New Teachers for Colombia: Is Quality Control Working?

Zelda Brutti*, Fabio Sánchez Torres°

*IEB and University of Barcelona; °CEDE and Universidad de Los Andes

ICFES International Seminar, Nov 3rd-4th 2016

Teacher quality

- Various strands of literature:
  - how much difference good teachers can make (Chetty et al. [2014], Rivkin et al. [2005], Rockoff [2005] ...)

  - what makes a good teacher (Rockoff et al. [2011], Kane et al. [2008], Gordon et al. [2006]...)

  - how do we make teachers better, or retain better ones (Hanushek et al. [2004], Figlio [2001]...)

- Consensus: teachers make a big difference for student learning and quality controls are desirable (even though no consensus on which ones exactly)

... and teacher quality in developing countries

- Several Latin American (and South-East Asian) countries are only *starting* interventions on the teacher quality front

... and teacher quality in developing countries

- Several Latin American (and South-East Asian) countries are only *starting* interventions on the teacher quality front

- Colombia is a country we can learn from, because it has already started a massive transition
    Colombia has made education a main priority to improve the economic and social prosperity of the country and pledged more resources to this sector than any other policy area [OECD, 2016]

... and teacher quality in developing countries

- Several Latin American (and South-East Asian) countries are only *starting* interventions on the teacher quality front

- Colombia is a country we can learn from, because it has already started a massive transition
    Colombia has made education a main priority to improve the economic and social prosperity of the country and pledged more resources to this sector than any other policy area [OECD, 2016]

    - Large reform in 2002: introduction of teacher quality requirements and incentives

This paper

- Evaluate how the 2002 Colombian reform of public teacher careers has impacted student performance
  - Has the new regulation succeeded at producing "better" teachers? (overall YES)

This paper

- Evaluate how the 2002 Colombian reform of public teacher careers has impacted student performance
    - Has the new regulation succeeded at producing "better" teachers? (overall YES)

- Investigate the potential channels of the effect
    - Which of the provisions of the reform have been the most effective?

This paper

- Evaluate how the 2002 Colombian reform of public teacher careers has impacted student performance

  - Has the new regulation succeeded at producing "better" teachers? (overall YES)

- Investigate the potential channels of the effect

  - Which of the provisions of the reform have been the most effective?

- Document significant (and worrying!) failure in compliance to the new rules

  - Consequent reduction in benefits for students

Closest literature

Effectiveness of teacher quality insurance policies
▸ Literature Teacher Quality

Teacher quality in Latin America
▸ Literature LA Teacher Quality

Evaluation of 2002 Colombian reform

- Ome (2012, 2013): school fixed effects (across-year variation), finds no effect on high school students, while finding effects at primary school level

Contribution

- Improve on past evaluations of effect of the Colombian teacher reform on student performance

- First attempt at investigating channels and heterogeneities in the effect

- Original analysis on relationship between characteristics of new teachers (esp. entry test scores) and student performance

- Novel documentation of important reform failures

Preview of findings

- Find positive and significant effect of quality-controlled teachers

Preview of findings

- Find positive and significant effect of quality-controlled teachers
  - Within-school fixed effects strategy

Preview of findings

- Find positive and significant effect of quality-controlled teachers
  - Within-school fixed effects strategy

- Effect seems to be driven by selection at entry, not along the career

Preview of findings

- Find positive and significant effect of quality-controlled teachers
  - Within-school fixed effects strategy

- Effect seems to be driven by selection at entry, not along the career

- New-type teachers seem to have decreasing marginal returns

The reform

(Before) the reform

- Entry process not transparent, teacher appointments considered "administrative acts"

- Automatic promotions (every 3 years)

- Excessive protection on the job (virtually impossible to fire)

(Before) the reform

- Entry process not transparent, teacher appointments considered "administrative acts"

- Automatic promotions (every 3 years)

- Excessive protection on the job (virtually impossible to fire)

- This system did not create incentives for teachers to invest in their teaching skills and subject knowledge

(After) the reform

1. Selective entry contest

2. Probation period

3. Permanent (yearly) performance evaluation
   3.1 Pre-requirement for being promoted

   3.2 Possibility of being fired

(Key aspects of) the reform

- The whole set of new regulation applied only to teachers hired after 2002 (deal with unions)

- Teachers already exercising the profession before the reform are still continuing their careers according to the old rules

(Key aspects of) the reform

- The whole set of new regulation applied only to teachers hired after 2002 (deal with unions)

- Teachers already exercising the profession before the reform are still continuing their careers according to the old rules

- We exploit the ensuing mix of new-type and traditional-type of teachers, to identify the effect of the new-type
    - controlling for selection (within school, across subjects variation)

1. The entry contest

- Each educational authority issues a public call for applications, specifying the number and type of vacancies

- Candidates choose educational authority they want to apply to

- Candidates enter the contest
  - Subject knowledge and teaching aptitude exam; evaluation of credentials; verbal interview

- Based on their global contest score, candidates get to choose their preferred vacancy (the remaining on a waiting list)

1. The entry contest
Entry contests so far

CONTESTS >	1st (2004)	2nd (2005)	3rd (2006)	4th (2009)	5th (2013)
N. of local authorities	69	66	49	66	92
Vacancies	50.947	23.355	14.579	25.392	?
Candidates to exam	140.541	134.090	109.487	228.985	301.589
Passed exam stage	60.078 (43%)	32.720 (24%)	27.931 (26%)	66.687 (29%)	54.906 (18%)
Assigned to vacancy	30.568 (22%)	14.092 (11%)	13.620 (12%)	39.468 (17%)	?

Note: all percentages are relative to 'Candidates to exam'
Source: MEN [2013]

2. and 3. Probation and permanent evaluation

2. Probation

- Candidates who obtain a vacancy first enter a probation period (min 4 months, max 1 academic year)
- School principal evaluates the new teacher using a standardized format and may veto his/her final employment

2. and 3. Probation and permanent evaluation

2. Probation

- Candidates who obtain a vacancy first enter a probation period (min 4 months, max 1 academic year)
- School principal evaluates the new teacher using a standardized format and may veto his/her final employment

3. Permanent evaluation

- School principal evaluates the teacher (like after probation) every year
- Teachers take tests similar to the entry exam when they apply for promotion (denied if unsatisfactory test result)

Data

- Teachers: administrative data (Ministry of Education), individual level, 2008–2013
  - 87.4% of new–type teachers matched to their entry test scores

Data

- Teachers: administrative data (Ministry of Education), individual level, 2008–2013
    - 87.4% of new-type teachers matched to their entry test scores

- Students: administrative data (ICFES) on final high school exam ('Saber11'), 2000 to 2013

Data

- Teachers: administrative data (Ministry of Education), individual level, 2008-2013

  - 87.4% of new-type teachers matched to their entry test scores

- Students: administrative data (ICFES) on final high school exam ('Saber11'), 2000 to 2013

- Match teachers and students at school-cohort level (no class information available)

  - e.g. match 2005 Saber11 Math results in school A to Math teachers teaching at high school level in school A in 2005

Descriptive statistics

Table A.4: Individual-level teacher descriptives

	All teachers	Old Regulation	All New Regulation	New Regul. Passed	New Regul. Not Passed
Age	45.80	50.22	37.26	37.09	37.82
	(10.06)	(7.59)	(8.65)	(8.74)	(8.35)
Experience	15.57	21.92	3.30	3.53	2.55
	(11.75)	(9.38)	(2.50)	(2.51)	(2.31)
Female	0.66	0.67	0.63	0.61	0.70
	(0.47)	(0.47)	(0.48)	(0.49)	(0.46)
Postgrad degree	0.20	0.26	0.09	0.10	0.04
	(0.40)	(0.44)	(0.28)	(0.30)	(0.19)
Experience 5-40yrs	0.75	0.98	0.31	0.35	0.20
	(0.43)	(0.15)	(0.46)	(0.48)	(0.40)
Age when hired	30.23	28.30	33.96	33.55	35.28
	(8.12)	(7.19)	(8.49)	(8.49)	(8.37)
Rural area	0.30	0.23	0.43	0.39	0.54
	(0.46)	(0.42)	(0.49)	(0.49)	(0.50)
Permanent position	0.85	0.99	0.58	0.70	0.19
	(0.36)	(0.09)	(0.49)	(0.46)	(0.39)
Temporary position	0.11	0.01	0.32	0.19	0.74
	(0.32)	(0.08)	(0.47)	(0.39)	(0.44)
Probation position	0.03	0.00	0.09	0.10	0.07
	(0.18)	(0.04)	(0.29)	(0.30)	(0.25)
Most recent test score				63.69	55.32
				(3.87)	(3.52)
N	1743,339	1149,239	594,100	452,493	141,607
N teachers	360,644	214,920	145,724	108,735	36,989

Note: Variable means and (standard deviations).

Descriptive statistics II

Table A.3: Descriptive statistics at school-year-subject level

	Total	Within	Between
Mean student score	43.33		
	(3.29)	(2.04)	(2.70)
Share New Regulation	0.47		
	(0.43)	(0.30)	(0.33)
Share New Regulation Passed	0.34		
	(0.40)	(0.31)	(0.28)
Share New Regulation Not Passed	0.13		
	(0.29)	(0.22)	(0.21)
Share Old Regulation	0.53		
	(0.43)	(0.30)	(0.33)
Mean age	44.20		
	(8.31)	(6.31)	(5.83)
Mean experience	12.49		
	(9.20)	(6.48)	(6.85)
Share postgraduate degree	0.21		
	(0.34)	(0.25)	(0.24)
N	151,178	151,178	151,178

Note: Variable means and (standard deviations). Total, within-group and between group standard deviations, where a group is a school-year cell.

The 'irregular' teachers I

Figure 1: Entry exam scores of teachers observed over 2008-2013
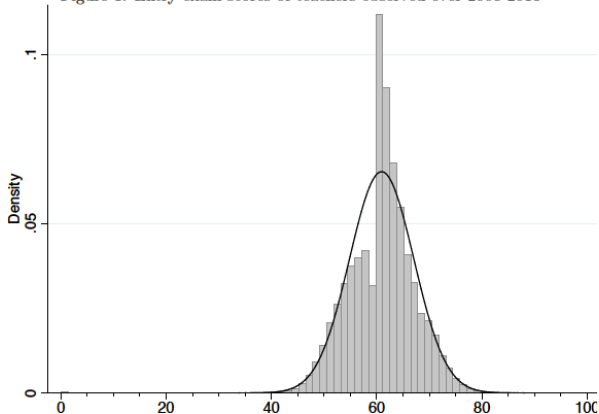
Density of entry test scores (most recent score of each teacher); bin width = 1.33 points; normal curve is overlaid.

Around 28% of all currently active new-type teachers has a highest exam score < the minimum

The 'irregular' teachers II

Table 2: Teachers with most recent score below 60

Type of position	Never above 60	At least once above 60	Total
Permanent	324	700	1024
	2.77	20.49	6.78
Temporary	11124	2456	13580
	95.11	71.90	89.86
Probation	248	260	508
	2.12	7.61	3.36
Total	11,696	3,416	15,112
	100.00	100.00	100.00

... and most of them never passed the entry exam, and operate in 'temporary positions'

The 'irregular' teachers II

- High heterogeneity in their distribution across the country
  - Probably correlated with other indicators, signal of inequality (work in progress)

The 'irregular' teachers II

- High heterogeneity in their distribution across the country
  - Probably correlated with other indicators, signal of inequality (work in progress)

- Most of these irregular teachers have been trying the entry contest multiple times (83%)

Identification

Identification

$$y_{its} = \beta_0 + \beta_1 SNP_{its} + \beta_2 SNNP_{its} + \boldsymbol{\beta_k X_{kits}} + \alpha_{it} + \alpha_s + e_{its}$$

- within school, within school year variation in 'share new type'
  and in student test scores

Identification

$$y_{its} = \beta_0 + \beta_1 SNP_{its} + \beta_2 SNNP_{its} + \beta_k X_{kits} + \alpha_{it} + \alpha_s + e_{its}$$

- within school, within school year variation in 'share new type' and in student test scores
  - e.g. In school A in school year 2005, there were 50% of Math teachers were 'new type', but only 25% of Spanish teachers were. Compare performance of the *same students* in Math and Spanish that year

Identification

$$y_{its} = \beta_0 + \beta_1 SNP_{its} + \beta_2 SNNP_{its} + \boldsymbol{\beta_k X_{kits}} + \alpha_{it} + \alpha_s + e_{its}$$

- within school, within school year variation in 'share new type' and in student test scores

  - e.g. In school A in school year 2005, there were 50% of Math teachers were 'new type', but only 25% of Spanish teachers were. Compare performance of the *same students* in Math and Spanish that year

- eliminate confounders such as school characteristics, management, temporary shocks to the school, etc.

Results

Effect of New teachers on student performance

Table 3: The effect of New Regulation teachers on student performance

	(1)
Share New Regulation Passed	−0.35***
	(0.05)
Share New Regulation Not Passed	−1.83***
	(0.06)
Age	
Age^2	
Experience	
Experience^2	
Share postgrad degree	
Subject FE	✓
School FE	
Year FE	✓
School-year FE	
Subject-specific trends	
Mean(y)	43.33
sd(y)	3.29
N.obs	151,178
N.groups	.
R-squared	0.19

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. No fixed effects in columns (1) and (2), school fixed effects in column (3), school-year fixed effects in columns (4) and (5). * p<0.05, ** p<0.01, *** p<0.001

Table 3: The effect of New Regulation teachers on student performance

	(1)	(2)
Share New Regulation Passed	−0.35***	0.89***
	(0.05)	(0.07)
Share New Regulation Not Passed	−1.83***	−0.36***
	(0.06)	(0.08)
Age		0.03*
		(0.01)
Age^2		−0.00*
		(0.00)
Experience		0.12***
		(0.01)
Experience^2		−0.00***
		(0.00)
Share postgrad degree		0.94***
		(0.06)
Subject FE	✓	✓
School FE		
Year FE	✓	✓
School-year FE		
Subject-specific trends		
Mean(y)	43.33	43.33
sd(y)	3.29	3.29
N.obs	151,178	151,178
N.groups	.	.
R-squared	0.19	0.20

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. No fixed effects in columns (1) and (2), school fixed effects in column (3), school-year fixed effects in columns (4) and (5). * p<0.05, ** p<0.01, *** p<0.001

Table 3: The effect of New Regulation teachers on student performance

	(1)	(2)	(3)
Share New Regulation Passed	−0.35***	0.89***	0.12***
	(0.05)	(0.07)	(0.03)
Share New Regulation Not Passed	−1.83***	−0.36***	−0.03
	(0.06)	(0.08)	(0.04)
Age		0.03*	0.02**
		(0.01)	(0.01)
Age^2		−0.00*	−0.00***
		(0.00)	(0.00)
Experience		0.12***	0.02***
		(0.01)	(0.00)
Experience^2		−0.00***	−0.00***
		(0.00)	(0.00)
Share postgrad degree		0.94***	0.01
		(0.06)	(0.02)
Subject FE	✓	✓	✓
School FE			✓
Year FE	✓	✓	✓
School-year FE			
Subject-specific trends			
Mean(y)	43.33	43.33	43.33
sd(y)	3.29	3.29	3.29
N.obs	151,178	151,178	151,178
N.groups	.	.	5,969
R-squared	.	0.19	0.20	0.68

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. No fixed effects in columns (1) and (2), school fixed effects in column (3), school-year fixed effects in columns (4) and (5). * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 3: The effect of New Regulation teachers on student performance

	(1)	(2)	(3)	(4)
Share New Regulation Passed	−0.35***	0.89***	0.12***	0.20***
	(0.05)	(0.07)	(0.03)	(0.04)
Share New Regulation Not Passed	−1.83***	−0.36***	−0.03	0.14***
	(0.06)	(0.08)	(0.04)	(0.04)
Age		0.03*	0.02**	0.03***
		(0.01)	(0.01)	(0.01)
Age^2		−0.00*	−0.00***	−0.00***
		(0.00)	(0.00)	(0.00)
Experience		0.12***	0.02***	0.03***
		(0.01)	(0.00)	(0.00)
Experience^2		−0.00***	−0.00***	−0.00***
		(0.00)	(0.00)	(0.00)
Share postgrad degree		0.94***	0.01	0.02
		(0.06)	(0.02)	(0.02)
Subject FE	✓	✓	✓	✓
School FE			✓	✓
Year FE	✓	✓	✓	✓
School-year FE				✓
Subject-specific trends				
Mean(y)	43.33	43.33	43.33	43.33
sd(y)	3.29	3.29	3.29	3.29
N.obs	151,178	151,178	151,178	151,178
N.groups	.	.	5,969	29,609
R-squared	.	0.19	0.20	0.68	0.79

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. No fixed effects in columns (1) and (2), school fixed effects in column (3), school-year fixed effects in columns (4) and (5). * p<0.05, ** p<0.01, *** p<0.001

Table 3: The effect of New Regulation teachers on student performance

	(1)	(2)	(3)	(4)	(5)
Share New Regulation Passed	−0.35***	0.89***	0.12***	0.20***	0.20***
	(0.05)	(0.07)	(0.03)	(0.04)	(0.04)
Share New Regulation Not Passed	−1.83***	−0.36***	−0.03	0.14***	0.14***
	(0.06)	(0.08)	(0.04)	(0.04)	(0.04)
Age		0.03*	0.02**	0.03***	0.03***
		(0.01)	(0.01)	(0.01)	(0.01)
Age^2		−0.00*	−0.00***	−0.00***	−0.00***
		(0.00)	(0.00)	(0.00)	(0.00)
Experience		0.12***	0.02***	0.03***	0.03***
		(0.01)	(0.00)	(0.00)	(0.00)
Experience^2		−0.00***	−0.00***	−0.00***	−0.00***
		(0.00)	(0.00)	(0.00)	(0.00)
Share postgrad degree		0.94***	0.01	0.02	0.02
		(0.06)	(0.02)	(0.02)	(0.02)
Subject FE	✓	✓	✓	✓	✓
School FE			✓	✓	✓
Year FE	✓	✓	✓	✓	✓
School-year FE				✓	✓
Subject-specific trends					✓
Mean(y)	43.33	43.33	43.33	43.33	43.33
sd(y)	3.29	3.29	3.29	3.29	3.29
N.obs	151,178	151,178	151,178	151,178	151,178
N.groups	.	.	5,969	29,609	29,609
R-squared	.	0.19	0.20	0.68	0.79	0.79

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. No fixed effects in columns (1) and (2), school fixed effects in column (3), school-year fixed effects in columns (4) and (5). * p<0.05, ** p<0.01, *** p<0.001

Table 3: The effect of New Regulation teachers on student performance

	(1)	(2)	(3)	(4)
Share New Regulation Passed	−0.35***	0.89***	0.12***	0.20***
	(0.05)	(0.07)	(0.03)	(0.04)
Share New Regulation Not Passed	−1.83***	−0.36***	−0.03	0.14***
	(0.06)	(0.08)	(0.04)	(0.04)
Age		0.03*	0.02**	0.03***
		(0.01)	(0.01)	(0.01)
Age^2		−0.00*	−0.00***	−0.00***
		(0.00)	(0.00)	(0.00)
Experience		0.12***	0.02***	0.03***
		(0.01)	(0.00)	(0.00)
Experience^2		−0.00***	−0.00***	−0.00***
		(0.00)	(0.00)	(0.00)
Share postgrad degree		0.94***	0.01	0.02
		(0.06)	(0.02)	(0.02)
Subject FE				
School FE				
Year FE				
School-year FE				
Subject-specific trends				
Mean(y)	43.33	43.33	43.33	43.33
sd(y)	3.29	3.29	3.29	3.29
N.obs	151,178	151,178	151,178	151,178
N.groups	.	.	5,969	29,609
R-squared	0.19	0.20	0.68	0.79

Increasing the share of 'new type' teachers by 10% in a given school, subject and year, raises student performance, on average, by 0.9% of a s.d.

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. No fixed effects in columns (1) and (2), school fixed effects in column (3), school-year fixed effects in columns (4) and (5). * p<0.05, ** p<0.01, *** p<0.001

How large is this effect?

Considering student performance as an outcome:

- In our sample and estimates:
    - Equivalent to raising average teacher experience by 7 years

How large is this effect?

Considering student performance as an outcome:

- In our sample and estimates:
  - Equivalent to raising average teacher experience by 7 years

- Compared to US literature on teacher fixed effects:
  - Going from 100% traditional-type teachers to 100% new-type teachers in a subject is comparable, in the US, to getting a teacher that has 1 sd higher quality

Falsification

Table 4: Falsification: Share of New Regulation teachers on pre-reform student test scores

Student test scores 2000	2008	2009	2010	2011	2012	2013
Share New Regulation Passed	0.06	0.01	0.01	−0.04	−0.01	−0.02
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)
Share New Regulation Not Passed	0.02	0.03	0.03	−0.05	0.01	0.00
	(0.11)	(0.10)	(0.10)	(0.10)	(0.09)	(0.08)
N.obs	11,715	12,395	12,395	13,366	14,188	15,204
N.groups	2,595	2,697	2,697	2,828	2,945	3,057
Student test scores 2001	2008	2009	2010	2011	2012	2013
Share New Regulation Passed	−0.05	−0.01	−0.03	−0.08	−0.09	−0.01
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)
Share New Regulation Not Passed	−0.01	0.03	−0.02	−0.03	−0.08	−0.02
	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.08)
N.obs	12,103	13,077	12,831	13,818	14,694	15,769
N.groups	2,696	2,783	2,798	2,932	3,054	3,171

Note: Pre-reform student test scores regressed on each post-reform year's share of New Regulation teachers. SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 2000 or 2001. School and subject fixed effects, and all controls of Table 3 - model (5) are also included. * $p<0.05$, ** $p<0.01$, *** $p<0.001$
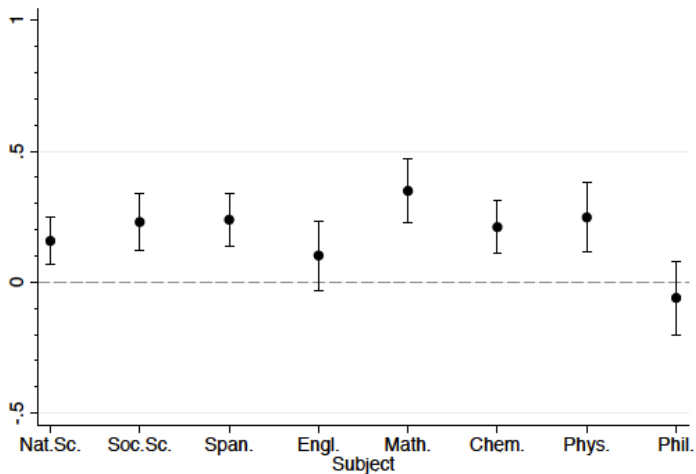
Exploring nonlinearities in the effect
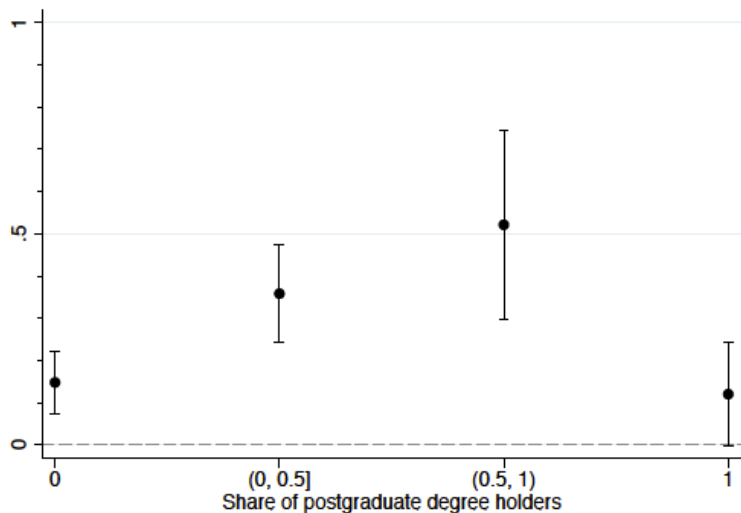
Nonlinearities 1
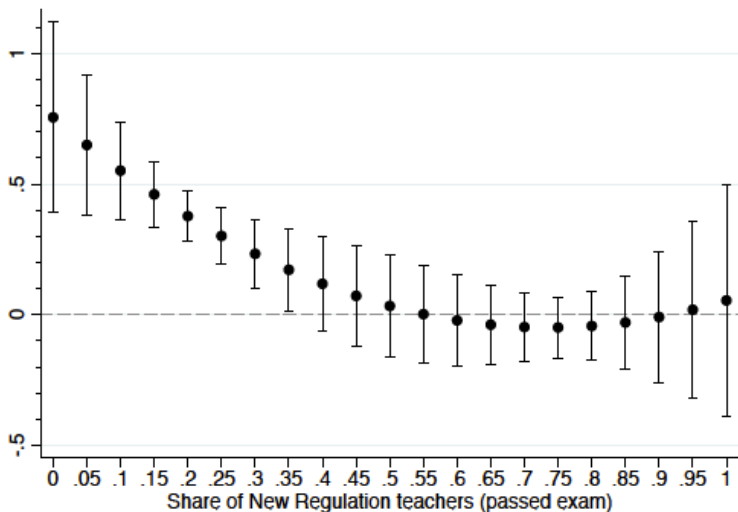
(a) Marginal effect at different levels of 'Share New Passed'

Nonlinearities 2

(e) Marginal effect at subjects

Nonlinearities 3

(d) Marginal effect at different shares of postgraduate degree holders

Nonlinearities summary

'New type' teachers seem to:

- yield decreasing marginal returns (largest effect for first entrants)
    - peer effects? spillovers?

- No significant differences by subject (Math a bit stronger, Philosophy and English a bit weaker)

- yield returns increasing in the share of postgraduate-holding colleagues
    - these colleagues are more receptive and open to novelties?

Exploring channels

Selection at entry

1) The exam is an active hurdle

Table A.1: Selectivity of the entry contests

CONTESTS >	1st (2004)	2nd (2005)	3rd (2006)	4th (2009)	5th (2013)
N. of local authorities	69	66	49	66	92
Vacancies	50.947	23.355	14.579	25.392	?
Candidates to exam	140.541	134.090	109.487	228.985	301.589
Passed exam stage	60.078 (43%)	32.720 (24%)	27.931 (26%)	66.687 (29%)	54.906 (18%)
Assigned to vacancy	30.568 (22%)	14.092 (11%)	13.620 (12%)	39.468 (17%)	?

Note: all percentages are relative to 'Candidates to exam'
Source: MEN [2013]

2) Teachers who score better at the exam 'make' better students

Table 5: Teacher's average entry exam score on student test scores

	(1)	(2)	(3)	(4)	(5)
Teacher exam score	0.083***	0.013***	0.017***	0.008***	0.008***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Share New					0.103*
					(0.048)
Experience				0.021***	0.027***
				(0.004)	(0.005)
Experience^2				−0.001***	−0.001***
				(0.000)	(0.000)
Postgrad degree (share)				0.012	0.013
				(0.034)	(0.034)
Subject-specific trends				✓	✓
Subject FE				✓	✓
School FE		✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
School-year FE			✓	✓	✓
Mean(y)	43.39	43.39	43.39	43.39	43.39
sd(y)	3.26	3.26	3.26	3.26	3.26
N.obs	106,552	106,552	106,552	106,552	106,552
N.groups		5,887	27,919	27,919	27,919
R-squared	0.03	0.52	0.66	0.81	0.81

Note: SE clustered by school in parentheses. Each observation is subject 's' in school 'i' in year 'y'. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

2) Teachers who score better at the exam 'make' better students + bias correction ▸ Bias problem

Table 6: Estimation of the true relationship between teacher quality and student performance

1) Lower and upper bound of the true β, based on the estimated $\hat{\beta}$

	(A) $var\left(\tilde{Q}^O\right) \simeq var\left(\tilde{Q}^N\right)$	(B) $var\left(\tilde{Q}^O\right) \approx 2var\left(\tilde{Q}^N\right)$	(C) $var\left(\tilde{Q}^O\right) \approx \frac{1}{2}var\left(\tilde{Q}^N\right)$
Lower bound	$\beta \approx 1.23\hat{\beta} \approx 0.010$	$\beta \approx 1.12\hat{\beta} \approx 0.009$	$\beta \approx 1.33\hat{\beta} \approx 0.011$
Upper bound	$\beta \approx 2.44\hat{\beta} \approx 0.020$	$\beta \approx 3.03\hat{\beta} \approx 0.024$	$\beta \approx 2.13\hat{\beta} \approx 0.017$

2) Estimated effect of increasing average teacher quality in a given school, year and subject

25th → 75th pctile	+3.2% to 6.4% s.d.	+ 2.9% to 7.7% s.d.	+ 3.5% to 5.5% s.d.
10th → 90th pctile	+ 7.0% to 13.9% s.d.	+ 6.3% to 16.7% s.d.	+ 7.7% to 11.9% s.d.

2) Teachers who score better at the exam 'make' better students + bias correction ▸ Bias problem

Table 6: Estimation of the true relationship between teacher quality and student performance

1) Lower and upper bound of the true β, based on the estimated $\hat{\beta}$

	(A) $var\left(\tilde{Q}^O\right) \simeq var\left(\tilde{Q}^N\right)$	(B) $var\left(\tilde{Q}^O\right) \approx 2var\left(\tilde{Q}^N\right)$	(C) $var\left(\tilde{Q}^O\right) \approx \frac{1}{2}var\left(\tilde{Q}^N\right)$
Lower bound	$\beta \approx 1.23\hat{\beta} \approx 0.010$	$\beta \approx 1.12\hat{\beta} \approx 0.009$	$\beta \approx 1.33\hat{\beta} \approx 0.011$
Upper bound	$\beta \approx 2.44\hat{\beta} \approx 0.020$	$\beta \approx 3.03\hat{\beta} \approx 0.024$	$\beta \approx 2.13\hat{\beta} \approx 0.017$

2) Estimated effect of increasing average teacher quality in a given school, year and subject

25th → 75th pctile	+3.2% to 6.4% s.d.	+ 2.9% to 7.7% s.d.	+ 3.5% to 5.5% s.d.
10th → 90th pctile	+ 7.0% to 13.9% s.d.	+ 6.3% to 16.7% s.d.	+ 7.7% to 11.9% s.d.

Selection on the probation period (?)

Table 5: Panel retention by type of position held (New Regulation teachers)

	Logit		RE Logit	
	(1)	(2)	(3)	(4)
Temporary position	0.25***	0.25***	0.19***	0.19***
	(0.00)	(0.00)	(0.00)	(0.00)
Probation period	0.77***	0.77***	0.78***	0.78***
	(0.03)	(0.03)	(0.03)	(0.03)
Age bins	No	Yes	No	Yes
N.obs	138,869	138,865	138,869	138,865
N.groups			48,172	48,171

Note: Odds ratios displayed. Outcome variable: Y=1 if the teacher is still in the panel the following year, 0 otherwise. Year 2013 excluded. Columns (1) and (2): SE clustered by individual in parentheses. Columns (3) and (4): Observed Information Matrix SE in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Turnover and discontinuation of employment (?)

Table 6: Panel retention per type of teacher regulation (permanent-position teachers)

	Logit			RE Logit		
	(1)	(2)	(3)	(4)	(5)	(6)
New Regulation	1.978***	1.516***	1.229***	2.409***	1.759***	1.313***
	(0.030)	(0.028)	(0.022)	(0.044)	(0.038)	(0.026)
Age		0.979***	0.972***		0.975***	0.971***
		(0.001)	(0.001)		(0.001)	(0.001)
Year dummies	No	No	Yes	No	No	Yes
N.obs	309,383	309,383	309,383	309,383	309,383	309,383
N.groups				94,285	94,285	94,285

Note: Outcome variable: Y=1 if the teacher is still in the panel the following year, 0 otherwise. Columns (1)-(3): SE clustered by individual in parentheses. Columns (4)-(6): Observed Information Matrix SE in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Conjectures on channels (so far)

- Selection at entry ✔

- Selection on probation ?

- Firing of permanent-position teachers ✘

Take-home message so far

- Colombian reform has yielded positive results in terms of student achievement
    - Around 0.9% within school-year standard deviations increase in test scores for every 10% extra new-type teachers

- National average might be affected by regional inequality

- Effects are likely to be driven by selection at entry, possibly by the probation period, but not by selection later in the career

- Teacher entry test scores seem to be a decent signal for teacher quality

- Large numbers of irregular teachers
    - Need to investigate why: real need or cheating?

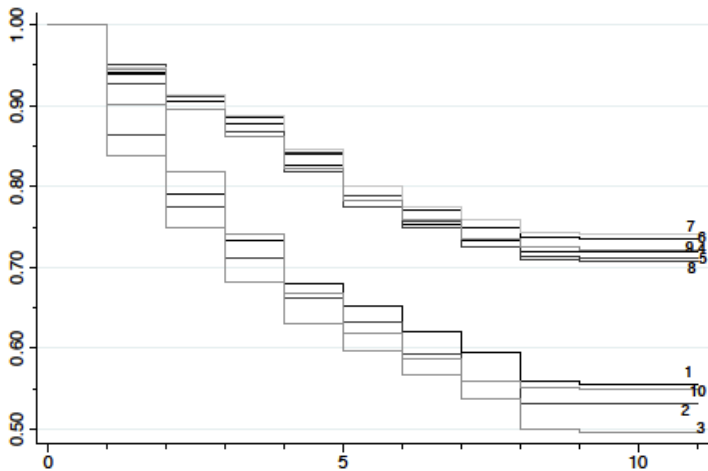Thank you!

Extras

Retention of new-type teachers

All teachers

Figure A.5: Kaplan Meier survival functions by entry test score deciles

(a) All teachers
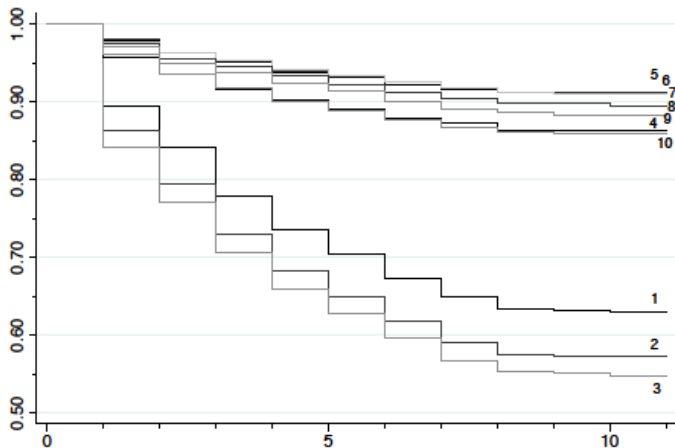
Retention of new–type teachers

Saber 11 teachers

(b) Saber 11 teachers

Literature review

Effectiveness of teacher quality insurance policies

- cite here

Teacher quality in Latin America

- cite here

Evaluation of 2002 Colombian reform

- Ome (2012, 2013): school fixed effects (across-year variation), finds no effect on high school students, while finding effects at primary school level

The bias problem
When trying to estimate the impact of (new-type) teacher quality on student performance

True population regression: student scores on total teacher quality

$$SS_i = \alpha + \beta Q_i^T + e_i$$

And $Q_i^T = Q_i^N \cdot S_i^N + Q_i^O \cdot \left(1 - S_i^N\right)$
What I am able to observe and estimate:

$$SS_i = \alpha + \beta Q_i^T + e_i$$

The bias problem

And we end up with:

$$\hat{\beta} = \beta \cdot \frac{cov\left(Q^N S^N, Q^N\right) + cov\left(Q^O\left(1 - S^N\right), Q^N\right)}{var\left(Q^N\right)}$$

Or, after controlling for the observables we have:

$$\hat{\beta} = \frac{cov\left(\widetilde{Q^N S^N}, \tilde{Q}^N\right) + cov\left(Q^O \widetilde{\left(1 - S^N\right)}, \tilde{Q}^N\right)}{var\left(\tilde{Q}^N\right)}$$

The bias problem

Bounds for the unobserved term

$$-1 \leq \frac{cov\left(Q^O \widetilde{(1 - S^N)}, \tilde{Q}^N \right)}{\sqrt{var\left(Q^O \left(1 - S^N\right)\right) var\left(\tilde{Q}^N \right)}} \leq 1$$

From which:

$$UB, LB = \pm \sqrt{\left[var\left(\tilde{Q}^O \right) var\left(\tilde{S}^N \right) + var\left(\tilde{Q}^O \right) E\left[\tilde{S}^N \right]^2 + var\left(\tilde{S}^N \right) E\left[\tilde{Q}^O \right]^2 \right] var\left(\tilde{Q}^N \right)}$$

The bias problem

Table A.12: Empirical values of bias components

	Component	Value in sample*
1.	$cov\left(\widetilde{Q^N S^N}, \tilde{Q}^N\right)$	≈ 26.22
2.	$var\left(\tilde{Q}^N\right)$	≈ 43.31
3.	$var\left(\tilde{S}^N\right)$	≈ 0.04
4.	$E\left[\tilde{S}^N\right]$	$\approx 2.61e^{-12}$
5.	$E\left[\tilde{Q}^N\right]$	$\approx -7.43e^{-10}$

* Sample used for main results (Table 3).