WPS7168

Policy Research Working Paper 7168

Broken Gears

The Value Added of Higher Education on Teachers' Academic Achievement

Carlos Felipe Balcázar
Hugo Ñopo

WORLD BANK GROUP
Poverty Global Practice Group
January 2015

Policy Research Working Paper 7168

Abstract

Good teachers are essential for high-quality educational systems. However, little is known about teachers' skills formation during college. By combining two standardized tests for Colombian students, one taken at the end of senior year in high school and the other when students are near graduation from college, this paper documents the extent to which education majors relatively improve or deteriorate their skills in quantitative reasoning, native language, and foreign language, in comparison to students in other programs. Teachers' skills vis-à-vis those in other majors deteriorate in quantitative reasoning and foreign language, although they deteriorate less for those in math-oriented and foreign language-oriented programs. For native language, there is no evidence of robust differences in relative learning mobility.

This paper is a product of the Poverty Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at cbalcazarsalazar@worldbank.org.

Broken Gears: The Value Added of Higher Education on Teachers' Academic Achievement

Carlos Felipe Balcázar
Poverty Global Practice
The World Bank Group
1818 H St. NW, Washington D.C.
United States of America

Hugo Ñopo[*]
Education Division
Inter-American Development Bank
1300 New York Ave. NW, Washington D.C.
United States of America

JEL Codes: I2, I21, J24

Key words: teacher performance, career choice, self-selection, relative learning mobility

[*] The comments of Mariana Alfonso, Leonardo Bonilla, Fernando Fernández, Diana Hincapie, Catherine Rodríguez and Emiliana Vegas are specially acknowledged. Daniel Alonso provided research assistance.

1. Introduction

This paper builds upon two literatures of high policy relevance: *teacher quality* and *value added of higher education.* It does so by providing evidence for Colombia on the difference in basic academic competencies between students in education majors and students in other majors. To do so, we compare test score-rankings between these two groups of students using a novel panel data set that combines two standardized tests, one before and one near the end of college studies. We measure test score-rankings using the z-scores of subjects assessed in both tests and compare changes in the relative position of students in the distributions of academic skills. We tackle the problem of endogeneity in the decision of becoming a teacher using instrumental variables and tackle the problem of differences in the distribution of observable characteristics between education majors and their peers using non-parametric matching.

A growing body of literature establishes that teachers are essential for a high-quality educational system (Goldhaber and Brewer, 1997; Hanushek, 2002; Eide et al., 2004; Rockoff, 2004; Rivkin et al., 2005; Hanushek and Rivkin, 2006; Clotfelter et. al., 2007; Kukla-Acevedo, 2009; De Paola, 2009). More recently, Chetty et al. (2013a and 2013b) find that, after controlling for socio-demographic characteristics, those students who had high-quality teachers during their school years are more likely to attend higher education, to go to a better university and to obtain higher labor earnings. Nonetheless, although the literature has extensively addressed the role of the quality of teachers on social and academic outcomes, we know little about the role of higher education on the production of those teachers.

Evidence from successful educational systems (e.g., Finland, Singapore, the Republic of Korea and China) highlights the importance of teachers and their formation. In Finland, for instance, only the best and brightest manage to become teachers after demanding undergrad and master's programs (OECD, 2011; Sahlberg, 2011). This has been claimed as an important stepping-stone for their impressive improvement between the 1960s and the early 21st century. From having educational outcomes comparable to those of developing economies, Finland now ranks among the top performers in educational achievement and attainment in the world (Sahlberg, 2009, 2011).

In most educational systems, however, the best and brightest do not choose education majors. Teachers' academic performance and potential is below than that of their peers, who choose to attend other programs (Giesen and Gold, 1993; Hanushek and Pace, 1995; Podgursky et al., 2004; Denzler and Wolter, 2008). The absence of competitive salaries and benefits, as well as the absence of a meaningful and challenging career path, might discourage talented people from choosing a career in education (Ballou and Podgursky, 1997; Odden and Kelley, 1997; Dolton, 1990, 2006; Dolton and van der Klaauw 1999; Dolton 2006; Mizala and Ñopo, 2014). In Colombia the situation is no different, teachers fare at the bottom of the distribution of standardized test scores (Barón and Bonilla, 2011; Barón et al., 2013). Talented people might get discouraged from choosing a career in education to avoid getting trapped in a profession with low social prestige, bad academic reputation and low salaries (García et al 2014).

There is solid evidence that self-selection plays a role on teachers' quality. Granted, but, what about the role of higher education? Even if prospective teachers start at disadvantage compared to their peers in other careers, does higher education level the playing field? To address these questions we analyze *relative learning mobility*, understood as the change in score-rankings between two periods, for students who are close to graduating from a university program in education (which we also call throughout the paper *teachers*) vis-à-vis students who are also about to graduate, but from other university academic programs (which we also call throughout the paper *other professionals*). We use data from two standardized tests: Saber 11 and Saber Pro, which provide information on educational outcomes and socio-demographic characteristics for students in Colombia at their senior year of high school and near the end of their college education.

Our results indicate that after around 5 years of academic training, teachers' skills vis-à-vis those in other professions relatively deteriorate (or do not improve as much) in quantitative reasoning and foreign language, although they deteriorate less for those in math-oriented and foreign language-oriented programs. We do not find evidence of statistically significant differences in relative learning mobility for native language (Spanish). In this sense this paper deals with teachers' academic competencies and also contributes to the *value added of higher education* literature (Klein et al. 2005; Liu 2011; Saavedra 2009; Saavedra and Saavedra 2011; Cunha and Miller 2014).

The rest of the paper proceeds as follows. In the next section we introduce the data sources and present some descriptive statistics. Section three introduces the concept of relative learning mobility. In the fourth section we present the main empirical analysis. In the fifth section we conclude.

2. Data and descriptive statistics

The data come from *Instituto Colombiano para la Evaluación de la Educación* (ICFES). We use data from two national standardized tests: Saber 11 (2002-2007) and Saber Pro (2011). *Saber 11* assesses senior year high school students' academic competencies in language (Spanish), mathematics, biology, chemistry, physics, social sciences, philosophy and foreign language (English). *Saber Pro* assesses the academic competences of higher education students who have completed at least 75% of their academic programs, on program specific areas (e.g., engineers are assessed in engineering; economists in economics; biologists in biology, and so on). The novelty that makes this paper feasible is that starting on the second semester of 2011, Saber Pro also assesses all students, regardless of their major, in quantitative reasoning, reading, writing and English as foreign language.[2]

Saber 11 is mandatory for all senior year high school students who wish to obtain their school degree, and serves as an input for college entrance. In fact, in many universities this is the sole admission criterion (Saavedra and Saavedra, 2011). Saber Pro is mandatory for those students who

[2] See: http://www.icfes.gov.co/examenes/.

wish to earn their college degrees. Furthermore, employers might use Saber Pro scores to screen applicants for a professional job position (Saavedra, 2009). Nonetheless, schools and universities cannot retain students based on their scores.

Saber 11 and Saber Pro share three components, which allow us to compare the “change” in students’ learning: native language, quantitative reasoning and foreign language. “Language” in Saber 11 is comparable to “Reading” in Saber Pro as both evaluate academic competences in Spanish (i.e., reading skills). In this paper we use “Native Language” to refer to both. “Mathematics” in Saber 11 is comparable to “Quantitative Reasoning” in Saber Pro as both evaluate basic competencies in quantitative problem solving. In this paper we use “Quantitative Reasoning” to refer to both. The foreign language component in both tests evaluates proficiency in English. For all three cases, the level of difficulty is higher in Saber Pro, capturing the fact that students should develop their basic academic skills during higher education. We do not compare scores in both tests in a direct way. Instead, we use the individuals’ position on the distribution of scores (z-scores), analyzing the changes from one test to the other.

ICFES provided researchers the possibility of linking both tests using a unique key that identifies individuals. These keys are available for Saber 11 from the first semester of 2002 to 2010, and for Saber Pro from the first semester of 2007 to 2011. However, we do not use data from Saber 11 from the second semester of 2007 on due to changes in the survey structure, and from Saber Pro from the first semester of 2011 backwards given that quantitative reasoning, native language and foreign language were not being assessed.

We restrict ourselves to students attending on-campus programs at their 4th year of college or above, who at the time of the test had completed at least 75% of their academic programs, whose identification key allows tracing them in both Saber 11 and Saber Pro, and with no missing values in the variables of interest. This comprises 63% of the full sample of college students who took Saber Pro in the second semester of 2011 (56% of the education majors and 64% of those in other majors). Those students for which tracing is possible and show no missing values on the variables of interest, that is, those in the 63%, scored above the remaining 37% in the three subjects (on average): 0.26 standard deviations in quantitative reasoning, 0.28 standard deviations in native language, and 0.29 standard deviations in foreign language. The sample of interest is then biased towards higher performing students, but such bias is similar for both education and other majors.

Table 1 shows the sample distribution of the 63% of students by the semester in which they took Saber 11, split by group: teachers (8% of the sample) and other professionals. The bulk of students in the second semester of each year is explained by the fact that most students at the national level (85%) enroll in the *A school calendar*, which goes from February to November. Students in the *B school calendar* (2% of the population of interest), which goes from August/September to June, take the test in the first semester of the year. The remaining 13% corresponds to *flexible calendar* students (sabbatine, fast-tracks, etc.). From these, 77% usually take the test in the second semester of each year and the remaining 23% in the first one.

Table 1. Sample size

Test	Year	Semester	Teachers		Other professionals	
			Number of observations	Percentage (%)	Number of observations	Percentage (%)
Saber 11	2002	I	36	0.87	393	0.84
		II	300	7.22	3038	6.52
	2003	I	55	1.32	585	1.25
		II	485	11.67	4362	9.36
	2004	I	58	1.40	830	1.78
		II	700	16.85	6773	14.53
	2005	I	101	2.43	1310	2.81
		II	958	23.06	10288	22.07
	2006	I	103	2.48	2133	4.58
		II	1278	30.76	14649	31.42
	2007	I	81	1.95	2256	4.84
Saber Pro	**2011**	**II**	**4155**	**100.00**	**46617**	**100.00**

Source: Authors' compilations based on ICFES data.

2.1. Differences in socio-demographic characteristics between teachers and other professionals

Table 2 presents socio-demographic characteristics for students of education majors and students of other programs. The share of females studying to become teachers is higher than in other programs. Also, students in education majors are more likely to come from larger families; they have less educated parents/guardians; they are less likely to migrate to another administrative unit to pursue their studies; they are more likely to have studied in public schools; and also they are more likely to enroll in public universities and in programs with lower tuition fees. This reaffirms the idea that teachers are more likely to come from a disadvantaged socioeconomic position compared to their peers (Denzler and Wolter, 2008; Aksu et al., 2010).

Table 2. Descriptive statistics

a. Characteristics at the student level

Variables	Other professionals	Teachers	Difference between groups
Socio-demographic characteristics (%)			
Gender (Female, as measured by Saber 11 and Saber Pro)	58.10	66.60	***
	(0.23)	(0.73)	
Familiy size (more than 5 persons, as measured by Saber Pro)	14.30	22.30	***
	(0.16)	(0.65)	
Max education of the parents/guardians (as measured by Saber Pro)			
Secondary incomplete or less	19.70	39.70	***
	(0.18)	(0.76)	
Secondary complete or tertiary incomplete	28.10	34.20	***
	(0.21)	(0.74)	
Technical or technician education complete	12.10	10.30	***
	(0.15)	(0.47)	
Universitary education complete	40.00	15.80	***
	(0.23)	(0.57)	
The student moved to another administrative unit for his/her higher education (as measured by a difference in the administrative unit of residence in Saber 11 and Saber Pro)	22.59	14.58	***
	(0.19)	(0.55)	
Semester of study in current university program (as measured by Saber Pro)			
7 or 8	11.20	9.10	***
	(0.15)	(0.45)	
9 or 10	77.70	81.80	***
	(0.19)	(0.6)	
11 or 12	11.10	9.10	***
	(0.15)	(0.45)	

b. Characteristics at the high school level

Variables	Other professionals	Teachers	Difference between groups
High school characteristics (%), as measured by Saber 11 unless otherwise noted			
School administration (public)	48.00	74.80	***
	(0.23)	(0.67)	
School type (mixed gender)	79.1	87.7	***
	(0.19)	(0.51)	
School day			
Complete	32.30	23.90	***
	(0.22)	(0.66)	
Morning	51.70	51.80	
	(0.23)	(0.78)	
Afternoon	14.50	21.10	***
	(0.06)	(0.26)	
Night	1.40	2.90	***
	(0.02)	(0.08)	
Weekend	0.10	0.30	
	(0.16)	(0.63)	
School Calendar (A calendar)†	82.10	87.30	***
	(0.18)	(0.52)	
Degree type (as measured by Saber Pro)			
Academic	78.40	69.40	***
	(0.15)	(0.45)	
Technical	19.10	21.40	***
	(0.19)	(0.6)	
Normal school††	2.60	9.20	***
	(0.18)	(0.64)	

c. Characteristics at the university level

Variables	Other professionals	Teachers	Difference between groups
Universtiy characteristics (%), as measured by Saber Pro			
University administration (public)	34.60	73.80	***
	(0.18)	(0.56)	
University fee (academic semester)†††			
None	1.00	1.40	*
	(0.05)	(0.19)	
Less than 1.000.000 COP	27.30	72.90	***
	(0.21)	(0.69)	
Between 1.000.000 and 3.000.000 COP	30.90	23.20	***
	(0.21)	(0.66)	
Between 3.000.000 and 5.000.000 COP	20.80	2.30	***
	(0.19)	(0.23)	
More than 5.000.000 COP	20.00	0.20	***
	(0.19)	(0.08)	

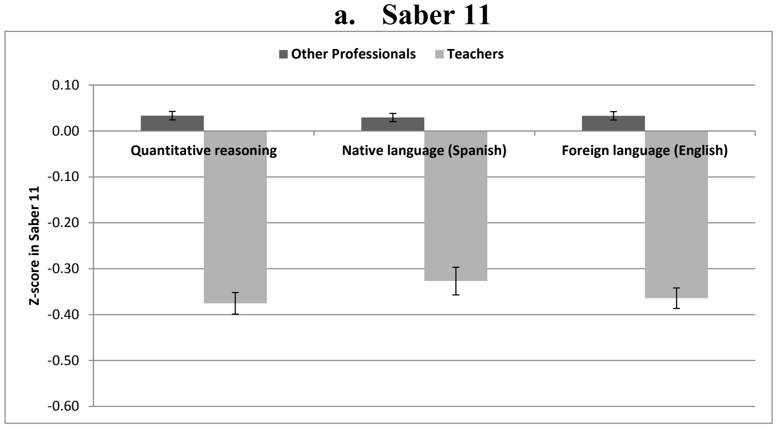
Standard errors in parentheses. * Significant at ten percent; ** significant at five percent; *** significant at one percent.
† Colombia has two regular school calendars: The "A calendar", which goes from February to November; and the "B calendar", which goes from August to June.
†† A "Normal school" is a school that trains high school students to become teachers.
††† COP: Colombian Peso. The average exchange rate for the second semester of 2011 was 1856 COP/USD.
Source: Authors' calculations based on ICFES data.

As mentioned above, the common components in Saber 11 and Saber Pro are not directly comparable. Although they evaluate the same subjects, the scores have different metrics. To perform the comparisons we use standardized test scores (z-scores), analyzing the changes in the students' rankings between the two tests. Figure 1 shows the average standardized scores (and confidence intervals for the means) for teachers and other professionals in both tests and the three subjects assessed. As it is clear from the figure, teachers underperform (on average) in comparison to their peers in all the three subjects in both tests.

Low scores in Saber 11 might also work as incentives for underachievers to choose the teaching profession, as admittance cut-off scores for teaching careers are low. Indeed, Figure A1 in the appendix shows that the probability of being a teacher is lower the higher the score attained in Saber 11 for our three common subjects. Therefore, low scores and socioeconomic disadvantage might narrow some students' possibilities when choosing a college program.

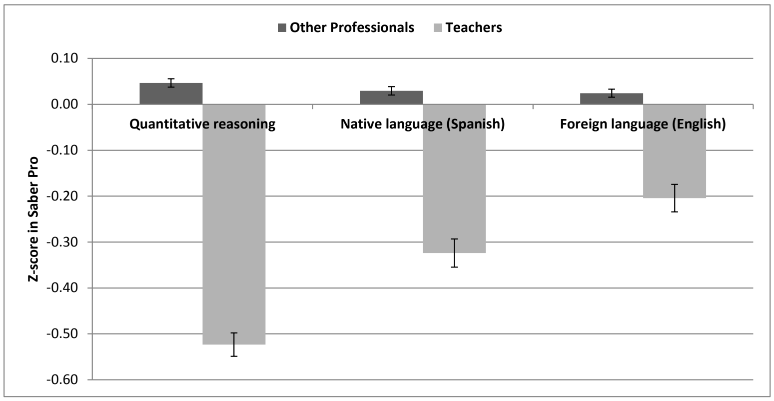
Figure 1. Average scores of teachers and non-teachers in Saber 11 and Saber Pro

a. Saber 11



(***Continues on next page***)

b. Saber Pro



Source: Authors' calculations based on ICFES data.
Note: Confidence intervals at 95% added.

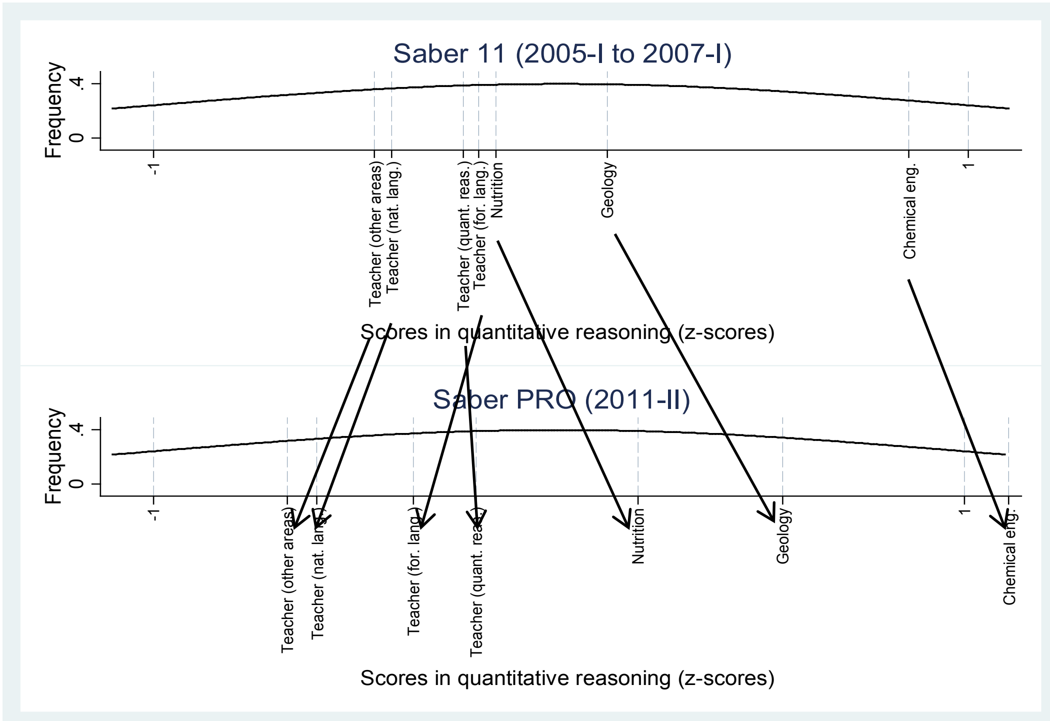
3. Relative learning mobility

Relative learning mobility is assessed by the change in z-scores in quantitative reasoning, native language and foreign language between Saber 11 and Saber PRO. Figure 2 illustrates such relative learning mobility for some selected university programs. The education majors are grouped into four categories according to their closest related emphasis: quantitative reasoning (mathematics, physics, electronics, etc.), native language (Spanish, social sciences, humanities, philology), foreign language (mainly English, although there are students of some other foreign languages), and others (preschool, arts, sports and others).[3] The figure illustrates that after around five years of college, gaps in quantitative reasoning between school teaching programs and other (selected) academic programs widened considerably.[4] That is, teachers' skills in quantitative reasoning either relatively deteriorate or, another way looking at it, do not improve as much as those of other professionals.

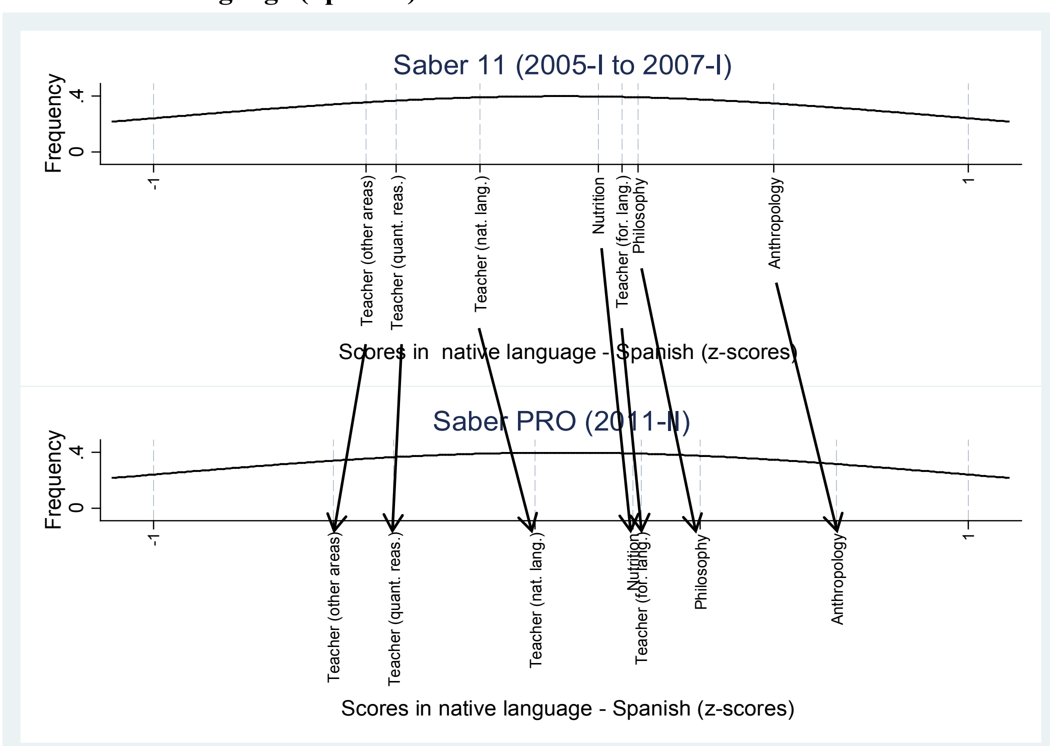
[3] See the online appendix for information on how we classified each major in education into each of these four categories: https://sites.google.com/site/cfbalcazars/misc
[4] We selected those programs with the highest changes in their z-scores between Saber 11 and Saber Pro for our illustrative purposes.

Figure 2. Distribution of standardized test scores in Saber 11 and Saber Pro, selected programs highlighted

a. Quantitative reasoning

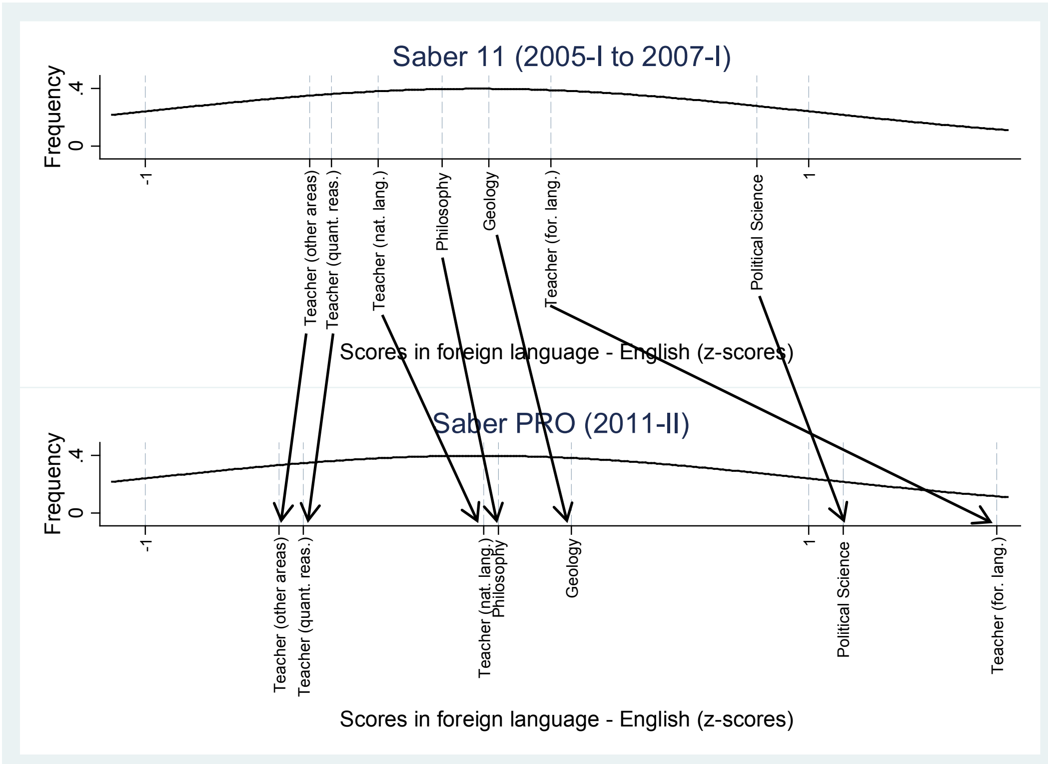


b. Native language (Spanish)



(*Continues on next page*)

c. Foreign language (English)



Source: Authors' calculations based on ICFES data.

To assess relative learning mobility we estimate:

$$y_{i,1} = \beta_0 + \beta_1 X_i + \beta_2 t_{i,1} + \beta_3 y_{i,0} + \beta_4 \left(t_{i,1} * y_{i,0}\right) + \varepsilon_{i,1}, \qquad (1)$$

where $y_{i,1}$ represents the z-score attained in Saber Pro by individual i; X_i is a vector of observable characteristics; $t_{i,1}$ is a dummy variable that takes the value 1 if the individual i is found in Saber PRO as an education major and 0 if she/he is in another program; $y_{i,0}$ is the z-score attained in Saber 11 by i; $\varepsilon_{i,1}$ is an idiosyncratic error term.

The vector of observable characteristics includes: gender; parents' education; family size; semester in which the student took Saber 11 and dummies by administrative unit where the student took Saber 11. We do not include observable school characteristics as they might act as confounding factors –we do not know if the student changed schools previous to his/her senior year of high school. Furthermore, university characteristics might act themselves as bad controls. They might be at the same time a result of a student's Saber 11 scores and observable socioeconomic conditions (e.g., parents'/guardians' education) that allow them to have a larger set of options for schooling to choose from. Hence, university characteristics are not included.

Given that some education majors emphasize on the subjects assessed, we proceed to capture this particularity by adding a dummy variable, $h_{i,1}$, that takes the value 1 if the individual i is studying an education major that makes emphasis in either quantitative reasoning, native language or foreign language –in accordance with the dependent variable. Therefore (1) becomes:

$$y_{i,1} = \beta_0 + \beta_1 X_i + \beta_2 t_{i,1} + \beta_3 y_{i,0} + \beta_4 \left(t_{i,1} * y_{i,0}\right) + \beta_5 h_{i,1} + \beta_6 \left(h_{i,1} * y_{i,0}\right) + \varepsilon_{i,1}. \quad (2)$$

This specification allows us to analyze whether teachers emphasizing the related subject perform better than those teachers who do not. That is, it allows us to do comparisons amongst education majors.

Both (1) and (2) suffer from selection bias. Indeed, we have shown that there are differences in the distribution of observable characteristics between teachers and other professionals. Furthermore, it is reasonable to think that selection is not only influenced by the social and economic background, but also by inclination and interest (Denzler and Wolter, 2008; Aksu et al., 2010; Mizala and Ñopo, 2014). Thus we cannot rule out the role of non-observable characteristics on the selection bias.

We address endogenity in the decision of studying a major in school education by using two approaches: non-parametric matching and Instrumental Variables (IV). The first approach allows us to address the differences in the distribution of observable characteristics in and out of the common support between teachers and other professionals (Ñopo, 2008). That is, we compare teachers and other professionals in the common support of the distribution of observable characteristics. The second approach allows us to correct for omitted variable bias. We describe both approaches next.

3.1 Non-parametric matching

There might be combinations of characteristics for which we can find a teacher, but not another professional. These differences in observables might lead to biased estimates given the differences in the supports of the distribution of observable characteristics. Ñopo (2008) tackles this problem by comparing individuals in the common support. The matching procedure goes as follows:

- Select one teacher from the sample (without replacement).
- Select all non-teachers that have the same observable characteristics as the teacher previously selected.
- Put the observations of both individuals (the synthetic teacher and non-teacher) in their respective new samples of matched individuals, reweighting the observations.
- Repeat until exhausting the original teachers' sample.

We use as variables for matching: gender; birth year; year and semester in which Saber 11 was taken; parents' education; administrative unit where the student's high school was located; administrative unit where the student's college was located; and scores in Saber 11 (as deciles of the distribution of scores).[5]

By matching on observables we obtain a new distribution of observable characteristics for other professionals that mimics the one for teachers. Therefore, we proceed to estimate:

$$y_{i,1}\, w_{matching} = [\beta_0 + \beta_1 X_i + \beta_2 t_{i,1} + \beta_3 y_{i,0} + \beta_4 (t_{i,1} \times y_{i,0}) + \varepsilon_{i,1}] w_{matching} \qquad (3)$$

and (subsequently)

$$y_{i,1}\, w_{matching} = [\beta_0 + \beta_1 X_i + \beta_2 t_{i,1} + \beta_3 y_{i,0} + \beta_4 (t_{i,1} \times y_{i,0}) + \beta_5 h_{i,1} + \beta_6 (h_{i,1} \times y_{i,0}) + \varepsilon_{i,1}] w_{matching}, \qquad (4)$$

[5] This combination of variables allow us to perform a matching exercise on most of our relevant observables by keeping a considerable number of teachers and other professionals in the common support.

where $w_{matching}$ denotes the weights after the differences in the distribution of observable characteristics vanish.

3.2 Instrumental variables

The instrumental variables methodology consists of a two-step regression. In the first step we regress our endogenous variables on all covariates and a vector of valid instruments,[6] which are correlated with the decision of being a teacher but uncorrelated with scores in Saber Pro. Then, in the second step we use the estimated values $\hat{t}_{i,1}$ and $\widehat{t_{i,1} * y_{i,0}}$, and estimate

$$y_{i,1} = \beta_0 + \beta_1 X_i + \beta_2 \hat{t}_{i,1} + \beta_3 y_{i,0} + \beta_4 \left(\widehat{t_{i,1} * y_{i,0}}\right) + \varepsilon_{i,1}. \qquad (5)$$

Analogously for Equation (2), we plug $\hat{t}_{i,1}$, $\widehat{t_{i,1} * y_{i,0}}$, $\hat{h}_{i,1}$ and $\widehat{h_{i,1} * y_{i,0}}$ –obtained from the corresponding first step regressions- into it and estimate

$$y_{i,1} = \beta_0 + \beta_1 X_i + \beta_2 \hat{t}_{i,1} + \beta_3 y_{i,0} + \beta_4 \left(\widehat{t_{i,1} * y_{i,0}}\right) + \beta_5 \hat{h}_{i,1} + \beta_6 \left(\widehat{h_{i,1} * y_{i,0}}\right) + \varepsilon_{i,1}. \quad (6)$$

Our instruments correspond to two dummy variables: The first one takes the value of 1 if the student moved to another administrative unit to pursue his/her college studies, 0 otherwise. The second one takes the value of 1 if the student studied in a superior normal school, 0 otherwise.

There is no reason to think that moving to another administrative unit is correlated to Saber Pro scores. In Colombia, although the main cities offer a broad spectrum of possibilities for higher education, there is high within-administrative unit variation in the quality of higher education programs.[7] Indeed, we find that the correlation between scores in quantitative reasoning, foreign language and native language and our instrument is below 0.008; using Saber 11 scores we find that this correlation is below 0.01.

[6] Given that $t_{i,1}$ is endogenous for modelling students' skills, $t_{i,1} * y_{i,0}$ is also endogenous. Also remember that any valid instrument for $t_{i,1}$ it can also be used as a valid instrument for $t_{i,1} * y_{i,0}$.

[7] See: http://www.icfes.gov.co/resultados/

Our first instrument embodies the process of comparing the costs and benefits of schooling alternatives. We showed that students that choose education majors are more likely to come from a disadvantaged socioeconomic background, which limits their options of schooling. Indeed, teachers are less likely to move to another administrative unit compared to their peers in other academic programs, probably, due to the costs that such decision implies. Similar variables, such as the proximity to college or differences in costs of schooling, have been used in the past as instruments for schooling decisions (Kane and Rouse, 1993; Connely and Uusitalo, 1997; Angrist and Krueger, 2001; Card, 1995, 2001).

Our second instrument acts as a proxy for inclination and interest. Superior normal schools correspond to vocational programs that train high school students to become school teachers. Students can choose to emphasize in school teaching during their high school education and obtain a high school diploma with emphasis in school teaching (*Law 115 of 1994, Republic of Colombia*). However, these are school programs of limited availability. Furthermore, it is important to consider that many of those who enroll in such programs do not enroll into school education majors. Hence, there is no reason to think of negative self-selection into normal schools. Indeed, we do not find evidence for negative self-selection given that the correlations between Saber 11 scores and having acquired a normal school degree are around 0.05 or lower, similar for Saber Pro.[8]

It is important to note that although it is reasonable to think that IV would also solve the problem of omitted variable bias in the matching estimator (i.e., applying IV on the matched sample), this might not be the case. Since instrumental variables estimates are consistent, but not unbiased, it is precise to have large samples to obtain unbiased estimators. Given that non-parametric matching reduces the size of the sample, combining both techniques would not be wise. Moreover, we do not

[8] To dissipate any doubts regarding our instruments, we also compute tests for underidentification, weak identification and overidentification.

know about the asymptotic properties of the non-parametric matching estimator when used alongside IV.

Our matching estimator, however, might reduce omitted variable bias if the matching variables are good proxies for an instrument because we eliminate any differences in the distribution of the proxies between teachers and other professionals. For that reason, we use as matching variables the student's administrative unit of residence before college and during college –an "instrument" similar to the first one we describe above. However, we do not use our second instrument as a control for matching because it reduces considerably the size of the common support and it does not add much to our identification strategy, given that it varies little.

Since our IV estimates give us a clear idea of the magnitude of the bias, we use our matching estimators as a benchmark for our argument on the relative deterioration of teachers' skills vis-à-vis those of other professionals. Indeed, under matching, we do not only compare individuals on the basis of the same observable characteristics but also on the same proxy of academic ability before college education, Saber 11 scores.

4. Results

Table 3 shows the estimates of relative learning mobility for 9 specifications; each triplet of columns corresponds to a subject: quantitative reasoning, native language and foreign language. The first column of each triplet shows the results obtained after estimating (1) through Ordinary Least Squares (OLS); the second column after estimating (3); that is, after matching on observable characteristics,[9] and the third column shows the results obtained from estimating our instrumental variables regression, Equation (5).[10]

[9] Table A1 in the appendix shows the size of the common support after adding each matching variable at a time. After controlling for the full set of observable characteristics we are left with 60% of the *teachers* sample

Our results show that, in general, being enrolled in an education major is negatively related to scores in quantitative reasoning and native language in Saber Pro. Also, performance in Saber 11 is a good predictor of performance in Saber Pro across the board, which comes as no surprise.

The interaction between scores in quantitative reasoning in Saber 11 and being a teacher shows a negative value on all the regressions. That is, on average, teachers' skills in quantitative reasoning relatively deteriorate compared to those of their peers in other university programs. Although this result does not hold under matching, it is precise to consider that our IV specification shows that β_4 is biased upwards under OLS.

For native language we do not find robust evidence of a relative change in skills of teachers vis-à-vis other professionals. For foreign language, on the other hand, we find evidence pointing towards a relative deterioration in teachers' skills –our IV estimate is negative and statistically significant.

We find evidence that our instrumental variables are good instruments –we reject the null hypotheses for underidentification (Kleibergen-Paap test), overidentification (Sargan-Hansen test) and weak identification (Angrist-Pischke test). In other words, we find evidence that β_2 is biased upwards in Equation (1), which also explains why we find that β_4 is also biased upwards.

All in all, although we find that β_2 under OLS appears to biased downwards compared to matching, we must keep in mind that the magnitude of the bias due to observable characteristics is considerably small compared to the magnitude of the bias due to unobservable characteristics. Considering that our matching estimators are also underestimated, and considering the magnitude of the bias of the OLS estimators, we can conjecture that after controlling for omitted variable bias our matching regressions would produce lower values for both β_2 and β_4.

---

and 21% of the *other professionals* sample. Kolmogorov-Smirnov tests for the equality of the distributions of teachers and other professionals fail to reject the null hypothesis.

[10] The first-step estimates are in Table A2 in the appendix.

Table 3. Relative learning mobility

Variable	Quantitative reasoning: Ordinary Least Squares	Quantitative reasoning: Non-parametric matching	Quantitative reasoning: Instrumental Variables	Native language (Spanish): Ordinary Least Squares	Native language (Spanish): Non-parametric matching	Native language (Spanish): Instrumental Variables	Foreign language (English): Ordinary Least Squares	Foreign language (English): Non-parametric matching	Foreign language (English): Instrumental Variables
	Dependent variable: Z-scores in Saber PRO								
Teacher	-0.4636***	-0.3810***	-1.2405***	-0.1441***	-0.1010***	-0.9698***	-0.0560***	-0.0003	-1.3436***
	(0.0161)	(0.0333)	(0.2353)	(0.0166)	(0.0316)	(0.2188)	(0.0184)	(0.0312)	(0.2209)
Saber 11 scores in									
Quantitative reasoning	0.5037***	0.4408***	0.5339***						
	(0.0047)	(0.0121)	(0.0106)						
Native language				0.4967***	0.4716***	0.4893***			
				(0.0043)	(0.0109)	(0.0101)			
Foreign language							0.6240***	0.6105***	0.6523***
							(0.0055)	(0.0146)	(0.0114)
Teacher * Saber 11 scores in									
Quantitative reasoning	-0.1554***	-0.0391	-0.8819***						
	(0.0173)	(0.0336)	(0.1873)						
Native language				0.0322**	0.0578*	0.0054			
				(0.0136)	(0.0323)	(0.1225)			
Foreign language							0.1025***	0.0812*	-0.8190***
							(0.0274)	(0.0454)	(0.2284)
Constant	0.5272*	-0.2659	0.5086*	0.5815**	0.3655	0.5849**	0.3366*	-0.2896	0.2944
	(0.2831)	(0.3837)	(0.2764)	(0.2518)	(0.2917)	(0.2533)	(0.1997)	(0.5698)	(0.1975)
Observable characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reliability of instruments									
Kleibergen-Paap rk Wald F statistic			193.1831			167.6556			130.4739
P-value			0.0000			0.0000			0.0000
Hansen J statistic			18.0750			40.0966			61.7321
P-value			0.0001			0.0000			0.0000
Angrist-Pischke multivariate F test of excluded instruments			99.4			102.22			84.34
P-value			0.0000			0.0000			0.0000
R-squared	0.3234	0.2751	0.2889	0.3045	0.3008	0.2807	0.5040	0.4327	0.4020
Observations	**50772**	**12431**	**50772**	**50772**	**12520**	**50772**	**50772**	**12463**	**50772**

Robust standard errors in parentheses. * Significant at ten percent; ** significant at five percent; *** significant at one percent.
Note: Observable characteristics include: gender; parents' education; family size; semester in which the student took Saber 11; and dummies by municipality where the student took Saber 11.
Source: Authors' calculations based on ICFES data.

Not all teachers receive the same training while in college. Some of them specialize in teaching math, some others in teaching Spanish, some others in teaching English and so on. It is reasonable to expect different learning gains in different subjects of Saber Pro according to their fields of specialization. Table 4 shows estimations exploring such differences in relative learning mobility. That is, when analyzing relative learning gains in quantitative reasoning we will pay special attention to those education majors emphasizing math. Likewise, when analyzing relative learning gains in native language (foreign language) we will pay special attention to those education majors emphasizing Spanish (foreign languages).

Similar to Table 3, Table 4 shows the estimates of relative learning mobility for 9 specifications: each triplet of columns corresponds to a subject; the first column of each triplet shows the results obtained after estimating Equation (2) through OLS; the second column after estimating Equation (4), and the third column shows the results obtained from estimating Equation (6).

As in the previous estimations, being a teacher is negatively correlated to performance in Saber Pro. Also, the predictive power of Saber 11 across the board (β_3) and the smaller learning mobility of teachers vis-à-vis other professionals in quantitative reasoning and foreign language (β_4), hold (under IV).

Nonetheless there are some important particularities that we find in these specifications. On the one hand, the coefficient corresponding to the $h_{i,1}$ dummy is positive and significant across the board, but the total effect $\beta_2+\beta_5$ is still negative for quantitative reasoning under OLS and IV. For foreign language we find statistically significant evidence for $\beta_2+\beta_5 > 0$. For native language we do not find robust evidence for $\beta_2+\beta_5 \neq 0$ given that under IV it is statistically zero.

We find that teachers' smaller learning mobility vis-à-vis other professionals gets counterbalanced among those teachers whose programs emphasized math and foreign languages (under IV), accordingly; that is, $\beta_6 > 0$. It is important, however, to note that their mobility is still smaller than that of other professionals ($\beta_4 + \beta_6 < 0$). Although these results do not hold under matching, if we consider the magnitude of the bias due to unobservables, our IV estimators provide suggestive evidence that these results would hold even under non-parametric matching.

The previous results also show that given the emphasis of the career, academic programs emphasizing a particular subject might obtain higher scores. This might explain why Chemical Engineering and Geology show high relative learning mobility in quantitative reasoning (Figure 2). However it is not a sufficient condition. It does not explain why Nutrition (a career that is not known for emphasizing math) shows high relative learning mobility in quantitative reasoning (Figure 2), or why β_6 is statistically zero for native language.

All in all, our results indicate that teachers' skills vis-à-vis those in other professions relatively deteriorate (or do not improve as much) in quantitative reasoning and foreign language, although they deteriorate less for those in math-oriented and foreign language-oriented programs respectively. For native language, on the other hand, we do not find evidence of statistically significant differences in relative learning mobility.

Table 4. Relative learning mobility controlling for career emphasis for teachers

	Dependent variable: Z-scores in Saber PRO								
	Quantitative reasoning			Native language (Spanish)			Foreign language (English)		
Variable	Ordinary Least Squares	Non-parametric matching	Instrumental Variables	Ordinary Least Squares	Non-parametric matching	Instrumental Variables	Ordinary Least Squares	Non-parametric matching	Instrumental Variables
Teacher	-0.4636***	-0.3810***	-1.2405***	-0.1441***	-0.1010***	-0.9698***	-0.0560***	-0.0003	-1.3436***
	(0.0161)	(0.0333)	(0.2353)	(0.0166)	(0.0316)	(0.2188)	(0.0184)	(0.0312)	(0.2209)
Saber 11 scores in									
Quantitative reasoning	0.5032***	0.4380***	0.5317***						
	(0.0047)	(0.0120)	(0.0107)						
Native language				0.4966***	0.4715***	0.4944***			
				(0.0043)	(0.0108)	(0.0112)			
Foreign language							0.6226***	0.6009***	0.6392***
							(0.0055)	(0.0136)	(0.0102)
Teacher * Saber 11 scores in									
Quantitative reasoning	-0.1983***	-0.0905**	-1.1270***						
	(0.0190)	(0.0399)	(0.2633)						
Native language				0.0398***	0.0587	-0.1560			
				(0.0155)	(0.0390)	(0.1882)			
Foreign language							-0.0456	0.0463	-0.9986***
							(0.0300)	(0.0518)	(0.2633)
Teacher program has emphasis in the assessed subject in Saber PRO	0.4183***	0.5551***	0.8530***	0.2612***	0.2989***	0.9865***	1.6177***	1.4955***	2.6941***
	(0.0330)	(0.0573)	(0.2306)	(0.0307)	(0.0630)	(0.2125)	(0.0488)	(0.0829)	(0.2299)

(***Continues on next page***)

Variable	Dependent variable: Z-scores in Saber PRO								
	Quantitative reasoning			Native language (Spanish)			Foreign language (English)		
	Ordinary Least Squares	Non-parametric matching	Instrumental Variables	Ordinary Least Squares	Non-parametric matching	Instrumental Variables	Ordinary Least Squares	Non-parametric matching	Instrumental Variables
Teacher program has emphasis in the assessed subject in Saber PRO * Saber 11 scores in									
Quantitative reasoning	0.0885**	0.0365	0.9171***						
	(0.0391)	(0.0664)	(0.2670)						
Native language				-0.0639**	-0.0391	0.2160			
				(0.0292)	(0.0660)	(0.1853)			
Foreign language							-0.1584***	-0.3296***	0.8286***
							(0.0546)	(0.0915)	(0.2707)
Constant	0.5234*	-0.0763	0.5001*	0.5810**	0.4350	0.5826**	0.3200	-0.2840	0.2879
	(0.2833)	(0.4170)	(0.2772)	(0.2517)	(0.2900)	(0.2537)	(0.1994)	(0.5065)	(0.1983)
Observable characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reliability of instruments									
Kleibergen-Paap rk Wald F statistic			145.7668			91.5435			105.2551
P-value			0.0000			0.0000			0.0000
Hansen J statistic			21.3165			46.2364			82.4951
P-value			0.0003			0.0000			0.0000
Angrist-Pischke multivariate F test of excluded instruments			35.64			35.93			39.79
P-value			0.0000			0.0000			0.0000
R-squared	0.3255	0.3026	0.2856	0.3057	0.3098	0.2724	0.5216	0.5233	0.4419
Observations	**50772**	**12431**	**50772**	**50772**	**12520**	**50772**	**50772**	**12463**	**50772**

Robust standard errors in parentheses. * Significant at ten percent; ** significant at five percent; *** significant at one percent.
Note: Programs with emphasis in mathematics are those related to the study of mathematics, physics, biology and the like; programs with emphasis in native language are those related to the study of social sciences, humanities, philology and the like; programs with emphasis in foreign language are those that are focused on the study, mainly, of a foreign language (including English). Observable characteristics include: gender; parents' education; family size; semester in which the student took Saber 11; and dummies by municipality where the student took Saber 11.
Source: Authors' calculations based on ICFES data

5. Conclusions

It is unquestionable that teachers are essential for a high-quality educational system. Nonetheless, there is evidence of negative selection into teaching as students from disadvantaged socioeconomic environments with low academic performance are more likely to enroll into programs in school education. This study builds onto that providing another piece of evidence. We find that the skills of students in education majors in quantitative reasoning and foreign language relatively deteriorate in comparison to those who enroll in other majors. After nearly 5 years of academic training, learning gaps between teachers and other professionals widen in favor of the latter. This raises an additional red flag regarding: *university school teaching programs might not be producing good teachers*.

What can be done?

On the one hand, the teaching bodies of the future teachers, or the curricula they follow, or their pedagogical approaches, might need some reforms. On the other hand, there is also room for action on the selection of students into teaching, such as stricter admission standards. This may work positively by two channels: a direct effect on the skills of students in teaching majors and an indirect effect through peers. Nonetheless, better teacher education programs and higher admissions standards, alone, most likely will have modest effects. To think that the solution to the problem of an inadequate teaching force lies only within the teaching community would be extremely myopic. It is necessary to push for ambitious policies aimed at making the teaching profession more attractive so that the most talented youngsters opt to teach and develop a good career path; e.g., higher salaries (Dolton & Marcenaro-Gutierrez, 2011).

The recent decades have seen innovation and progress in teacher policies both in the developed and developing world. Many of them have been fortunately accompanied by rigorous impact evaluations. Thus, although perhaps we are still far from answering the question of the design of optimal policies for the teaching profession in different contexts, knowledge is in the making (Vegas and Ganimian, 2013).

References

Aksu, M., Demir, C., Daloglu, A., Yildirim, S., & Kiraz, E. (2010). "Who are the future teachers in Turkey? Characteristics of entering student teachers," International Journal of Educational Development 30: 91-101.

Angrist, J., & Krueger, A. (2001). "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," Journal of Economic Perspectives, 15 (4): 69-85

Ballou, D., & Podgursky, M. (1997). "Teacher Pay and Teacher Quality," W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.

Barón, J. D., & Bonilla, L. (2011). "La calidad de los maestros en Colombia: Desempeño en el examen de Estado ICFES y la probabilidad de graduarse en el área de educación," Documentos de Trabajo sobre Economía Regional 152, Banco de la República, Colombia.

Barón, J. D., Bonilla, L., Cardona-Sosa, L., & Ospina, M. (2013). "Quénes eligen la disciplina de la educación en Colombia? Caracterización desde las pruebas Saber 11," Borradores de Economía 785, Banco de la República, Colombia.

Card, D. (1995). "Earnings, Schooling and Ability Revisited," Research in Labor Economics 14: 23-48

Card, D. (2001). "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," Econometrica 69(5): 1127-1160.

Chetty, R., Friedman, J., & Rockoff, J. (2013a). "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," NBER Working Papers 19423, National Bureau of Economic Research, Inc.

Chetty, R., Friedman, J., & Rockoff, J. (2013b). "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," NBER Working Papers 19424, National Bureau of Economic Research, Inc.

Clotfelter, T., Ladd, H. F., & Vigdor, J. L. (2007). "Teacher credentials and student achievement: Longitudinal analysis with student fixed effects," Economics of Education Review 26(6): 673-682.

Conneely, K. & UusitaLo, R. (1997): "Estimating Heterogeneous Treatment Effects in the Becker Schooling Model," Unpublished Discussion Paper, Princeton University Industrial Relations Section.

Cunha, J., & Miller., T. (2014). "Measuring value-added in higher education: Possibilities and limitations in the use of administrative data," Economics of Education Review 42: 64–77

Denzler, S., & Wolter. S. (2008). “Selbstselektion bei der Wahl eines Lehramtsstudiums: Zum Zusammenspiel individueller und institutioneller Faktoren,” Beiträge zur Hochschulforschung, 30(4): 112-141.

De Paola, M. (2009). “Does Teacher Quality Affect Student Performance? Evidence From An Italian University,” Bulletin of Economic Research, 61(4): 353–377.

Dolton, P. (1990). “The economics of UK teacher supply,” Economic Journal, 100: 91–104.

Dolton, P. (2006). “Teacher supply,” chapter 19 in E. Hanushek and F. Welch (eds.), Handbook of the Economics of Education vol. 2, Elsevier, The Netherlands.

Dolton, P., & Marcenaro-Gutierrez, O. (2011). “If you pay peanuts do you get monkeys? A cross-country analysis of teacher pay and pupil performance,” Economic Policy, 26(65): 5-55.

Dolton, P., & van der Klaauw, W. (1999). “The turnover of teachers: A competing risks explanation,” Review of Economics and Statistics, 81(3): 543–550.

Eide, E., Goldhaber, D., & Brewer, D. (2004). “The Teacher Labour Market and Teacher Quality,” Oxford Review of Economic Policy, 20(2): 230–244.

García, S., Maldonado, D., Perry, G., Rodríguez, C., & Saavedra., J. (2014). “Tras la excelencia docente: Cómo mejorar la calidad de la educación para todos los colombianos,” Fundación Compartir. Bogotá, Colombia.

Giesen, H., & Gold, A. (1993). “Leistungsvoraussetzungen und Studienbedingungen bei Studierenden verschiedener Lehrämter,” Psychologie in Erziehung und Unterricht, 40: 111–124

Goldhaber, D., & Brewer, D. (1997). “Why don’t schools and teachers seem to matter?: Assessing the impact of unobservables on educational productivity”, Journal of Human Resources 32(3): 505–523.

Hanushek, E. A., & Pace, R. (1995). “Who Chooses To Teach And Why?” Economics of Education Review, 14(2): 101–117

Hanushek, E. A. (2002). “Publicly provided education,” chapter 30 in A. J. Auerbach and M. Feldstein (eds.), Handbook of Public Economics, vol. 4, Elsevier, The Netherlands.

Hanushek, E. A., & Rivkin, S. G. (2006). “Teacher Quality,” chapter 18 in E. Hanushek and F. Welch (eds.), Handbook of the Economics of Education, vol. 2, Elsevier, The Netherlands.

Kane, T., & Rouse, C. (1993). “Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?” NBER Working Paper 4268.

Klein, S. P., Kuh, G., Chun, M., Hamilton, L., & Shavelson, R. (2005). “An approach to measuring cognitive outcomes across higher education institutions,” Research in Higher Education, 46: 251–276.

Kukla-Acevedo, S. (2009). “Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement,” Economics of Education Review 28(1): 49-57.

Liu, O. (2011). “Value-added assessment in higher education: A comparison of two methods,” Higher Education, 61(4): 445–461.

Mizala, A., & Ñopo, H. (2014). “Measuring the Relative Pay of Latin American School Teachers at the turn of the 20th Century,” Working Papers 2014-15, Peruvian Economic Association.

Murnane, R., Singer, J., Willett, J., Kemple, J., & Olson, R. (1991). “Who Will Teach? Policies that Matter,” Harvard University Press, Cambridge, MA.

Ñopo, H. (2008). “Matching as a Tool to Decompose Wage Gaps,” Review of Economics and Statistics, 90(2): 290-299.

Odden, A., & Kelley, C. (1997). “Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools,” Corwin Press, Thousand Oaks, CA.

OECD (2011). “Finland: Slow and steady reform for consistently high results” in Strong Performers and Succesful Reformes in Education, Lessons from PISA for the United States, OECD Publishing.

Podgursky, M., Monroe, R., & Watson, D. (2004). “The academic quality of public school teachers: an analysis of entry and exit behavior,” Economics of Education Review, 23(5): 507–518

Rivkin, S., Hanushek, E., & Kain, J. (2005). “Teachers, Schools, and Academic Achievement,” Econometrica, 73(2): 417-458.

Rockoff, J. E. (2004). “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data,” The American Economic Review, 94(2): 247–252.

Saavedra, J. (2009). “The learning and early labor market effects of college quality: a regression discontinuity analysis,” Mimeo, Harvard University.

Saavedra, A., & Saavedra, J. (2011). “Do colleges cultivate critical thinking, problem solving, writing and interpersonal skills?” Economics of Education Review, 30(6): 1516–1526.

Sahlberg, P. (2009). “Educational Change in Finland” in A. Hargreaves, M. Fullan, A. Lieberman, and D. Hopkins (Eds.), International Handbook of Educational Change, Springer, The Netherlands.

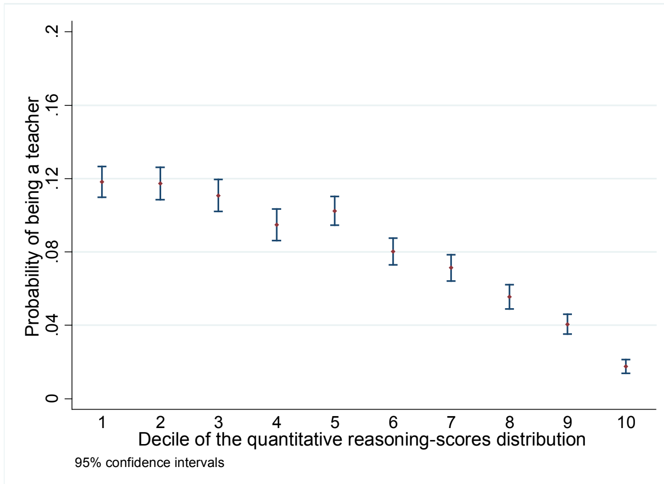
Sahlberg, P. (2011). “Finnish lessons. What can the world learn from educational change in Finland?” Teachers College Press, New York.

Vegas, E., & Ganimian, A. (2013). “Theory and Evidence on Teacher Policies in Developed and Developing Countries,” IDB Working Paper Series No. IDB-WP-438.

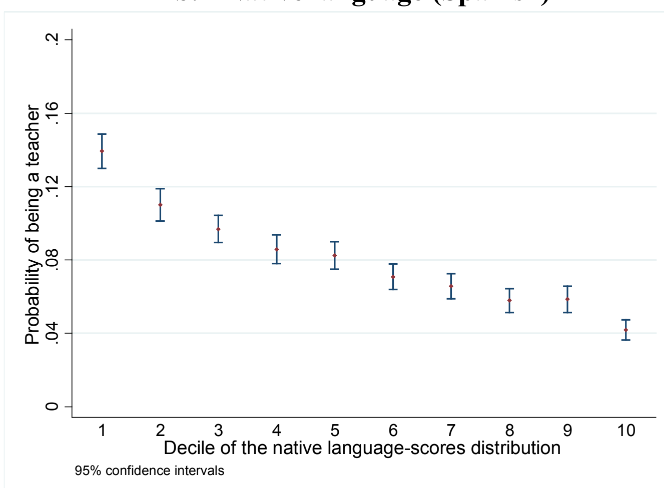
Appendix A

Figure A1. Probability of being a teacher by deciles of the z-score distribution

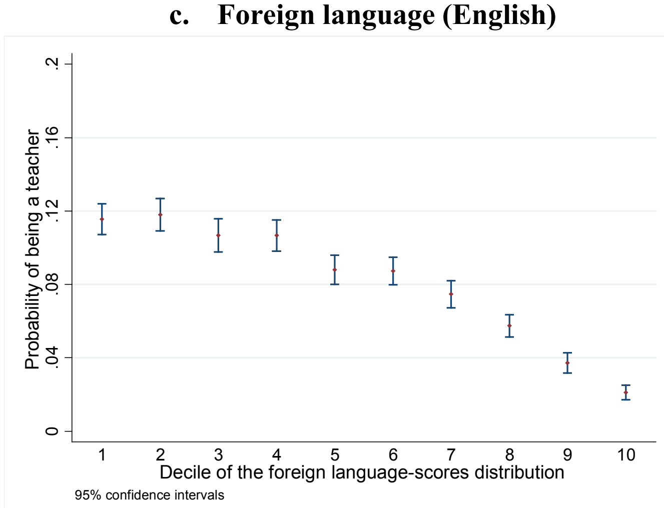
a. Quantitative reasoning



b. Native language (Spanish)



c. Foreign language (English)



Source: Authors' calculations based on ICFES data.

Table A1. Size of common support and Kolmogorov-Smirnov tests (p-values after adding observables characteristics)

Variables	Percentage of students		Kolmogorov-Smirnov P-value
	Teachers	Other professionals	
Gender	100.00	100.00	0.00
+ year of birth	99.99	99.98	0.00
+ semester and year in which Saber 11 was taken	98.75	99.70	0.00
+ parents' max education	94.69	99.26	0.00
+ administrative unit of residence in senior year of high school	65.94	92.61	0.00
+ administrative unit of residence in college	54.58	85.27	0.00
& Quantitative reasoning scores deciles (Saber 11)	52.17	81.04	0.87
& Native language scores deciles (Saber 11)	54.58	85.27	0.66
& Foreing language scores deciles (Saber 11)	65.94	92.61	0.38

The Kolmogorov-Smirnov test corresponds to the test of equality of two distributions.
Source: Authors' calculations based on ICFES data.

Table A2. First stage regressions for the decision of being a teacher

A. First step regressions, Equation (5)

Variable	Dependent variable: 1 if student studies a program in education, 0 if not		
	Quantitative reasoning	Native language	Foreign language
The student moved to another administrative unit†	-0.0386***	-0.0384***	-0.0384***
	(0.0032)	(0.0032)	(0.0032)
The student studied in a normal school††	0.1387***	0.1426***	0.1308***
	(0.0106)	(0.0106)	(0.0108)
The student moved to another administrative unit * Saber 11 scores in			
Quantitative reasoning	-0.0288**		
	(0.0114)		
Native language		-0.0252**	
		(0.0114)	
Foreign language			-0.0505***
			(0.0136)
The student studied in a normal school * Saber 11 scores in			
Quantitative reasoning	0.0123***		
	(0.0021)		
Native language		0.0128***	
		(0.0026)	
Foreign language			0.0147***
			(0.0021)
Constant	0.0567***	0.0529**	0.0567***
	(0.0209)	(0.0219)	(0.0207)
Saber 11 scores in respective subject	Yes	Yes	Yes
Observable characteristics	Yes	Yes	Yes
R-squared	0.0656	0.0651	0.0648
Observations	**50772**	**50772**	**50772**

B. First step regressions, Equation (6)

	Dependent variable: 1 if student studies a program in education, 0 if not		
	Quantitative reasoning	**Native language**	**Foreign language**
The student moved to another administrative unit†	-0.0454*** (0.0031)	-0.0491*** (0.0030)	-0.0430*** (0.0031)
The student studied in a normal school††	-0.0454*** (0.0031)	-0.0491*** (0.0030)	-0.0430*** (0.0031)
The student moved to another administrative unit * Saber 11 scores in			
Quantitative reasoning	-0.0283*** (0.0107)		
Native language		-0.0253** (0.0110)	
Foreign language			-0.0577*** (0.0133)
The student studied in a normal school * Saber 11 scores in			
Quantitative reasoning	0.0145*** (0.0020)		
Native language		0.0140*** (0.0025)	
Foreign language			0.0143*** (0.0020)
The student moved to another administrative unit * Teacher program has emphasis in the assessed subject in Saber PRO	0.6744*** (0.0256)	0.6893*** (0.0266)	0.5834*** (0.0559)
The student studied in a normal school * Teacher program has emphasis in the assessed subject in Saber PRO	0.8535*** (0.0194)	0.8658*** (0.0188)	0.8687*** (0.0253)

(***Continues on next page***)

	Dependent variable: 1 if student studies a program in education, 0 if not		
	Quantitative reasoning	**Native language**	**Foreign language**
The student moved to another administrative unit * Teacher program has emphasis in the assessed subject in Saber PRO			
Quantitative reasoning	0.1202***		
	(0.0256)		
Native language		0.0497**	
		(0.0236)	
Foreign language			0.1219**
			(0.0511)
The student studied in a normal school * Teacher program has emphasis in the assessed subject in Saber PRO			
Quantitative reasoning	0.0150		
	(0.0140)		
Native language		-0.0107	
		(0.0191)	
Foreign language			0.0278
			(0.0218)
Constant	-0.0278	0.1871	0.1568
	(0.0738)	(0.1285)	(0.1104)
Saber 11 scores in respective subject	Yes	Yes	Yes
Observable characteristics	Yes	Yes	Yes
R-squared	0.0986	0.1110	0.0818
Observations	**50772**	**50772**	**50772**

Standard errors in parentheses. * Significant at ten percent; ** significant at five percent; *** significant at one percent.
† The student moved to another administrative unit for his/her higher education.
†† A "Normal school" is a school that trains high school students to become teachers.
Source: Authors' calculations based on ICFES data.