Abstract

We investigate whether firms use university prestige to statistically discriminate among college graduates. The test is based on the employer learning literature which suggests that if firms 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 to estimate a 19% wage premium for recent graduates of two 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’ quality when they leave school. Nevertheless, as workers reveal their productivity throughout their careers, they become rewarded based on their true quality rather than the prestige of their college.

JEL Classification: I21, J31, J71

Keywords: Employer Learning, Statistical Discrimination, College Selectivity.

*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, 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: pbordont@gmail.com and bbraga@urban.org.
1 Introduction

Labor markets are characterized by incomplete information on workers’ productivity (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, firms often have to make judgments on workers’ unobservable quality 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 quality. 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 schooling (Farber and Gibbons, 1996, Altonji and Pierret, 2001 and Lange, 2007) or information on lay-offs (Gibbons and Katz, 1991 and Hu and Taber, 2011) to statistically discriminate workers.

In this paper we study a new dimension of statistical discrimination: we investigate if firms use the prestige of the university attended by a worker to predict his or her unobservable labor market quality. We believe that college prestige satisfies the typical features of group identity that might be used for statistical discrimination for two main reasons. First, this information is easily accessible to firms: workers use the university name 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 firms use university prestige in order to infer the unobservable labor market quality of workers.

In order to test if employers use university prestige as a signal of workers’ unobservable quality, we rely on the statistical discrimination and employer learning (EL-SD) literature (Altonji and

\(^1\)As it will become clear later, the underlying assumption is that universities are better at screening candidates than firms.
The underlying assumption is that the imperfect information about a worker’s quality tends to disappear with time. At the early stages, firms assess workers on the basis of easily observable variables that are correlated with their unobservable quality. As a worker gains experience in the labor market, employers weigh these characteristic 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 information should become less important for earnings as a worker reveals his true productivity with time.

We use data from *Futuro Laboral* of the Chilean Ministry of Education to test if 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 college graduate workers from Chile in their first years in the labor market, the period in which most of the employer learning happens (Lange, 2007). Second, the data presents information on labor market outcomes such as earnings from administrative data and we can identify workers that graduate from elite universities. Finally, the data contains information on the scores of the centralized admission test to universities in Chile. This information will be useful for providing both a measure of a worker’s ability not easily observed by firms and as a running variable in the regression discontinuity test we propose.

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 firms statistically discriminate among workers on the basis of college prestige, then as firms learn about productivity, the coefficients on college prestige should fall, and the coefficients on hard-to-observe ability measures should rise with experience. We use information on math and reading scores of the admission to university test as the measure of a worker’s ability not easily observed by firms. We present some further evidence that math and reading scores are good measures of hard-to-observe

---

2 Other important papers in this literature include Lange (2007), Schönb erg (2007), Arcidiacono et al. (2010), and Mansour (2012).

3 Kaufmann et al. (2012) and Hastings et al. (2013) are two recent papers that have also explored the regression discontinuities generated by the centralized admission process to universities in Chile. Kaufmann et al. (2012) looks at effect of graduating from a elite university on marriage outcomes and Hastings et al. (2013) studies labor market returns to college admission. None of these papers explore how the selective university wage premium changes throughout a worker’s career, that is the main object of interest of this paper.
correlates of productivity.

In addition to the traditional EL-SD test, we take advantage of the centralized admission process to college in Chile to propose a statistical discrimination test based on regression discontinuity design (RD). Using information on 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 an EL-SD test that compares the earnings’ dynamics between these two group of workers as they gain experience in the labor market. The test assumption is that works below and above the admission thresholds are similar in terms of their pre-college productivity correlates.

The RD statistical discrimination test predicts that if firms use university prestige to statistic 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 group of workers should shrink as individuals progress in their career. Different than the traditional EL-SD test, the estimation of the regression discontinuity test can be interpreted as the casual effect of attending a prestigious university on earnings for those around the admission cut-off. We discuss that the regression discontinuity test is not biased by other individual’s characteristics that might be used by employers for statistical discrimination, such as family social-economic status.

We find evidence that firms statistically discrimination using both the traditional EL-SD and regression discontinuity test. Following the traditional EL-SD test we estimate that college graduates from the two most selective universities from Chile earn on average 26% after graduation. However, we find that this wage differential decreases by 1.6 percentage points by year of experience. We also estimate that the Math component of admission test, our measure of ability not observed by firms, increases on 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 firms use selectivity of universities to statistically discriminate workers.

The regression discontinuity design approach provides further evidence for statistical discrimination. We estimate a 19% wage premium for recent graduates of the two most prestigious university
in Chile, but this wage premium decreases by 3 percentage points per year of experience, to the point that we cannot reject a zero earnings differential between these two groups of workers 5 years after their graduation. We interpret the difference between the traditional EL-SD test and the RD test as evidence that there individual characteristics used for statistical discrimination by the firms that are correlated with university selectivity but are not available in the data.

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 an extensive literature that analyzes the use of race, gender, and schooling, we are one of the first papers to study whether firms use prestige of college to statistically discriminate workers. To the best of our knowledge, Lang and Siniver (2011) and Hershbein (2013) are the two other papers that have addressed this issue. Lang and Sniver have a similar approach to estimate how returns to attending a elite university in Israel changes with labor market experience. However, the authors are unable to properly exploit the regression discontinuity in the college admission process.

Furthermore, to the best of our knowledge, our paper is the first to propose an employer learning-statistical discrimination test based on a regression discontinuity design. We demonstrate that the traditional test proposed by Altonji and Pierret (2001) is biased if employers statistically discriminate workers on the basis of characteristics that are not present in the data and are correlated with graduating from a prestigious university, such as family social-economic background.

Second, we contribute to the literature which 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, Hoxby, 1998, Dale and Krueger, 2002, and Black and Smith, 2006), including papers that have used a regression discontinuity design (Saavedra, 2008 and Hoekstra, 2009). The overall finding is that there is a positive effect of graduating from a prestigious college on earnings.4 While there is big effort in

---

4The only exception is Dale and Krueger (2002) who find no wage premium from attending a selective college. It is interesting to note that the authors estimates the 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 advanced age does not contradict the empirical findings of this paper.
the literature to overcome the selection bias associated to attending a prestigious university, little attention has been given to the mechanisms that generate the college selectivity wage premium.\textsuperscript{5}

Different from past work, in this paper we shed some light on the reasons for 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 to employers an unobservable inherent ability of a worker. In this context, the extra value added from a selective college education might not be significantly higher than that from a less prestigious university. Our finding of a rapid decrease in the elite college premium for workers with similar pre-university characteristics is evidence that signaling mechanisms are stronger than productivity mechanisms. In particular, the fact that we cannot reject a wage differential between workers just above and below the admission cutoff after 4 years in the labor market suggests that the value added from the two most prestigious university in Chile is not significantly different from the less prestigious schools.

2 Institutional Framework

Higher education in Chile comprises three types of institutions: Universities, Professional Institutes (IPs), and Technical Formation Centers (CFTs). 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 are in charge of granting professional degrees other than those awarded by universities, and they are also authorized to grant higher education technical degrees in areas where this is required. Technical Formation Centers are intended to equip higher level technicians with the competencies and skills needed to respond to

\textsuperscript{5}Pop-Eleches and Urquiola (2013) is one of the few papers that have discussed the benefits from attending a higher quality school. Their paper address the behavior effect of students, parents and teachers in response to a student admission to better secondary school. Nevertheless, the paper has little to say about the impact of attending a higher quality school on labor market outcomes.
the needs of industry in the public and private sectors.

Universities 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 that derived from the old universities (created after 1980). Traditional establishments consist of 25 fully autonomous universities coordinated by the Council of Chancellors of Chilean Universities (CRUCH) and are eligible to obtain partial funding from the state. They employ a single admission process: the University Selection Test (PAA)\textsuperscript{6}. This test is made up of three compulsory sub-tests including language, mathematics, and history and geography of Chile. Additionally, depending on which programs they are planning to apply to, students may be required to take the following specific PAA tests: advanced mathematics, physics, chemistry, biology, and history.

The time-line of the admission process into traditional universities happens as described in figure 1. First, students take the PAA test and after receiving their score they make their application choices. Students apply to a major and university (or program) simultaneously and can only apply to 8 programs, ranking them up by preferences. The only criterion used for admission in the traditional universities is the score in the PAA. This final admission scores consists of a weighted average of the compulsory and major specific tests and high school GPA, with each program setting its specific PAA weights.\textsuperscript{7} 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 and it is not known before the application decisions and therefore students cannot manipulate which side of the cutoff on they fall on.\textsuperscript{8} Non-traditional universities were created after 1981, have no state financial support and might not necessarily use the PAA score to select their incoming students. Nevertheless, the anecdotal evidence is that the majority of students willing to attend

\textsuperscript{6}In 2004 the university selection test was modified and it is now called PSU.

\textsuperscript{7}For example, engineering in a prestigious university requires 20% of mathematics, 10% of language, 10% of history, 20% high school GPA, 30% specifics mathematics, and 10% physics. The final score to the same major in a different university might requires different weights.

\textsuperscript{8}Students could use the admission score cutoff of previous years as a reference. Given the variation of the admission cutoff overtime and the possibility to apply to 8 different programs, we believe that students with marginal scores to be admitted in prestigious university tend to to apply to these competitive programs.
higher education in Chile take the PAA at the end of high school independent of the university they are planning to attend. The test is relatively inexpensive and administrated throughout the country.

All higher education institutions charge tuition and fees. However, for those students enrolled in one of the traditional universities, solidarity credits and scholarships are available. Competition in these markets, particularly for undergraduates, is often geographically circumscribed to local and regional markets, and it can be more or less intense depending on the institution. As of 2001, the Chilean higher education system consisted of 60 universities (25 traditional universities and 35 new private universities without direct public subsidy), 42 professional institutes (all of them private), and 117 private technical formation centers.

3 Data

The data to be used in the study comes from Futuro Laboral, a project of the Ministry of Education of Chile that follows individuals over the first years of labor market experience after graduating from higher education programs. The panel data set matches tax returns with transcripts of students’ majors and the institutions they graduated from. The unit of analysis concerns only those who graduate from both traditional and non-traditional universities; those who have stopped studying or did not continue their studies after graduating from high school 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.\textsuperscript{9}

The information provided by the Internal Revenue Service (SII) comprises age, sex, name of the institution that individuals graduated from, major, the year of graduation, annual income reported in tax returns, city or cities of employment, number of employers and economic sector. The raw data contains every worker in Chile that had positive earnings between 1996 and 2005, even those

\textsuperscript{9}Note that the cohorts are observed for different length of time. For example, while we observe 10 years of labor market experience for the 1995 graduation class, we only 4 years of labor market experience for the 2001. Unfortunately, the project was deactivated and the income data for more recent years was not not collected.
who exempt from tax.\footnote{Note that in Chile, married couples must fill their taxes separately.}  A concern is that part of the individual from prestigious universities might go to graduate school after finishing their baccalaureate studies and therefore would be omitted in the earnings sample. However, the fraction of workers that 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 years old individuals with a bachelor degree were enrolled in graduate school or had obtained a graduate degree. 

\footnote{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 week of work. Nevertheless, workers with a bachelor 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 individuals with a bachelor degree work are employed in the period of the interview and from those, 88% work more than 35 hours per week.}

For a random sample, the Ministry of Education gathers information about the PAA score, high school grades and the institutions students graduated from high school. As the PAA scores have an important role in regression discontinuity analysis, we restrict our study to this sample.

The wage measured in the sample is the annual income that comes from jobs and services provided by the individual and does not include self-employment income.\footnote{A concern is that part of the individual from prestigious universities might go to graduate school after finishing their baccalaureate studies and therefore would be omitted in the earnings sample. However, the fraction of workers that 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 years old individuals with a bachelor degree were enrolled in graduate school or had obtained a graduate degree.} We use consumer price index (IPC) as a deflator to compute real wages. The experience variable is computed as the number of years an individual has income and has paid taxes after graduation. The final sample consists of 58,616 individuals and 348,531 observations.

We divide universities into two groups: selective and non-selective universities. The selective universities comprises two of the oldest and most prestigious universities in the country. These 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. Due to a confidentiality agreement with the Ministry of Education, we cannot provide the name of these two institutions.

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

4 Employer Learning Statistical Discrimination Model

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

$$y_{it} = r_{si} + \alpha_1 q_i + \lambda z_i + \eta_i + H(t)$$

(1)

where $s_i$ captures information that is available to both employers and researchers. In this paper, $s_i$ is defined as an indicator if a worker graduated from a prestigious university or not. The variable $q_i$ describes information available to employers and not present in the data, such as family social economic background, $z_i$ is a characteristic present in the data but not available to employers and $\eta_i$ 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 and does not depend on the other variables of the model.

In the absence of information on $z_i$ and $\eta_i$, employers form expectations based on other observed characteristics of workers. Altonji and Pierret (2001) assume that these conditional expectations are linear on $s$ and $q$:13

$$z = E[z|s, q] + v = \gamma_1 q_i + \gamma_2 s + v$$

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

where $v$ and $e$ are scalar with mean zero and uncorrelated with $s$ and $q$ by construction. Under this assumption, one can characterize the expected value of $y$ given information on $s$ and $q$:

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

---

13 A normalization allows suppressing $q$ in the second expectation. We omit the subscript $i$ henceforth.
In the traditional EL-SD model, employers have access to a noisy measure of a worker’s productivity after each period that an individual spend in the labor market:

\[ \tilde{y}_{it} = y_{it} + \varepsilon_{it} \]

where the noise \( \varepsilon_{it} \) is independent of all the variables of 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.

\[ W_{it} = \mathbb{E} [\exp(y_{it}) | s, q, \tilde{y}_{i0}, ..., \tilde{y}_{i(t-1)}] \]

Lange (2007) assumes that \( \varepsilon_{it} \) is independently, identically and normally distributed with a finite variance. Under this assumption, the process of updating the expectations of employers have a very simple structure an the log-wage process can be represent by:

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

where \( \tilde{H}(t) \) is a linear transformation of \( H(t) \) and \( \theta_t \) is a function of the variances of and \( \varepsilon_{it\tau}, s \) and \( q \). Furthermore \( \theta_0 = 0 \) and \( \theta_t \) strictly increases with \( t \) converging to 1 as \( t \) goes to infinite.\(^{14}\) This expression demonstrates that as a worker progress in his career, employer weight less their initial believe on a worker’s productivity based on \( s \) and \( q \), and weight more the new information that becomes available during a 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 \( t \).

\[ \mathbb{E}^* [w_{it} | s, z, x] = b_{sz} s + b_{zz} z + \tilde{H}(t) \]

\(^{14}\)See Lange (2007) for the formal derivations of these parameters.
Without lost of generality, one can define the the projections of the unobservable variables \((q, \eta)\) on the observable variables \((s, z)\):

\[
q = \gamma_3 s + \gamma_4 z + u_1
\]

\[
\eta = \gamma_5 s + \gamma_6 z + u_2
\]

Using the independence of \(\varepsilon_{it}\) to all the variables of the model, Lange (2007) show that the coefficients of the projections:

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

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

where, as discuss before, \(\theta_0 = 0\) and \(\lim_{t \to \infty} \theta_t = 1\). The traditional EL-SD test consists in estimating how \(b_{st}\) and \(b_{zt}\) change with experience level \(t\). Indeed, Altonji and Pierret (2001) propose that if firms 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 \(b_{zt}\) should rise with \(t\).

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

\[
b_{s0} = r_A + \alpha_1 \gamma_3 + \alpha_2 + \lambda(\gamma_2 + \gamma_1 \gamma_3) \tag{5}
\]

\[
b_{z0} = (\alpha_1 + \lambda \gamma_1) \gamma_4 \tag{6}
\]

where the coefficient \(b_{s0}\) represents the relation between graduating from a prestigious university and wages in the beginning of a workers career. The first term \(A\) captures the direct effect of attending a prestigious university on productivity. The second 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 relation between any variable that affects wages, is correlated to graduating from a prestigious university and it is not present in the data. Finally, the term $C$ reflects the fact that employers do not observe $\eta$ and $z$ in the beginning of a worker’s career, but are aware of their relation with $s$. Therefore, employer 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 in 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.

$$
\begin{align*}
  b_{s\infty} &= \frac{r}{E} + \alpha_1 \gamma_3 + \gamma_5 \\
  b_{z\infty} &= \frac{\lambda}{G} + \alpha_1 \gamma_4 + \gamma_6
\end{align*}
$$

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 $\eta$ and $q$ have an impact on long-run wages, 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 firms only learn $\eta$ 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 $\eta$ and $q$.

One important issue that has been not discussed in 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 statistical discrimination test. This issue arises if firms statistically discriminate workers on the basis of variables that are not observed in the data, such as family social economic background, that are correlated to graduating from prestigious university. In this situation, 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 social
economic status as a signal of a worker’s unobservable characteristics.

For instance, we analyze the extreme case where \( s \) is not correlated to \( \eta \) and \( z \) (\( \alpha_2 = 0, \gamma_2 = 0 \) and \( \gamma_5 = 0 \)). In this situation, employers should 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 \( \eta \) and \( z \) (\( \gamma_4 \neq 0 \)), and therefore \( q \) is used by employer 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 \( z \), which is capture by the term \( \lambda \gamma_1 \gamma_3 \) in equation (6).

### 4.2 Regression Discontinuity EL-SD Test

The object of interest of the EL-SD 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 \( \text{Dist.Cutoff}_i \) 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 graduate from this university, such that \( s_i = 1 \) if \( \text{Dist.Cutoff}_i \geq 0 \) and \( s_i = 0 \) otherwise.\(^{15}\)

The parameter of interest in the paper is:

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

(9)

that represents local average difference of log-wages by experience levels at the admission cutoff. The employer learning statistical discrimination RD test consists in testing if \( \tau_t \) decreases with \( t \).

Note that by definition, we have that:

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

\(^{15}\)As it will be clear later, this assumptions is not confirmed in the data because some students admitted to a prestigious university decide to attend a less prestigious university (fuzzy regression discontinuity). For simplicity, we ignore this possibility here.
Furthermore, the distribution of the other variables of the model \( \{ z_i, q_i, \eta_i, \varepsilon_{itr} \} \) is continuous around the admission cut-offs. In this case, the expected values of these variables just above and just below the admission cutoff are the same:

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

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

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

\[
= \underbrace{r + (1 - \theta_t)(\alpha_2 + \lambda \gamma_2)}_{L}
\]

where \( \theta_t \) is defined in the same way as above. The regression discontinuity effect of graduating from a prestigious university on wages at experience level \( x \) is composed by two terms. The first term \( I \) represents the direct effect of \( s \) on the workers productivity. The second term \( L \) represents the fact that employers do not observe \( \eta \) and \( z \) and use \( s \) as a signal for these two variables. In other words, if firms 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, \( \tau_t \) decreases with \( t \) and converges to \( r \) as \( \theta_t \) goes to 1.

There are is an important difference between the regression discontinuity test we propose and the traditional employer learning test: the parameter \( \tau_t \) does not depend on the relation between \( s \) and \( q \). In other words, the regression discontinuity test is robust to the existence of characteristics that could be used for statistic 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 \( \tau_t \) converges to the \( r \) as \( t \) increases, which represents the actual effect of graduating from a prestigious university on 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 math and reading components of the PAA as a measure of ability correlates not easily observed by firms. We have a couple 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 capacity in the admission process to traditional universities. Second, there is evidence that employers do not have access to the PAA score at the moment of setting 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 the PAA score in the resume 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. Therefore, the absolute value of the PAA score for a single worker might be not very informative to a firm.

We present the estimation of equation (1) in table 3. All the standard errors are clustered at the individual level and we use White-Huber corrections for possible heteroscedasticity. Actual 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 selective university dummy are excluded. It turns out that going to a traditional, prestigious and very selective university is associated with higher earnings. In exact terms, it increases log wage in 0.207. Notice also that our proxy for innate ability, PAA has a positive and statistically significant effect on earnings. An increase in one standard deviation in the language PAA test increases wages in 3.8%, whereas an increase in one
standard deviation in the math test rises wages in 7%.

In column 2 we introduce the interaction between the selective university dummy and experience. If university selectivity provides a signal of workers ability to employers, we should expect that 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 workers gains experience if employers learn about a worker’s true ability across time. This is in fact the finding of the equation presented in column 2, where we estimate a coefficient of -0.005 for the interaction between the selective university and experience.\(^\text{16}\)

The important result of table 3 comes from column 3 where we add PAA scores interacted with experience to capture the idea of statistical discrimination and employer learning. The coefficient on selective universities is 0.261, which is large and statistically significant. That is, the premium to a selective university on earnings with respect to a less selective one at the beginning of the worker’s career is 26% in term of log wages, but it decreases with experience. Therefore, the effect of attending the most selective and oldest universities on earnings decreases over time supporting the theory of the EL-SD model rather than human capital mechanisms. The reason is that human capital models and on-the-job training empirical evidence suggest that education and ability make workers most likely to be trained and that more educated and more able workers receive more training. If this is the case, then we expect that the effect of going to college and the PAA score on wages would both increase over time.

Finally, the coefficient of 0.011 on standard math PAA interacted with experience suggests that the effect of a shift in standard math PAA score changes significantly as worker accumulates 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 and the hiring process at the beginning of a worker’s career. Therefore there is not much learning to happen regarding this

\(^{16}\)Both Farber and Gibbons (1996) and Altonji and Pierret (2001) estimate insignificant effects of interaction between schooling and experience using the the same specification. Our estimates are small but marginally significant what might be justified by the grater number of observations we have.
attribute with experience.

Figure 4 presents further 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 8% after the 8-th year of labor market experience. One could interpret this trend as that it takes about 8 years for employers to learn a worker’s ability and the wage premium remaining after that represents the permanent productivity differences between those that attended a selective university in Chile. Note however that, as presented in equation 7, this coefficient represents both the direct effect of graduating from a prestigious university and the indirect effect of $\eta$ and $q$ on wages.

We present a robustness check for the evidence that firms statistically discriminate workers on the bases of university selectivity. In table 4, rather than using an indicator that a workers graduated from one of the two most prestigious universities from Chile, we use a continuous measure of prestige for all colleges. Precisely, we assign to each university its quality score as presented in the “Que Passa” college ranking, one of the most widely recognized in the country, for 2011. In this framework, we estimate how earnings vary with this score, defined as university quality index in the table, its interactions with experience and the remaining controls. We loose 5,977 observations of those individuals whose colleges do present an available “Que passa” 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 evidence for statistical discrimination based on college prestige, we use a regression discontinuity (RD) design. The test consists on comparing how the earnings of those just above and just below the cutoff for admission to the most selective universities in Chile change as workers accumulate experience in the labor market. 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.\textsuperscript{17} As discussed in section 4, different than the traditional EL-SD test presented in the past section, the estimation of the regression discontinuity test can be interpreted as causal effect for those individuals around the admission cutoff and is not biased by other factors observed by firms 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, the college he or she graduated from and the major. We do not observe application decisions and therefore have to make extra assumptions and sample restrictions to perform the regression discontinuity design. 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 least 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 highest admission cutoffs and therefore should be top choices of applicants. Second, there is a positive wage

\textsuperscript{17}Students can retake the test next year, but they cannot retake the test after they got their results the same year, which decreases the probability of manipulation.
differential between workers with the competitive majors in less prestigious university and workers with less competitive major in prestigious university. We interpret this as evidence that students have incentives to study engineering, business, medical or law degree at a less prestigious rather than other major in a prestigious university. 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.\(^{18}\) 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 in two, one or neither of the prestigious universities, we define the running variable used in the RD as follows:

\[
\text{Dist.Cutoff}_i = \max\{Univ1.Score_i - Univ1.Cutoff_i, Univ2.Score_i - Univ2.Cutoff_i\}
\]

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. 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 reject by both schools.\(^{19}\)

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

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

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

where \(g_i\) is an indicator if worker \(i\) graduated from an elite university, \(t\) measures years of experience in the labor market, and \(w_{it}\) is the \(\log(\text{wages})\) after \(t\) years of experience. Note that the parameter \(\tau_t\) represents the local average treatment effect on earnings after \(t\) years of experience.

\(^{18}\)We were only able to obtain PAA weights for years starting in the year 2000. In order to construct final scores for individuals that too the PAA prior to 2000, we assume that programs used the same weights for previous years. The evidence is that programs do not change weights over time.

\(^{19}\)Information on program admission cutoffs were collected at the universities websites (late application years) and newspapers (early application years). 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 weights used in the paper or transfers from less prestigious universities. We drop these individuals from the sample used in the RD analysis.
for workers around the admission cutoffs that would enroll in a prestigious university if they were admitted (intent-to-treat effect).  

The employer learning-statistical discrimination RD test we propose consists of estimating if $\tau_t$ decreases with $t$. The test is based on the assumption that the unobserved ability ($\eta_i$) is positively correlated to graduating from a selective university but is continuous around the admission cutoff. In this framework, assuming that firms 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 $\eta_i$. However, the wage differential between those above and below the cutoff should decline if firms learn the true distribution of $\eta_i$ as workers gain experience and therefore should rely less on college prestige to set wages.

5.3 Results

We first address the empirical question if 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 graduation from a prestigious university is approximately 60 percentage. This means that around 60% of the individuals with PAA scores just sufficiently high 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, which is 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 cut-off is random, we should observe a zero treatment effect on the 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 dependent variable, we cannot reject at reasonable levels of significance that there are zero effects of being above the cut-off on these per-treatment outcomes.\footnote{Due to space constraint we omit the tests here, but they are available under request.}

In order to present evidence of the effects of admission to a selective university on earning, we plot in figure 8 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 tend to decrease to the point that there is no apparent difference in terms of earnings between workers just above and just below the cutoffs four years after graduation. In addition to that, 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 specification
that allows that the return to being approved at a selective university to change along a worker’s career. Under this specification, we estimate a 10%-14% of 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 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 casual 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 a robustness checks for the main RD findings, we present in table 6 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 Conclusion

This paper tests whether firms statistically discriminate based on the selectivity of the university attended by workers. We first follow the employer learning statistical discrimination test suggested by Altonji and Pierret (2001) and show that the returns to graduating from an elite university in Chile decreases with experience and that the returns to hard-to-observe ability correlates increase with experience. We discuss that traditional employer learning statistical discrimination test might be biased by if employers use individual’s characteristics that are not available in the data, such as family social-economic status, to predict a worker’s unobservable ability.

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 consist in comparing the earning of those just below and just above the admission cutoff to the tho most prestigious universities in Chile. We show that recent graduates just above the admission cutoff have significantly higher earnings than those just below the cutoff. However, as workers gain labor market experience, the earnings gap between these two groups decreases to the point that we cannot reject zero wage differentials 5 years after graduation. We interpret this result as firms paying workers in accordance with the selectivity of their college when they graduate from school, but rewarding them based on their true productivity as they reveal their quality to employers.

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 firms 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. In addition, the fact that we cannot find any significant difference on earnings between workers that are similar in terms of their pre-college characteristics 5 years after graduation suggests that the value added from a prestigious to an individual worker productivity is not significant higher than the value added from second tier universities.
References


25


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.


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.


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

<table>
<thead>
<tr>
<th>Variables</th>
<th>Selective Universities</th>
<th>Non-selective Universities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Female</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>Language PAA Score</td>
<td>680.6</td>
<td>61.3</td>
</tr>
<tr>
<td>Math PAA Score</td>
<td>715.9</td>
<td>68.7</td>
</tr>
<tr>
<td>High School Grade</td>
<td>644.5</td>
<td>78.7</td>
</tr>
<tr>
<td>Private High School</td>
<td>0.11</td>
<td>0.32</td>
</tr>
<tr>
<td>Number of Individuals</td>
<td>11,615</td>
<td></td>
</tr>
</tbody>
</table>

Note: Math and Language PAA scores are components of the centralized test for admission in University in Chile. See section 3 for definition of selective university.

Table 2: Earnings for Selective and Non-Selective Universities

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Selective Universities</th>
<th>Non-selective Universities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of Annual Wage (in 1999 Pesos)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>15.46</td>
<td>15.07</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.18</td>
<td>1.25</td>
</tr>
<tr>
<td>Observations</td>
<td>69,597</td>
<td>278,934</td>
</tr>
</tbody>
</table>

Note: See section 3 for definition of selective university.
Table 3: Traditional EL-SD Regression

<table>
<thead>
<tr>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective University</td>
<td>0.207</td>
<td>0.225</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td>(0.009)***</td>
<td>(0.012)***</td>
<td>(0.013)***</td>
</tr>
<tr>
<td>Selective University x Experience</td>
<td>-0.0050</td>
<td>-0.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)***</td>
<td>(0.003)***</td>
<td></td>
</tr>
<tr>
<td>PAA Language</td>
<td>0.038</td>
<td>0.038</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.004)***</td>
<td>(0.004)***</td>
<td>(0.006)***</td>
</tr>
<tr>
<td>PAA Math</td>
<td>0.070</td>
<td>0.070</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.005)***</td>
<td>(0.005)***</td>
<td>(0.007)***</td>
</tr>
<tr>
<td>PAA Language x Experience</td>
<td></td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>PAA Math x Experience</td>
<td></td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)***</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.463</td>
<td>12.457</td>
<td>12.684</td>
</tr>
<tr>
<td></td>
<td>(0.110)***</td>
<td>(0.110)***</td>
<td>(0.113)***</td>
</tr>
<tr>
<td>Observations</td>
<td>348,531</td>
<td>348,531</td>
<td>348,531</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.274</td>
<td>0.274</td>
<td>0.275</td>
</tr>
</tbody>
</table>

**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. See section 3 for definition of selective university.
Table 4: Traditional EL-SD Regression - University Quality Index

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>University Quality Index</td>
<td>0.092</td>
<td>0.094</td>
<td>0.115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)***</td>
<td>(0.005)***</td>
<td>(0.006)***</td>
<td></td>
</tr>
<tr>
<td>University Quality Index x Experience</td>
<td>-0.001</td>
<td>-0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAA Language</td>
<td>0.034</td>
<td>0.034</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)***</td>
<td>(0.004)***</td>
<td>(0.006)***</td>
<td></td>
</tr>
<tr>
<td>PAA Math</td>
<td>0.053</td>
<td>0.052</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)***</td>
<td>(0.005)***</td>
<td>(0.007)**</td>
<td></td>
</tr>
<tr>
<td>PAA Language x Experience</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAA Math x Experience</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>12.018</td>
<td>12.004</td>
<td>12.101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.110)***</td>
<td>(0.113)***</td>
<td>(0.114)***</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>342,554</td>
<td>342,554</td>
<td>342,554</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.275</td>
<td>0.275</td>
<td>0.275</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Dependent Variable: Log Annual Wage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Reduced Form</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Bandwidth (Points from Cutoff)</strong></td>
</tr>
<tr>
<td>125</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>Approved at Selective University</td>
</tr>
<tr>
<td>0.081</td>
</tr>
<tr>
<td>(0.0287)** (0.0308)** (0.0342)*</td>
</tr>
<tr>
<td>Approved at Selective Univ.* Experience</td>
</tr>
<tr>
<td>-0.027</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>39,748</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>0.135</td>
</tr>
<tr>
<td><strong>Panel B</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2 Stages Least Square</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Bandwidth (Points from Cutoff)</strong></td>
</tr>
<tr>
<td>125</td>
</tr>
<tr>
<td>Model</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>Graduated from Selective University</td>
</tr>
<tr>
<td>0.133</td>
</tr>
<tr>
<td>(0.0470)** (0.0512)** (0.0576)*</td>
</tr>
<tr>
<td>Graduated from Selective Univ.* Experience</td>
</tr>
<tr>
<td>-0.037</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>39,748</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>0.140</td>
</tr>
</tbody>
</table>

**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 instrumented 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 degree (see section 5 for details).
Table 6: **EL-SD Regression Discontinuity Test - Robustness Checks**

<table>
<thead>
<tr>
<th>Regression Specification</th>
<th>Additional Controls</th>
<th>Function of Points from the Cutoff</th>
<th>Flexible Coefficient?</th>
<th>Sample</th>
<th>Estimated Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Approved at Selective University</td>
</tr>
<tr>
<td>(1) No</td>
<td>Linear</td>
<td>Yes</td>
<td>All</td>
<td></td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0354)**</td>
</tr>
<tr>
<td>(2) Yes</td>
<td>Cubic</td>
<td>No</td>
<td>All</td>
<td></td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0426)**</td>
</tr>
<tr>
<td>(3) Yes</td>
<td>Cubic</td>
<td>Yes</td>
<td>All</td>
<td></td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0425)**</td>
</tr>
<tr>
<td>(4) Yes</td>
<td>Linear</td>
<td>Yes</td>
<td>Males</td>
<td></td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0427)**</td>
</tr>
<tr>
<td>(5) Yes</td>
<td>Linear</td>
<td>Yes</td>
<td>Females</td>
<td></td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.0599)**</td>
</tr>
</tbody>
</table>

All specifications include Cubic Experience Polynomial.

**Approved at Selective Univ.**: Points from the Cutoff $\leq 0$

**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 degree (see section 5 for details).
Figure 1: Application Process to Traditional Universities

Taking the PAA Test → Result of the PAA Test → Application to Programs → Results and Enrollment

1st Week of Dec. → 1st Week of Jan. → 2nd Week of Jan. → 3rd Week of Jan.

Figure 2: Smoothed Language PAA Score Distribution

Note: Language PAA is a component of the centralized test for admission to university in Chile. See section 3 for definition of selective university.
Figure 3: Smoothed Math PAA Score Distribution

Note: Math PAA is a component of the centralized test for admission to university in Chile. See section 3 for definition of selective university.
Figure 4: 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 Tables 3 (including Math and Language PAA scores). 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

- Earnings - 1st Year after Graduation
- Earnings - 2 years after Graduation
- Earnings - 3 years after Graduation
- Earnings - 4 years after Graduation
- Earnings - 5 or more years after Graduation

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 degree (see section 5 for details).