CrossMark

The impact of teacher credentials on student achievement in China

Jessica Hsiaochieh CHU [a], Prashant LOYALKA [a,e,*], James CHU [a], Qinghe QU [b], Yaojiang SHI [c], Guirong LI [d]

[a] Rural Education Action Program, Stanford University, United States
[b] Institute for Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, China
[c] Center for Experimental Economics for Education, Shaanxi Normal University, China
[d] Henan University, China
[e] National Research University Higher School of Economics, Russia

ARTICLE INFO

ABSTRACT

Teacher quality is an important factor in improving student achievement. As such, policymakers have constructed a number of different credentials to identify high quality teachers. Unfortunately, few of the credentials used in developing countries have been validated (in terms of whether teachers holding such credentials actually improve student achievement). In this study, we employ a student-fixed effects model to estimate the impact of teacher credentials on student achievement in the context of the biggest education system in the world: China. We find that having a teacher with the highest rank (a credential based on annual assessments by local administrators) has positive impacts on student achievement relative to having a teacher who has not achieved the highest rank. We further find that teacher rank has heterogeneous impacts, benefiting economically poor students more than non-poor students. However, whether a teacher attends college or holds teaching awards does not appear to provide additional information on teacher quality (in terms of improving student achievement).

A key finding in the economics literature is that education is important for economic growth (Hanushek & Woessmann, 2010; Schultz, 1961). Educational attainment (as measured by average number of years of schooling in a population) is consistently and positively related to economic growth (Aghion & Howitt, 1998; Romer, 1990). Studies have also documented a larger, positive relationship between student achievement (reflecting cognitive ability) and economic growth (Hanushek & Woessmann, 2008; Hanushek & Kimko, 2000). Indeed, researchers have argued that failing to raise national levels of student achievement can ultimately cause economies to stagnate (Hanushek & Woessmann, 2010, 2012).

An important factor in raising student achievement is teacher quality (Hanushek & Rivkin, 2010; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Studies in the United States have shown that a student improves three times more in his or her academic achievement when taught by a high quality teacher (relative to a low quality teacher—Hanushek, 2011). Although fewer in number, studies from developing countries have also documented how variation in teacher quality can lead to substantial differences in student achievement (e.g. Kingdon & Teal, 2010).

Recognizing the importance of teacher quality in raising student achievement, policymakers and researchers in developed countries have sought to identify the specific teacher credentials that signal teacher quality (and raise student achievement). For example,

* Corresponding author at: 5th Floor, Encina Hall, Stanford University, Stanford, CA 94305, United States. Tel.: +1 650 724 5302; fax: +1 650 725 1992.
E-mail addresses: jessica.chu1127@gmail.com (J.H. Chu), loyalka@stanford.edu (P. Loyalka), jchu1225@stanford.edu (J. Chu), qhqu.ccap@igsnrr.ac.cn (Q. Qu), syj8882002@yahoo.com.cn (Y. Shi), guirong1965@gmail.com (G. Li).

researchers have established that educational background (Harris & Sass, 2011; Kukla-Acevedo, 2009), professional certifications (for example, fulfilling professional teaching requirements set by a national agency—Clotfelter, Ladd, & Vigdor, 2007; Harris & Sass, 2009; Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006), and teaching experience (Rockoff, 2004; Ferguson & Ladd, 1996) raise student achievement in primary schools.

In contrast to research from developed countries, much less is known about the types of teacher credentials that positively impact student achievement in developing countries, their impact on different types of students, and how they translate to improvements in student achievement. Existing studies have estimated the impact of teacher subject-specific knowledge and union membership on student achievement (in Peru and India—Metzler & Woessmann, 2012; Kingdon & Teal, 2010). However, these two studies did not explicitly focus on those teacher credentials that are more often used by policymakers or school administrators in teacher hiring, assignment, and compensation decisions (such as educational background or professional certification). Moreover, to our knowledge, there have not been any studies in the context of developing countries that focus on measuring the differential impact of teacher credentials on the achievement of disadvantaged (poor or low-performing) students. Finally, even less is known about *how* such credentials might translate to improved student achievement (e.g. what the credentials tell us about the actual behavior of the teacher).

One country that especially could benefit from an evaluation of teacher credentials is China. China has one of the largest education systems, and hence one of the largest teacher labor forces, in the world. In 2012, approximately 5.6 million teachers taught nearly 100 million primary school students (MOE, 2014). About 42% of these teachers worked in rural areas (not including rural–urban transition areas—MOE, 2014). The size of the rural teacher labor force in China is noteworthy as it exceeds that of entire developed countries such as the United States (which had approximately 2 million primary school teachers in 2011–2012—NCES, 2013).

Surprisingly, almost no studies examine which teacher credentials actually improve student achievement in China. With the exception of one unpublished study (from over a decade ago—Park & Hannum, 2001), we know of no study from China that systematically examines which teacher credentials raise student achievement and to what degree. Moreover, we know of no study that examines the heterogeneous impacts of teacher credentials or how such credentials lead to improvements in student achievement. Lacking empirical evidence on which teacher credentials matter for student achievement, policymakers and school administrators in China (and other developing countries) may make inefficient decisions about teacher hiring, assignment, and compensation.

The goal of this paper is to examine which teacher credentials improve student achievement in China. As part of this goal, we have three specific objectives. First, we identify the impact of three commonly used teacher credentials (introduced in the section below) on the achievement of the average primary school student. Second, we consider whether there are heterogeneous impacts of teacher credentials on the achievement of economically poor and low-performing students (as well as by gender). Third, we explore whether "additional teacher effort" (the total amount of non-class time that teachers spend on their teaching activities, both inside and outside of school) is a plausible mechanism through which teacher credentials impacts student achievement.

To estimate the impacts of teacher credentials, we analyze data on approximately 4000 students from 70 rural primary schools in one province of Northwest China.[1] Similar to a number of recent studies on the effects of teacher credentials, our identification strategy relies on a cross-subject student-fixed effects model (Kingdon & Teal, 2010; Van Klaveren, 2011; Clotfelter, Ladd, & Vigdor, 2010; Dee, 2005, 2007). The model removes potential sources of bias that can result from the non-random sorting of students to teachers and schools. Our analytical models also control for a large set of cross-subject student and teacher covariates.

The rest of the paper is organized as follows. Section 2 provides a background on policies underlying teacher credentials in China, describes our data, and outlines our analytical approach. Section 3 presents results of the impact of teacher credentials on student achievement. In the analysis we examine the effect of teacher credentials on the average student as well as on disadvantaged (poor, female, or low-achieving) students. In Section 3, we also explore one potential mechanism by which teacher credentials may affect student achievement. Section 4 discusses the results and concludes.

1. Research design.

1.1. Teacher credentials in China

To identify teacher quality (and thus facilitate the effective hiring, assignment, and compensation of a large teacher labor force), policymakers and school district administrators in China have traditionally relied on three teacher credentials: teacher education background, teaching awards, and teacher rank (MoE, 1986). First, policymakers have traditionally used teacher education background to screen individual teachers for prerequisite training and ability (MoE, 1993a). For example, elementary school teachers are required to hold at least a vocational college degree as a pre-requisite to teach (with college degrees or higher considered favorably in hiring and promotion decisions—MoE, 1993a).

Second, policymakers have developed a system where teachers bestow teaching awards to one another based on results from classroom audits. Multiple award categories (such as excellence in classroom management or pedagogy) are usually given on an annual basis. Administrators and principals consider teachers who garner teaching awards more favorably when making promotion and hiring decisions (MoE, 1993b).

---

[1] We choose to focus on rural schools because improving student achievement in rural China is relevant for the country's future economic development. Rural China will provide a large share of the labor force in the coming years and there is a danger that China could fail to make the transition from developing to developed country if the average rural child is not being educated to support a modern, industrial society. Unfortunately, a number of studies already show that rural students are far behind their urban counterparts in a variety of educational outcomes (Yi et al., 2012; Mo, Zhang, Yi, Luo, Rozelle, & Brinton, 2013).

Third, policymakers rely on a system of teacher ranks (*zhicheng* or 职称) to make hiring and assignment decisions (MoE, 1986). Teacher rank (among primary school teachers) consists of four levels (in descending order of prestige): senior primary, level one, level two, level three. Teachers are routinely evaluated once every year (as part of standard practice in the education system in China) on whether they should be promoted in rank. As of 2011, 54% of primary school teachers had the senior primary rank. The remaining 46% of teachers either did not have a rank or were ranked at level one, two, or three (MoE, 2011).

To increase in rank, teachers must fulfill two sets of requirements. First, teachers must possess certain characteristics that are easily observed and measured. These are characteristics that are routinely recorded in administrative databases (sometimes termed "common measures"—Hanushek, 2002). Such characteristics include awards received, educational degree, years of teaching experience, and experience as a homeroom teacher (in years). Second, teachers are evaluated on their performance and behavior, which is less easy to observe and measure. For example, teachers are evaluated on whether they are able to employ a variety of pedagogical techniques, hold special talents or skills that are relevant in the classroom, demonstrate "exemplary moral and civic behavior," and carry a "persevering and disciplined" attitude (MoE, 1986).

As with education level and teaching awards, teacher rank is linked to monetary incentives. In much of the United States, salaries increases are mostly based on experience and educational degree (Podgursky & Springer, 2007). However, in China, beyond experience and one's education degree, salary increases are also based, in no small part on teacher rank (MoE, 1986). Teachers thus have (monetary) incentives to engage in activities that will increase their rank.

Policymakers and school district administrators in China rely on all three types of credentials—teacher education, teacher awards, and teacher rank—to make hiring and assignment decisions. However, it is an open empirical question and one of considerable policy relevance to assess whether these credentials positively impact student achievement. For example, it could very well be that some or all three credentials have no positive impact on student achievement and therefore should not be used to identify teacher quality in China (and should not be used in hiring, assignment and promotion decisions).

Moreover, we know little about how these credentials reflect quality (if at all). For example, as it turns out, teacher rank is an ambiguous measure. It is an aggregate of multiple factors (teacher characteristics and behaviors). We do not know how each factor is weighted when determining teacher rank in rural China. Thus, even if teacher rank does positively impact student achievement, our study further assesses whether it is an indicator that can be used systematically (on a system-wide basis) to compare teachers across rural China. If the positive impacts of teacher rank are based on factors that are observable to the policymaker, then hiring and assignment decisions can be made more reliably and on a system wide basis. If, on the other hand, the positive impacts of teacher rank rely on unobservable factors (such as school principals subjective assessments of the way in which teachers affect student outcomes), the indicator is much less reliable and cannot be used to compare teachers on a system-wide basis.

1.2. Sampling

The data for our study is from a sample of rural schools in Shaanxi province (a northwestern province of China). Specifically, our study sample was selected from a population frame of all primary schools in Ankang Prefecture in Shaanxi. We chose Ankang Prefecture primarily because we are interested in understanding what teacher credentials benefit students in poor, rural areas. Ranked 9th out of 10 prefectures in Shaanxi in terms of GDP per capita (2300 US dollars—NBS, 2013), Ankang is one of the poorest regions of China.

We sampled schools from the population frame in several steps. We first excluded 152 (out of a total of 252) elementary schools that did not have grades 1 through 6 or enrolled fewer than 30 students in total. These were often village schools without sufficient enrollment to open grades 5 or 6. We excluded such schools because policymakers informed us that these schools were at high risk of being closed or merged during the school year. We then excluded all schools located in county seats, as these tend to be richer schools that enroll non-rural students. The number of excluded schools enrolled less than 15% of the students in Ankang Prefecture. We then enrolled the remaining 70 schools to our sample.

In each of the 70 sample schools, we surveyed all students from grades 3 and 5. In particular, we administered a baseline survey to grade 3 and 5 students in June 2011 (in fact, when the students were finishing grades 2 and 4). One year later (at the end of grades 3 and 5, in June 2012), we conducted a follow-up survey with the same set of students. At this time, we also identified and surveyed all of the grade 3 and 5 mathematics and Chinese teachers. In other words, we surveyed all of the math and Chinese language teachers of the sample students. Altogether, we surveyed 3947 students (clustered in 127 classes) and 287 of their teachers.

1.3. Data collection

We collected information on our main outcome measure (student achievement) during the follow-up survey (in June 2012). In particular, we measured student achievement in math and Chinese subjects in both the baseline and follow-up waves by administering grade and subject-specific standardized tests.

To ensure the validity of the test items, these grade and subject-specific tests were created in the following three steps. First, we collected a pool of subject exam items from nationally and provincially mandated curricula and verified by experts in math and Chinese. Second, we piloted a pool of these exam items with more than 300 students. Third, drawing on results from the pilot and in consultation with experts, the final sets of test items for each test were selected to reflect the appropriate range and level of content for each subject and grade. In total, we created 8 standardized tests: grade 3 start-of-the-year (baseline) math and Chinese; grade 3 end-of-the-year (endline) math and Chinese; grade 5 start-of-the-year (baseline) math and Chinese; grade 5 end-of-the-year (endline) math and Chinese.

Our estimates show that each test is reliable. The Cronbach alpha coefficients for each test are high: .72 and .83 for the grade 3 base-line and follow up math tests; .81 and .83 for the grade 3 baseline and follow up Chinese tests; .85 and .81 for the grade 5 baseline and follow up math tests; and .83 and .69 for the grade 5 baseline and follow up Chinese tests.

We also administered the standardized exams ourselves. Students were given 30 min to finish each examination, which was proctored by our enumerators. For the purpose of our analysis, we standardized the resulting scores (by subtracting the mean and dividing by the standard deviation of the score distribution for each wave-grade-subject specific test).[2]

To obtain information about teacher credentials, we surveyed all mathematics and Chinese teachers that taught our sample students during the 2011–2012 school year. Specifically, we asked teachers to indicate which of four possible ranks: senior primary (the highest), level one, level two, or level three (the lowest) they held at the beginning of the school year. Because only a very small percentage of the teachers in our sample held the two lowest ranks, we created a dummy variable equal to 1 if a teacher held the highest rank and 0 otherwise. Note that the distribution of teacher rank in our sample is similar to the distribution of teacher rank in the population of teachers in China (NBS, 2014). In addition to teacher rank, we collected information on two other teacher credentials: whether a teacher ever received a teaching award (a dummy variable = 1, if the teacher had ever received a teaching award) and whether a teacher had a bachelor's degree (i.e. the variable = 1, if the teacher attended college).

We also collected information on student and teacher covariates. We first asked students to report the number of hours per week they spent in mathematics and Chinese courses in school (*class hours*). We also measured student achievement in both mathematics and Chinese during the baseline survey (in June 2011). Similar to the follow-up survey, we gave students 30-min tests in each subject and normalized the achievement measures (separately for each subject) as z-scores. Finally, we asked students about their gender and their families' ownership of different household assets. We used the information on household assets to construct a poverty index to distinguish between (economically) poor and non-poor students.[3]

We asked teachers about their work and background characteristics, such as age (in years), gender (whether female or not), county of origin (whether teachers attended primary school in the same county that they were currently teaching),[4] political affiliation (whether the teacher was a member of the Communist Party), teaching experience (a series of dummy variables indicating experience between two to five years, five to seven years, 10 to 15 years, 15 to 20 years, or 20 or more years of experience), and home-room teacher experience (in years). We also asked teachers to report hours worked on weekdays and weekends (beyond designated course hours) on teaching-related activities. Reported hours worked serves as our measure of additional teacher effort in our subsequent analyses.

1.4. Analytic strategy

Accurately estimating the causal impact of teacher credentials on student achievement is a challenge because the assignment of students to their teachers is non-random. This selection bias could lead researchers to under or overestimate the impact of teaching credentials. For example, higher-achieving students might be more likely to be placed with teachers with higher credentials, which would result in an upward bias when estimating the effect of teacher credentials on student achievement. If instead, lower-achieving students are more likely to be placed with teachers of higher credentials, an underestimation of the impact of teaching credentials could be expected Thus while the direction of the selection bias is ambiguous, the problem of estimating unbiased impacts of teacher credentials on student achievement remains significant.

In an attempt to estimate unbiased causal effects of teacher credentials on student achievement, researchers have used student fixed effects models. In particular, researchers have relied on a "cross-subject" student fixed effects model that removes the potentially confounding influence of unobservable, subject-invariant characteristics (e.g. Kingdon & Teal, 2010; Van Klaveren, 2011; Clotfelter et al., 2010; Dee, 2005, 2007). In effect, the method compares the same student's achievement level between two subjects (taught by two different teachers). If one teacher has a certain teacher credential and the other does not, the difference between a student's achievement in one subject versus the other may be considered a less biased estimator of the impact of the credential (as it controls away any individual characteristics of the student).

We use this type of model to estimate the impact of teacher credentials on student achievement. Following the approach of Clotfelter et al. (2010), we use within-student variation across language and mathematics subjects to identify the causal impacts. To illustrate how the cross-subject student fixed effects model removes the potentially confounding effects of unobservable, subject-invariant characteristics, we first examine the relationship between student achievement and teacher credentials using a standard linear regression model:

$$A_{is,t} = \alpha + \pi A_{is,t-1} + \beta T_{is} + \delta C_{is} + \lambda_i + \varepsilon_{is} \tag{1}$$

where $A_{is,t}$ is the (endline) achievement of student i in subject s at time t, $A_{is,t-1}$ is the (baseline) achievement of student i in subject s at time $t-1$; T_{is} is our treatment indicator for whether the teacher was of the highest teaching rank (1 = yes), had a teaching award (1 = yes), or attended college (1 = yes); and C_{is} is a vector of student and teacher covariates that vary across subjects s (within

---

[2] Notably, prefecture-level standardized tests are not administered by the local department of education (in our sample) to measure achievement in grades 3 and 5. Thus, none of the credentials that policymakers rely upon to make hiring and assignment decisions are based on standardized measures of student achievement.

[3] We created the poverty index using polychoric principal components analysis (Kolenikov and Angeles, 2009).

[4] Hannum and Park (2007) find that whether a teacher attended elementary school in the same area he or she teachers has strong positive correlations with their teaching quality and motivation.

Table 1
Average math and Chinese language teacher characteristics.
Source: *Author's Survey* (*Ankang Prefecture Teacher Survey*).

	(1) Math teachers	(2) Chinese teachers	(3) Difference (1) − (2)
Highest rank (y/n)	0.32	0.23	0.09
Received teacher award (y/n)	0.44	0.44	0.00
Attended college (y/n)	0.22	0.25	−0.03
Class hours	8.26	9.87	−1.61***
Teacher Age (in years)	37.53	35.59	1.94*
Female	0.52	0.66	−0.13**
Home county (y/n)	0.99	0.95	0.04*
Communist Party Member (y/n)	0.47	0.41	0.07
Experience (2–5 years)	0.05	0.05	−0.00
Experience (5–10 years)	0.14	0.17	−0.03
Experience (10–15 years)	0.34	0.38	−0.04
Experience (15–20 years)	0.14	0.12	0.02
Experience (20+ years)	0.3	0.24	0.06
Home room teacher (years)	11.81	12.49	−0.69
Number of teachers	141	146	

Notes:
1. Test of differences (column 3) are based on a two-tailed t-test
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

students) and that serve as control variables. In our study, C_{is} includes dummy indicators for class hours (number of hours per week spent in with a particular teacher), teacher age (in years), gender (1 = female), whether the teacher's county of origin was the same as his or her student's (1 = yes), whether the teacher was a Communist Party member (1 = yes), teacher experience (a series of dummy variables indicating experience between two to five years, five to seven years, 10 to 15 years, 15 to 20 years, or 20 or more years of experience), and homeroom teacher experience (in years).

There is no need to include any student individual characteristics because we can proceed to include the term λ_i, a student fixed effect. The student fixed effect term controls all student characteristics that do not vary across subjects. For example, the student characteristics also include family, school, and broader contextual characteristics (associated with the student) that do not differ across subjects. These effects are controlled for by the fixed effect. Finally, the symbol, ε_{is}, represents an error term that varies across both students and subjects.

The other terms (π, α, β, and δ) in Eq. (1) are coefficients (or vectors of coefficients) to be estimated. The coefficients reflect the relationship between the variables on the right hand side and student achievement on the left hand side. We are most interested in β, which identifies the relationship between teacher credentials and student achievement.

Because the student fixed effects (λ_i) are equivalent across both subjects, differencing Eq. (1) for the two subjects yields an equivalent Eq. (2) as follows:

$$\left(A_{i1,t} - A_{i2,t}\right) = \pi\left(A_{i1,t-1} - A_{i2,t-1}\right) + \beta(T_{i1} - T_{i2}) + \delta(C_{i1} - C_{i2}) + (\varepsilon_{i1} - \varepsilon_{i2}) \tag{2}$$

Unobserved student, teacher, and classroom characteristics that vary across subjects are captured in the differenced error term ($\varepsilon_{i1} - \varepsilon_{i2}$). To obtain unbiased estimates of β, this model relies on the assumption that the error term ($\varepsilon_{i1} - \varepsilon_{i2}$) in Eq. (2) must be uncorrelated with the treatment across the two subjects ($T_{i1} - T_{i2}$) and achievement($A_{i1} - A_{i2}$). We include ($A_{i1,t-1} - A_{i2,t-1}$) and the vector ($C_{i1} - C_{i2}$) to account for this possibility.[5]

That is, it could also be true that student characteristics vary by subject and correlate with the sorting of students and teachers. We make two adjustments to control for this selection problem. First, as noted in our model above, we conducted two waves of data collection and thus were able to control for baseline student achievement and the class average achievement of each individual student.

Second, we control for a series of teacher observable characteristics. Table 1 details the mean characteristics of math and Chinese teachers that are controlled for in our model. Mathematics (as a subject) tends to be taught for fewer times per week than Chinese. Mathematics teachers also tend to be about 2 years older than Chinese teachers (significant at the 10% level). Chinese teachers are more likely to be female (by 13 percentage points—significant at the 5% level), and math teachers are more likely to have grown up in the same county that they are teaching in currently (significant at the 10% level). Aside from these four differences, there are no other statistically significant differences across the other characteristics between the mean of language teachers and the mean of math teachers. This set of results in Table 1 suggests there is modest systematic sorting of teachers across language and math

---

[5] To further control for potential subject-specific characteristics in mathematics or Chinese that might be correlated with teacher rank (and student achievement), we also re-ran the model adding an indicator variable controlling for whether a teacher was a math or Chinese teacher (1 = math). The results were essentially identical to the ones presented in this paper.

subjects. Although the nature of differences between the teachers of the math and Chinese subjects is narrow, we control for all of these observable teacher characteristics in our analysis (since it is possible or even likely that the variables are systematically correlated with both teacher credentials and student achievement).

Finally, there are also some contextual reasons to believe that there is little sorting by subjects or schools: students usually go to local schools in rural areas (e.g. the distance to other schools reduces their ability to select other options). The rural schools we surveyed tended to have only a single class of students (they had smaller enrollments). As such, the entire class would be exposed to the same math and Chinese teacher, and we would not have to worry about teachers of higher rank having a larger say in the assignment process within school.

Taken together, this estimation strategy (like other studies employing cross subject variation across students, see e.g. Dee, 2005, 2007; Clotfelter et al., 2010; Kingdon & Teal, 2010; Van Klaveren, 2011; Schwerdt & Wupperman, 2011; Zakharov, Carnoy, & Loyalka, 2014) attempts to address endogeneity in three ways. First, it controls for a rich set of baseline (wave 1) characteristics including pre-test achievement scores to determine the impact on student test outcomes in the endline (wave 2). In this way, our strategy exploits the two-wave nature of the data. Second, the cross-subject nature of our data enables us to control for every observable and unobservable aspect of the student (including their parents, peers, school, and locality) that does not differ across subjects. Third, to further account for unobservable factors that may differ across subjects, even with the same student, we control for teacher and class (peer pre-test scores) characteristics that differ across subjects.

To examine the heterogeneity of treatment effects (between disadvantaged and advantaged students), we interact the treatment variables in Eq. (2) with student baseline characteristics. The student baseline characteristics include whether the student is:

Table 2
The impact of teacher rank on student achievement.
Source: *Author's Survey* (*Ankang Prefecture Teacher Survey*).

Dependent variable = student academic achievement, in standard deviations	(1)	(2)	(3)	(4)
Highest rank (y/n)	0.23***			0.20**
	(0.09)			(0.09)
Received teacher award (y/n)		0.07		0.05
		(0.06)		(0.06)
Attended college (y/n)			0.06	0.04
			(0.09)	(0.09)
Student baseline score	0.07	0.06	0.07	0.06
	(0.07)	(0.08)	(0.08)	(0.08)
Class hours	0.00	−0.00	−0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Teacher Age (in years)	0.06	0.02	0.06	0.06
	(0.20)	(0.20)	(0.21)	(0.20)
Teacher Gender (1 = female)	0.09	0.04	0.11	0.09
	(0.19)	(0.19)	(0.22)	(0.20)
Home county (1 = yes)	0.10	0.06	0.15	0.12
	(0.20)	(0.20)	(0.22)	(0.21)
Communist party affiliation (1 = yes)	−0.20	−0.18	−0.08	−0.17
	(0.23)	(0.23)	(0.27)	(0.26)
Experience 2 to <5 years	0.21***	0.21***	0.21***	0.21***
	(0.02)	(0.02)	(0.02)	(0.02)
Experience 5 to <10 years	0.02*	0.02*	0.02**	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
Experience 10 to <15 years	0.00	0.01*	0.01*	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Experience 15 to <20 years	0.16***	0.17***	0.16***	0.15***
	(0.06)	(0.06)	(0.06)	(0.06)
Experience 20+ years	0.12	0.09	0.07	0.12
	(0.18)	(0.17)	(0.17)	(0.18)
Homeroom Experience (years)	−0.21	−0.19	−0.09	−0.18
	(0.28)	(0.27)	(0.29)	(0.28)
Student Fixed Effects	YES	YES	YES	YES
Constant	−0.62*	−0.83**	−0.89**	−0.67*
	(0.33)	(0.34)	(0.34)	(0.35)
Observations	7894	7894	7816	7816
R-squared	0.076	0.068	0.068	0.078

Notes:
1. Standard errors (adjusted for clustering at the school level) in parentheses.
2. Student baseline score is measured by a 30-min standardized test (for mathematics and Chinese) administered and proctored by the research team. Number of class hours refers to the number of hours per week students spend in mathematics or Chinese courses in school (self-reported by students). Home county indicates that the teacher attended primary school in the same county that he or she is currently teaching.
*** p < 0.01.
** p < 0.05.
* p < 0.1.

(a) economically poor (ranking at the bottom third in terms of the household wealth index we constructed); (b) female; or (c) low-achieving (ranking at the bottom third in terms of achievement according to our tests). The coefficient on the interaction term reflects the degree to which the achievement of disadvantaged and advantaged students is differentially affected by teacher rank.

Finally, to estimate the relationship between teacher effort and teacher credentials, we employ a standard linear regression model:

$$E_t = \alpha + \pi T_t + \beta C_t + \varepsilon_t \tag{1}$$

where E_t is the total amount of non-class time that a teacher t spent on his or her teaching activities, both inside and outside of school; T_t is our treatment indicator for whether the teacher was of the highest teaching rank (1 = yes), had a teaching award (1 = yes), or attended college (1 = yes); and C_t is a vector of teacher characteristics. The vector C_t includes all of the same control variables that were used in our models above. Because our goal is to measure the relationship between teacher characteristics and effort, this model only relies on data from our teachers (and not students).

2. Results

According to our results, teacher rank has a positive and significant impact on the achievement of the average student (Table 2, Column 1). Specifically, the results show that teacher rank increases student achievement by 0.23 standard deviations (Table 2, Column 1). The result is statistically significant at the 1% level. By contrast, the results show that the separate, direct impacts of teaching awards and attending college are not statistically different from zero.

When the three treatment credentials (teacher rank, teaching awards, and attending college) are included together, the results show that only teacher rank increases student achievement (by 0.20 standard deviations—Table 2, column 4). The adjusted result is significant at the 5% level. By contrast, teacher awards and educational degrees have no statistically significant effect on student achievement.

Taken together, the results suggest that teachers with the highest rank are able to increase student learning more than teachers with lower ranks. In addition, whether a teacher attends college or holds teaching awards does not appear to provide additional information on teacher quality (in terms of improving student achievement). In other words, the significant impact of teacher rank

Table 3
Impact of teacher credentials on poor versus non-poor student achievement.[***]
Source: Author's Survey (*Ankang Prefecture Teacher Survey*).

Dependent variable = student academic achievement, in standard deviations	(1)	(2)	(3)	(4)
Highest rank (y/n)	0.19[**]			0.16
	(0.09)			(0.10)
Highest rank * poorest (y/n)	0.12[**]			0.12[*]
	(0.06)			(0.07)
Received teacher award (y/n)		0.07		0.06
		(0.06)		(0.06)
Teacher award * poorest (y/n)		0.02		−0.02
		(0.05)		(0.06)
Attended college (y/n)			0.08	0.06
			(0.09)	(0.09)
Attended college * poorest (y/n)			−0.05	−0.05
			(0.06)	(0.06)
Covariates controlled	Yes	Yes	Yes	Yes
Student fixed effects	Yes	Yes	Yes	Yes
Constant	−0.61[*]	−0.82[**]	−0.90[**]	−0.67[*]
	(0.33)	(0.34)	(0.34)	(0.35)
Observations	7894	7894	7816	7816
R-squared	0.077	0.068	0.068	0.080

Notes:
1. Standard errors (adjusted for clustering at the school level) in parentheses.
2. Covariates controlled are the same as those in Table 2 (student baseline score, teacher number of class hours, teacher age (in years), teacher gender (1 = female), whether the teacher is in his or her home county (1 = yes), teacher Communist party affiliation (1 = yes), Experience 2 to <5 years, Experience 5 to <10 years, Experience 10 to <15 years, Experience 15 to <20 years, Experience 20+ years, number of years the teacher has been a homeroom teacher).
3. Student baseline score is measured by a 30-min standardized test (for mathematics and Chinese) administered and proctored by the research team. Number of class hours refers to the number of hours per week students spend in mathematics or Chinese courses in school (self-reported by students). Home county indicates that the teacher attended primary school in the same county that he or she is currently teaching.
[***] p < 0.01.
[**] p < 0.05.
[*] p < 0.1.

holds even after controlling for whether a teacher attends college or holds teaching awards. Similarly, there are no significant impacts of a teacher attending college or holding teaching awards, whether or not we control for teacher rank.

While these results hold for the sample on average, do certain subgroups of students benefit differentially from teacher rank, awards, or college attendance? As it turns out, teachers with highest rank appear to have a positive and statistically significant impact on economically poor students. When all three treatment variables are considered together in the full model, the results indicate that teacher rank improves student achievement by 0.12 standard deviations (significant at the 10% level—Table 3, column 4). In sum, the results suggest that teachers with highest rank have a significant impact on the achievement of poor as opposed to non-poor students. However, there is no statistically significant evidence to suggest that teaching awards and having a college degree have an impact on the achievement of poor students.

In terms of gender, neither teacher rank, teaching awards nor college degree have any additional positive impact on female student achievement. Having a teacher with the highest rank, in fact, seems to reduce female student achievement (relative to male student achievement) by 0.07 standard deviations (Table 4, column 1). Importantly, however, this finding is not statistically significant. The same is true for teaching awards and college attendance—there are no statistically significant differential impacts of any teacher credentials on female student achievement (columns 2–3). When the credentials are all included in the model together, the results remain consistent: none of the credentials seem to yield differential, positive benefits on female student achievement.

As with gender, the analysis shows that none of the teacher credentials yield additional benefits for low achieving students. That is, the unadjusted results continue to indicate that having a teacher with the highest teacher rank increases the achievement of students by 0.23 standard deviations (significant at the 5% level—Table 5, column 1, row 1). However, the coefficient for the interaction term between teacher rank and the dummy variable indicating low-achieving students suggests that teacher rank may reduce modestly the achievement of low-achieving students by 0.02 standard deviations (Table 5, column 1). The story remains consistent for teacher awards and college attendance. None of the coefficients are statistically significant (columns 2–3). In addition, even when the three treatment variables are considered together, none of the results are statistically different from zero (column 4).

In summary, the evidence suggests that the only teacher credential that increases the academic achievement of the average student is teacher rank. Economically poor students especially benefit from teachers who have high teacher rank. By contrast, none of the other two credentials (teaching awards and a college degree) appear to have a positive impact on student achievement.

Why is it that teacher rank has an impact but not the other two credentials examined? In an attempt to understand how teacher rank translates into student achievement, we further examine the relationship between teacher rank and additional teacher effort (Table 6). We do so because we hypothesize that teachers who hold the highest rank are also those who are particularly diligent.

Table 4

Impact of teacher credentials on female versus male student achievement.
Source: *Author's Survey* (*Ankang Prefecture Teacher Survey*).

Dependent variable = student academic achievement, in standard deviations	(1)	(2)	(3)	(4)
Highest rank (y/n)	0.26***			0.24**
	(0.09)			(0.10)
Highest rank ∗ female student (y/n)	−0.07			−0.09
	(0.05)			(0.06)
Received teacher award (y/n)		0.06		0.03
		(0.06)		(0.06)
Teacher award ∗ female student (y/n)		0.03		0.06
		(0.06)		(0.07)
Attended college (y/n)			0.03	0.01
			(0.09)	(0.09)
Attended college ∗ female student (y/n)			0.08	0.07
			(0.07)	(0.07)
Covariates controlled	Yes	Yes	Yes	Yes
Student fixed effects	Yes	Yes	Yes	Yes
Constant	−0.62*	−0.83**	−0.89**	−0.68*
	(0.33)	(0.34)	(0.34)	(0.35)
Observations	7894	7894	7816	7816
R-squared	0.076	0.069	0.069	0.080

Notes:
1. Standard errors (adjusted for clustering at the school level) in parentheses.
2. Covariates controlled are the same as those in Table 2 (student baseline score, teacher number of class hours, teacher age (in years), teacher gender (1 = female), whether the teacher is in his or her home county (1 = yes), teacher Communist party affiliation (1 = yes), Experience 2 to <5 years, Experience 5 to <10 years, Experience 10 to <15 years, Experience 15 to <20 years, Experience 20+ years, number of years the teacher has been a homeroom teacher).
3. Student baseline score is measured by a 30-min standardized test (for mathematics and Chinese) administered and proctored by the research team. Number of class hours refers to the number of hours per week students spend in mathematics or Chinese courses in school (self-reported by students). Home county indicates that the teacher attended primary school in the same county that he or she is currently teaching.
*** p < 0.01.
** p < 0.05.
* p < 0.1.

Table 5
Impact of teacher credentials on low–achieving students.
Source: Author's Survey (Ankang Prefecture Teacher Survey).

Dependent variable = student academic achievement, in standard deviations	(1)	(2)	(3)	(4)
Highest rank (y/n)	0.23**			0.21**
	(0.09)			(0.10)
Highest rank ∗ low achievement (y/n)	− 0.02			− 0.05
	(0.08)			(0.09)
Received teacher award (y/n)		0.06		0.03
		(0.06)		(0.06)
Teacher award ∗ low achievement (y/n)		0.05		0.09
		(0.08)		(0.08)
Attended college (y/n)			0.04	0.02
			(0.08)	(0.08)
Attended college ∗ low achievement (y/n)			0.08	0.09
			(0.10)	(0.09)
Covariates controlled	Yes	Yes	Yes	Yes
Student fixed effects	Yes	Yes	Yes	Yes
Constant	− 0.62*	− 0.83**	− 0.92***	− 0.72**
	(0.33)	(0.34)	(0.35)	(0.35)
Observations	7894	7894	7816	7816
R-squared	0.076	0.069	0.069	0.079

Notes:
1. Standard errors (adjusted for clustering at the school level) in parentheses.
2. Covariates controlled are the same as those in Table 2 (student baseline score, teacher number of class hours, teacher age (in years), teacher gender (1 = female), whether the teacher is in his or her home county (1 = yes), teacher Communist party affiliation (1 = yes), Experience 2 to <5 years, Experience 5 to <10 years, Experience 10 to <15 years, Experience 15 to <20 years, Experience 20+ years, number of years the teacher has been a homeroom teacher).
3. Student baseline score is measured by a 30-min standardized test (for mathematics and Chinese) administered and proctored by the research team. Number of class hours refers to the number of hours per week students spend in mathematics or Chinese courses in school (self-reported by students). Home county indicates that the teacher attended primary school in the same county that he or she is currently teaching.

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

That is, they put in additional hours of work during the weekends or after their formal duties conclude during the day. If it is true that teachers with the highest rank do in fact spend more time on their work than other teachers, teacher rank may serve as a signal for hard-working teachers who invest their energies into improving student achievement.

When we examine results from our regression between teacher credentials and effort, we find that teacher rank has no significant relationship with additional teacher effort. Compared with teachers who have not yet achieved highest rank, teachers with the highest rank appear to put in 1.26 extra hours of work per week (Table 6, Row 1). However, this difference is not statistically significant. Although the point estimates suggest that teachers with a teaching award or a college degree reduced their effort by 0.23 or 0.44 h (Row 2 and Row 3, respectively), neither result is statistically significant. In other words, there is no evidence that any of the teacher characteristics considered in our study positively relate to teacher effort.

While we cannot be sure, one potential implication of the null relationship between increased effort and teacher rank is that having the highest rank signals more effective teaching practices (relative to teaching awards or a college degree). That is, teachers with the highest rank do not seem to invest increased effort in their students, but their students still demonstrate increased gains in achievement. If so, teachers with high ranks are more effective at translating their existing efforts into student gains in achievement.

3. Discussion and conclusion.

In this study, we used a student-fixed effects model to estimate the impact of teacher credentials on student achievement in the context of the biggest education system in the world—China. We find that having a teacher with the highest rank in fact does improve student achievement (relative to having a teacher who has not achieved the highest rank) by approximately 0.20 standard deviations. We further find that teacher rank especially benefits economically poor students. However, we found no evidence that teacher rank is associated with increased teacher effort (which we hypothesized was the channel by which teacher rank would improve student achievement). In addition, once accounting for teacher rank, whether a teacher attends college or holds teaching awards does not appear to provide additional information on whether a teacher can improve student achievement.

Granted, there is considerable regional variation in the standards on which the granting of teacher credentials (and teacher ranks in particular) is based. For example, some places may account for academic degrees and student test scores while others may not. Moreover, the weights attached to each item may differ by township or county. Our sample is representative of poor prefecture in Northwest China. As such, we do not have sufficient regional variation in evaluation standards. Nevertheless, this study is the first

Table 6
The association between teacher credentials and (extra) teacher effort.
Source: *Author's Survey* (*Ankang Prefecture Teacher Survey*).

Dependent variable: extra work hours	(1)
Highest rank (y/n)	1.26
	(0.83)
Received teacher award (y/n)	−0.23
	(0.53)
Attended college (y/n)	−0.44
	(0.61)
Student baseline score	−0.89
	(0.79)
Class hours	−0.26***
	(0.07)
Teacher Age (in years)	−0.11
	(0.08)
Teacher Gender (1 = female)	−1.53**,*
	(0.59)
Home county (1 = yes)	−1.57
	(1.04)
Communist party affiliation (1 = yes)	0.10
	(0.64)
Experience 2 to <5 years	1.89
	(1.29)
Experience 5 to <10 years	−0.11
	(1.07)
Experience 10 to <15 years	1.19
	(0.99)
Experience 15 to <20 years	1.89
	(1.35)
Experience 20+ years	0.08
	(1.80)
Homeroom Experience (years)	0.14***
	(0.05)
Constant	13.45***
	(2.76)
Observations	266
R-squared	0.139

Notes:
1. Standard errors (adjusted for clustering at the school level) in parentheses.
2. The dependent variable (extra work hours) refers to the total number of non-class hours that teachers report spending on their teaching activities, both inside and outside of school.
3. Student baseline score is measured by a 30-min standardized test (for mathematics and Chinese) administered and proctored by the research team. Number of class hours refers to the number of hours per week students spend in mathematics or Chinese courses in school (self-reported by students). Home county indicates that the teacher attended primary school in the same county that he or she is currently teaching.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

to examine the causal impacts of teacher qualifications on student achievement in China. Future work should consider regional variation in evaluation standards (and the cases under which the standards lead to increased student achievement).

One significant implication of our results is that local policymakers and school administrators in China appear to be able to (at least to a degree) identify less observable components of teacher quality (such as teacher attitudes, practices, or performance). That is, even after controlling for other teacher credentials and characteristics in our adjusted model (e.g. teaching awards and education level), teacher rank continues to have a significant positive impact on student achievement. As noted above, teacher rank is evaluated on two dimensions: (a) observable teacher credentials (such as teacher awards, education level, and teaching experience) and (b) less readily observable signs of teacher quality including teacher performance and behavior in the classroom. Because our analyses control for (observable) teacher credentials and characteristics, it appears that the impact of teacher rank on student achievement is primarily due to (unobservable) teaching performance and behavior. Indeed, we also discovered that teacher rank is uncorrelated with increased effort (also an observable behavior). This finding is consistent with the interpretation that teacher rank captures other important—but less readily observable—components of teacher quality.

What does this mean for policymakers? Policymakers/administrators must make hiring, assignment, and compensation decisions based on the information that is available. The teacher rank system appears, at least to some degree, to distinguish between higher and lower quality teachers (who are improving the achievement of their students). However, the fact that teacher rank reflects

unobservable factors (such as principals' subjective assessments) means that it cannot be used to compare teachers on a system-wide basis. Nonetheless, the fact that school administrators have access to other information that can be used to determine teacher quality to make human resource decisions is nontrivial. Indeed, this information can be used within the school or locally. To the best of our knowledge, our study is the first to explore this important issue in China.

However, there are also clear drawbacks to using teacher rank as a primary indicator of teacher quality. For example, we note that the majority of our sample of rural teachers was categorized into only one of two ranks (either the highest rank or the next highest rank). We conjecture that two ranks provide little information to help policymakers/administrators differentiate between teachers of varying quality, reducing the usefulness of rank as a signal of quality teachers.

References

Aghion, P., & Howitt, P. (1998). *Endogenous growth theory*. Cambridge, MA: MIT Press.
Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy, 1*(2), 176–216.
Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources, 45*(3), 655–681.
Clotfelter, C. 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.
Dee, T. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review, 95*(2), 158–165.
Dee, T. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources, 42*(3), 528–554.
Ferguson, R., & Ladd, H. (1996). How and why money matters: An analysis of Alabama schools. In H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC: Brookings Institution.
Hannum, E., & Park, A. F. (2007). *Education and reform in China*. New York, NY: Routledge.
Park, A., & Hannum, E. (2001). Do teachers affect learning in developing countries? *Paper presented at the rethinking social science research on the developing world in the 21st century conference for the social science research council, salt lake city, UT, 2001.*
Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor force quality, and the growth of nations. *American Economic Review, 90*(5), 1184–1208.
Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 267–271.
Hanushek, E A and Woessmann L. (2010). Education and economic growth. In: Penelope Peterson, Eva Baker, Barry McGaw, (Editors), International encyclopedia of education. volume 2, pp. 245–252. Oxford: Elsevier.
Hanushek, E. A., & Woessmann, L. (2012). Schooling, educational achievement, and the Latin American growth puzzle. *Journal of Development Economics, 99*(2), 497–512.
Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature, 46*(3), 607–668.
Hanushek, E. A. (2002). Teacher quality. In L. T. Izumi, & W. M. Evers (Eds.), *Teacher quality* (pp. 1–12). Stanford, Calif.: Hoover Institution Press.
Hanushek, E.A. (2011). The economic value of higher teacher quality. Economics of Education Review 30(30), 466–479.
Harris, D. N., & Sass, T. R. (2009). The effects of NBPTS-certified teachers on student achievement. *Journal of Policy Analysis and Management, 28*(1), 55–80.
Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*(7), 798–812.
Kingdon, G., & Teal, F. (2010). Teacher unions, teacher pay and student performance in India: A pupil fixed effects approach. *Journal of Development Economics, 91*(2), 278–288.
Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth, 55*(1), 128–165.
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.
Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics, 99*(2), 486–496.
Ministry of Education of the People's Republic of China. (1986). Regarding the Implementation of Credentials and Ranks in Hiring Practices, State Council Document #27. Last Accessed Feb., 2014. <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s7567/201308/156436.html>
Ministry of Education of the People's Republic of China. (1993a). Criteria for Teacher Awards, Ministry of Education Document #38. Last Accessed Feb., 2014. <http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_619/200407/1314.html>
Ministry of Education of the People's Republic of China. (1993b). National Legal Standards for Public Teachers in the People's Republic of China. Last Accessed Feb., 2014. <http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s7382/201305/152500.html>
Ministry of Education of the People's Republic of China (2011). *Compulsory Education Math Curriculum Standard*. Beijing: Beijing Normal University Press.
Ministry of Education of the People's Republic of China. (2014). Number of Full-time Teachers in Regualr Primary Schools by Rank and Age. Last Accessed Feb., 2014. <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s7567/201308/156436.html >.
Mo, D., Zhang, L., Yi, H., Luo, R., Rozelle, S., & Brinton, C. (2013). School Dropouts and Conditional Cash Transfers: Evidence from a Randomised Controlled Trial in Rural China's Junior High Schools. *The Journal of Development Studies, 49*(2), 190–207.
China National Bureau of Statistics (2013). *China Statistical Yearbook*. Beijing: China Statistics Press.
China National Bureau of Statistics (2014). *China Statistical Yearbook*. Beijing: China Statistics Press.
National Center for Education Statistics, U.S. Department of Education (2013d). *Digest of Education Statistics, 2012 (NCES 2014-015)*.
Podgursky, M. J., & Springer, M. G. (2007). Teacher Performance Pay: A Review. *Journal of Policy Analysis and Management, 26*(4), 909–949.
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (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.
Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, S71–S102.
Schultz, T. W. (1961). Investment in human capital. The American Economic Review, 51(1), 1–17
Schwerdt, G., & Wupperman, A. C. (2011). Is traditional teaching all that bad? A within-student between subject approach. *Economics of Education Review, 30*(2), 365–379.
Van Klaveren, C. (2011). Lecturing style teaching and student performance. *Economics of Education Review, 30*(4), 729–739.
Yi, H., Zhang, L., Luo, R., Shi, Y., Mo, D., Chen, X., ... Rozelle, S. (2012). Dropping out: Why are students leaving junior high in China's poor rural areas? *International Journal of Educational Development, 32*(4), 555–563.
Zakharov, A., Carnoy, M., & Loyalka, P. (2014). Which teaching practices improve student performance on high-stakes exams? Evidence from Russia. *International Journal of Educational Development, 26*, 13–21.