of signaling and human capital accumulation on the determination of

[29]We thank an anonymous referee for suggesting this section.

earnings. We show that, based on the assumptions from the regression discontinuity EL-SD test, we can identify the signaling and human capital components of the college selectivity premium.

It is important to emphasize that these two mechanisms are only identifiable under the assumptions of the employer learning-statistical discrimination model. In particular, the EL-SD model imposes that human capital accumulation from school and on-the-job training are independent from each other. In terms of the model presented in section 5.1, the function $H(t)$ is independent of any other variables in the model. This might not be true if some workers benefit more from on-the-job training than others. Indeed, if workers from prestigious universities benefit more from on-the-job training than graduates from less prestigious universities, then our model will underestimate the impact of signaling on wages. On the other hand, if workers from less prestigious university benefit more from on-the-job training, then our model will overestimate the impact of signaling on wages.[30]

6.1 Parameter of Interest and Estimation Procedure

The objective is to identify the components of the college wage premium that are determined from signaling and from human capital accumulation. From equation 10, we show that the college selectivity premium for workers with same pre-college characteristics and experience t is:

$$\tau_t = r + (1 - \theta_t)(\alpha_2 + \lambda\gamma_2) \tag{11}$$

The parameter r is the direct effect from graduating from a prestigious university on a worker's productivity, which can be generated by the better instruction provided by prestigious universities and having more accomplished peers. We therefore define r as the human capital effect of college selectivity on earnings. Next, we define the parameter $\varpi = \alpha_2 + \lambda\gamma_2$. As described in the section 5.1, this parameter is the employers prediction of a worker's unobservable characteristics based on college prestige. We define this parameter as the signaling effect of graduating from a prestigious

[30]Moreover, in order to recover the parameter of the employer learning model, we need to assume that the expectation errors by employers follow a normal distribution.

university on earnings. Finally, the parameter $1-\theta_t$ is the weight that employers give to college prestige in the prediction of unobservable characteristics. Under the assumptions of the employer learning model described in 5.1, Lange (2007) showed that:

$$\theta_t = \frac{tK_1}{1+(t-1)K_1}$$

where K_1 is the speed of employer learning parameter[31]. Note that $\theta_0 = 1$, and θ_t goes to zero over time as employers learn a worker's true ability.

The goal is to estimate the parameters $\{r, \varpi, K_1\}$. For this purpose, we follow the estimation procedure presented in Lange (2007): we first estimate the RD coefficients $\{\hat{\tau}_t\}_{t=1}^{10}$for each experience level t[32]. The coefficients are estimated using a two-stage least square method within a 100 points bandwidth, as presented in Panel B of Table 5., where graduating from a selective university and its interactions with experience dummies are instrumented with approved at prestigious university and its interactions with experience dummies.[33].

Next, we will find parameters that best fit the non-linear equation 11 given the RD estimated coefficients. For this purpose, we treat each estimated coefficient $\{\hat{\tau}_t\}_{t=1}^{10}$as an observation and uses a non-linear least squares method to estimate $\{r, \varpi, K_1\}$.

$$\min_{r,\varpi,K_1} \sum_{t=1}^{T} (\hat{\tau}_t - r - (1 - \frac{tK_1}{1+(t-1)K_1})\varpi)^2$$

6.2 Results

Figure 9 shows the estimated coefficients $\{\hat{\tau}_t\}_{t=1}^{10}$ as well a quadratic function to fit the observations. The figure clearly indicates a monotonic decrease in the university prestige wage premium as workers progress in their career, which is consistent with the findings of Tables 5 and 6, as well as Figure 8. Furthermore, the decrease of the wage gap is more significant in the beginning of a workers career and the college selectivity premium tends to stabilize later in a a worker's career. This result is in accordance with. the employer learning model, where wage premium reflects both human capital

[31]The parameter K_1 is a function of the variances of $\varepsilon_{i\tau}$, s and q. See Lange (2007) for its derivation.

[32]We use the same specification and two stage least square estimation procedure as presented in Panel B from Table 5, but interacting the a complete set of experience dummies with the graduate from selective university variable.

[33]Recall that we are using the variable approved as an instrument for the variable selective.

and signaling effect when workers start their careers, but only reflect differences in productivity later in a worker's life cycle.

We report the estimated coefficients $\{r, \varpi, K_1\}$ as well as their 95% confidence intervals in Column (1) of Table 7. The estimate of the speed of learning parameter K_1 is 0.1654, which implies that initial expectation errors by employers decline by 16% during the first period and 37% after 3 periods. This speed of learning is slower than what is found in the literature for the U.S., as Lange (2007) estimates a 26% decline in employer expectations error after the first year of a worker's career. In addition, we estimate a negative human capital effect r of -0.0676 and the large signaling effect of ϖ of 0.3166. In both cases we find big standard errors and we cannot reject that there are differences in the human capital effect and signaling effects at a 95% significance level.

Based on these findings, we suspect that due to data limitations, our estimation is underestimating the speed of learning and the human capital effects on earning and overestimating the effect of signaling on the college wage premium. The reason is that our data only follows workers for a maximum of 10 years of labor market experience.[34] As a result, we are oversampling the periods in a worker's career when there is a decline in the college selectivity wage premium and undersampling the period where there is a stable wage premium, which might lead to the estimation of slower speed of learning and negative long-term college selectivity wage premium.

Due to this data limitation, we decided to estimate a model imposing a speed of learning parameter that is more in line with the literature (Lange, 2007). We set $K_1$to be equal to 0.26, which implies a 26% decline in employer expectations error after the first year of a worker's career. Under this assumption, we find more plausible signaling and human capital effects as well as smaller confidence intervals, presented in Column (2). While we still estimate negative coefficients for r, the confidence intervals are narrower to the point that we cannot reject with 95% of confidence that the human capital effect on the college wage premium is equal to 0.029 log points, which resembles the trend to long run wage premium from figure 9. In addition, we estimate with 95% confidence that the signaling effect on the college selectivity wage premium is at least 0.1829 log points. Based on

[34]As a reference, Lange (2007) follows workers for up to 17 years of experience.

this estimation, we conclude that under a more plausible speed of learning, the signaling effects on the college selectivity wage premium is at least 6.3 times higher than the human capital effects. This implies that the human capital effect represents at most 14% of the college prestigious premium for recent university graduates, with the remaining 86% being driven by signaling effects for people with similar admission cutoffs scores.

7 Conclusion

This paper tests whether employers statistically discriminate based on the selectivity and prestige of the university attended by workers. We first follow the employer learning statistical discrimination test suggested by Altonji and Pierret (2001); our results indicate that the returns to graduating from an elite university in Chile decrease with experience and that the returns from hard-to-observe ability correlates increase with experience. In addition, we show that the traditional employer learning statistical discrimination test is biased on the basis of the treatment variable if employers use characteristics of individuals that are not available in the data, such as family socioeconomic status, to predict a worker's unobservable productivity.

For this reason, we take advantage of the centralized admission process of traditional universities in Chile to propose a statistical discrimination test based on a regression discontinuity design. The test consists of comparing the earnings of those just above and just below the admission cutoff to the most prestigious universities in Chile. We show that recent graduates with weighted PAA scores just above the admission cutoff have significantly higher earnings than those with weighted PAA scores just below the cutoff. However, as workers gain labor market experience, the earnings gap between these two groups decreases with labor market experience. We interpret this result as employers paying workers in accordance with the selectivity of their college when they first graduate from university, but rewarding them based on their true productivity as they reveal their quality to employers over time.

These results shed some light on the benefits of graduating from a selective university. We interpret our findings as evidence that attending a prestigious university has a significant impact on signaling to employers a worker's unobservable quality. However, employers learn fast and individuals tend to be paid in accordance with their true ability as they gain experience in the labor market. We interpret the erosion of the earnings premium for workers with similar pre-university characteristics as evidence of signaling. In addition, under some stronger assumptions we can disentangle these two mechanisms and estimate that the human capital effect represents at most 14% of the college prestige premium for recent university graduates, with the remaining 86% being driven by signaling effects.

While this interpretation is consistent with the employer learning literature, there are other possible theories that could explain this pattern. For example, it is possible that individuals who graduated from a less prestigious university learn faster on-the-job than those who went to a prestigious university. In this framework, the wages of individuals who graduated from less selective schools could catch up with the wages of those who graduated from selective schools. While traditional human capital theory suggests that investment at school and on-the-job training are complements, we cannot rule out alternative theories where these two types of investments are substitutes. We hope that our new empirical findings may lead to further research on the topic.

Finally, it is important to keep in mind that our results are obtained using data from Chile. Future research should apply our framework to test employer learning, statistical discrimination and university prestige in other countries, particularly in those countries where students use a university selection test and apply directly to a major and a school at the time of application. The fact students in Chile apply to a program, that is, a major and a university simultaneously is the most notable difference when compared to the admission process for schools in the US.

References

Aigner, Dennis J. and Glen G. Cain (1977), 'Statistical theories of discrimination in labor markets', *Industrial and Labor Relations Review* **30**(2), pp. 175–187.

Altonji, Joseph G. and Charles R. Pierret (2001), 'Employer learning and statistical discrimination', *The Quarterly Journal of Economics* **116**(1), pp. 313–350.

Araki, Shota, Daiji Kawaguchi and Yuki Onozuka (2015), University prestige, performance evaluation, and promotion: Estimating the employer learning model using personnel datasets, Technical report, Research Institute of Economy, Trade and Industry (RIETI).

Arcidiacono, Peter, Patrick Bayer and Aurel Hizmo (2010), 'Beyond signaling and human capital: Education and the revelation of ability', *American Economic Journal: Applied Economics* pp. 76–104.

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), pp. 701–728.

Bordon, Paola (2014), 'The effects of private high schools, university rankings and employer learning on wages in Chile', Working Paper 133, Estudios Publicos.

Brewer, Dominic J, Eric R Eide and Ronald G Ehrenberg (1999), 'Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings', *Journal of Human Resources* pp. 104–123.

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', *The Quarterly Journal of Economics* **117**(4), 1491–1527.

Farber, Henry S. and Robert Gibbons (1996), 'Learning and wage dynamics', *The Quarterly Journal of Economics* **111**(4), pp. 1007–1047.

Hastings, Justine S, Christopher A Neilson and Seth D Zimmerman (2013), Are some degrees worth more than others? Evidence from college admission cutoffs in Chile, Technical report, National Bureau of Economic Research.

Hershbein, Brad (2013), 'Worker signals among new college graduates: The role of selectivity and GPA', *Working Paper, Upjohn Institute* .

Hoekstra, Mark (2009), 'The effect of attending the flagship state university on earnings: A discontinuity-based approach', *The Review of Economics and Statistics* **91**(4), pp.717–724.

Hoxby, Caroline M (1998), 'The return to attending a more selective college: 1960 to the present', *Unpublished manuscript, Department of Economics, Harvard University, Cambridge, MA* .

Imbens, Guido W and Thomas Lemieux (2008), 'Regression discontinuity designs: A guide to practice', *Journal of Econometrics* **142**(2), 615–635.

Kaufmann, Katja Maria, Matthias Messner and Alex Solis (2012), Returns to elite higher education in the marriage market: Evidence from Chile, Technical report, Bocconi University working paper.

Lang, Kevin and Erez Siniver (2011), 'Why is an elite undergraduate education valuable? Evidence from Israel', *Labour Economics* **18**(6), pp. 767 – 777.

Lange, Fabian (2007), 'The speed of employer learning', *Journal of Labor Economics* **25**(1), pp. 1–35.

Lee, David S. and Thomas Lemieux (2010), 'Regression discontinuity designs in economics', *Journal of Economic Literature* **48**(2), pp. 281–355.

Mansour, Hani (2012), 'Does employer learning vary by occupation?', *Journal of Labor Economics* **30**(2), pp. 415–444.

McCrary, Justin (2008), 'Manipulation of the running variable in the regression discontinuity design: A density test', *Journal of Econometrics* **142**(2), pp. 698–714.

Phelps, Edmund S. (1972), 'The statistical theory of racism and sexism', *The American Economic Review* **62**(4), pp. 659–661.

Saavedra, Juan E. (2008), 'The returns to college quality: A regression discontinuity analysis', Unpublished manuscript, Harvard University, Cambridge, MA.

Schönberg, Uta (2007), 'Testing for asymmetric employer learning', *Journal of Labor Economics* **25**(4), pp. 651–691.

Spence, Michael (1973), 'Job market signaling', *The Quarterly Journal of Economics* **87**(3), pp. 355–374.

Zimmerman, Seth (2013), 'Making top managers: The role of elite universities and elite peers', *Link: https://sites. google. com/site/sethdavidzimmerman/research* .

Table 1: **Descriptive Statistics for Selective and Non-Selective Universities**

Variables	Selective Universities		Non-selective Universities	
	Mean	Std. Dev.	Mean	Std.Dev.
Female	0.53	0.50	0.56	0.50
Language PAA Score	680.6	61.3	588.8	82.8
Math PAA Score	715.9	68.7	609.4	103.4
High School Grade	644.4	78.7	575.1	94.1
Private High School	0.11	0.32	0.07	0.25
Number of Individuals	11,495		46,684	

Note: Math and Language PAA scores are components of the centralized test for admission to university in Chile. Selective universities are defined as the two most prestigious universities in Chile. See section 3 for details.

Table 2: **Earnings for Selective and Non-Selective Universities**

Statistic	Selective Universities	Non-selective Universities
Log of Annual Wage (in 1999 Pesos)		
Mean	15.58	15.19
Std. Deviation	1.09	1.16
Observations	61,844	251,233

Note: Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Table 3: **Traditional EL-SD Regression**

Dependent Variable: Log Annual Wage

Model	(1)	(2)	(3)
Variables			
Selective University	0.196	0.229	0.259
	(0.009)***	(0.013)***	(0.014)***
Selective University x Experience		-0.009	-0.017
		(0.003)***	(0.003)***
PAA Language	0.031	0.031	0.028
	(0.004)***	(0.004)***	(0.007)***
PAA Math	0.075	0.075	0.047
	(0.005)***	(0.005)***	(0.007)***
PAA Language x Experience			0.001
			(0.001)
PAA Math x Experience			0.007
			(0.001)***
Constant	12.726	12.716	12.905
	(0.115)***	(0.115)***	(0.119)***
Observations	307,864	307,864	307,864
R-squared	0.244	0.245	0.245

Controls: Female, Private High School,Cubic Experience Polynomial, Major Dummies, and Year Dummies.

White/Huber standard errors clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Math and Language PAA are standardized by test year. Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Table 4: **Traditional EL-SD Regression - University Quality Index**

Dependent Variable: Log Annual Wage

Model	(1)	(2)	(3)
Variables			
University Quality Index	0.089	0.102	0.121
	(0.004)***	(0.006)***	(0.006)***
University Quality Index x Experience		-0.003	-0.008
		(0.001)***	(0.001)***
PAA Language	0.028	0.027	0.024
	(0.004)***	(0.004)***	(0.007)***
PAA Math	0.057	0.057	0.025
	(0.005)***	(0.005)***	(0.007)***
PAA Language x Experience			0.001
			(0.001)
PAA Math x Experience			0.008
			(0.001)***
Constant	12.304	12.206	12.299
	(0.115)***	(0.119)***	(0.120)***
Observations	302,417	302,417	302,417
R-squared	0.245	0.245	0.245

Controls: Female, Private High School,Cubic Experience Polynomial, Major Dummies, and Year Dummies.

White/Huber standard errors clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: University quality index is the score awarded to colleges by the "*Que Pasa*" ranking of 2011 and is measured in standard deviations. Math and Language PAA are standardized by test year. The sample is restricted to individuals whose colleges were assigned a score by the "Que Pasa" ranking in 2011.

Table 5: **EL-SD Regression Discontinuity Test**

Dependent Variable: Log Annual Wage

Panel A	Reduced Form					
	Bandwidth (Points from Cutoff)			Bandwidth (Points from Cutoff)		
	125	100	75	125	100	75
Model	(1)	(2)	(3)	(4)	(5)	(6)
Approved at Selective University	0.081	0.073	0.065	0.143	0.119	0.099
	(0.0287)***	(0.0308)**	(0.0342)*	(0.0330)***	(0.0348)***	(0.0382)***
Approved at Selective Univ.* Experience				-0.027	-0.020	-0.015
				(0.0071)***	(0.0073)***	(0.0078)*
Observations	39,748	36,639	31,843	39,748	36,639	31,843
R-squared	0.135	0.128	0.120	0.135	0.128	0.120

Panel B	2 Stages Least Square					
	Bandwidth (Points from Cutoff)			Bandwidth (Points from Cutoff)		
	125	100	75	125	100	75
Model	(1)	(2)	(3)	(4)	(5)	(6)
Graduated from Selective University	0.133	0.122	0.110	0.222	0.189	0.160
	(0.0470)***	(0.0512)**	(0.0576)*	(0.0525)***	(0.0565)***	(0.0629)**
Graduated from Selective Univ.* Experience				-0.037	-0.028	-0.021
				(0.0098)***	(0.010)***	(0.0109)*
Observations	39,748	36,639	31,843	39,748	36,639	31,843
R-squared	0.140	0.133	0.125	0.140	0.134	0.126

Approved at Selective Univ.: Points from the Cutoff>= 0

Controls: Points from the Cutoff, and Interaction of Points from the Cutoff with Approved at Prestigious Univ., Female, Cubic Experience Polynomial, Major Dummies, and Year Dummies.

Instrument in Panel B: In columns (1)-(6) the endogenous variables are intrumented with Approved at Prestigious University and in columns (4)-(6) also with its interaction with experience

White/Huber standard errors accounting clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details). Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Table 6: **EL-SD Regression Discontinuity Test - Robustness Checks**

Dependent Variable: Log Annual Wage, Applicants within 100 points from Cutoff

Regression Specification	Additional Controls	Function of Points from the Cutoff	Flexible Coefficient?	Sample	Estimated Coefficients: Approved at Selective University	Estimated Coefficients: Approved at Selective Univ.* Experience
(1)	No	Linear	Yes	All	0.130 (0.0354)***	-0.019 (0.0075)**
(2)	Yes	Cubic	No	All	0.103 (0.0426)**	-0.020 (0.0073)***
(3)	Yes	Cubic	Yes	All	0.109 (0.0425)**	-0.020 (0.0073)***
(4)	Yes	Linear	Yes	Males	0.100 (0.0427)**	-0.018 (0.0086)**
(5)	Yes	Linear	Yes	Females	0.153 (0.0599)**	-0.022 (0.0135)

All specifications include Cubic Experience Polynomial.

Approved at Selective Univ.: Points from the Cutoff>= 0

Additional Controls: Female, Major Dummies, and Year Dummies.

Flexible coefficient indicates whether the estimated coefficients of points from cutoff was allowed to differ on each side of the admission cutoff

White/Huber standard errors clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details). Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Table 7: **Human Capital and Signaling Parameters**

Parameter	(1)	(2)
Speed of learning K_1	0.1654	0.2600
	[-0.3667, 0.2316]	
Human Capital Effect r	-0.0676	-0.0241
	[-0.2530, 0.5837]	[-0.0770, 0.0290]
Signaling Effect ω	0.3166	0.3080
	[0.1192, 0.5140]	[0.1829, 0.4330]

Note: The reported parameters are estimated by nonlinear least squares using the RD coefficient estimates at different experience levels with with 95 percent confidence intervals in brackets. Section 6 describes the link between the parameters reported here and the estimated coefficients. In Column (1) we estimate the speed of learning parameter and in Column (2) we impose a coefficient of 0.26, which implies a speed of learning consistent with the literature. The sample is restricted to individuals with engineering, business, medical and law degree (see section 5 for details). Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Figure 1: **Application Process to Traditional Universities**

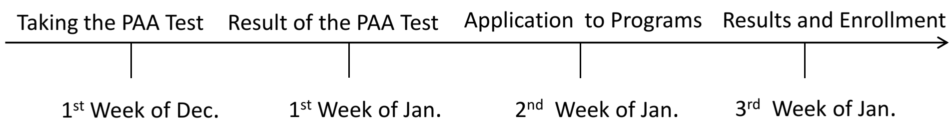
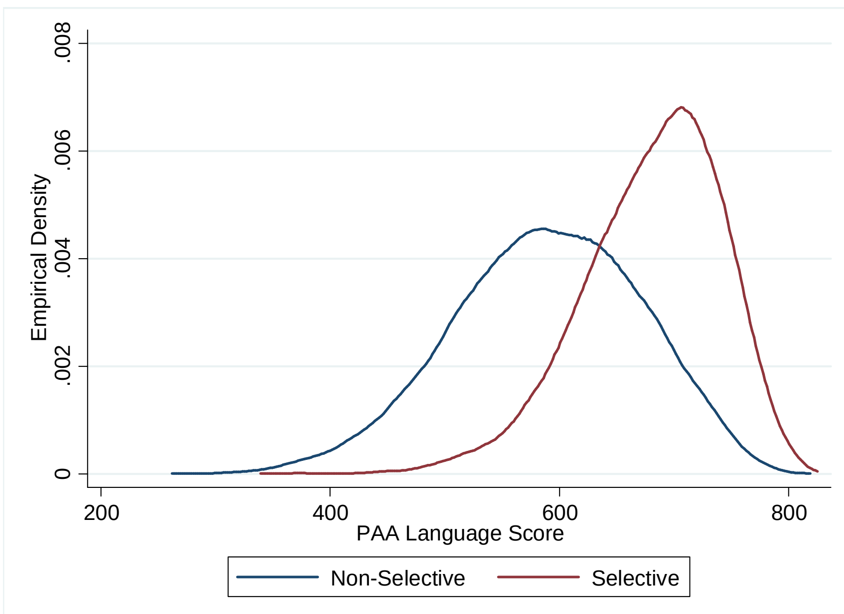
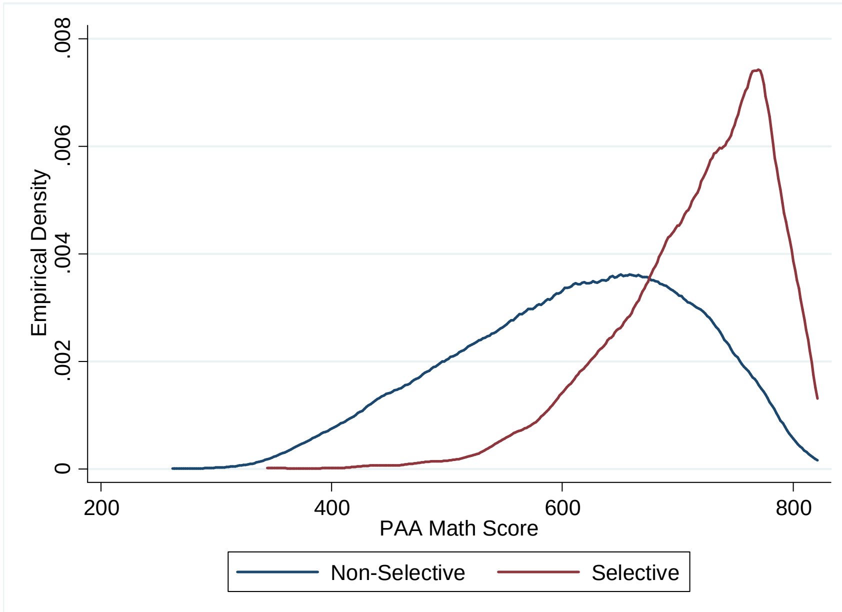


Figure 2: **Smoothed Language PAA Score Distribution**



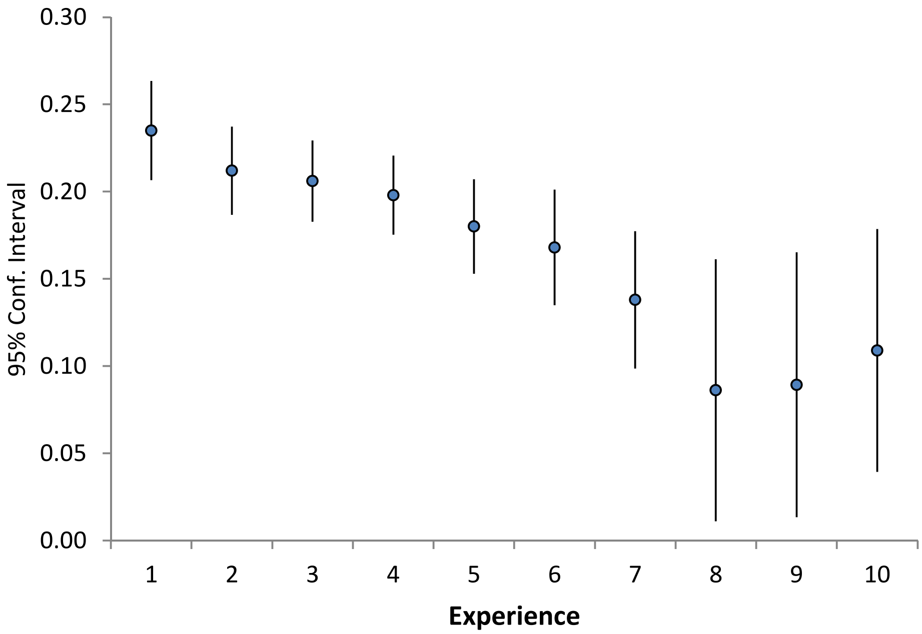
Note: Language PAA is a component of the centralized test for admission to university in Chile. Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Figure 3: **Smoothed Math PAA Score Distribution**



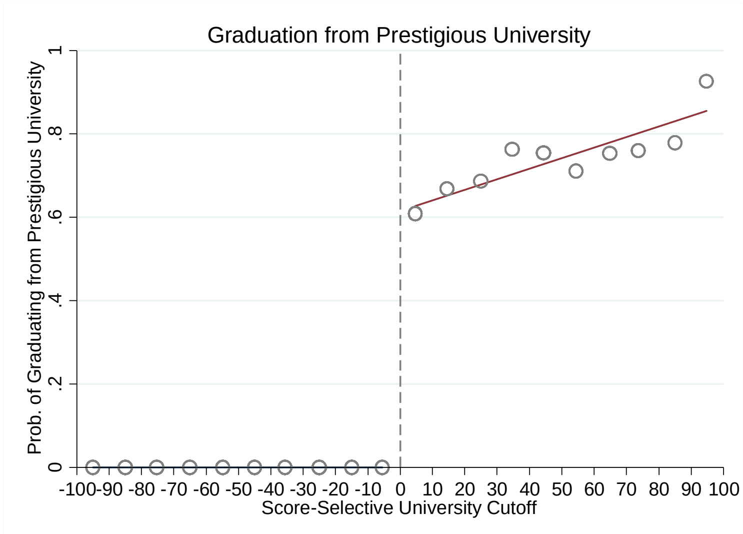
Note: Math PAA is a component of the centralized test for admission to university in Chile. Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Figure 4: **Traditional EL-SD Selective University Coefficient by Experience Level**



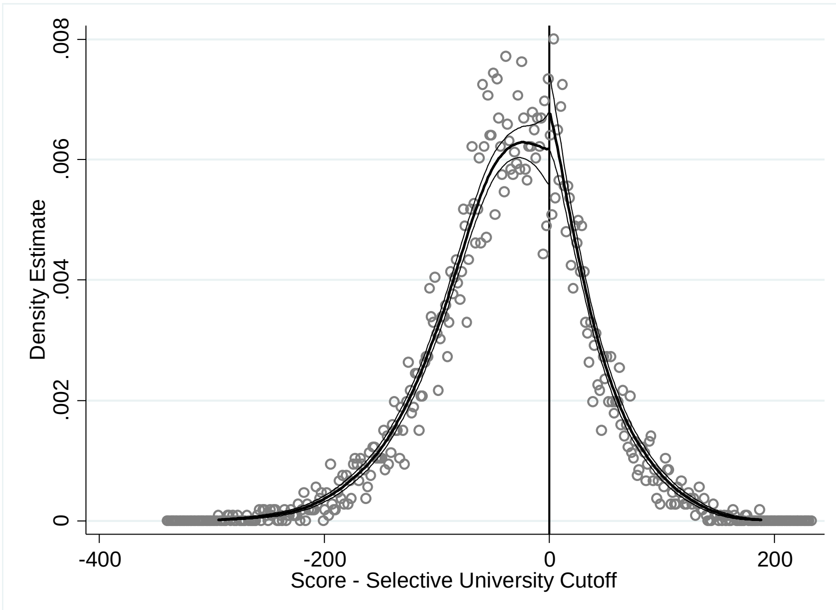
Note: Each circle represents the effect of the selective university dummy estimated by linear least squares within each of the 10 experience groups. The controls used in the regressions are the same as those presented in Table 3 (including Math and Language PAA scores). Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details). Confidence intervals are calculated using White/Huber heteroscedasticity standard errors.

Figure 5: **Graduation from Selective University Discontinuity**



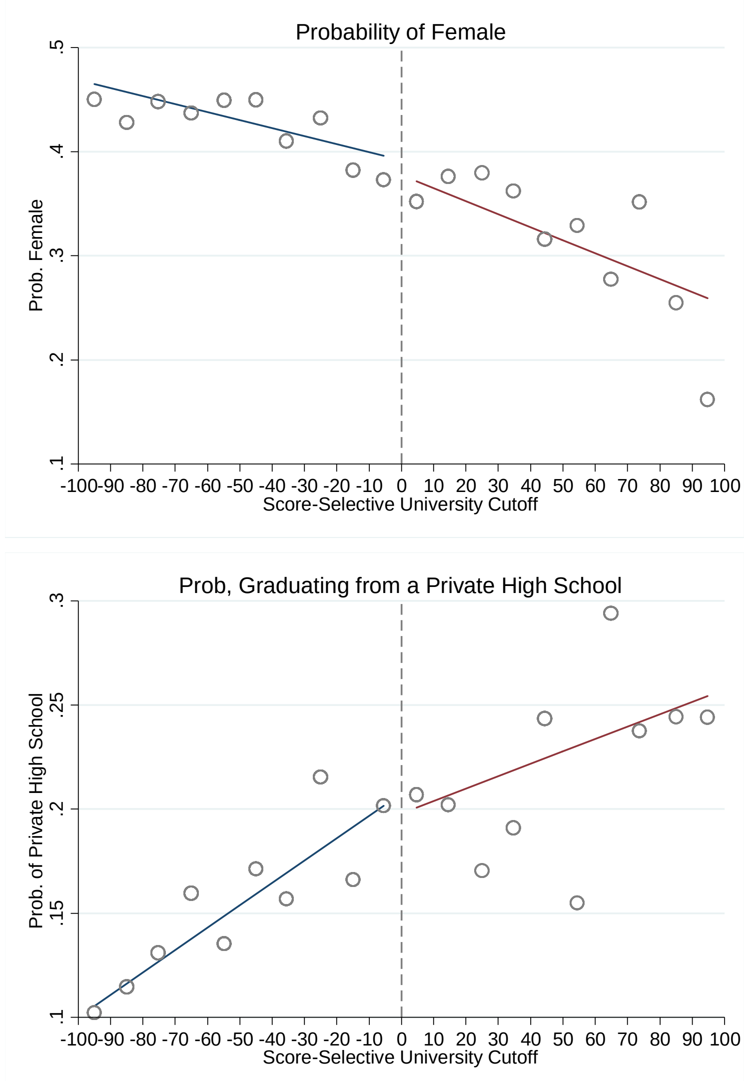
Note: Open circles represent 10 points local averages and the lines are local linear fits below and above the admission cutoff. The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details).

Figure 6: **McCrary Density Tests**



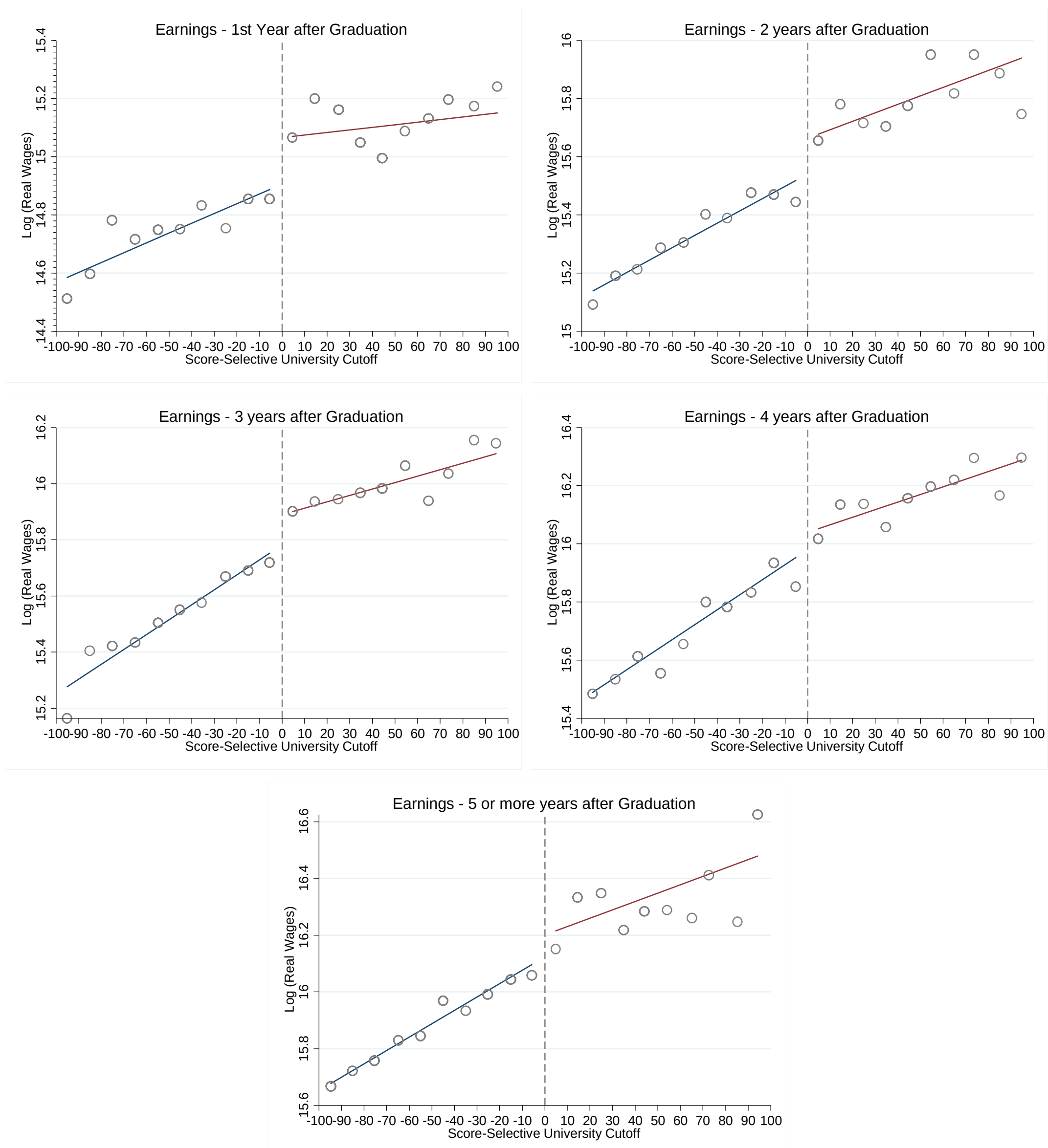
Note: Weighted kernel estimation of the log density of the distance to admission cutoff performed separately on either side of the admission threshold. Optimal binwidth and binsize as in McCrary (2008). The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details).

Figure 7: **Discontinuity at Pre-treatment Outcomes**



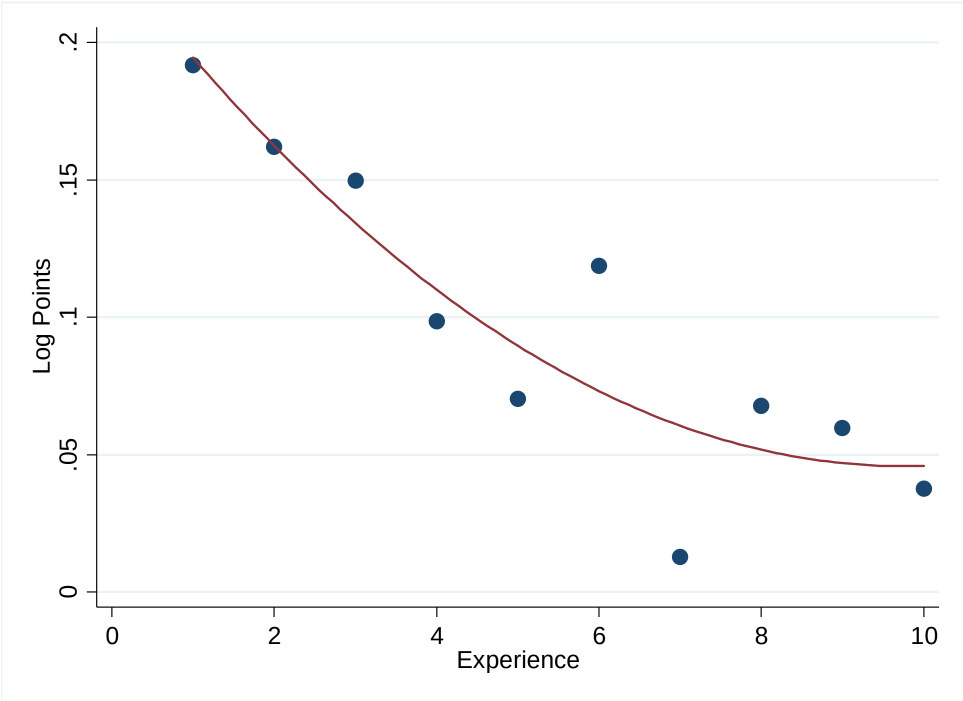
Note: Open circles represent 10 points local averages and the lines are local linear fits below and above the admission cutoff. The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details).

Figure 8: **Earnings Discontinuity by Experience**



Note: Earnings are defined as log annual wages measured in real Chilean pesos. Open circles represent 10 points local averages and the lines are local linear fits below and above the admission cutoff. The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details).

Figure 9: **RD Coefficients by Experience Level**



Note: Each circle represents the effect of the selective university dummy estimated by two stage least square within each of the 10 experience groups. The line is a quadratic fit between these coefficients.

Appendix (for online publication)

College Selectivity Measures

In the paper, we defined the selective universities as the two of the oldest and most prestigious universities in the country and non-selective are all other universities. Using the 2002 data from the Chilean Department of Evaluation and Educational Testing Service (DEMRE) and the yearly reports of the Council of Chancellors of Chilean Universities (CRUCH) we provide further evidence that these two universities are distinguishably more selective than all other universities. In Table 1 we rank the 10 most selective universities of Chile by their average PAA cutoff score. As expected, the two prestigious universities have higher cutoffs than any other universities in the country, demonstrating they select the best candidates. We also show evidence that the two most selective universities have considerably higher instructional spending per student than any other school in the country. Finally, we also provide statistics on admission rates for the most selective universities in country, as defined by number of students accepted divided by the number of applicants. While the two selective universities also have low acceptance rates compared to most of the other schools, these application rates should be interpreted with cautions given the centralized admission process to universities in Chile. As described on section 2, students only apply to university after receiving their PAA scores, ranking 8 programs by preference. As a result, students with low scores are less likely to apply to the two most selective schools in the country, as they might consider the impossibility of admission to a prestigious school.

Traditional EL-SD Test - 1995 Cohort

Different from Altonji and Pierret (2001), the panel used in this paper is unbalanced, since the length of follow-up period is not the same for all cohorts. In order to make the data consistent with the past literature, we report results using only the first cohort of workers in Table 2. These are workers who graduated in 1995 and for whom we observe 10 years of labor market experience. Similar to the results presented in Tables 3, we estimate a significant effect of graduating from a prestigious university for recent graduates, but the selectivity premium decreases over time. Note however, that for this sample we find that the returns to standard math PAA score decreases with experience while the returns to standard verbal PAA score increases with experience.

Major Choice Statistics

While we do not have information on application decisions for workers in our sample, we were able to obtain data for applications for the year 2004 from Chilean Department of Evaluation and Educational Testing Service, which contains up to eight preferences of the students who apply each year.[35] Based on this data, we can test our hypothesis that students who choose engineering, business, medical and law degrees (competitive majors) in a selective university as their first application choice also apply to these same major as their second and third options (Table 3) . Overall, the results from this table are consistent with the assumption of the paper. For instance, 96.3% of students that choose Business as their first major option in a selective university also choose business as their major in their second application option. In addition, 93.9% of students that that choose Business as their first major option in a selective university also choose Business as their major in their third application option. The probability of second and third choices are very high for the other majors as well. This result is not surprising given the significant wage premium associated with a competitive major.

University Prestige Wage Premium vs Regional Wage Premium

A potential concern on the mechanisms driving the results of the paper is that the most prestigious universities are located in Santiago, the capital and biggest city of Chile. On one hand, it is possible that students barely above and below the cutoff face different wage prospects if they start their career in the city where their college is located. On the other hand, most managerial positions are located in Santiago, and it is a natural career path to move to the capital of the country if a worker reveals that he or she is a high ability type. While we cannot distinguish these two mechanisms

[35] The 2004 data is the oldest application data we could have access to.

in our data, we can control our regressions for region of employment to show that the employer learning mechanism holds for workers within a region of employment.

We report the results of the RD test controlling for a indicator whether the worker is employed in Santiago on Table 4. This estimation shows that there is a significant wage premium for workers employed in Santiago. Nonetheless, the inclusion of this geographic covariate barely affects the results of the RD test. We estimate very similar college selectivity premium at the beginning of a workers career in Tables 4and 5. We also show that this wage premium decreases as a worker gains labor market experience even when controlling for region of employment, showing that differences in wage prospects in the city which a worker graduates from does not explain the results of this paper.

Retaking the Admissions Exam and Discontinuity of Age at College Entry

Another potential threat to the validity of the regression discontinuity is that rejected students can retake the test until they score enough to be admitted in one of the two elite universities. If those who retake the test are different than those who do not retake, it is possible that individuals above the cutoff are not different in terms of their unobservables characteristics. We do not have application data for this time period to show that individuals barely rejected by a prestigious university are unlikely to apply again to college in the next year. Nevertheless, we can use information of age at college entry to show that individuals above the cutoff are not significantly older than those below the admission cutoff. Indeed, Table 5 shows that individuals above the cutoff are not significantly older at college entry than those below the cutoff. This result is not surprising given the main finding of the paper that returns to graduating from an elite university in Chile rapidly decrease with time for those around the admission cutoff.

Table 1: **Selective University Measure**

Rank	University	Average Cutoff PAA Score	Invest/student*	Admission Rate
1	Selective University	671.5	3,870	23.5%
2	Selective University	667.4	3,326	22.1%
3	Non-Selective University	638.6	1,297	38.1%
4	Non-Selective University	626.7	1,280	21.2%
5	Non-Selective University	619.6	1,064	14.3%
6	Non-Selective University	600.1	900	21.8%
7	Non-Selective University	598.2	1,079	32.0%
8	Non-Selective University	597.9	1,139	23.5%
9	Non-Selective University	597.2	1,109	26.3%
10	Non-Selective University	591.7	1,195	38.7%

* Instructional spending per student, in millions of Chilean pesos

Note: Data from DEMRE and CRUCH statistics 2002.

Table 2: **Traditional EL-SD Regression - 1995 cohort**

Dependent Variable: Log Annual Wage			
Model	(1)	(2)	(3)
Variables			
Selective University	0.151	0.274	0.272
	(0.023)***	(0.034)***	(0.036)***
Selective University x Experience		-0.021	-0.021
		(0.005)***	(0.005)***
PAA Language	0.012	0.012	-0.028
	(0.012)	(0.012)	(0.019)
PAA Math	0.085	0.085	0.130
	(0.014)***	(0.014)***	(0.021)***
PAA Language x Experience			0.007
			(0.002)***
PAA Math x Experience			-0.007
			(0.002)***
Constant	11.259	11.223	11.242
	(0.307)***	(0.307)***	(0.316)***
Observations	47,891	47,891	47,891
R-squared	0.334	0.335	0.335

Controls: Female, Private High School,Cubic Experience Polynomial, Major Dummies, and Year Dummies.

White/Huber standard errors clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Sample restricted to individuals who graduated in 1995. Math and Language PAA are standardized by test year. Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details).

Table 3: **Major Choices**

	Probability of choosing the same major again in their ..						
Major of First Choice	2nd choice	3rd choice	4th choice	5th choice	6th choice	7th choice	8th choice
Business	96.3	93.9	89.8	86.1	76.5	66.8	47.0
Engineering	98.4	97.8	97.1	97.9	95.9	94.4	90.2
Law	95.9	87.4	78.9	76.5	64.6	51.1	38.8
Medicine	96.7	89.9	79.5	80.7	65.9	54.4	41.2

Note: Data from DEMRE 2004. The sample is restricted to individuals that chose a program in a selective university as their first application choice.

Table 4: **EL-SD Regression Discontinuity Test - Regional Control**

Dependent Variable: Log Annual Wage

Panel A

	Bandwidth (Points from Cutoff)		
	125	100	75
Model	(4)	(5)	(6)
Approved at Selective University	0.134	0.105	0.084
	(0.040)***	(0.041)**	(0.045)*
Approved at Selective Univ.* Experience	-0.026	-0.019	-0.016
	(0.008)***	(0.008)**	(0.008)*
Santiago	0.349	0.340	0.330
	(0.019)***	(0.020)***	(0.021)***
Observations	32,547	29,961	25,973
R-squared	0.167	0.159	0.151

Approved at Selective Univ.: Points from the Cutoff>= 0

Controls: Points from the Cutoff, and Interaction of Points from the Cutoff with Approved at Prestigious Univ., Female, Cubic Experience Polynomial, Major Dummies, and Year Dummies. White/Huber standard errors accounting clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is restricted to individuals with engineering, business, medical and law degrees (see section 5 for details). Selective universities are defined as the two most prestigious universities in Chile (see section 3 for details). Santiago indicates whether the worker is employed in Santiago.

Table 5: **Age at College Entry Discontinuity Test**

Dependent Variable: Age at College Entry

	Reduced Form		
	Bandwidth (Points from Cutoff)		
	125	100	75
Model	(1)	(2)	(3)
Approved at Selective University	-0.168	-0.169	-0.183
	(0.120)	(0.133)	(0.149)
Observations	34,990	32,025	27,580
R-squared	0.044	0.044	0.047

Approved at Selective Univ.: Points from the Cutoff>= 0

Controls: Points from the Cutoff, and Interaction of Points from the Cutoff with Approved at Prestigious Univ., Female, Cubic Experience Polynomial, Major Dummies, and Year Dummies.

White/Huber standard errors accounting clustered at the individual level are reported in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: The sample is restricted to individuals with engineering, business, medical and law degree (see section 5 for details). Number of observation is different from the main tables of the paper due to individuals with missing year of college entry.