The Han-Minority Achievement Gap, Language, and Returns to Schools in Rural China

YUNFAN YANG
Institute for Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences; and University of the Chinese Academy of Sciences

HUAN WANG
Northwest University of Xi'an

LINXIU ZHANG
Institute for Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences

SEAN SYLVIA
Renmin University of China

RENFU LUO
Institute for Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences

YAOJIANG SHI
Center for Experimental Economics for Education, Shaanxi Normal University

WEI WANG
Center for Chinese Agricultural Policy, Chinese Academy of Sciences

SCOTT ROZELLE
Stanford University

I. Introduction

Over the past 3 decades, China's rural population has experienced rapid income growth and a dramatic reduction in poverty (Ravallion and Chen 2007;

The authors acknowledge financial support from the National Institutes of Health (R01HL106023-03), the Chinese Academy of Sciences (KZZD-EW-06-02), the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences (2012ZD008), and the Stanford Center for International Development. Contact the corresponding author, Renfu Luo, at luorf.ccap@igsnrr.ac.cn.

Electronically published November 18, 2014

Chen and Ravallion 2010). Although the well-being of the population as a whole has risen sharply, the average economic standing of country's nearly 114 million ethnic minorities has improved relatively less than that of the Han majority (Gustafsson and Li 2003; Gustafsson and Sai 2009a, 2009b). Between 1988 and 1995 the average per capita income of Han living in rural areas increased by more than 52%, while incomes of the rural minority population only grew by nearly 22% (Gustafsson and Li 2003). Over this period, the Han-minority income gap nearly doubled from 19.2% to 35.9% (Gustafsson and Li 2003). In 2002, rural minorities remained more than one and a half times as likely as the rural Han majority to be in poverty and twice as likely to have experienced poverty in the past 2 years (Gustafsson and Sai 2009a; Hannum and Wang 2012).

Lagging educational attainment among minorities has undoubtedly played a significant role in the persistence of the Han-minority income gap (Hannum and Wang 2012). Education is an increasingly important determinant of wages and access to off-farm employment (Zhang et al. 2005; De Brauw and Rozelle 2008). At the same time, educational attainment among minorities lags (Hannum 2002; Hannum et al. 2008; Hannum and Wang 2012). Analyzing mid-census survey data from 2005, Hannum and Wang (2012) find that—among 16–21-year-olds—minorities were nearly one-third as likely as Han to have attained 9 years of compulsory schooling. Minorities are also significantly less likely to enroll at the tertiary level. In a 2008 census of entering freshman at four tier 1 universities in western China, only 4% were non-Han, while minorities comprise approximately 11% of the population cohort (Wang et al. 2013). The same survey shows female minority students to be at a particular disadvantage in college admissions: in this entering class, female minority students were only 25% of their population share.

If the Han-minority differences in educational attainment persist, the relative well-being of minority populations is likely to continue to fall as China's economy increasingly demands a higher-quality workforce. Tightening demographics and a nearly complete transition into off-farm labor in China (more than 80% of 16–30-year-olds are now employed off farm) are driving up wages for unskilled labor at close to 10% per year (Park, Cai, and Du 2010; Zhang et al. 2013). As unskilled wages rise and low-paying basic manufacturing jobs are replaced with jobs involving more sophisticated tasks, China's economy will increasingly demand a high-quality, educated workforce (Zhang et al. 2013). Educationally disadvantaged minorities will find it more difficult to participate in this new labor market and benefit from the higher wages that will come with it.

In the context of rural China, poor academic performance in school may play a significant role in reducing educational attainment or years of schooling (Yi et al. 2012). In competitive educational systems—such as China's—lower expectations of poorly performing students to thrive in the system may discourage continued enrollment (Chuang 1997; Clarke, Haney, and Madaus 2000; Reardon and Galindo 2002; Rumberger and Lim 2008). Heavy emphasis on testing may further lead teachers to direct more attention to higher-performing children and even lead schools to push at-risk students out in an effort to raise overall test scores (Vickers 1994; Vélez and Saenz 2001; Fortin et al. 2006). These influences are compounded by rising unskilled wages, which drive up the opportunity costs of schooling (Angrist and Lavy 2009; Fiszbein, Schady, and Ferreira 2009). Indeed, the available evidence highlights the correlation between poor performance and dropout among poor students in western China (Yi et al. 2012). Thus, if minority students perform worse than their Han peers, they are likely to attain fewer years of schooling as they forgo school and opt to enter the labor force in unskilled jobs.

Despite the implications of an achievement gap between Han and minority students, no study that we know of has compared their achievement (either grades or test scores). Likewise, we find almost no research on the how the determinants of achievement may vary between the two groups. Existing empirical work on the disparity between Han and minority educational outcomes has focused on attainment. Hannum (2002), for example, using a 1992 national survey of children in China, finds large differences in enrollment between Han and minority children of primary school age, with enrollment rates lowest among minorities in western China. She concludes that much of this difference is attributable to geographic composition and family background. Research like this, however, is focused on attainment and, presumably due to the absence of data, has not examined achievement.

The overall goal of this article is to document and analyze the achievement gap between Han and minority students in rural China. To meet this goal we have two specific objectives. First, we estimate the overall achievement gap (henceforth, the "Han-minority achievement gap"). We also measure two other subgaps: the gap between Han and minority students that speak Mandarin as a first language and the gap between Han and minority student that speak Mandarin as a second language. Second, we assess what factors contribute most to these achievement gaps. To do this, we first decompose the achievement gap into two parts: one part representing the portion of the gap due Han-minority differences in endowments of student, household, peer, teacher, and school characteristics and a second part due to differences in returns to these char-

acteristics. We then asses what effect specific schools have on the Han-minority achievement gap and what types of schools narrow or widen this gap. That is, we analyze how returns to attending specific schools (school fixed effects) differ between Han and minority students and what school characteristics are most strongly associated with these Han-minority differences in returns to specific schools.

To achieve these objectives, we draw on a large-scale survey of schools sampled from across rural Shaanxi, Gansu, and Qinghai provinces covering nearly 21,000 students, approximately 13% of whom are minorities. We measure achievement of Han and minority students using curriculum-based standardized exams in math and Chinese given as part of the survey. To assess factors that contribute to the Han-minority achievement gap, we use detailed information on students, households, teachers, and schools and apply decomposition methods pioneered by Oaxaca (1973) and Blinder (1973). Oaxaca-Blinder type decomposition, originally used to analyze wage differences between groups, has now been applied in a wide variety of contexts. In education, previous research has used this approach to analyze differences in academic achievement across countries (e.g., McEwan and Marshall 2004; Ammermueller 2007), across time (Barrera-Osorio et al. 2011), and between indigenous and nonindigenous students (McEwan 2004; McEwan and Trowbridge 2007; Sakellariou 2008).

Our analysis yields three primary findings. First, we find that minority students in our sample score significantly below Han students on standardized exams in math and Chinese. The Han-minority achievement gap is nearly 0.3 standard deviations (SD) in math and more than 0.2 SD in Chinese. Among minorities in our sample whose primary language is not standard Mandarin (Salar and Tibetan—henceforth, "Non-Mandarin minorities"), the achievement gap is even more striking: these students score 0.62 SD lower than Han in math and 0.65 SD lower than Han in Chinese.

Second, our decomposition analysis suggests that the Han-minority achievement gap for Mandarin-speaking minority students (Hui and Tu—henceforth, "Mandarin minorities") is almost fully explained by differences in student, peer, teacher, and school characteristics. Of these, the largest contributor is student and family background. Differences in school quality play a relatively small role. Endowments, however, explain very little of the achievement gap between Han students and non-Mandarin minorities.

Third, we find that—in "mixed" schools with both Han and minority students—the effects of individual schools play a role in widening the Han-minority achievement gap. In these mixed schools, returns to Han students of (observed and unobserved) specific school attributes are higher than those for

similar minority students attending the same school. Teachers appear to play a central role in affecting the relative returns of Han and minority students.

The rest of this article is organized as follows. Section II reviews the background of minorities in China. Sections III and IV describe the survey and data that we use for the analysis. Sections V and VI discuss the empirical approach and results. The final section concludes and discusses the policy implications of our findings.

II. Background: The Education of Minorities in China

In addition to the Han majority, there are 55 officially recognized minority nationalities in China. According to the 2010 census, minorities comprised 8.5% of the total national population, approximately 114 million people (Cherng, Hannum, and Lu 2012). Geographically, minorities in China are concentrated in relatively poor regions of western China: 71.6% of the minority population lives in western provinces, and 91.6% of ethnic autonomous counties are located in western China. Approximately 40% of these autonomous counties are nationally designated poverty counties (Hannum and Wang 2012).

Beyond geographically targeted antipoverty funds (from which minorities disproportionately benefit due to concentration in poor areas; Park, Wang, and Wu 2002), a number of policies and programs have aimed to expand access to education among minority groups. For example, the 1980 Law on Regional and Ethnic Autonomy recommended subsidization of education in minority areas beyond standard educational funding (Cherng et al. 2012). More recently, as part of the Tenth 5-Year Plan (2001–5), the central government invested approximately ¥34.2 billion for boarding schools and ethnic universities in western China and minority areas (Cherng et al. 2012). A number of affirmative action policies have also been implemented in higher education, such as university admissions spots reserved for minority students and acceptance of minority students with lower entrance exam scores (Hannum and Wang 2012).

Although certain policies have been designed to improve educational attainment among minorities, the structure of education for minority groups is largely similar to the rest of the country (Cherng et al. 2012). Curriculum and assessment are generally the same for minority and Han students (Cherng et al. 2012). One exception is the language of instruction. While official policy regarding language of instruction emphasizes the use of Mandarin, schools with more than 50% minority students who speak a local language are permitted to use the local language (Cherng et al. 2012). In practice, however, there are significant challenges to bilingual instruction. For example, some

minority groups with their own language do not have a written language. In addition, there seldom are financial resources available to develop a local language curriculum. Schools also are often integrated with students attending class with Han students or students belonging to other minority groups (Hannum and Wang 2012). In our survey of schools in northwest China (described in the next section), only 5% have no Han students. No schools provide instruction or teaching material in minority languages.

III. Survey Design

The data used in this study come from a survey of 300 schools in Shaanxi, Gansu, and Qinghai provinces in western China during the 2011/2012 academic year. Schools were sampled as follows. We first obtained a list of all schools in the following regions: Haidong (in Qinghai), Longnan, Dingxi, Tianshui (in Gansu), and Ankang (in Shaanxi). A map of these regions is provided in figure 1. In total, 26 counties were included in the sampling frame. Within each township located in these five regions, one school was selected from among all schools with 150–300 students as reported by the local education ministry. The survey is thus roughly representative of primary schools in these regions of northwestern China.

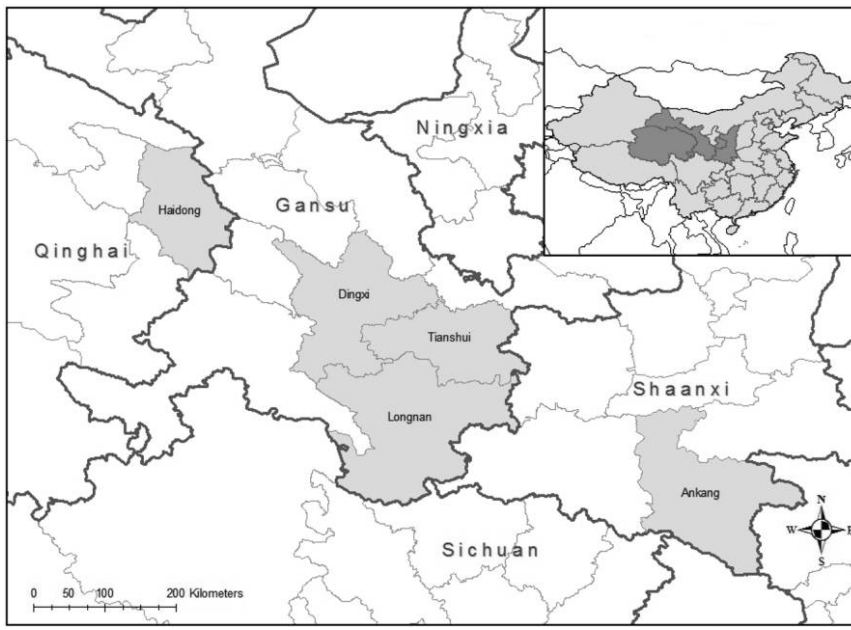

Figure 1. Survey regions

Due to the survey's geographical coverage, our sample includes both completely Han (37%) and completely minority schools (5%). A significant number of schools (58%) are mixed Han and minority schools. We focus most of the study on the full sample but, in some parts of the analysis, restrict the sample to mixed schools only.[1] In the analysis, we define "mixed" schools as schools with at least two minority and two Han students. By restricting the analysis to mixed schools, we are better able to pick up differences between Han and minority students not confounded by differences in location; however, limiting the sample both limits variation in the data (leading to less precise parameter estimates) and reduces representativeness.

Within each school, we collected information on all fourth and fifth graders (more than 21,000 students in total). A survey questionnaire administered to students collected detailed information on students and their families. Table 1 lists all additional variables that we use in our analysis and provides descriptions of each. All of these variables were asked or measured at the beginning of the school year.

As our measure of academic achievement, we use student scores on standardized exams in math and Chinese administered by the survey team at the end of the school year. Within each classroom, half of the students were randomly assigned to take a math exam, and the rest took a Chinese exam. To ensure coherence with the national curriculum, the tests were developed with assistance from local department/bureaus of education. Questions used in the math exam were drawn from the question bank of the Trends in International Mathematics and Science Study, an international assessment of mathematics and science knowledge of primary and lower-secondary school students. Questions used in the Chinese exam were taken from national fourth or fifth grade textbooks. To minimize cheating, two versions of each exam (with reordered questions) were randomly assigned to students. The exam also was proctored closely by the enumerators. For analysis, scores for both subject tests are normalized by the distribution of scores in each grade. Exams were given in Mandarin, just like year-end tests usually given in the schools in our sample.

IV. Characteristics of Students, Peers, Teachers, and Schools

A. Minority Status

We solicited minority status directly from students as part of the survey. Out of the full sample, 12.5% of students identified themselves as belonging to a

[1] Similar strategies of dealing with differences in location in decomposition analysis has been used in previous studies (see, e.g., van de Walle and Gunewardena 2001).

TABLE 1

VARIABLE DESCRIPTIONS

Variable	Description
Student and household characteristics:	
Standardized math exam score	Normalized score on standardized math exam. Exam designed using grade-appropriate questions from the Trends in International Mathematics and Science Study with assistance from the Chinese Ministry of Education.
Standardized Chinese exam score	Normalized score on standardized Chinese exam. Exam designed using questions from national curriculum with assistance from the Chinese Ministry of Education.
Female (0/1)	Student is female.
Boarding student (0/1)	Student boards at school.
Age (years)	Student age in years
Household size	Total number of individuals living the in the student's household
Travel time to school (minutes)	Travel time from student's home to school in minutes
Mother has lower secondary degree or above (0/1)	Student's mother has completed middle school education or above.
Father has lower secondary degree or above (0/1)	Student's father has completed middle school education or above.
Father at home (0/1)	Father currently living at home (has not migrated for work)
Mother at home (0/1)	Mother currently living at home (has not migrated for work)
Household asset index (0/1)	Index of household durable assets. Constructed using first principal component of motorbike, tractor, car, van, refrigerator, air conditioning, computer, laundry machine, and dummy variables for type of housing (cave house, packed earth, brick, apartment building, other)
Class peer characteristic:	
Proportion of peers' mothers with lower secondary degree or above	Class level mean of "mother has lower secondary degree or above" excluding student *i*
Proportion of class peers of same ethnicity	Proportion of students in class of same ethnic background as student *i*
Peer average household asset index	Class level mean of "household asset index" excluding student *i*
Teacher characteristic:	
Female teacher (0/1)	Teacher is female
Han teacher (0/1)	Teacher is Han majority
Teacher has higher education degree (0/1)	Teacher has completed college or above
Teacher attended normal college (0/1)	Teacher attended normal school
Teacher has received provincial or national teaching award (0/1)	Teacher has received a provincial or national level teaching award
Gongban teacher (0/1)	Teacher is a regular teacher, not on a short-term contract.
Teacher experience (years)	Teacher years of teaching experience
School characteristic:	
School size (students)	Number of students in school
Student-teacher ratio	Student-teacher ratio
Distance to farthest village served by school (minutes)	Travel time to the farthest village in school's catchment area

TABLE 1 (*Continued*)

Variable	Description
School has provided teacher training in past year (0/1)	School has provided training to teachers in past year
School infrastructure index	Index of school infrastructure constructed using first principal component of number of classrooms, library, garden, school wall, cafeteria, playground, number of computers for student use

minority group. Table 2 shows the distribution of each of the five main ethnic groups in the sample by province. Of all minority students, 56.5% are Hui, 12.3% are Salar, 13.5% are Tibetan, and 17.0% are Tu.[2]

For comparison, table 2 also gives the ethnic composition of 10-year-olds for the counties in our sample from the 2010 national census. The composition found in the school survey largely mirrors census data. Our survey covers a slightly larger proportion of minorities overall (0.8% more) and a slightly larger proportion of Tu and Salar and smaller proportion of Hui and Tibetans. Note that some difference is to be expected given time trends and that the school survey covers a wider age range.

B. The Achievement Gap

According to our data, there is a significant achievement gap between Han and minority students (fig. 2). Leftmost bars show the mean standardized exam scores in math (dark gray) and Chinese (light gray) for Han students; the next pair of bars shows mean scores for all minority students; and the remaining bars show mean scores by minority group. The gap between all minority students (all ethnic groups pooled together) and Han students is substantial: 0.29 SD in math and 0.25 SD in Chinese.

The data also show a striking amount of heterogeneity in exam scores among individual minority groups. For example, students from the Tu minority perform comparably to Han students. In contrast, the scores of Salar students are nearly 0.75 SD below those of the Han students. Importantly, figure 2 suggests that language may be a factor contributing to China's Han-minority achievement gap. The students from the two minority groups that typically speak non-Mandarin languages (Salar and Tibetan) perform much worse than Han students. At the same time, the achievement gap between students from the two minority groups that generally speak Mandarin as their primary language (Tu and Hui) and Han students is much narrower. Given substantial differences between the achievement of Mandarin-speaking and

[2] And 0.75% belong to other minority groups. We exclude the other category from the analysis given their small number.

TABLE 2
SAMPLE COUNTY ETHNIC COMPOSITION IN SCHOOL SURVEY AND 2010 CENSUS (%)

	Han	Hui	Tibetan	Tu	Salar
Census data:					
Gansu	95.74	3.95	.28	.00	.00
Shaanxi	99.16	.80	.00	.00	.00
Qinghai	40.62	23.44	13.00	7.72	15.00
Total	88.33	7.96	1.79	1.43	.45
School survey data:					
Gansu	92.50	5.65	.65	.24	1.09
Shaanxi	99.25	.63	.06	.00	.00
Qinghai	46.38	21.47	9.97	16.46	5.72
Total	87.53	7.04	1.69	2.12	1.54

Sources. 2010 census data (China Statistics Press, 2012) and authors' survey.

non-Mandarin-speaking minority students, we analyze these two minority groups separately in addition to analyzing the pooled sample of all minority students.

C. *Endowments of Background Characteristics*

The statistics in table 3 highlight some significant differences between Han and minority students in terms of student and household characteristics. First, minority students from both categories are significantly older than Han stu-

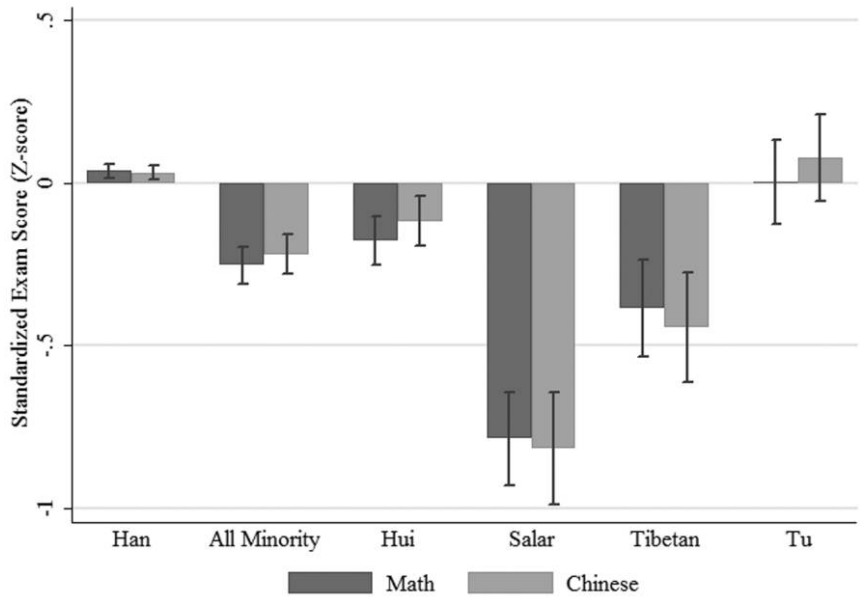

Figure 2. Standardized exam results by ethnic group. Uses all observations in the data set. "Other minority" group excluded from graph due to small sample size. Error bars give 95% confidence intervals constructed using 500 bootstrap replications accounting for clustering at the school level.

dents by around 0.2–0.3 years (table 3). In our sample, Han students were more likely to have repeated a grade compared to all minority groups; thus, this age difference likely reflects longer delays in primary school enrollment on the part of minorities.[3] Available evidence from other countries suggests that delayed enrollment may have a positive influence on academic achievement (Glewwe, Jacoby, and King 2001; McEwan and Shapiro 2008); however, delayed enrollment may be due to malnutrition in early childhood (Glewwe and Jacoby 1995; Glewwe et al. 2001). Second, minority students live in significantly larger households, likely a reflection of differential treatment under family planning policies. Given evidence that there is a strong quality-quantity trade-off in rural China, having more siblings may disadvantage minority students (Li, Zhang, and Zhu 2008). The third significant difference between Han and minority students in our sample is that parents of minority children are significantly less educated themselves. Numerous studies from a variety of contexts have shown evidence that parental education—particularly the mother's—has a causal influence on the academic achievement of children. Interestingly the one area in which minority students appear to be unambiguously better off is in terms of household asset ownership (although this could reflect cheaper prices in regions where minorities are likely to live).

Table 3 also shows differences in class peer characteristics. Minorities attend classes with peers whose mothers are significantly less educated but whose families possess more durable household assets compared to Han students. They also attend classes with a significantly smaller proportion of peers of the same ethnicity as their own. To examine the distribution of Han and minority students across schools in more detail, figure 3 plots kernel density estimates of this variable. These plots clearly show that Han students are much more concentrated in ethnically homogenous schools than are minority students. Nearly 33% of minority students are in the ethnic minority of their class, while this figure is only 1% for Han students.

A priori it is unclear what affect peer ethnic composition may have on student achievement for minority and Han students. Minorities (and their minority peers) are of generally lower socioeconomic status; however, there may be advantages to attending school with peers of the same ethnicity. Beyond theories related to social identity (Akerlof and Kranton 2002), non-Mandarin minority students may benefit from classes in which teachers are more likely to teach (entirely or partly) in the local language.[4]

[3] In our sample, 39.5% of Han students repeated a grade. This is significantly more likely than for Hui (7 percentage points, $p = .04$) and Tu (11 percentage points, $p = .01$) students.
[4] In our sample, no teachers report doing so.

TABLE 3
SUMMARY STATISTICS

	Full Sample (1)	Han Students (2)	Minority Students: All (3)	Minority Students: Mandarin Speaking (4)	Minority Students: Non-Mandarin Speaking (5)	Difference (4) − (2) [*P*-Value]	Difference (5) − (2) [*P*-Value]
Student and household characteristics:							
Standardized math exam score	.00	.04	−.25	−.14	−.58	−.17	−.62
	(1.00)	(1.00)	(.98)	(.97)	(.93)	[.00]	[.00]
Standardized Chinese exam score	.00	.03	−.22	−.07	−.62	−.10	−.65
	(1.00)	(.99)	(1.06)	(1.00)	(1.11)	[.11]	[.00]
Female (0/1)	.49	.49	.48	.48	.50	−.01	.01
	(.50)	(.50)	(.50)	(.50)	(.50)	[.45]	[.62]
Boarding student (0/1)	.07	.07	.10	.08	.16	.02	.09
	(.26)	(.25)	(.30)	(.28)	(.36)	[.58]	[.19]
Age (years)	10.83	10.80	11.03	11.01	11.08	.21	.28
	(1.13)	(1.14)	(1.04)	(1.08)	(.90)	[.00]	[.00]
Household size	5.32	5.27	5.64	5.63	5.68	.35	.40
	(1.55)	(1.52)	(1.71)	(1.64)	(1.92)	[.00]	[.01]
Travel time to school (minutes)	25.72	25.91	24.34	23.85	25.63	−2.07	−.29
	(26.26)	(26.55)	(24.03)	(22.90)	(26.71)	[.29]	[.94]
Mother has lower secondary degree or above (0/1)	.27	.28	.19	.21	.14	−.08	−.15
	(.44)	(.45)	(.39)	(.40)	(.34)	[.00]	[.00]
Father has lower secondary degree or above (0/1)	.46	.48	.36	.37	.32	−.11	−.16
	(.50)	(.50)	(.48)	(.48)	(.47)	[.00]	[.00]
Father at home (0/1)	.58	.58	.55	.55	.52	−.03	−.06
	(.49)	(.49)	(.50)	(.50)	(.50)	[.27]	[.18]
Mother at Home (0/1)	.68	.69	.66	.67	.63	−.02	−.05
	(.46)	(.46)	(.47)	(.47)	(.48)	[.48]	[.14]
Household asset index (0/1)	.00	−.05	.36	.30	.53	.35	.58
	(1.44)	(1.43)	(1.48)	(1.45)	(1.53)	[.00]	[.00]

Class peer characteristic:							
Proportion of peers' mothers with lower secondary degree or above	.27	.28	.20	.22	.14	−.06	−.14
	(.17)	(.17)	(.16)	(.17)	(.10)	[.02]	[.00]
Proportion of class peers of same ethnicity	.93	.96	.68	.70	.64	−.26	−.32
	(.18)	(.10)	(.36)	(.36)	(.36)	[.00]	[.00]
Peer average household asset index	.00	−.05	.35	.28	.56	.33	.61
	(.63)	(.62)	(.61)	(.54)	(.74)	[.00]	[.00]
Teacher characteristic:							
Female teacher (0/1)	.40	.39	.49	.48	.51	.09	.11
	(.49)	(.49)	(.50)	(.50)	(.50)	[.10]	[.18]
Han teacher (0/1)	.90	.96	.51	.51	.48	−.44	−.48
	(.30)	(.20)	(.50)	(.50)	(.50)	[.00]	[.00]
Teacher has higher education degree (0/1)	.81	.80	.89	.89	.89	.09	.09
	(.39)	(.40)	(.32)	(.31)	(.31)	[.00]	[.05]
Teacher attended normal college (0/1)	.79	.78	.88	.88	.89	.10	.11
	(.40)	(.41)	(.32)	(.32)	(.31)	[.01]	[.01]
Teacher has received provincial or national teaching award (0/1)	.08	.08	.10	.11	.10	.03	.02
	(.27)	(.26)	(.30)	(.30)	(.30)	[.38]	[.72]
Gongban teacher (0/1)	.87	.88	.81	.82	.78	−.06	−.09
	(.34)	(.33)	(.39)	(.39)	(.41)	[.32]	[.09]
Teacher experience (years)	13.39	13.53	12.45	12.16	13.11	−1.38	−.42
	(10.96)	(11.11)	(9.86)	(10.16)	(8.79)	[.29]	[.74]
School characteristic:							
School size (students)	221.05	221.47	218.08	224.03	201.71	2.59	−19.74
	(59.25)	(58.65)	(63.25)	(65.51)	(53.59)	[.85]	[.10]
Student-teacher ratio	17.41	17.29	18.28	18.50	17.63	1.21	.34
	(5.11)	(5.17)	(4.55)	(4.57)	(4.36)	[.14]	[.72]
Distance to farthest village served by school (minutes)	66.83	67.70	60.77	57.56	69.62	−10.14	1.92
	(50.63)	(50.75)	(49.39)	(39.84)	(68.78)	[.17]	[.92]

TABLE 3 (*Continued*)

	Full Sample (1)	Han Students (2)	Minority Students			Difference (4) − (2) [*P*-Value]	Difference (5) − (2) [*P*-Value]
			All (3)	Mandarin Speaking (4)	Non-Mandarin Speaking (5)		
School has provided teacher training in past year (0/1)	.92	.95	.77	.77	.77	−.17	−.18
	(.26)	(.23)	(.42)	(.42)	(.42)	[.04]	[.12]
School infrastructure index	.00	.01	−.07	.13	−.61	.12	−.62
	(1.21)	(1.18)	(1.38)	(1.16)	(1.77)	[.55]	[.22]
Sample size:							
Total number of students	19,129	16,741	2,388	1,753	617		
Number of students—math sample	9,468	8,286	1,182	871	301		
Number of students—Chinese sample	9,661	8,455	1,206	882	316		
Number of schools	300	285	191	167	64		

Note. Variables are as described in table 1. Standard errors (in parentheses) account for clustering at the school level.

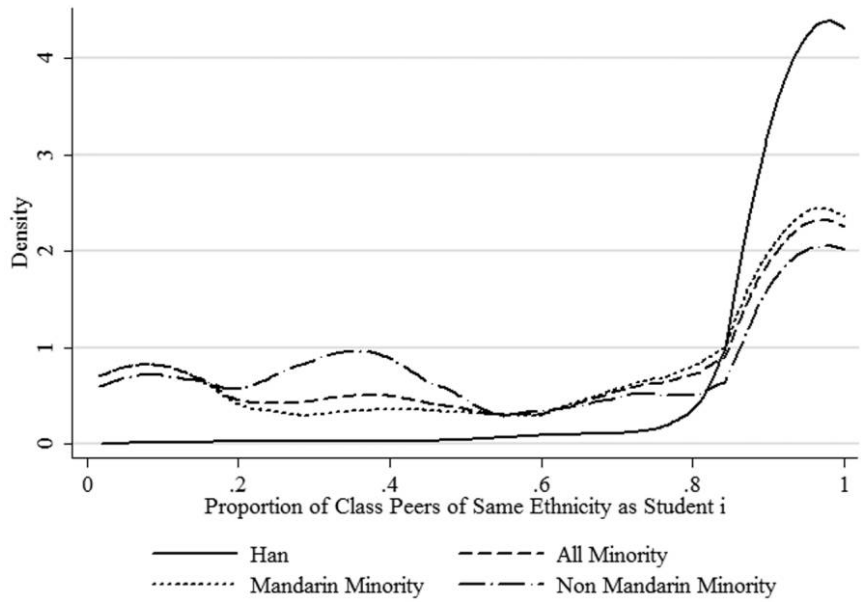

Figure 3. Distribution of class peer ethnic composition by ethnic group. Kernel density estimated using a bandwidth of 0.07.

In terms of teacher and school quality (table 3), minorities appear to be, if anything, better off than their Han counterparts. For example, teachers of minority students are significantly more likely to have a higher education degree and to have attended a specialized teaching college. Schools attended by Han and minority students are similar in terms of size, student-teacher ratio, remoteness, and infrastructure (although minority schools are slightly less likely to have provided teacher training in the past year). This may be a result of significant government educational investment focused on minority areas.

Characteristics of students in mixed Han and minority schools (with at least two Han and two minority students) are given in the appendix (table A1). We construct mixed school samples for both types of minority students. Mixed Mandarin minority schools have at least two Mandarin minority students, and mixed non-Mandarin minority schools have at least two non-Mandarin minority students. For the most part, mean differences in characteristics between Han and minority students attending the same schools are less significant than the full sample, as would be expected.

V. Returns to Minority Status

We take a first look at the relationship between the achievement gap and observed characteristics directly by estimating how the Han-minority achieve-

ment gap changes as we adjust for characteristics collected as part of our survey. That is, we estimate variants of the following regression

$$Y_{is} = \alpha + \beta_1 \text{Mandarin Minority} + \beta_2 \text{Non\_Mandarin Minority} + X'\theta + \varepsilon_{is}, \quad (1)$$

where Y_{is} is the normalized test score of student i in schools; *Mandarin Minority* is a dummy variable equal to 1 if the student is Hui or Tu; *Non_Mandarin Minority* is a dummy variable equal to 1 if the student is Tibetan or Salar; X' is a vector of student and household, peer, teacher, and school characteristics; and ε_{is} is an error term possibly correlated at the school level. The coefficients of interest are β_1 and β_2. How these two coefficients change as we add characteristics to the X' vector from the error term provides a first look at the ability of these characteristics to account for differences in achievement between Han students and Mandarin and non-Mandarin minority students.

The results of this analysis for standardized math scores are shown in table 4.[5] The raw mean differences are −0.17 SD for Mandarin minority students and −0.62 SD for non-Mandarin minority students (col. 1). Controlling for student and household characteristics reduces the size of these estimates to −0.12 and −0.51 SD, respectively (col. 2). Sequentially adding peer, teacher, and school characteristics (cols. 3–5) shows that once student and peer characteristics are controlled for, the Mandarin minority coefficient decreases in size and becomes insignificant. The coefficient on non-Mandarin minorities remains large (−0.2 SD) and significant even after controlling for school fixed effects (col. 6). In other words, Mandarin minority students score an average of 0.2 SD less than Han students with similar individual, peer, and teacher characteristics in the same schools. Adding school fixed effects (i.e., controlling for all observed and unobserved school-level characteristics) does reduce the estimated gap for this group by more than half, which suggests that—despite detailed controls—unobserved school-level heterogeneity is an important factor.

VI. Decomposing the Han-Minority Achievement Gap

To decompose the Han-minority achievement gap, we first estimate educational productions functions, or *achievement regressions*, that quantify returns to individual, family, teacher, and school-level characteristics for each of our student classifications (Han, Mandarin minority, and non-Mandarin minority). We then use the traditional Oaxaca-Blinder decomposition method

[5] Results for Chinese scores are similar. These results are in table A2.

TABLE 4
MATH ACHIEVEMENT REGRESSIONS (POOLED, FULL SAMPLE)

	(1)	(2)	(3)	(4)	(5)	(6)
Mandarin-speaking minority	−.17***	−.12**	−.09	−.08	−.06	−.03
	(.058)	(.056)	(.063)	(.065)	(.063)	(.066)
Non-Mandarin-speaking minority	−.62***	−.51***	−.47***	−.46***	−.46***	−.19*
	(.097)	(.096)	(.101)	(.093)	(.090)	(.111)
Student and household characteristics		Yes	Yes	Yes	Yes	Yes
Class peer characteristics			Yes	Yes	Yes	Yes
Teacher characteristics				Yes	Yes	Yes
School characteristics					Yes	
School fixed effects						Yes
Constant	.04	1.92***	1.27**	1.18*	1.57**	1.13*
	(.027)	(.623)	(.631)	(.647)	(.643)	(.628)
Adjusted R^2	.013	.091	.100	.101	.114	.209

Note. Each column represents a separate regression. Standard errors (in parentheses) account for clustering at the school level. Student and household characteristics, class peer characteristics, teacher characteristics, and school characteristics include those in table 1. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in math). $N = 9{,}468$.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

(Blinder 1973; Oaxaca 1973) to decompose the achievement gap—between Han students and both types of minority students. We decompose the gap into two components. First, there is a component that can be explained by differences in student, peer, teacher, and school characteristics. In the rest of the analysis, we refer to this component as that due to "differences in characteristics." The second component is due to between-group differences in returns to characteristics.

The achievement regressions that we use in the decomposition are based on the following linearized specification of the educational production function:

$$Y_{is} = \alpha + \beta_1 I_{is} + \beta_2 P_{is} + \beta_3 T_{is} + \beta_4 S_{is} + \varepsilon_{is}, \tag{2}$$

where, as above, Y_{is} is the observed test score of student i in schools, I_{is} is a vector of individual student and household variables, P_{is} is a vector of peer group variables, T_{is} is a vector of teacher characteristics, S_{is} is a vector of school variables, and ε_{is} is an error term. The error term is allowed to be correlated at the school level to account for clustering effects. In some specifications, we substitute S_{is} for school fixed effects (γ_s) to control for unobserved heterogeneity at the school level.

The Han-minority achievement gap (difference in test scores) can be expressed as

$$\left(Y_H^* - Y_M^*\right) = \left(X_H^* - X_M^*\right)\beta_H + X_M^*(\beta_H - \beta_M), \tag{3}$$

where Y_H^* and Y_M^* are the predicted mean standardized test scores of Han and minority students, X_H^* and X_M^* are the mean characteristics of Han and minority students (I_{is}, P_{is}, T_{is}, and S_{is}); and β_H and β_M are the returns to characteristics for Han and minority students estimated using equation (2) above. Note that, because individual school fixed effects cannot be estimated for minority (Han) students in schools where no minority (Han) students attend, we restrict the sample to only mixed schools with at least two Han and two minority students in analysis that includes school fixed effects.

The overall difference in exam scores can, therefore, be decomposed into two components. One is the portion attributable to differences in the quantity of characteristics, evaluated using Han returns: $\beta_H(X_H^* - X_M^*)$. The other portion, $X_M^*(\beta_H - \beta_M)$, is that attributable to differences in returns to the characteristics of Han and minority students.

A. Returns to Characteristics by Ethnic Group

Table 5 reports the results of separate math achievement regressions for Han students, Mandarin minority students, and non-Mandarin minority students.[6] The odd-numbered columns in the table include all characteristics in table 3; even-numbered columns substitute school characteristics for school fixed effects. The coefficients from these regressions (which are the measured achievement returns to the characteristics) are used in the Oaxaca-Blinder decompositions below.

A few insights emerge from comparing the estimated returns to inputs across groups. First, the pattern of returns for Han and Mandarin minority students are similar (comparing the coefficients in table 5 cols. 1 and 2 for the Han students with the coefficients in cols. 3 and 4 for the Mandarin minority students). While some coefficient estimates for Mandarin minorities are not significant, point estimates largely coincide. One exception is the coefficient on age: after controlling for school-level fixed effects, it appears that Mandarin-speaking minority students benefit from delayed school enrollment.

But, some estimated returns for non-Mandarin minority students differ from the other two groups. For example, non-Mandarin students appear to be strongly and negatively affected by a larger proportion of classmates of the same ethnicity (cols. 5 and 6). The differences in estimated returns to class peer ethnic composition are highlighted in figure 4. While both Han

[6] Results for Chinese scores are in table A3. Because power is reduced by separating the two minority groups (Mandarin and non-Mandarin minorities), we also conducted all analyses pooling students who were given the Chinese exam and students who were given the math exam to estimate returns with more precision. Qualitative results of the analyses do not change substantially when using the pooled sample.

students and Mandarin minority students benefit slightly from being in classes with more students of their same ethnicity, there is a negative correlation among non-Mandarin minority students even after controlling for fixed school-level factors. In other words, having more class peers of a students' same ethnicity has a large, negative relationship with achievement of non-Mandarin minority students. Given the large degree of underperformance of students from the non-Mandarin minority group, this correlation may be in part due to the effect of having lower-achieving peers. Non-Mandarin-speaking students also appear to be strongly influenced by the quality of teaching. Both the coefficient on having a teacher who has received a teaching award (cols. 3 and 6) and the coefficient on the school having provided teacher training are large and significant for this group but not in others.

Table 6 repeats these regressions for the sample of mixed schools.[7] Compared to the full sample, estimated returns are much more similar for Han and minority students (of both types) attending the same schools. This suggests that the large differences in returns observed in the full sample are largely due to differences between Han students in Han-only schools and minorities in minority-only schools.

B. Returns to Schools by Ethnic Group

While estimated returns to observed school characteristics are similar for Han and minority students, there may still be differences in estimated school fixed effects for Han and minority students. That is, returns to specific schools (accounting for observed and unobserved characteristics) may differ between Han and minority students. To examine this in more detail, we estimate the school fixed effect version of equation (2) for Han and minority students separately, using the sample of mixed schools with at least two Han students and two minority students:[8]

$$Y_{is}^{H} = \alpha^{H} + X^{H'}\beta^{H} + \gamma_{s}^{H} + \varepsilon_{is}^{H}, \tag{4a}$$

$$Y_{is}^{M} = \alpha^{M} + X^{M'}\beta^{M} + \gamma_{s}^{M} + \varepsilon_{is}^{M}, \tag{4b}$$

where X includes the same student, peer, and teacher characteristics as above, and γ_s is a vector of school dummy variables. We interpret the estimated school fixed effects for Han students ($\hat{\gamma}_{s}^{H}$) and minority students ($\hat{\gamma}_{s}^{M}$) as the return of attending a specific school for Han and minority students, respectively,

[7] Results for Chinese scores are in table A4.

[8] For this part of the analysis, we pool both types of minority students.

TABLE 5

MATH ACHIEVEMENT REGRESSIONS BY ETHNICITY

	Han Students		Mandarin-Speaking Minority		Non-Mandarin-Speaking Minority	
	(1)	(2)	(3)	(4)	(5)	(6)
Student and household characteristics:						
Female (0/1)	−.22***	−.22***	−.25***	−.24***	−.10	−.13
	(.021)	(.020)	(.061)	(.072)	(.076)	(.085)
Boarding student (0/1)	−.17***	−.15***	−.34*	−.28	−.24	.05
	(.042)	(.046)	(.173)	(.198)	(.148)	(.101)
Age (years)	−.16	−.13	.49	1.09*	.77	−.72
	(.113)	(.107)	(.533)	(.578)	(.855)	(1.099)
Age^2	.00	−.00	−.02	−.05**	−.03	.03
	(.005)	(.005)	(.023)	(.025)	(.038)	(.047)
Household size	−.01**	−.01	−.02	−.03	.02	.01
	(.007)	(.006)	(.020)	(.023)	(.021)	(.025)
Travel time to school (minutes)	.00***	.00***	.00	.00	.00	.00
	(.000)	(.001)	(.001)	(.002)	(.002)	(.002)
Mother has lower secondary degree or above (0/1)	.03	.03	.04	.08	−.35**	−.23
	(.025)	(.025)	(.084)	(.093)	(.141)	(.139)
Father has lower secondary degree or above (0/1)	.25***	.20***	.19***	.21***	.13	.14
	(.024)	(.024)	(.071)	(.080)	(.145)	(.166)
Father at home (0/1)	−.00	.01	−.13*	−.11	−.11	−.07
	(.022)	(.022)	(.069)	(.083)	(.140)	(.169)
Mother at home (0/1)	−.00	−.03	.12	.07	−.13	−.08
	(.026)	(.024)	(.076)	(.094)	(.104)	(.115)
Household asset index (0/1)	.02**	.02**	.02	.02	−.00	.01
	(.008)	(.008)	(.024)	(.029)	(.030)	(.036)
Class peer characteristic:						
Proportion of peers' mothers with lower secondary degree or above	.17	−.43**	.23	.60	−.46	1.81***
	(.154)	(.201)	(.246)	(.537)	(.867)	(.533)
Proportion of class peers of same ethnicity	.20	.09	.10	.44	−.66**	−3.28***
	(.171)	(.520)	(.109)	(.536)	(.272)	(1.207)
Peer average household asset index	.08**	−.06	.16*	−.05	.13	−.23
	(.039)	(.062)	(.084)	(.153)	(.155)	(.144)

Teacher characteristic:						
Female teacher (0/1)	.01	−.01	−.05	−.13	−.28*	.08
	(.047)	(.047)	(.083)	(.139)	(.164)	(.133)
Han teacher (0/1)	.04	.02	−.05	−.13	−.26	−.84***
	(.094)	(.146)	(.077)	(.111)	(.168)	(.216)
Teacher has higher education degree (0/1)	−.01	.03	.19	−.38*	.12	.20
	(.059)	(.076)	(.132)	(.215)	(.202)	(.120)
Teacher attended normal college (0/1)	−.03	−.02	.03	−.32	.13	−.28
	(.049)	(.050)	(.165)	(.199)	(.196)	(.174)
Teacher has received provincial or national teaching award (0/1)	−.05	.03	−.04	−.24	1.03***	1.14***
	(.057)	(.065)	(.119)	(.176)	(.243)	(.213)
Gongban teacher (0/1)	−.01	.07	.25	.38	−.11	.05
	(.059)	(.053)	(.169)	(.343)	(.138)	(.085)
Teacher experience (years)	−.00	−.00	−.00	−.02***	.01	.03***
	(.002)	(.003)	(.005)	(.008)	(.010)	(.008)
School characteristic:						
School size (students)	−.00		.00*		.00	
	(.000)		(.001)		(.002)	
Student-teacher ratio	−.02***		−.00		−.01	
	(.004)		(.010)		(.012)	
Distance to farthest village served by school (minutes)	.00		.00*		−.00	
	(.000)		(.001)		(.002)	
School has provided teacher training in past year (0/1)	.02		−.01		.61***	
	(.083)		(.110)		(.193)	
School infrastructure index	.01		.12***		.08**	
	(.024)		(.031)		(.040)	
School fixed effects	No	Yes	No	Yes	No	Yes
Constant	1.81***	1.59	−3.27	−5.53	−5.31	5.78
	(.676)	(1.044)	(3.072)	(3.356)	(4.817)	(6.541)
Observations	8,286		871		301	
Adjusted R^2	.109	.205	.107	.153	.078	.172

Note. Each column represents a separate regression. Standard errors (in parentheses) account for clustering at the school level. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in math).

* $p < .1$.

** $p < .05$.

*** $p < .01$.

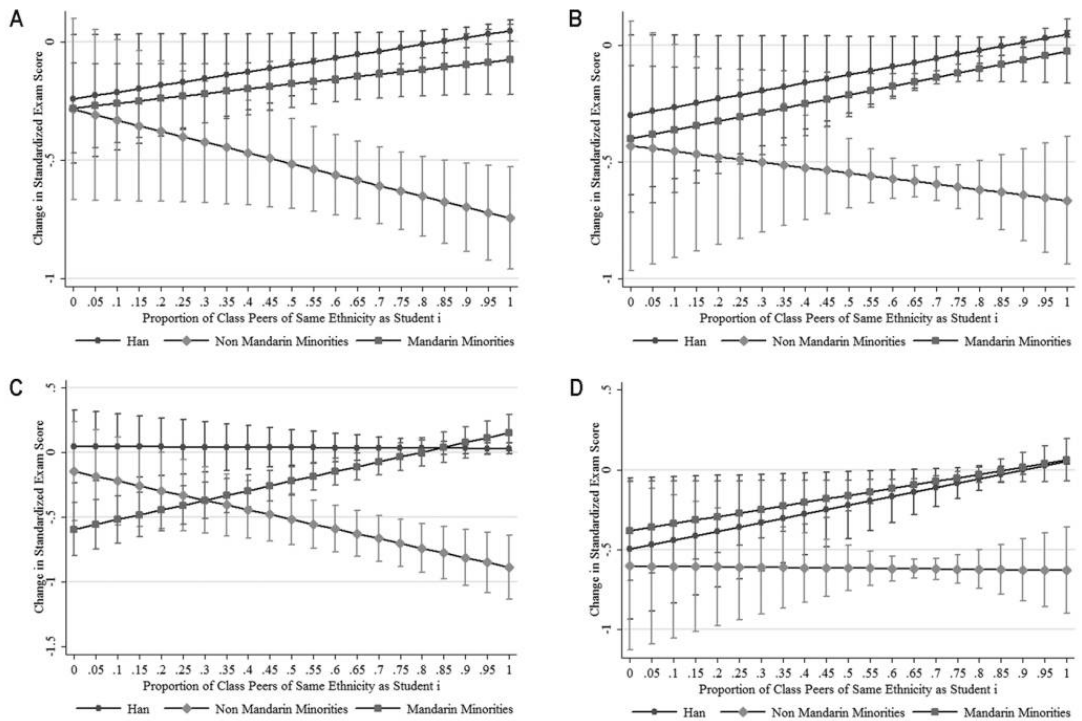

Figure 4. Predicted returns to class peer ethnic composition by ethnic group. *A*, Math, no school fixed effects; *B*, math with school fixed effects; *C*, Chinese, no school fixed effects; *D*, Chinese with school fixed effects.

relative to a reference school (the school whose dummy variable is omitted from the regressions).[9]

We estimate that, on average, school fixed effects estimated for Han students are 0.3 SD higher in math and 0.39 SD higher in Chinese compared to those for minority students. Both of these differences are significant at 1%. To compare the effects of a specific school on Han and minority students directly, figure 5 plots the school coefficients for Han ($\hat{\gamma}_s^H$) against those estimated for minority students ($\hat{\gamma}_s^M$). Figure 5*A* does this for math scores, and 5*B* for Chinese scores. In these figures, the majority of schools (63% of schools for math and 72% for Chinese) lie below the 45° line (where $\hat{\gamma}_s^H$ and $\hat{\gamma}_s^M$ are equal). Individual schools tend to generate larger returns for Han compared to similar minority students attending the same school. In other words, the benefits that Han students receive from (observed and unobserved) attributes of individual schools tend to be larger than the benefits received by minority students.

What types of schools have larger differences in their effect on Han and minority students? We examine school-level differences in Han and minority effects by estimating the following regression:

$$\left(\hat{\gamma}_s^H - \hat{\gamma}_s^M\right) = \alpha + \beta X_s + \varepsilon_s, \tag{5}$$

where X_s is a vector of school-level characteristics and ε_s is an error term. Here, the X_s vector includes the same teacher characteristics (aggregated to the school level) and school characteristics as above, as well as the proportion of students belonging to a Mandarin minority group and the proportion belonging to a non-Mandarin minority group. We use White-Huber standard errors to account for heteroskedasticity.

The results of this analysis are in table 7. In the full models for math and Chinese, observed covariates explain more than 25% of the variation of the difference between the return of school characteristics to Han students and to minority students. Focusing on the results for math, it appears that teachers play the most significant role in reducing the Han-minority difference in returns. Coefficients on variables related to teachers' education and experience are negative and highly significant. Assuming that these variables (having a higher education degree, attending a normal college, and teaching experience) reflect teaching quality, these results suggest that pedagogical practice in the classroom highly influences how much Han and minority students benefit from specific schools.

[9] This analysis is similar to that used in Meng (2004) to examine the effect of firm-level wage policies on gender wage gaps.

TABLE 6

MATH ACHIEVEMENT REGRESSIONS BY ETHNICITY (MIXED SCHOOLS ONLY)

	Mixed Mandarin Minority Schools				Mixed Non-Mandarin Minority Schools			
	Han Students		Mandarin-Speaking Minority Students		Han Students		Non-Mandarin-Speaking Minority Students	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Student and household characteristics:								
Female (0/1)	−.20***	−.17***	−.40***	−.31***	−.13	−.12	−.25	−.22
	(.054)	(.058)	(.094)	(.103)	(.088)	(.089)	(.162)	(.176)
Boarding student (0/1)	−.30***	−.26**	−.70***	−.56***	.13	.10	.17	.14
	(.084)	(.112)	(.186)	(.179)	(.163)	(.209)	(.117)	(.090)
Age (years)	.18	−.12	.24	.31	1.10**	1.34***	.92	.43
	(.262)	(.251)	(.638)	(.675)	(.421)	(.446)	(1.086)	(1.271)
Age^2	−.01	−.00	−.01	−.02	−.06***	−.07***	−.03	−.02
	(.011)	(.011)	(.027)	(.028)	(.017)	(.019)	(.048)	(.056)
Household size	−.02	−.02	.03	.00	−.00	−.01	.04	.05
	(.018)	(.018)	(.032)	(.036)	(.045)	(.047)	(.040)	(.039)
Travel time to school (minutes)	.00	−.00	.00*	.00**	−.00	−.00	.00	.00
	(.002)	(.002)	(.001)	(.001)	(.003)	(.003)	(.002)	(.002)
Mother has lower secondary degree or above (0/1)	.06	.05	.04	.04	−.05	−.08	−.38	−.18
	(.063)	(.065)	(.114)	(.119)	(.157)	(.167)	(.228)	(.228)
Father has lower secondary degree or above (0/1)	.32***	.30***	.32***	.39***	.26**	.27**	.22	.21
	(.064)	(.065)	(.114)	(.129)	(.109)	(.115)	(.219)	(.228)
Father at home (0/1)	.06	.08	−.16	−.11	−.09	−.10	−.16	−.10
	(.060)	(.066)	(.106)	(.118)	(.115)	(.115)	(.235)	(.216)
Mother at home (0/1)	.05	.06	.20*	.13	−.12	−.11	−.17	−.09
	(.068)	(.070)	(.108)	(.142)	(.109)	(.115)	(.245)	(.287)
Household asset index (0/1)	.03	.03	−.01	−.03	.03	.03	.03	.02
	(.021)	(.021)	(.034)	(.036)	(.052)	(.055)	(.056)	(.065)
Class peer characteristic:								
Proportion of peers' mothers with lower secondary degree or above	.07	−.10	.08	.65	.05	.98	.02	2.59
	(.304)	(.566)	(.370)	(.876)	(.843)	(1.154)	(2.162)	(3.253)
Proportion of class peers of same ethnicity	−.20	.13	.38	.61	.06	.92	.38	.24
	(.263)	(.592)	(.241)	(.454)	(.328)	(1.026)	(.844)	(1.802)
Peer average household asset index	−.06	−.36**	.30**	−.10	−.22	−.06	−.40	−.53
	(.100)	(.162)	(.135)	(.340)	(.181)	(.203)	(.325)	(.395)

Teacher characteristic:								
Female teacher (0/1)	−.18**	−.22**	−.25**	−.27	.12	−.29	.03	.03
	(.078)	(.107)	(.110)	(.224)	(.139)	(.303)	(.267)	(.469)
Han teacher (0/1)	.06	.09	−.03	−.16	−.41**	−.20	−.75***	−.50
	(.158)	(.182)	(.149)	(.186)	(.148)	(.175)	(.249)	(.304)
Teacher has higher education degree (0/1)	−.04	−.09	−.27	−1.06***	.23	.47**	−.02	−.29
	(.105)	(.125)	(.210)	(.308)	(.157)	(.217)	(.273)	(.341)
Teacher attended normal college (0/1)	−.00	.05	.01	−.37	−.29	.03	−.54	−1.34***
	(.083)	(.124)	(.244)	(.262)	(.187)	(.289)	(.342)	(.296)
Teacher has received provincial or national teaching award (0/1)	−.03	.01	−.19	−.20	−.65	−.85***	.40	.47
	(.097)	(.142)	(.126)	(.173)	(.404)	(.231)	(.398)	(.466)
Gongban teacher (0/1)	.23***	.18	.77***	.74	.60**	.58	.41	1.70
	(.072)	(.135)	(.239)	(.475)	(.253)	(.526)	(.731)	(1.357)
Teacher experience (years)	−.01*	−.01	−.01	−.03***	−.01	.01	−.01	−.04
	(.004)	(.006)	(.006)	(.008)	(.008)	(.016)	(.020)	(.037)
School characteristic:								
School size (students)	−.00		.00		−.00**		−.00	
	(.001)		(.001)		(.001)		(.003)	
Student-teacher ratio	−.03***		−.03*		.00		.06*	
	(.011)		(.018)		(.019)		(.027)	
Distance to farthest village served by school (minutes)	.00		.00**		.00		−.00	
	(.001)		(.002)		(.001)		(.002)	
School has provided teacher training in past year (0/1)	−.23		−.28		−1.07***		−.16	
	(.174)		(.255)		(.278)		(.239)	
School infrastructure index	.09**		.03		−.16**		−.09	
	(.041)		(.047)		(.062)		(.128)	
School fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Constant	.50	1.26	−1.47	−.80	−4.05	−8.13***	−6.09	−2.85
	(1.566)	(1.518)	(3.782)	(4.038)	(2.515)	(2.738)	(6.261)	(7.364)
Observations	1,338		401		391		139	
Adjusted R^2	.149	.187	.163	.234	.092	.094	.082	.193

Note. Each column represents a separate regression. Standard errors (in parentheses) account for clustering at the school level. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in math) in mixed schools.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

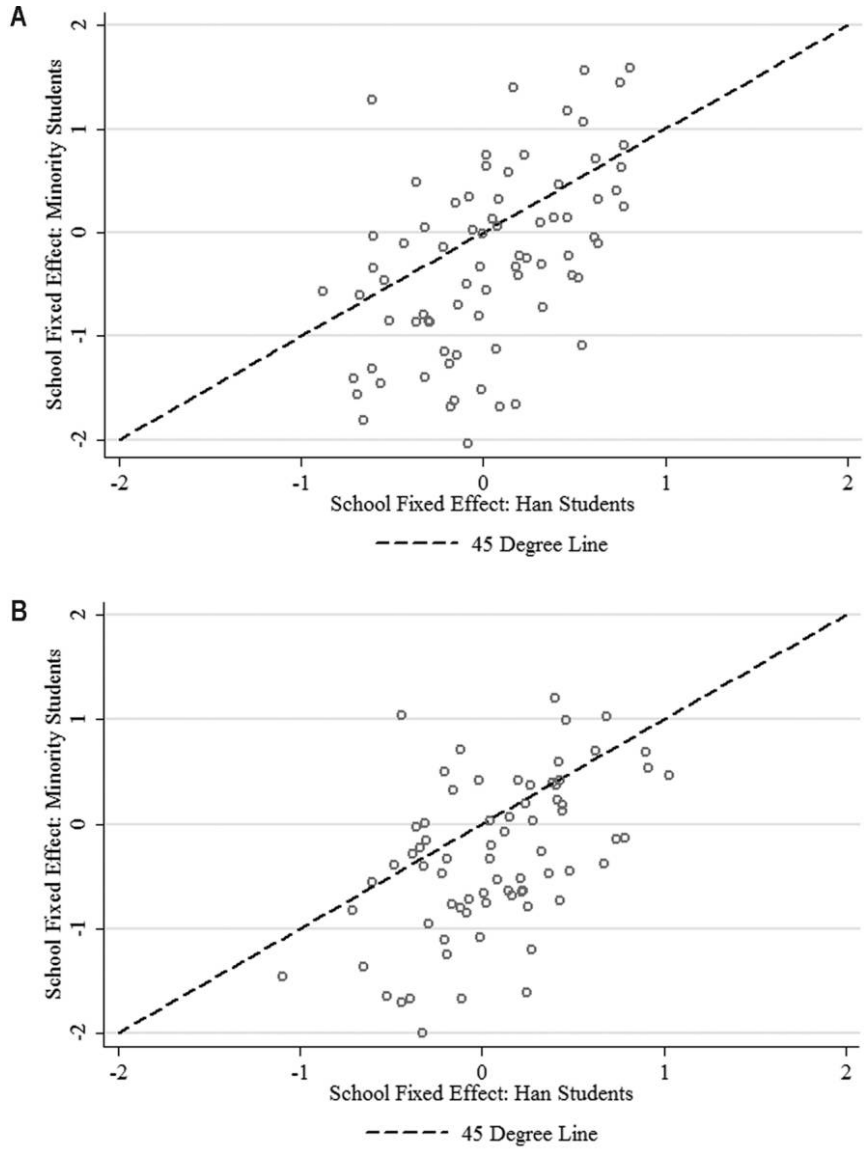

Figure 5. School fixed effects by ethnicity. *A*, Math; *B*, Chinese. Estimated using all mixed schools with more than two minority students and two Han students.

C. *Decomposition Results*

The results of the Oaxaca-Blinder decomposition for math are presented in table 8.[10] The first three columns show results for the full sample, the next three for the mixed school sample without including school fixed effects, and

[10] Results for Chinese are shown in table A5.

TABLE 7

CORRELATES OF DIFFERENCES IN RETURNS TO SCHOOL CHARACTERISTICS (SCHOOL FE) BETWEEN HAN AND MINORITY STUDENTS

	Math			Chinese		
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion Mandarin minority students	−.11	.12	.12	−.26	.22	.22
	(.328)	(.304)	(.417)	(.230)	(.262)	(.375)
Proportion non-Mandarin minority students	.55*	.54	.69	1.11***	1.36***	.22
	(.328)	(.433)	(.558)	(.310)	(.340)	(.571)
Proportion of female teachers		−.06	−.08		.16	.19
		(.175)	(.205)		(.202)	(.200)
Proportion of Han teachers		−.31	−.29		.43**	.37
		(.314)	(.419)		(.206)	(.292)
Proportion of teachers with higher education degree		−.81***	−.92***		−.19	−.23
		(.294)	(.308)		(.313)	(.303)
Proportion of teachers who attended normal college or university		−1.48***	−1.45***		−.36	−.35
		(.242)	(.254)		(.250)	(.230)
Proportion of teachers who have received provincial or national teaching awards		.25	.28		−.51*	−.63*
		(.251)	(.312)		(.303)	(.347)
Proportion of *Gongban* teachers		.34	.23		−.58	−.33
		(.390)	(.389)		(.505)	(.471)
Average teacher experience (years)		−.03**	−.03***		−.01	−.01
		(.012)	(.011)		(.013)	(.011)
School size (students)			.00			−.00
			(.002)			(.001)
Student-teacher ratio			−.02			.02
			(.016)			(.013)
Distance to farthest village served by school (minutes)			−.00			.00***
			(.002)			(.001)
School has provided teacher training in past year (0/1)			−.10			−.44
			(.250)			(.278)
School infrastructure index			−.07			−.01
			(.091)			(.079)
Constant	.46***	2.59***	3.17***	.52***	1.19**	1.24*
	(.123)	(.598)	(.859)	(.102)	(.560)	(.713)
Observations	75	75	75	73	73	73
Adjusted R^2	.080	.297	.260	.147	.193	.254

Note. Dependent variable is the difference between the estimated school fixed effect (FE) for Han students and the estimated school fixed effect for minority students. School fixed effects used to construct the dependent variable were estimated using ordinary least squares regressions of student standardized exam scores on all student, peer, and teacher characteristics in table1 and school dummy variables using the sample of mixed schools only. All covariates are at the school level.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

TABLE 8

OAXACA DECOMPOSITIONS (MATH)

	Full Sample			Mixed Schools Only					
				No School FE			School FE		
	All Minority	Mandarin Minority	Non-Mandarin Minority	All Minority	Mandarin Minority	Non-Mandarin Minority	All Minority	Mandarin Minority	Non-Mandarin Minority
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Total gap	.29***	.17***	.62***	.10	.01	.24*	.10	.01	.24
	(.058)	(.054)	(.084)	(.066)	(.068)	(.129)	(.073)	(.067)	(.150)
Different characteristics	.15***	.15**	.18**	−.02	−.00	−.09	.10	.15	−.10
	(.059)	(.061)	(.070)	(.078)	(.105)	(.094)	(.146)	(.164)	(.216)
Student and household	.06***	.05***	.10***	.05**	.03	.06	.04*	.03	.06
	(.015)	(.017)	(.019)	(.023)	(.027)	(.068)	(.024)	(.029)	(.073)
Peer	.04	.03	.04	−.03	−.06	.05	−.03	.14	.35
	(.050)	(.046)	(.064)	(.080)	(.081)	(.126)	(.241)	(.205)	(.340)
Teacher	.02	.02	.03	−.00	.05	.12	.01	.07	.23
	(.046)	(.047)	(.046)	(.045)	(.064)	(.151)	(.051)	(.064)	(.206)
School	.03	.04	.02	−.04	−.02	−.32*			
	(.023)	(.026)	(.034)	(.031)	(.047)	(.189)			
School fixed effects							.08	−.09	−.75*
							(.272)	(.290)	(.412)
Different returns to characteristics	.14**	.03	.44***	.12	.01	.33***	−.00	−.14	.34
	(.065)	(.063)	(.099)	(.077)	(.097)	(.115)	(.153)	(.151)	(.259)
Observations	9,468	9,157	8,587	2,582	1,739	530	2,582	1,739	530

Note. First three columns use all schools in sample; last six columns use sample of mixed schools only. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in math). Standard errors (in parentheses) account for clustering at the school level. FE = fixed effects.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

the last three for the mixed sample with school fixed effects included. Within each set of columns we give results for the comparison between Han and (1) all minority students, (2) Mandarin minority students, and (3) non-Mandarin minority students. The first row shows the estimated total gap. The second row gives the total portion of the gap estimated to be due to differences in Han and minority characteristics. Estimated subtotals for each category of included characteristic (student and household characteristics, peer characteristics, teacher characteristics, and school characteristics or fixed effects) add up to row 2. The penultimate row gives the portion of the gap due to differences in returns to characteristics.

The first key result of the decomposition analysis is that, for both groups, differences in student and household endowments are the largest explained contributor to the Han-minority achievement gap (first row). For Mandarin minorities, differences in these variables account for 29.4% of the gap in math. Likewise, for non-Mandarin minority students these variables explain 16.1% of the math gap—more than any of the other explained components.

The second key finding is that a much larger portion of the gap can be explained for Mandarin minority students compared to non-Mandarin minority students. We estimate that differences in endowments explain 88% of the math gap for Mandarin minority students while only explaining 29% of the gap for non-Mandarin minority students (cols. 2 and 3, second row).

A third key finding is that the gap between Han and Mandarin minority students disappears when we restrict the sample to schools with both Han and minority students but remains large for non-Mandarin minority students (cols. 4–9). The achievement gap between Han and Mandarin minority students is thus nearly entirely due to the high performance of Han students in schools without minority students. The math gap for non-Mandarin minority students (cols. 6 and 9) is reduced by 0.38 SD (61%) yet remains large with non-Mandarin minority students scoring 0.24 SD below their Han counterparts. This gap is fully due to differences in returns to characteristics. Results change little when we substitute school characteristics for school fixed effects.

Following this set of findings, the decomposition analysis implies that differences in characteristics are unable to explain 0.44 SD (71%) of the gap in math[11] between Han students and non-Mandarin minority students in the full sample and none of the gap after restricting the sample to mixed schools only. This unexplained gap has several possible interpretations. First, it may indicate that some inputs that are important determinants of learning for

[11] Note that 0.64 SD (98.5%) of the gap in Chinese scores is unexplained by differences in characteristics.

these students are omitted. However, it is likely that the influence of these is limited given the large portion of the gap explained between Han and Mandarin minority students.

Another explanation is that, even when given similar educational resources (or inputs), non-Mandarin minority students benefit less from these inputs. This could be due to these students facing a different schooling environment, even when in the same class as Han students. For example, lower teacher expectations could lead them to focus instruction on Han who they believe may benefit more from their instruction (cf. McEwan and Trowbridge 2007). This could also be due to students having difficulty comprehending instruction in Mandarin. Even though instruction in local ethnic languages is permitted in China, this is often difficult in practice (Cherng et al. 2012). For example, instruction in Salar and Tibetan languages is not feasible when these students attend school with students of other ethnicities, a common occurrence in our sample.

VII. Summary and Conclusions

The goal of this article was to document and explain the gap in educational achievement between Han and minority students in primary schools in western China. In our survey of 300 schools in Shaanxi, Gansu, and Qinghai provinces (involving nearly 21,000 fourth- and fifth-grade students), we find large differences in achievement on standardized exams between Han and minority students. On average, minority students perform 0.25 SD lower in math and 0.22 SD lower in Chinese. Most strikingly, minority students who do not generally speak Mandarin as their primary language score 0.62 SD lower than Han in math and 0.65 SD lower than Han in Chinese.

Using decomposition methods pioneered by Oaxaca (1973) and Blinder (1973), we find that most of the achievement gap between Han and minority students with no alternative ethnic language can be explained by differences in endowments of student, family, and school characteristics. Of these, differences in students and family characteristics appear to contribute the most to differences in achievement. Little of the gap between Han students and non-Mandarin minority students (Salar and Tibetan in our sample), however, can be explained by endowment differences. Comparing these students only to Han students in the same schools significantly reduces the size of the achievement gap, yet a difference of more than 0.2 SD persists. None of this remaining gap is explained by differences in endowments. Although several explanations are possible, we believe that a likely explanation is that the ability of students to learn may be hindered by difficulty comprehending instruction

in Mandarin (given that no schools in our sample provided instruction or texts in minority languages). While we cannot say with certainty why these students may benefit less from a given amount of schooling inputs, our analysis suggests that teachers play a significant role.

While we believe that the findings of this article are important, admittedly, the study has a number of limitations. First, although our sample contains sufficient numbers of minority students to conduct analyses, studies involving a larger sample of minority students (particularly non-Mandarin minority students) would provide further insight into the achievement gap. Second, our survey did not collect information on the Mandarin ability of individual students (although we tested students on the Chinese curriculum, this may be distinct from pure language ability). Future studies should employ such information to assess to what degree language is contributing to the underperformance of students belonging to groups that do not speak Mandarin as their primary language.

Despite these limitations, however, our results call for the attention of policy makers to approaches to address the underperformance of minority students in China's rural areas. Given the large and increasing importance of educational attainment to economic well-being, addressing the large achievement gap between Han and minority students may help to mitigate economic disparities in the future. On the basis of our results, promising approaches to address the achievement gap would include those focused on improving the returns to minority students of given schooling inputs (e.g., through pedagogical practice). Further, if future studies show language to contribute significantly to the gap, interventions such as remedial tutoring in Mandarin may also yield large benefits.

Appendix

TABLE A1
SUMMARY STATISTICS FOR MIXED SCHOOL SAMPLE

	Mixed Mandarin Minority Schools			Mixed Non-Mandarin Minority Schools		
	Han Students (1)	Mandarin Minority Students (2)	Difference (2) − (1) [*P*-value]	Han Students (3)	Non-Mandarin Minority Students (4)	Difference (4) − (3) [*P*-value]
Student and household characteristics:						
Standardized math exam score	−.08	−.06	.03	−.16	−.43	−.27
	(1.00)	(.97)	[.68]	(.95)	(.94)	[.08]
Standardized Chinese exam score	−.04	−.05	−.01	−.02	−.39	−.37
	(1.02)	(1.01)	[.95]	(.99)	(1.1)	[.03]
Female (0/1)	.48	.45	−.03	.5	.5	0
	(.5)	(.5)	[.13]	(.5)	(.5)	[.94]
Boarding student (0/1)	.08	.09	.01	.08	.26	.18
	(.27)	(.28)	[.77]	(.27)	(.44)	[.13]
Age (years)	10.85	10.92	.07	10.72	11.19	.47
	(1.08)	(1.1)	[.31]	(1.08)	(1.06)	[.00]
Household size	5.24	5.32	.08	5.19	5.33	.14
	(1.46)	(1.54)	[.38]	(1.39)	(1.71)	[.36]
Travel time to school (minutes)	23.52	22.43	−1.08	25.53	31.4	5.87
	(25.7)	(25.53)	[.67]	(22.99)	(32.27)	[.29]
Mother has lower secondary degree or above (0/1)	.27	.23	−.04	.2	.15	−.05
	(.44)	(.42)	[.25]	(.4)	(.36)	[.22]
Father has lower secondary degree or above (0/1)	.46	.4	−.06	.38	.34	−.04
	(.5)	(.49)	[.16]	(.49)	(.48)	[.42]
Father at home (0/1)	.59	.57	−.03	.52	.59	.07
	(.49)	(.5)	[.51]	(.5)	(.49)	[.30]
Mother at Home (0/1)	.69	.69	0	.68	.61	−.07
	(.46)	(.46)	[.94]	(.47)	(.48)	[.34]
Household asset index (0/1)	.01	.4	.39	.13	.18	.06
	(1.37)	(1.43)	[.00]	(1.38)	(1.5)	[.80]
Class peer characteristic:						
Proportion of peers' mothers with lower secondary degree or above	.26	.26	−.01	.2	.16	−.03
	(.17)	(.17)	[.82]	(.11)	(.11)	[.25]

Proportion of class peers of same ethnicity	.9	.52	−.38	.81	.43	−.38
	(.15)	(.33)	[.00]	(.21)	(.25)	[.00]
Peer average household asset index	.02	.35	.33	.12	.22	.1
	(.56)	(.53)	[.00]	(.63)	(.67)	[.55]
Teacher characteristic:						
Female teacher (0/1)	.36	.49	.12	.4	.48	.07
	(.48)	(.5)	[.08]	(.49)	(.5)	[.54]
Han teacher (0/1)	.92	.62	−.3	.86	.61	−.25
	(.27)	(.49)	[.00]	(.35)	(.49)	[.05]
Teacher has higher education degree (0/1)	.78	.93	.15	.89	.9	.01
	(.41)	(.25)	[.00]	(.31)	(.3)	[.86]
Teacher attended normal college (0/1)	.79	.91	.12	.77	.83	.05
	(.41)	(.28)	[.00]	(.42)	(.38)	[.46]
Teacher has received provincial or national teaching award (0/1)	.09	.19	.1	.03	.15	.12
	(.29)	(.39)	[.05]	(.18)	(.36)	[.26]
Gongban teacher (0/1)	.9	.91	.01	.97	.83	−.14
	(.3)	(.28)	[.62]	(.18)	(.38)	[.07]
Teacher experience (years)	13.61	13.47	−.15	15.73	15.13	−.61
	(11.28)	(9.84)	[.90]	(10.08)	(9.62)	[.70]
School characteristic:						
School size (students)	225.77	212.43	−13.28	218.74	206.27	−12.47
	(61.88)	(66.01)	[.43]	(58.76)	(57.87)	[.47]
Student-teacher ratio	18.06	17.15	−.92	17.39	17.36	−.04
	(5.32)	(4.26)	[.30]	(4.39)	(5.66)	[.98]
Distance to farthest village served by school (minutes)	68.34	60.12	−8.21	65.51	106.05	40.55
	(59.43)	(42.92)	[.40]	(48.98)	(82.67)	[.16]
School has provided teacher training in past year (0/1)	.93	.95	.01	.96	.72	−.24
	(.25)	(.23)	[.76]	(.19)	(.45)	[.12]
School infrastructure index	.15	.18	.03	.03	.53	.5
	(1.2)	(1.17)	[.91]	(1.35)	(1.31)	[.21]
Sample size:						
Total number of students	4,682	932		1,036	294	
Number of students—math sample	2,330	471		508	142	
Number of students—Chinese sample	2,352	461		528	152	
Number of schools	86	86		23	23	

Note. Mixed Mandarin minority schools are schools with at least two Han students and two Mandarin minority students. Mixed non-Mandarin minority schools are schools with at least two Han students and two non-Mandarin minority students. Variables are as described in table 1. Standard errors (in parentheses) account for clustering at the school level.

TABLE A2

CHINESE ACHIEVEMENT REGRESSIONS (POOLED, FULL SAMPLE)

	(1)	(2)	(3)	(4)	(5)	(6)
Mandarin-speaking minority	−.10	−.05	−.02	−.05	−.01	−.25***
	(.064)	(.065)	(.071)	(.071)	(.067)	(.062)
Non-Mandarin-speaking minority	−.65***	−.59***	−.54***	−.57***	−.55***	−.20*
	(.106)	(.104)	(.120)	(.121)	(.118)	(.111)
Student and household characteristics		Yes	Yes	Yes	Yes	Yes
Class peer characteristics			Yes	Yes	Yes	Yes
Teacher characteristics				Yes	Yes	Yes
School characteristics					Yes	
School fixed effects						Yes
Constant	.03	.22	−.39	−.74	−.74	−.43
	(.026)	(.678)	(.682)	(.653)	(.660)	(.656)
Adjusted R^2	.014	.068	.077	.088	.099	.201

Note. Each column represents a separate regression. Standard errors (in parentheses) account for clustering at the school level. Student and household characteristics, class peer characteristics, teacher characteristics, and school characteristics include those in table 1. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in Chinese). $N = 9{,}661$.
* $p < .1$.
*** $p < .01$.

TABLE A3

CHINESE ACHIEVEMENT REGRESSIONS BY ETHNICITY

	Han Students		Mandarin-Speaking Minority		Non-Mandarin-Speaking Minority	
	(1)	(2)	(3)	(4)	(5)	(6)
Student and household characteristics:						
Female (0/1)	.10***	.08***	.14*	.18*	.20*	.17
	(.021)	(.020)	(.077)	(.090)	(.115)	(.146)
Boarding student (0/1)	−.14**	−.15***	−.37**	−.32*	−.03	.06
	(.056)	(.055)	(.172)	(.187)	(.285)	(.461)
Age (years)	.16	.15	−.43	−.59*	1.28	.83
	(.120)	(.124)	(.344)	(.332)	(.780)	(.980)
Age^2	−.01**	−.01**	.02	.02	−.06*	−.04
	(.005)	(.006)	(.015)	(.014)	(.033)	(.042)
Household size	−.02**	−.01	−.01	−.01	.06*	.08***
	(.008)	(.008)	(.018)	(.020)	(.029)	(.022)
Travel time to school (minutes)	.00	.00	−.00	−.00	−.00	−.00
	(.001)	(.001)	(.003)	(.003)	(.003)	(.004)
Mother has lower secondary degree or above (0/1)	−.02	−.02	.01	.05	−.04	−.03
	(.025)	(.025)	(.105)	(.124)	(.205)	(.222)
Father has lower secondary degree or above (0/1)	.20***	.17***	.14**	.13*	−.15	−.21
	(.023)	(.021)	(.068)	(.071)	(.154)	(.171)
Father at home (0/1)	−.04*	.00	−.15***	−.03	−.11	−.23
	(.023)	(.021)	(.055)	(.052)	(.154)	(.158)
Mother at home (0/1)	.02	−.03	.05	−.06	.13	.17
	(.027)	(.025)	(.067)	(.072)	(.167)	(.165)
Household asset index (0/1)	.02***	.03***	.09***	.06*	.10**	.10**
	(.008)	(.008)	(.030)	(.032)	(.047)	(.042)

TABLE A3 (*Continued*)

	Han Students		Mandarin-Speaking Minority		Non-Mandarin-Speaking Minority	
	(1)	(2)	(3)	(4)	(5)	(6)
Class peer characteristic:						
Proportion of peers' mothers with lower secondary degree or above	.16	−.19	−.30	.03	.50	3.70*
	(.135)	(.209)	(.377)	(.473)	(.995)	(1.873)
Proportion of class peers of same ethnicity	−.07	.34	.73***	−.06	−.60	3.42
	(.168)	(.557)	(.141)	(.530)	(.376)	(2.356)
Peer average household asset index	.08**	.09	.08	−.04	−.05	.07
	(.035)	(.064)	(.105)	(.152)	(.159)	(.292)
Teacher characteristic:						
Female teacher (0/1)	.14***	.08*	.14	.29***	−.11	−.05
	(.041)	(.044)	(.106)	(.105)	(.226)	(.456)
Han teacher (0/1)	.06	−.11	.12	.16	−.07	−.47**
	(.084)	(.095)	(.092)	(.116)	(.117)	(.182)
Teacher has higher education degree (0/1)	.08	.01	.27	−.05	−.24	.12
	(.060)	(.086)	(.203)	(.227)	(.269)	(.417)
Teacher attended normal college (0/1)	.07*	.01	.17	.28**	−.15	−.52
	(.045)	(.050)	(.140)	(.129)	(.250)	(.459)
Teacher has received provincial or national teaching award (0/1)	−.01	.18**	−.02	−.24	−.21	−.11
	(.055)	(.075)	(.105)	(.147)	(.207)	(.448)
Gongban teacher (0/1)	.05	.01	.23	−.01	.11	.36**
	(.061)	(.070)	(.145)	(.157)	(.175)	(.159)
Teacher experience (years)	.00	−.00	−.00	−.01	.01	.05**
	(.002)	(.003)	(.005)	(.009)	(.011)	(.017)
School characteristic:						
School size (students)	−.00**		.00*		.00**	
	(.000)		(.001)		(.002)	
Student-teacher ratio	−.01**		−.01		.00	
	(.005)		(.009)		(.017)	
Distance to farthest village served by school (minutes)	.00***		.00		−.00	
	(.000)		(.001)		(.002)	
School has provided teacher training in past year (0/1)	.01		.15		−.14	
	(.076)		(.130)		(.445)	
School infrastructure index	.04*		.11**		.05	
	(.021)		(.044)		(.074)	
School fixed effects	No	Yes	No	Yes	No	Yes
Constant	−.31	−.39	1.22	3.60*	−7.42	−7.78
	(.679)	(.865)	(2.015)	(1.911)	(4.545)	(6.560)
Observations	8,455	8,455	882	882	316	316
Adjusted R^2	.092	.187	.155	.258	.109	.243

Note. Each column represents a separate regression. Standard errors (in parentheses) account for clustering at the school level. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in Chinese).

* $p < .1$.

** $p < .05$.

*** $p < .01$.

TABLE A4

CHINESE ACHIEVEMENT REGRESSIONS BY ETHNICITY (MIXED SCHOOLS ONLY)

	Mixed Mandarin Minority Schools				Mixed Non-Mandarin Minority Schools			
	Han Students		Mandarin-Speaking Minority Students		Han Students		Non-Mandarin-Speaking Minority	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Student and household characteristics:								
Female (0/1)	.15***	.15***	.14	.19	.10	.09	.25	.09
	(.043)	(.043)	(.113)	(.138)	(.061)	(.065)	(.228)	(.279)
Boarding student (0/1)	−.50***	−.52***	−.91***	−.76***	.17	.24	.19	.46
	(.140)	(.167)	(.253)	(.255)	(.214)	(.216)	(.503)	(.431)
Age (years)	−.01	.09	−.71*	−.82*	−.89	−.71	1.29	1.57
	(.323)	(.319)	(.406)	(.419)	(.692)	(.755)	(.941)	(1.216)
Age^2	−.00	−.01	.03	.03*	.04	.03	−.06	−.08
	(.014)	(.014)	(.017)	(.017)	(.031)	(.034)	(.039)	(.050)
Household size	.02	.02	−.02	−.04	.04	.04	.15**	.12*
	(.021)	(.020)	(.030)	(.034)	(.027)	(.027)	(.060)	(.056)
Travel time to school (minutes)	.00	−.00	.00	.00	−.00	−.00	−.00	−.00
	(.001)	(.001)	(.003)	(.003)	(.002)	(.002)	(.006)	(.006)
Mother has lower secondary degree or above (0/1)	−.02	−.04	.14	.21	.12	.12	−.17	−.37
	(.066)	(.068)	(.133)	(.169)	(.140)	(.149)	(.238)	(.288)
Father has lower secondary degree or above (0/1)	.09*	.06	.09	.15	.18**	.19**	−.10	−.06
	(.050)	(.048)	(.108)	(.114)	(.078)	(.074)	(.201)	(.226)
Father at home (0/1)	−.05	−.02	−.12*	−.02	−.08	−.10	.21	.05
	(.053)	(.060)	(.072)	(.088)	(.100)	(.117)	(.242)	(.234)
Mother at home (0/1)	−.03	−.10	−.06	−.20*	.06	.05	−.21	.14
	(.057)	(.060)	(.093)	(.113)	(.069)	(.078)	(.195)	(.208)
Household asset index (0/1)	.02	.03	.09**	.07	.03	.03	.04	.02
	(.020)	(.019)	(.046)	(.057)	(.047)	(.049)	(.060)	(.062)
Class peer characteristic:								
Proportion of peers' mothers with lower secondary degree or above	−.52*	.04	−.83*	.25	−1.80**	−.83	−1.65	4.82
	(.282)	(.553)	(.419)	(.720)	(.800)	(2.121)	(1.910)	(4.262)
Proportion of class peers of same ethnicity	−.23	−.33	.41*	−.06	−.46	1.29	2.06	3.24
	(.266)	(.798)	(.231)	(.776)	(.350)	(1.694)	(1.332)	(2.992)
Peer average household asset index	.06	.28	.24	.01	.47***	−.24	−.59	−.69
	(.103)	(.201)	(.149)	(.243)	(.095)	(.496)	(.516)	(.468)

Teacher characteristic:								
Female teacher (0/1)	.11	.16	.07	.44***	−.18**	−.57	.25	−1.79**
	(.098)	(.134)	(.122)	(.145)	(.067)	(.351)	(.159)	(.772)
Han teacher (0/1)	−.02	−.15	−.02	.27	.03	−.14	−.38	−.99**
	(.089)	(.111)	(.138)	(.205)	(.120)	(.288)	(.269)	(.362)
Teacher has higher education degree (0/1)	.10	.12	.40*	.16	.45***	1.23	−.34	2.89**
	(.162)	(.177)	(.211)	(.419)	(.127)	(.805)	(.356)	(1.305)
Teacher attended normal college (0/1)	.09	.23	.19	−.02	−.30	−1.21*	−.19	−6.60**
	(.117)	(.149)	(.250)	(.297)	(.175)	(.641)	(.361)	(3.037)
Teacher has received provincial or national teaching award (0/1)	.08	.00	−.02	−.28**	−1.32***	−1.43**	.96	−3.41
	(.172)	(.202)	(.118)	(.137)	(.357)	(.608)	(1.193)	(2.069)
Gongban teacher (0/1)	.08	−.22	.09	.12	.73***	.04	.17	. . .
	(.180)	(.242)	(.179)	(.262)	(.126)	(.287)	(.505)	. . .
Teacher experience (years)	.00	.01	−.00	−.00	−.02*	.03	.04**	.29**
	(.004)	(.005)	(.008)	(.015)	(.009)	(.035)	(.016)	(.123)
School characteristic:								
School size (students)	.00		−.00*		.00		−.00	
	(.001)		(.001)		(.001)		(.002)	
Student-teacher ratio	−.03***		−.02		.00		.02	
	(.011)		(.018)		(.008)		(.029)	
Distance to farthest village served by school (minutes)	.00***		−.00		.00**		−.01***	
	(.001)		(.001)		(.001)		(.004)	
School has provided teacher training in past year (0/1)	.31*		.10		.26		−1.82***	
	(.160)		(.279)		(.218)		(.532)	
School infrastructure index	.02		.11		.00		−.49***	
	(.051)		(.069)		(.031)		(.162)	
School fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Constant	.52	.17	4.35	4.73*	4.28	3.16	−5.54	−9.97
	(1.891)	(2.001)	(2.638)	(2.633)	(3.948)	(4.665)	(5.803)	(7.624)
Observations	1,429		387		344		122	
Adjusted R^2	.102	.170	.215	.288	.076	.075	.233	.280

Note. Each column represents a separate regression. Standard errors (in parentheses) account for clustering at the school level. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in Chinese) in mixed schools.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

TABLE A5

OAXACA DECOMPOSITIONS (CHINESE)

	Full Sample			Mixed Schools Only					
				No School FE			School FE		
	All Minority (1)	Mandarin Minority (2)	Non-Mandarin Minority (3)	All Minority (4)	Mandarin Minority (5)	Non-Mandarin Minority (6)	All Minority (7)	Mandarin Minority (8)	Non-Mandarin Minority (9)
Total gap	.25***	.10	.65***	.10	.02	.27	.10	.02	.27
	(.070)	(.063)	(.102)	(.089)	(.086)	(.181)	(.091)	(.083)	(.192)
Different characteristics	.02	.03	.01	−.20**	−.21**	−.14	−.18**	−.19**	−.08
	(.054)	(.053)	(.072)	(.091)	(.091)	(.127)	(.091)	(.092)	(.211)
Student and household	.05***	.05***	.06**	.02	−.01	.06	.02	−.02	.08
	(.014)	(.015)	(.022)	(.020)	(.023)	(.043)	(.021)	(.024)	(.053)
Peer	−.04	−.03	−.05	−.18*	−.12	−.29*	−.07	−.23	.43
	(.049)	(.046)	(.060)	(.099)	(.103)	(.154)	(.350)	(.325)	(.668)
Teacher	−.01	−.01	−.01	−.02	−.05	.09	−.04	−.07*	.28
	(.036)	(.035)	(.042)	(.031)	(.044)	(.161)	(.034)	(.045)	(.241)
School	.02	.03	.01	−.02	−.04	−.00			
	(.022)	(.021)	(.045)	(.041)	(.060)	(.102)			
School fixed effects							−.09	.13	−.88
							(.286)	(.312)	(.698)
Different returns to characteristics	.23***	.08	.64***	.30***	.23***	.42***	.28***	.21**	.36
	(.077)	(.071)	(.113)	(.091)	(.083)	(.144)	(.087)	(.093)	(.223)
Observations	9,661	9,337	8,771	2,571	1,816	532	2,571	1,816	532

Note. First three columns use all schools in sample; last six columns use sample of mixed schools only. Estimation sample includes a randomly chosen half of all sample students (those who were given a standardized exam in Chinese). Standard errors (in parentheses) account for clustering at the school level. FE = fixed effects.

* $p < .1$.

** $p < .05$.

*** $p < .01$.

References

Akerlof, George, and Rachel Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature* 40, no. 4:1167–1201.

Ammermueller, Andreas. 2007. "PISA: What Makes the Difference?" *Empirical Economics* 33, no. 2:263–87.

Angrist, Joshua, and Victor Lavy. 2009. "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial." *American Economic Review* 99, no. 4:1384–1414.

Barrera-Osorio, Felipe, Vicente Garcia Moreno, Harry A. Patrinos, and Emilio Ernesto Porta. 2011. "Using the Oaxaca-Blinder Decomposition Technique to Analyze Learning Outcomes Changes over Time: An Application to Indonesia's Results in PISA Mathematics." Unpublished manuscript, World Bank, Washington, DC.

Blinder, Alan. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8, no. 4:436–55.

Chen, Shaohua, and Martin Ravallion. 2010. "China Is Poorer than We Thought, but No Less Successful in the Fight against Poverty." In *Debates on the Measurement of Global Poverty*, ed. Sudhir Anand, Paul Segal, and Joseph E. Stiglitz. Oxford: Oxford University Press.

Cherng, Hua-Yu, Emily Hannum, and Chunping Lu. 2012. "Sociological Perspectives on Ethnicity and Education in China: Views from Chinese and English Literatures." Unpublished manuscript, University of Pennsylvania.

Chuang, Hwei-Lin. 1997. "High School Youths' Dropout and Re-enrollment Behavior." *Economics of Education Review* 16, no. 2:171–86.

Clarke, Marguerite, Walter Haney, and George Madaus. 2000. *High Stakes Testing and High School Completion*. Chestnut Hill, MA: National Board on Educational Testing and Public Policy.

De Brauw, Alan, and Scott Rozelle. 2008. "Reconciling the Returns to Education in Off-Farm Wage Employment in Rural China." *Review of Development Economics* 12, no. 1:57–71.

Fiszbein, Ariel, Norbert Schady, and Francisco Ferreira. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington, DC: World Bank.

Fortin, Laurier, Dianne Marcotte, Pierre Potvin, Egide Royer, and Jacques Joly. 2006. "Typology of Students at Risk of Dropping out of School: Description by Personal, Family and School Factors." *European Journal of Psychology of Education* 21, no. 4: 363–83.

Glewwe, Paul, and Hanan Jacoby. 1995. "An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition." *Review of Economics and Statistics* 77, no. 1:156–69.

Glewwe, Paul, Hanan Jacoby, and Elizabeth King. 2001. "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis." *Journal of Public Economics* 81, no. 3:345–68.

Gustafsson, Björn, and Shi Li. 2003. "The Ethnic Minority-Majority Income Gap in Rural China during Transition." *Economic Development and Cultural Change* 51, no. 4:805–22.

Gustafsson, Björn, and Ding Sai. 2009a. "Temporary and Persistent Poverty among Ethnic Minorities and the Majority in Rural China." *Review of Income and Wealth* 55:588–606.

———. 2009b. "Villages Where China's Ethnic Minorities Live." *China Economic Review* 20, no. 2:193–207.

Hannum, Emily. 2002. "Educational Stratification by Ethnicity in China: Enrollment and Attainment in the Early Reform Years." *Demography* 39, no. 1:95–117.

Hannum, Emily, Jere Behrman, Meiyan Wang, and Jihong Liu. 2008. "Education in the Reform Era." In *China's Great Economic Transformation*, ed. Loren Brandt and Thomas Rawski. Cambridge: Cambridge University Press.

Hannum, Emily, and Meiyan Wang. 2012. "China: A Case Study in Rapid Poverty Reduction." In *Indigenous People, Poverty, and Development*, ed. Gillette Hall and Harry Patrinos. Cambridge: Cambridge University Press.

Li, Hongbin, Junsen Zhang, and Yi Zhu. 2008. "The Quantity-Quality Trade-Off of Children in a Developing Country: Identification Using Chinese Twins." *Demography* 45, no. 1:223–43.

McEwan, Patrick. 2004. "The Indigenous Test Score Gap in Bolivia and Chile." *Economic Development and Cultural Change* 53, no. 1:157–90.

McEwan, Patrick, and Jeffery Marshall. 2004. "Why Does Academic Achievement Vary across Countries? Evidence from Cuba and Mexico." *Education Economics* 12, no. 3:205–17.

McEwan, Patrick, and Joseph Shapiro. 2008. "The Benefits of Delayed Primary School Enrollment Discontinuity Estimates Using Exact Birth Dates." *Journal of Human Resources* 43, no. 1:1–29.

McEwan, Patrick, and Marisol Trowbridge. 2007. "The Achievement of Indigenous Students in Guatemalan Primary Schools." *International Journal of Educational Development* 27, no. 1:61–76.

Meng, Xin. 2004. "Gender Earnings Gap: The Role of Firm Specific Effects." *Labour Economics* 11, no. 5:555–73.

Oaxaca, R. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14, no. 3:693–709.

Park, Albert, Fang Cai, and Yang Du. 2010. "Can China Meet Her Employment Challenges?" In *Growing Pains: A Study of China's Growth and Development*, ed. Jean Oi, Xueguang Zhou, and Scott Rozelle. Washington, DC: Brookings.

Park, Albert, Sangui Wang, and Guobao Wu. 2002. "Regional Poverty Targeting in China." *Journal of Public Economics* 86:123–53.

Ravallion, Martin, and Shaohua Chen. 2007. "China's (Uneven) Progress against Poverty." *Journal of Development Economics* 82, no. 1:1–42.

Reardon, Sean, and Claudia Galindo. 2002. "Do High-Stakes Tests Affect Students' Decisions to Drop out of School? Evidence from NELS." Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Rumberger, Russel, and Sun Ah Lim. 2008. *Why Students Drop out of School: A Review of 25 Years of Research*. Policy Brief no. 15, California Dropout Research Project.

Sakellariou, Chris. 2008. "Peer Effects and the Indigenous/Non-indigenous Early Test-Score Gap in Peru." *Education Economics* 16, no. 4:371–90.

Van de Walle, Dominique, and Dileni Gunewardena. 2001. "Sources of Ethnic Inequality in Vietnam." *Journal of Development Economics* 65, no. 1:177–207.

Vélez, William, and Rogelio Saenz. 2001. "Toward a Comprehensive Model of the School Leaving Process among Latinos." *School Psychology Quarterly* 16, no. 4: 445–67.

Vickers, Harleen. 1994. "Young Children at Risk: Differences in Family Functioning." *Journal of Educational Research* 87, no. 5:262–70.

Wang, Xiaobing, Chengfang Liu, Linxiu Zhang, Renfu Luo, and Scott Rozelle. 2013. "College Is a Rich, Han, Urban, Male Club: Research Notes from a Census Survey of Four Tier One Colleges in China." *China Quarterly* 214:456–70.

Yi, Hongmei, Linxiu Zhang, Renfu Luo, Yaojiang Shi, Di Mo, Xinxin Chen, Carl Brinton, and Scott Rozelle. 2012. "Dropping Out: Why Are Students Leaving Junior High in China's Poor Rural Areas?" *International Journal of Educational Development* 32, no. 4:555–63.

Zhang, Linxiu, Xiaofei Li, Scott Rozelle, and Jikun Huang. 2013. "China's Labor Transition and the Future of China's Rural Wages and Employment." Unpublished manuscript, Stanford University.

Zhang, Junsen, Yaohui Zhao, Albert Park, and Xiaoqing Song. 2005. "Economic Returns to Schooling in Urban China, 1988 to 2001." *Journal of Comparative Economics* 33, no. 4:730–52.