Published in the International Journal of Educational Development, Vol. 27, No. 1, 2007.

The achievement of indigenous students in Guatemalan primary schools*

Patrick J. McEwan[1]
Wellesley College

Marisol Trowbridge[2]
Bates White

March 2006

* We are grateful to Horacio Alvarez and Jeff Marshall for their assistance in obtaining the data, and to Wellesley College for providing financial support. Joe Shapiro and two anonymous referees provided helpful comments. This research does not represent the views of Bates White, an economic consulting firm.
[1] Department of Economics; Wellesley College; 106 Central St.; Wellesley MA 02481; USA; (781) 283-2987; pmcewan@wellesley.edu.
[2] marisol.trowbridge@bateswhite.com

Abstract

This paper analyzes the difference in academic achievement between indigenous and nonindigenous children that attend rural primary schools in Guatemala. The gap ranges between 0.8 and 1 standard deviation in Spanish, and approximately half that in mathematics. A decomposition procedure suggests that a relatively small portion of the achievement gap is explained by differences in the socioeconomic status of indigenous and nonindigenous families. Other results are consistent with the notion that school attributes play an important role in explaining the achievement gap. The paper discusses several explanations—such as the lack of bilingual education—that are consistent with the empirical findings.

Keywords: Guatemala, indigenous groups, achievement, educational policy.

1. Introduction

In 2000, 33% of Guatemalans ages seven and older reported speaking at least one indigenous language, usually a Mayan one.[3] A growing literature describes the severe and intertwined deficits that confront all Guatemalans—but especially indigenous groups—in multiple indicators of human development (Shapiro, 2006; World Bank, 2003, 2004). For example, a 2000 household survey found that 79% of the indigenous population fell below the poverty line, compared to the 42% of the nonindigenous population.[4]

Among many causes, the difference in poverty is influenced by disparities in formal schooling.[5] More than half of the earnings gap between indigenous and nonindigenous adults can be explained by differences in years of schooling (de Ferranti et al., 2004; Shapiro, 2006; Steele, 1994). The unexplained gap can be interpreted as an upper-bound estimate of labor market discrimination, since observationally similar indigenous workers earn less (Psacharopoulos & Patrinos, 1994). Another plausible explanation is that the gap is due to differences between the two groups that are not well-observed by researchers, such as the quality of schooling.[6] Yet, there is little research that compares educational quality between indigenous and nonindigenous Guatemalans (for an exception, see Hernandez-Zavala, 2006).

[3] The estimate is obtained from a 2000 household survey, the *Encuesta de Condiciones de Vida* (ENCOVI). The most commonly spoken indigenous languages are Mayan, including Ki'che, Kaqchikel, Mam, and Q'eqchi.

[4] The estimate is obtained from the 2000 ENCOVI household survey.

[5] Others include differences in health and nutrition; the distribution of agricultural land and other capital; access to basic services; vulnerability to civil conflict, natural disasters, and other shocks; and a pervasive discrimination. For a broad summary of the empirical evidence, see Edwards (2003), Shapiro (2006), and World Bank (2004).

[6] There is no research on whether quality differences explain ethnic earnings gaps in Latin America. For limited evidence on the overall returns to measures of school quality in Honduras, see Bedi and Edwards (2002). A U.S. literature suggests that black-white test score differences may explain a large proportion of the black-white earnings gap, and that convergence in school quality is partly responsible for reductions in earnings inequality (Card & Krueger, 1992; Neal, 1996).

This paper describes differences in academic achievement between indigenous and nonindigenous children that attend rural primary schools in Guatemala. There are three reasons to expect that indigenous children have lower achievement. First, indigenous parents in Guatemala and other Latin American countries have less schooling and lower incomes, two indicators of the quality of the educational environment in the home (Hernandez-Zavala et al., 2006; Psacharoupoulos & Patrinos, 1994; Hall & Patrinos, 2006).

Second, indigenous families attend schools with fewer instructional materials, lower-quality infrastructure, and less qualified teachers (Hernandez-Zavala et al., 2006; World Bank, 2004). Third, schools rarely address linguistic diversity among indigenous children, creating challenges for indigenous children whose dominant language is not Spanish (McEwan, 2004). Some indigenous children have access to bilingual education, but coverage is far from universal (Shapiro, 2006). Thus, indigenous children may perform worse, even if their family and school environments are otherwise similar to those of nonindigenous children.

This paper uses data from Guatemala's PRONERE (*Programa Nacional de Evaluación del Rendimiento Escolar*) survey to compare academic achievement between indigenous and nonindigenous children. The 2001 PRONERE survey collected data on the Spanish and mathematics achievement of third- and sixth-graders in a representative sample of more than 500 rural schools (PRONERE, 2001). The data contain limited information on the background characteristics of students and families. The empirical analysis shows that the difference in average achievement between indigenous and nonindigenous children is substantial, as high as one standard deviation in language test scores and approximately half that in mathematics.[7] Guatemala's language test score gaps are among the largest of any country in the Western Hemisphere, while its mathematics gaps fall in the middle of the distribution. These countries include Bolivia (McEwan, 2004; Vera, 1998), Canada (Ma & Klinger, 2000), Chile (McEwan, 2004, 2006), Ecuador (Garcia Aracil & Winkler, 2004), Mexico (Hernandez-Zavala et al., 2006), Peru (Hernandez-Zavala et al., 2006; World Bank, 2001), and the U.S. (Freeman & Fox, 2005).[8]

The paper further decomposes the Guatemalan gap into three portions, following McEwan (2004): (1) a portion due to differences in observed family variables, such as parental schooling; (2) a portion due to differences in the quality of schools attended by indigenous and nonindigenous children; and (3) a portion that remains unexplained, even when comparing similar children within schools. It concludes that at least half of the gap can be explained by differences in the quality of schools, broadly conceived, that are attended by each group. At least another quarter of the gap is unexplained within schools, even when children have similar socioeconomic status. The paper reviews several explanations—such as the lack of bilingual education—that are consistent with the empirical findings.

2. Background on the indigenous population and schooling

[7] The measured gaps are similar to estimates from other Guatemalan data (Hernandez-Zavala et al., 2006; Wu, Goldschmidt, & Hara, 2001).

[8] The literature on indigenous and nonindigenous test score gaps has been informed by the voluminous research of sociologists and economists on the black-white test score gap in the U.S. For recent evidence, see Cook and Evans (2000), Fryer and Levitt (2004), Jencks and Phillips (1998), and the citations therein. This empirical evidence suggests that black-white test score gaps, typically as high as one standard deviation, fall towards the upper end of the gaps cited in this paper.

This section describes three essential facts concerning Guatemala's education system. First, the system serves a diverse and poor population in which at least one-third of individuals speak an indigenous language other than Spanish. Second, the population of indigenous children is less likely to attend formal schooling. Third, there is limited but consistent evidence that indigenous students attend primary schools of lower quality, though recent education reforms have attempted to remedy such gaps.

The size of the indigenous population

As a proportion of its total population, Guatemala's indigenous population is one of the largest in Latin America, alongside that of Bolivia (Layton & Patrinos, 2006). Approximately 33% of the population ages 7 and older reports speaking an indigenous language, usually a Mayan one (the estimates are from a nationally-representative household survey, the ENCOVI, collected in 2000). These individuals are disproportionately located in rural areas, where 43% report speaking an indigenous language.

Self-reported language competence is only one possible dimension of indigenous status.[9] The household survey also inquired whether individuals identify themselves as indigenous. By this measure, the population is even larger—39% nationally and 49% in rural areas. It is probable that both measures of indigenous status—based on individual responses to household surveys—represent a lower bound to the true proportion, given the lower social status of indigenous languages (Albó, 1995; McEwan, 2004).

---

[9] For a discussion of the size of indigenous populations across Latin America, using multiple definitions when available, see Layton and Patrinos (2006).

The indigenous population is disproportionately impacted by poverty, noted in the introduction. In urban areas, 56% and 20% of indigenous and nonindigenous individuals, respectively, are in poverty.[10] In rural areas, the equivalent percentages (86% and 62%) are more pronounced for both groups. The poverty comparisons are based upon the first definition of indigenous status, using self-reported language competence, though they are similar when self-identification is used instead.

School attendance

Guatemala's education indicators place it behind every country in the Western Hemisphere except Haiti (Shapiro, 2006; World Bank, 2004). Within Guatemala, indigenous adults and children are at an even greater disadvantage. In 2000, indigenous adults had an average of 2.5 years of schooling, compared to 5.7 among nonindigenous adults (Shapiro, 2006). The inequities persist among younger cohorts. Among children ages 7-16, 77% of nonindigenous children attend school, compared to 63% of indigenous children (see Table 1).[11] Enrollment rates are relatively higher in urban areas, but so too is the gap between indigenous and nonindigenous children.

Figure 1 displays attendance rates that are disaggregated by age. They resemble an "inverted U", characteristic of enrollment patterns in much of Latin America (Urquiola & Calderón, 2005). The lower enrollment rates among seven year-olds are consistent with delayed primary school enrollment (World Bank, 2004). Enrollment rates rise as children enter primary schools, though attendance is only close to universal among nonindigenous

[10] The poverty line is defined as the annual cost of a minimal caloric requirement and an allotment for minimal non-food expenditures (4,319 quetzales, or US$561). World Bank (2004) provides a detailed explanation of poverty line calculations.
[11] Guatemala's official age of primary school enrolllment is seven.

children in urban areas. Enrollment rates then decline steadily after 12 or 13 years of age. The gap between indigenous and nonindigenous attendance rates persists at most ages, though it is somewhat lower in rural areas.

Notwithstanding these deficits, primary school coverage has increased since the 1996 Peace Accords and the official cessation of civil conflict (Anderson, 2003; Valerio & Rojas, 2003). At least half the increase has occurred through PRONADE schools. PRONADE schools are decentralized and community-managed, although the government pays for teacher salaries and other operational expenses. They are disproportionately located in rural and indigenous communities. Table 2 summarizes the type of school attended by children aged 7-16 in 2000. Not surprisingly, just 4% of nonindigenous students attend the predominantly rural PRONADE schools, compared to 16% of indigenous students. The opposite pattern is true of private schools, which are predominantly located in urban areas, and more likely to be attended by nonindigenous students.

School quality

There is scattered evidence that indigenous children attend primary schools of lower quality. Rates of age-grade distorsion are higher among indigenous children in Guatemala (Patrinos & Psacharopoulos, 1996; Shapiro, 2006). That is, many indigenous children are older than expected for their grade. This reflects both delayed enrollment by indigenous students—already evident in Figure 1—as well as grade repetition. Indeed, Shapiro (2006) cites data that 19% of indigenous students repeat the first grade, compared with 25% of nonindigenous students.

One probable cause of relatively higher grade repetition, beyond poverty, is low school quality. Summarizing evidence from a 2002 survey of Guatemalan schools, Hernandez-Zavala et al. (2006) show that indigenous children are less likely to have access to textbooks, and that their teachers have fewer years of experience.[12] They are more likely to attend a multigrade classroom, consistent with a disproportionately rural population.[13]

The Ministry of Education has pursued two large reforms in the past decade—PRONADE and bilingual schooling—that could disproportionately affect the quality of schools attended by indigenous students. Some evidence indicates that PRONADE schools provide greater time in the classroom than standard public schools, and that rates of parental participation are higher (Anderson, 2003; Valerio & Rojas, 2003). Nonetheless, the infrastructure in PRONADE schools is relatively worse, among other deficiencies, perhaps because poor communities must provide their own financing.

Bilingual education has received national support in Guatemala since 1984, first though PRONEBI (the *Programa Nacional Educación Bilingüe Intercultural*) and since 1995 through a unit of the Ministry of Education, DIGEBI (*Dirección General de Educación Bilingüe Intercultural*).[14] The 2000 household survey provides family-reported data on the coverage of bilingual schooling, though it does not directly assess whether students attend officially-designated bilingual schools. Students report whether they receive at least some instruction in a Mayan language (see Table 3). As expected,

[12] World Bank (2004) reports evidence from the 2000 household survey confirming that indigenous families are less likely to have educational materials such as textbooks, and that they spend fewer private resources on such materials.
[13] Multigrade schools include students of multiple ages, grades, and abilities. Sometimes they are supported with an explicit instructional model, such as the the *Nueva Escuela Unitaria* (Anderson, 2003).
[14] For an overview of bilingual education in Guatemala, see Rubio (2004).

nonindigenous children receive very little bilingual instruction. But even among indigenous students, it is far from universal. Nationally, 33% receive some bilingual instruction, but it is more common in rural areas.

There have been few attempts to infer the impact of bilingual schooling on student achievement, usually by comparing students' achievement to those attending monolingual comparison schools (Enge & Chesterfield, 1996; Patrinos & Velez, 1996; Rubio, 2004). This often indicates that students attending bilingual schools score somewhat higher. However, there is little information on how comparison schools were selected or on the attributes of families attending bilingual and comparison schools. Thus, it is challenging to separate the causal impact of bilingual schooling from the effects of unobserved differences among families attending schools.

3. Data on third- and sixth-graders

The 2001 PRONERE survey

To describe differences in academic achievement, this paper uses a sample of third- and sixth-graders collected in 2001 by PRONERE, a joint project of the Ministry of Education and Universidad del Valle. The 2001 survey is a representative sample of rural children, conducted in two stages that would allow for a self-weighting sample.[15] In the first stage, rural schools were apparently selected with probability proportional to size. In the second, 30 students from each grade (third and sixth) were randomly chosen from within each school.

[15] There are few details available on the exact sampling procedures (PRONERE, 2001). It appears that at least a few schools were purposively sampled. It is also not certain whether schools were selected with probability proportional to size. No sampling weights appear to have been calculated or included in the data file. In lieu of further information, all estimates in this paper are unweighted.

All sampled students were administered tests of Spanish and mathematics achievement.[16] In the third grade, 10 of the 30 students within each school were randomly chosen to participate in an oral background questionnaire. In the sixth grade, all 30 students participated in a written questionnaire. The questionnaire provides a limited number of variables on gender, parental schooling, and some additional proxies of socioeconomic status, such the presence of books in the home and television viewing (see Table A1 for variable definitions). Finally, the questionnaire also asks students whether they speak a Mayan language (*MAYA*), which constitutes the key independent variable in this paper.

Descriptive statistics

The working samples for third- and sixth-graders are described in Tables 4 and 5, respectively. Each sample includes student observations with nonmissing values of the achievement tests and *MAYA*. This only involved large sample exclusions for third-grade students, given that the background questionnaire was not applied to two-thirds of the sample. However, the exclusions were apparently random.[17]

Tables 4 and 5 report means within the full sample, as well as separately for subsamples of nonindigenous and indigenous children, as defined by *MAYA*. The Spanish and mathematics tests were standardized to a mean of zero and a standard deviation of 1. This will facilitate comparisons of mean test score differences across tests. It will further permit comparisons of Guatemala's test score gaps with those of other Latin American countries.

[16] The reliabilities of each test, gauged by Cronbach's alpha, are over 0.82 (PRONERE, 2001).
[17] Consistent with this, the mean Spanish and mathematics test scores are similar between the larger and smaller third-grade samples.

In the third- and sixth grade samples, 32% and 37% of the sample report speaking an indigenous language, respectively.[18] Household survey data suggested a slightly higher proportion (41%) among all rural children ages 7-16. The discrepancy is partly due to the slightly lower likelihood that rural indigenous children attend school. Indeed, data from the ENCOVI household survey further show that 38% of rural children *that attend school* are indigenous. This helps to emphasize an important caveat: the subsequent estimates of achievement gaps can be generalized to the population of children attending school. To the extent that indigenous children are less likely to attend school than nonindigenous children, and are perhaps also lower achievers, then achievement gaps within the entire population may be understated.

Finally, Tables 4 and 5 report means for a limited array of student and family variables that measure socioeconomic status (see Table A1 for variable definitions). They indicate, as expected, that indigenous children have lower levels of parental schooling. For example, the dummy variable *MOTHSCH1* indicates that a child's mother has received *no* formal schooling. In the third grade, the variable mean indicates that 53% of indigenous mothers have not received schooling, compared with 32% of nonindigenous mothers, a large difference that is statistically significant at 1% (see Table 4). Similar differences are evident for fathers' schooling, in both the third and sixth grades.

If such variables influence achievement, then their uneven distribution across indigenous and nonindigenous children may partially explain differences in achievement.

[18] It is not clear why the percentage of indigenous children is slightly higher in the sixth grade. One might have supposed that indigenous children would be more likely drop out early, and thus become underrepresented in higher grades. The result may simply reflect sampling noise, or the different survey techniques (oral vs. written) that were used in the third- and sixth-grade background questionnaires.

In a later section, this paper describes a straightforward method of decomposing the relative contribution of such variables to the achievement gap.

4. The magnitude of achievement gaps

Guatemalan results

The 2001 PRONERE data show large achievement gaps between indigenous and nonindigenous children in rural areas (see Table 6). The gap in third-grade Spanish is the most substantial, just over 1 standard deviation in test scores, favoring nonindigenous students. The achievement gap in math scores is approximately half the size. Finally, sixth-grade gaps are somewhat smaller than third-grade gaps. However, it is difficult to interpret this as evidence that gaps change during schooling, since the estimates are drawn from different cohorts. In fact, the smaller sixth-grade gap may be partly attributable to sample selection, if lower-achieving indigenous students are more likely to leave primary school early. Similar PRONERE surveys were conducted in earlier school years (1997, 1999, and 2000). Drawing upon published summaries, Table 6 reports standardized achievement gaps which are generally consistent with those in 2001.

Comparison with other countries

Table 7 reports estimates of indigenous achievement gaps from other countries in the Western Hemisphere. The estimates are taken from tests of language and mathematics, within samples of students attending grade 1 to 8. All estimates were obtained by subtracting mean nonindigenous achievement from mean indigenous achievement, and dividing by the pooled standard deviation of the sample. Compared to other countries in the Western Hemisphere, the gaps in Guatemala are among the largest, particularly in language tests.[19] The causes of Guatemala's larger gaps are likely related to larger differences across indigenous and nonindigenous students in variables that affect test scores, such as family income or school resources.

As one example, 79% of indigenous Guatemalans fell below the poverty line in 2000, compared to 42% of the nonindigenous. In comparison, 32% of indigenous Chileans were poor in the same year, compared to 20% of the nonindigenous (MIDEPLAN, 2002). The group difference in poverty is much smaller in Chile and, perhaps not coincidentally, so too is the indigenous test score gap. Similarly, it is possible that Guatemala has a wider gap than other countries in the quality of schooling provided to indigenous and nonindigenous students. There is no data to compare such gaps across countries, although recent evidence has documented that Chile's policy of "positive discrimination" has disproportionately targeted indigenous students with reforms such as remedial tutoring and a lengthened school day (McEwan, 2006).

5. Explaining achievement gaps

Empirical approach

The essential goal is to decompose the mean difference in achievement into several portions: a portion due to differences in the family variables of indigenous and nonindigenous children; a portion due to differences in the schools attended by each group; and a portion due to other, unobserved differences. We begin by estimating a

[19] Hernandez-Zavala et al. (2006) used a different, nationally-representative Guatemalan dataset collected in 2002. Despite the use of different tests, the magnitude of test score gaps is similar, especially in Spanish.

regression by ordinary least squares, taking advantage of the few independent variables available in the 2001 data:

$$(1) \qquad T_{ij} = \beta_0 + \beta_1 MAYA_{ij} + X_{ij}\beta_2 + e_{ij}$$

where T_{ij} is the test score of student *i* in school *j*, $MAYA_{ij}$ is a dummy variable indicating whether a student speaks a Mayan language, and X_{ij} is a vector of observed characteristics such as parental schooling. The β's are coefficients to be estimated, and e_{ij} is an error term. All standard errors are adjusted for student clustering within schools.

Given the coefficient estimates, the mean difference between indigenous and nonindigenous test scores—the achievement gap—can be written as

$$(2) \qquad \overline{T}^{I} - \overline{T}^{N} = \hat{\beta}_1 + (\overline{X}^{I} - \overline{X}^{N})\hat{\beta}_2$$

where the I and N superscripts indicate indigenous and nonindigenous samples, respectively. The first term on the right side—simply the coefficient estimate for $MAYA_{ij}$—is the "unexplained" portion of the gap, even after controlling for background variables. The second term is the portion of the gap attributable to different endowments of the background variables, such as parental schooling.

This simple decomposition is based upon a regression that omits many variables, particularly school variables that are probably correlated with both achievement and the independent variables. This is likely reflected in biased estimates of the coefficients on family variables (for an additional discussion of bias in the same research context, see McEwan, 2004). A particular concern is that regression coefficients on family variables will be overestimated because they are confounded with unobserved school variables. In such a case, the decomposition will overstate the importance of family variables.

Because multiple students attend each school, it is possible to estimate a second regression specification:

$$T_{ij} = \beta_0 + \beta_1 MAYA_{ij} + X_{ij}\beta_2 + \mu_j + e_{ij} \quad (3)$$

where μ_j are fixed effects for each school. They capture the influence of any variables that are constant across schools, such school resources that shared across classrooms. It may further reflect other school-constant variables such as peer-group characteristics (McEwan, 2004).[20] The new decomposition can be written as

$$\begin{aligned} \overline{T}^I - \overline{T}^N &= \hat{\beta}_1 + (\overline{X}^I - \overline{X}^N)\hat{\beta}_2 \\ &+ \left(\frac{1}{N^I} \sum_{j=1}^{J} \sum_{i=1}^{I_j} MAYA_{ij}\hat{\mu}_j - \frac{1}{N^N} \sum_{j=1}^{J} \sum_{i=1}^{I_j} (1 - MAYA_{ij})\hat{\mu}_j \right) \end{aligned} \quad (4)$$

where N^I and N^N are the total number of indigenous and nonindigenous students in the sample, respectively.[21] The total number of schools is J and the total number of students in school j is I_j.

The first term again indicates the unexplained portion of the gap. In this case, it is the gap that exists within schools, even after conditioning on family background variables. The second term, as before, is the portion attributable to varying endowments of family variables. The final term is the portion due to the varying quality of schools attended by each group, where school quality is interpreted as the influence of both school resources and peer groups.[22]

[20] A substantial literature in the U.S. has searched—with mixed success—for evidence that the socioeconomic or ethnic composition of a school can affect student outcomes, independently of individual student characteristics (Angrist & Lang, 2004; Hoxby, 2000). In Chile, McEwan (2003) finds that the socioeconomic status of classroom peer groups—but not ethnic composition—has strong effects on student achievement, even after making attempts to control for the endogeneity of peer-group variables.
[21] For similar applications, see Cook and Evans (2000) and McEwan (2004).
[22] In practice, it is simply the difference between the mean school fixed effect of indigenous students and the mean fixed effect of nonindigenous students.

Regression results

Tables 8 and 9 report the regression results for third- and sixth-grade, respectively. Columns (1) and (4) report the simplest specification, controlling only for the dummy variable *MAYA*. The coefficients replicate the previously reported estimates of achievement gaps.

The next specifications, in columns (2) and (5), add controls for student and family variables. There are consistent female disadvantages in achievement, especially math, of between 0.12-0.27 standard deviations. Mothers' and fathers' schooling, as expected, has positive and large associations with test scores (all coefficients are interpreted relative to the excluded category of no schooling). Some variables available for third-graders suggest that the presence of books in the household and television viewing—both simple proxies for the home educational environment and income—have positive and large associations. The addition of family variables reduces the estimated magnitude of the achievement gap, though not substantially.

Finally, columns (3) and (6) report specifications that include school fixed effects. The coefficients on *MAYA* decline substantially, suggesting that school quality, broadly defined, is contributing to achievement gaps. However, the coefficients are still positive and in all regressions and statistically significant in three of four regressions. In regressions with *SPANISH* as the dependent variables, the coefficients on *MAYA* are still around one-third of a standard deviation. Thus, substantial test score gaps persist within Guatemalan schools, even among students of the same gender and with similar parental schooling.

The addition of school fixed effects also influences other coefficient estimates. For example, coefficients on parental schooling variable generally decline in magnitude and, in a few cases, become statistically indistinguishable from zero. This is consistent with the notion that such variables are picking up some of the effects of omitted school variables.

Decompositions

Table 10 presents results of the two decompositions that were suggested in equations (2) and (4). The first column presents the decomposition of regressions that exclude school fixed effects. The most striking finding is that differences in background variables such as parental schooling explain a very small proportion of the overall achievement gap (between 10% and 17%). This small proportion may even be overstated if observed family variables are partially reflecting the influence of unobserved school variables, as discussed above. In fact, family variables can only explain 3-8% of achievement gaps when the decomposition is based upon regressions that control for school fixed effects (see column 2 in Table 10).

In contrast, 50%-69% of the gaps are explained by differences in the average fixed effects of schools attended by indigenous and nonindigenous students. This can be broadly interpreted as school quality, a point further discussed in the next section. Lastly, a substantial amount of the achievement gap—24-45%, depending on the grade and test—persists within schools among students with similar backgrounds. It cannot be explained by either students' observed characteristics or the school that they attend.

Interpreting the unexplained achievement gap

The large unexplained portion in the fixed-effects decomposition has at least three interpretations. First, it may indicate that some family and student variables are omitted, such as income, parenting behaviors, and other features of the home educational environment. If omitted variables are correlated with indigenous status and achievement within schools, then their influence is reflected in the unexplained portion of the gap. This is certainly plausible, given the relative scarcity of control variables in the 2001 PRONERE data. Evidence from other Guatemalan school surveys suggests that a richer set of family variables can explain a somewhat larger portion of the gap. The decomposition in Hernandez-Zavala et al. (2006), applying more detailed controls for socioeconomic status, suggests that family variables explain 23%-33% of the large achievement gaps.

Second, it is possible that indigenous and nonindigenous students, even within the same school, receive instruction or classroom resources of varying quantity and quality. The U.S. literature on the black-white test score gap has explored a variety of explanations, including differences in teacher expectations and overt discrimination that may lead teachers to provide better instruction to white students (Ferguson, 1998; McEwan, 2004). In Chile, ethnographic evidence suggests that gaps within schools might be traced to low teacher expectations for indigenous students (Herrera Lara, 1999). In the Guatemalan context, it is possible that indigenous students are subject to lower teacher expectations, or assigned to classrooms within a given school that have fewer instructional resources, but the PRONERE data do not allow good tests of these hypotheses.

Third, indigenous students' achievement may be lower even if given similar instructional resources within a given school. The most plausible example is that of monolingual and bilingual instruction. Table 3 showed that about one-third of indigenous students report receiving at least some instruction in a Mayan language. Thus, the majority of indigenous students continue to receive monolingual instruction in Spanish. It would not be surprising if they derive fewer benefits from such instruction than their monolingual Spanish classmates in the same school.

Interpreting other portions of the achievement gap

Between 50% and 69% of the gap is explained by differences in school quality. However, it bears emphasis that school quality, as captured by school fixed effects, subsumes the effects of any variable that is constant across schools, and not just those observed by the researcher. This includes two categories of variables. First, it includes features of the instructional environment that are constant across a school, such as textbooks, teacher quality, and school type (e.g., a private or PRONADE school). Such variables can be affected by policymakers, and indeed have been the focus of recent educational policy.

The second category of variables includes the school's social context, such as characteristics of a student's peer group in the classroom and school that influence academic achievement. In Guatemala, there is substantial variation in peer-group characteristics across schools. For example, primary schools are highly segregated by indigenous status.[23] The contribution of school fixed effects to the achievement gap

[23] The existence of segregation can be confirmed by the calculation of dissimilarity indexes. A value of 0 indicates a perfectly representative school; that is, a school in which the mix of indigenous students and

could reflect, in part, the influence of segregation and peer-group effects. Such variables are less amenable to policy interventions, though attempts to diminish segregation by race have been a hallmark of education policy in the U.S. (e.g., Angrist & Lang, 2004).

From a policy perspective, it is vital to distinguish between these explanations. Some evidence from other sources suggests that the impact of school variables like textbooks may be substantial. Using different Guatemalan data, Hernandez-Zavala et al. (2006) estimated regressions that included a range of school variables such as teacher characteristics and textbook availability. Decompositions showed that such variables explain 17%-22% of the indigenous test score gap. The proportion is perhaps understated in their decomposition, given that many school variables are unobserved and thus excluded from regressions. To summarize, a narrow category of school variables can explain *at least* 17% of the gap in their study. In the present study, school fixed effects capture a broad category of school variables—including peer-quality—that can explain *at most* 69% of the gap. The true contribution of school resources to the test score gap may lie between the two extremes. Regardless, it suggests a considerable scope for direct investments in schools that might diminish the indigenous test score gap.

The evidence is not able to determine which school investments might yield the greatest impact. There are a few attempts, cited in section 2, to evaluate the impact of policies such as PRONADE and bilingual schooling. However, they are severely limited in their ability to draw valid conclusions about the causal impact of programs on the achievement of primary students, indigenous or otherwise. There is a similar dearth of

---

nonindigenous students reflects the proportion found in the total population. A value of 1 represents a perfectly segregated school in which all students are either indigenous or nonindigenous. In the 2001 data, the third-grade dissimilarity index is 0.87, implying that 87% of indigenous third-graders would have to switch schools in order for each school in the country to have a representative proportion of indigenous and non-indigenous students. In sixth grade, the country wide dissimilarity index is 0.92.

information on the distribution of students across primary schools and classrooms, and how patterns of segregation may affect student achievement.

6. Conclusions

This paper has used the data from Guatemala's 2001 PRONERE survey to describe the achievement gap between indigenous and nonindigenous students in a sample of rural primary schools in Guatemala. It found that indigenous achievement is lower, on average, than nonindigenous achievement. The gap is quite large in both the 2001 data and other Guatemalan school surveys. It ranges between 0.8-1 standard deviations in Spanish, and approximately half that in mathematics. The estimated gaps are among the largest of any country in the Western Hemisphere, including countries with large indigenous populations such as Bolivia and Mexico. The literature on poverty and inequality in Guatemala has emphasized that the quantity of formal schooling is a key determinant of ethnic earnings inequality. The descriptive findings in this paper are suggestive that the quality of schooling, though less studied, is an important topic for research.

A relatively small portion of the achievement gap can be explained by observed differences in the socioeconomic status of indigenous and nonindigenous families, though the PRONERE data do not contain an especially rich collection of variables. A larger portion of the gap (24%-45%) persists between indigenous and nonindigenous students within schools, even when they have similar parental characteristics. Several plausible explanations were suggested, including low teacher expectations for indigenous students and the frequent use of monolingual Spanish instruction in linguistically diverse classrooms.

Finally, 50%-69% of the gap is explained by the varying quality of schools that are attended by indigenous and nonindigenous children. In the present empirical framework, school quality may reflect instructional resources, teacher quality, government interventions, or a range of other inputs. It may also reflect peer-group attributes that are unevenly distributed across schools, given intense levels of segregation by social class and ethnicity.

In sum, the decomposition results are consistent with the notion that schools play a key role in explaining the large achievement gap. The data prevent more specific inferences about the specific aspects of formal schooling that contribute most to enlarging or reducing the gap. The current base of empirical research also provides few compelling clues. There are some mixed attempts, cited throughout the paper, to evaluate the impact of recent primary education reforms, such as PRONADE and bilingual instruction. Since this research relies on imperfect quasi-experimental approaches, it is difficult to ascertain whether it identifies the causal effect of reforms on student achievement. Thus, a priority for future research is the careful evaluation of the impact of education reforms—particularly bilingual curriculum and instruction—on student achievement. Ideally, such evaluations would involve the randomized assignment to treatment and controls groups (Kremer, 2003). Alternatively, they might apply a more transparent quasi-experiment, perhaps based on a regression-discontinuity assignment (Chay, McEwan, & Urquiola, 2005).

Finally, it should be emphasized that this paper has explored only two measures of student outcomes: Spanish and mathematics test scores. An important goal of recent education policy in Guatemala and other countries is to forestall the decline of indigenous languages and cultural heritage. In Guatemala, Edwards and Winkler (2004) have shown that Mayan languages are gradually disappearing among younger generations.[24] In particular, the children of more educated Mayan parents are less likely to speak Mayan. Thus, it is plausible that some interventions in school quality could have negative effects on Mayan language competence, despite positive effects on Spanish, math, or other school outcomes. Other interventions, such as bilingual education, could lessen such tradeoffs.

[24] McEwan and Jiménez (2002) report similar evidence for Bolivia.

References

Albó, X. (1995). *Bolivia plurilingüe: Guía para planificadores y educadores* (Vol. 1). La Paz: UNICEF-CIPCA.

Anderson, M. E. (2003). Overview of the education sector. In *Guatemala: Equity and student achievement in primary education* (Vol. 2: Background Papers). Washington, DC: World Bank.

Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's Metco program. *American Economic Review, 94*(5).

Bedi, A. & Edwards, J. (2002). The impact of school quality on earnings and educational returns: evidence from a low-income country. *Journal of Development Economics, 68,* 157-185.

Card, David, & Krueger, Alan B. (1992). School quality and black-white relative earnings: A direct assessment. *Quarterly Journal of Economics, 107,* 151-200.

Chay, K. Y., McEwan, P. J., & Urquiola, M. (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review, 95*(4), 1237-1258.

Cook, M. D., & Evans, W. N. (2000). Families or schools? Explaining the convergence in white and black academic performance. *Journal of Labor Economics, 18*(4), 729-754.

de Ferranti, D., Perry, G. E., Ferreira, F., & Walton, M. (2004). *Inequality in Latin America: Breaking with history?* Washington, DC: World Bank.

Edwards, J. (2003). Education and poverty in Guatemala. In *Guatemala: Equity and student achievement in primary education* (Vol. 2: Background Papers). Washington, DC: World Bank

Edwards, J. H. Y., & Winkler, D. R. (2004). Capital humano, globalización y asimilación cultural: Un studio aplicado a los Mayas de Guatemala. In D. R. Winkler & S. Cueto (Eds.), *Etnicidad, raza, género y educación en América Latina* (pp. 133-177). Santiago: PREAL.

Enge, K., & Chesterfield, R. (1996). Bilingual education and student performance in Guatemala. *International Journal of Educational Development, 16*(3), 291-302.

Ferguson, R. F. (1998). Teachers' perceptions and expectations and the black-white test score gap. In C. Jencks & M. Phillips (Eds.), *The black-white test score gap.* Washington, DC: Brookings Institution Press.

Freeman, C., & Fox, M. (2005). *Status and trends in the education of American Indians and Alaska Natives* (NCES 2005-108). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Fryer, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics, 86*(2), 447-464.

Garcia Aracil, A., & Winkler, D. R. (2004). Educación y etnicidad en Ecuador. In D. R. Winkler & S. Cueto (Eds.), *Etnicidad, raza, género y educación en América Latina* (pp. 55-92). Santiago: PREAL.

Hall, G., & Patrinos, H. A. (2006). *Indigenous people, poverty and human development in Latin America: 1994-2004.* Palgrave.

Hernandez-Zavala, M., Patrinos, H. A., Sakellariou, C., & Shapiro, J. (2006). *Quality of schooling and quality of schools for indigenous students in Guatemala, Mexico and Peru.* Unpublished manuscript, World Bank.

Herrera Lara, R. H. (1999). *Cultural images, education and domination: The case of the Mapuche of southern Chile.* Unpublished Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Hoxby, C. (2000). *Peer effects in the classroom: Learning from gender and race variation.* Working Paper No. 7867. Cambridge, MA: National Bureau of Economic Research.

Jencks, C., & Phillips, M. (Eds.). (1998). *The black-white test score gap.* Washington, DC: Brookings Institution Press.

Kremer, M. (2003). Randomized evaluations of educational programs in developing countries: Some lessons. *American Economic Review, 93*(2), 102-106.

Layton, H. M., & Patrinos, H. A. (2006). Estimating the number of indigenous people in Latin America. In G. Hall & H. A. Patrinos (Eds.), *Indigenous people, poverty and human development in Latin America: 1994-2004.* Palgrave.

Ma, X., & Klinger, D. A. (2000). Hierarchical linear modeling of student and school effects on academic achievement. *Canadian Journal of Education, 25*(1), 41-55.

McEwan, P. J. (2003). Peer effects on student achievement: Evidence from Chile. *Economics of Education Review, 22,* 131-141.

McEwan, P. J. (2004). The indigenous test score gap in Bolivia and Chile. *Economic Development and Cultural Change, 53,* 157-190.

McEwan, P. J. (2006). Can schools reduce the test score disadvantage of ethnic minorities? Evidence from Chile. Unpublished manuscript, Wellesley College.

McEwan, P. J., & Jiménez, W. (2002). *Indigenous students in Bolivian primary schools: Patterns and determinants of inequities.* Working paper. Washington, DC: World Bank.

Ministerio de Planificación y Cooperación (MIDEPLAN). (2002). *Análisis de la VIII Encuesta de Caracterización Socioeconómica Nacional (CASEN 2000).* Documento No. 14: Etnias y pobreza en Chile. Santiago: Ministerio de Planificación y Cooperación.

Neal, Derek, & Johnson, William R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy, 104,* 869-895.

Patrinos, H. A., & Psacharopoulos, G. (1996). Socioeconomic and ethnic determinants of age-grade distorsion. *International Journal of Educational Development, 16*(1), 3-14.

Patrinos, H. A., & Velez, E. (1996). *Costs and benefits of bilingual education in Guatemala: A partial analysis.* Human Capital Development Working Paper 74. Washington, DC: World Bank.

PRONERE. (1997). *Resultados del Programa Nacional de Evaluación del Rendmiento Escolar.* Guatemala City: PRONERE.

PRONERE. (1999). *Informe de resultados del Programa Nacional de Evaluación del Rendimiento Escolar.* Guatemala City: PRONERE.

PRONERE. (2000). *Informe de resultados del Programa Nacional de Evaluación del Rendimiento Escolar.* Guatemala City: PRONERE.

PRONERE. (2001). *Informe de resultados del Programa Nacional de Evaluación del Rendimiento Escolar.* Guatemala City: PRONERE.

Psacharopoulos, G., & Patrinos, H. A. (Eds.). (1994). *Indigenous people and poverty in Latin America: An empirical analysis.* Washington, DC: World Bank.

Rubio, F. E. (2004). *Educación bilingüe en Guatemala: Situación y desafíos.* Unpublished manuscript.

Shapiro, J. (2006). Guatemala. In G. Hall & H. A. Patrinos (Eds.), *Indigenous people, poverty and human development in Latin America: 1994-2004.* Palgrave.

Steele, D. (1994). Guatemala. In G. Psacharopoulos & H. A. Patrinos (Eds.), *Indigenous people and poverty in Latin America: An empirical analysis.* Washington, DC: World Bank.

Urquiola, M., & Calderón, V. (2005). *Apples and oranges: Educational enrollment and attainment across countries in Latin America and the Caribbean.* Washington, DC: Inter-American Development Bank.

Valerio, A., & Rojas, C. (2003). Educational decentralization and school management by local communities: The case of PRONADE. In *Guatemala: Equity and student achievement in primary education* (Vol. 2: Background Papers). Washington, DC: World Bank.

Vera, M. (1998). *Bilingüismo y rendimiento escolar en Bolivia.* M.A. thesis, ILADES-Georgetown.

World Bank. (2001). *Peruvian education at a crossroads: Challenges and opportunities.* Washington, DC: World Bank.

World Bank. (2003). *Guatemala: Equity and student achievement in primary education* (Vol. 2: Background Papers). Washington, DC: World Bank.

World Bank. (2004). *Poverty in Guatemala.* Washington, DC: World Bank.

Wu, K. B., Goldschmidt, P., & Hara, M. (2003). The determinants of student achievement in primary education. In *Guatemala: Equity and student achievement in primary education* (Vol. 2: Background Papers). Washington, DC: World Bank.

Table 1
School enrollment among children ages 7-16, 2000

	Percent enrolled in school	
	Nonindigenous	Indigenous
Nationwide	76.6%	63.4%
Urban	85.7%	66.6%
Rural	69.9%	62.7%

Source: *Encuesta de Condiciones de Vida* (ENCOVI), 2000 and authors' calculations.
Notes: All calculations used ENCOVI sampling weights. Indigenous status is based upon self-reported knowledge of an indigenous language.

Table 2
School type among children ages 7-16, 2000

	Percent attending four school types	
	Nonindigenous	Indigenous
Panel A: Nationwide		
Public	73.0%	74.8%
PRONADE	4.3%	16.0%
Private	18.6%	6.0%
Other	4.2%	3.2%
Panel B: Urban		
Public	64.4%	72.1%
PRONADE	1.2%	1.6%
Private	30.6%	22.3%
Other	3.7%	4.0%
Panel C: Rural		
Public	80.8%	75.3%
PRONADE	7.1%	19.3%
Private	7.4%	2.3%
Other	4.7%	3.0%

Source: *Encuesta de Condiciones de Vida* (ENCOVI), 2000 and authors' calculations.
Notes: All calculations used ENCOVI sampling weights. Indigenous status is based upon self-reported knowledge of an indigenous language. The category "Other" includes a range of school types, including multigrade schools under the *Nueva Escuela Unitaria* program.

Table 3
Instruction in indigenous language among children ages 7-16, 2000

	Percent receiving at least partial instruction in Mayan language	
	Nonindigenous	Indigenous
Nationwide	0.8%	33.3%
Urban	1.0%	18.3%
Rural	0.6%	36.7%

Source: *Encuesta de Condiciones de Vida* (ENCOVI), 2000 and authors' calculations.
Notes: All calculations used ENCOVI sampling weights. Indigenous status is based upon self-reported knowledge of an indigenous language.

Table 4
Means and standard deviations, grade 3

	Full sample	(a) Non-indigenous	(b) Indigenous	(b) - (a)
SPANISH	-0.02	0.31	-0.71	-1.03**
	(1.01)	(0.90)	(0.87)	
MATH	-0.01	0.17	-0.38	-0.55**
	(1.01)	(1.01)	(0.93)	
MAYA	0.32	0.00	1.00	
FEMALE	0.48	0.48	0.49	0.01
MOTHSCH1	0.39	0.32	0.53	0.21**
MOTHSCH2	0.38	0.41	0.32	-0.09**
MOTHSCH3	0.12	0.14	0.05	-0.09**
MOTHSCH4	0.02	0.03	0.01	-0.02**
MOTHSCH5	0.09	0.09	0.09	-0.01
FATHSCH1	0.26	0.23	0.33	0.11**
FATHSCH2	0.36	0.35	0.40	0.05**
FATHSCH3	0.16	0.18	0.12	-0.06**
FATHSCH4	0.04	0.05	0.03	-0.03**
FATHSCH5	0.17	0.19	0.12	-0.07**
BOOKS1	0.28	0.27	0.30	0.03
BOOKS2	0.71	0.72	0.69	-0.03
BOOKS3	0.00	0.00	0.01	0.01*
TV1	0.32	0.25	0.46	0.21**
TV2	0.68	0.75	0.54	-0.21**
TV3	0.00	0.00	0.01	0.00
Number of students	4,504	3,047	1,457	
Number of schools	521	432	236	

Source: 2001 PRONERE survey and authors' calculations.
Note: Standard deviations (in parentheses) are not reported for dummy variables. ** (*) indicates statistical significance at 1% (5%). The standard errors used in hypothesis tests are adjusted for school-level clustering.

Table 5
Means and standard deviations, grade 6

	Full sample	(a) Non-indigenous	(b) Indigenous	(b) - (a)
SPANISH	0.00	0.30	-0.53	-0.83**
	(1.00)	(0.95)	(0.85)	
MATH	0.00	0.14	-0.25	-0.39**
	(1.00)	(0.99)	(0.96)	
MAYA	0.37	0.00	1.00	
FEMALE	0.44	0.45	0.42	-0.03*
MOTHSCH1	0.31	0.23	0.46	0.23**
MOTHSCH2	0.39	0.41	0.34	-0.08**
MOTHSCH3	0.11	0.14	0.05	-0.09**
MOTHSCH4	0.04	0.05	0.02	-0.03**
MOTHSCH5	0.16	0.17	0.14	-0.03*
FATHSCH1	0.19	0.15	0.27	0.12**
FATHSCH2	0.38	0.36	0.41	0.05*
FATHSCH3	0.16	0.19	0.12	-0.07**
FATHSCH4	0.07	0.08	0.05	-0.03**
FATHSCH5	0.20	0.23	0.16	-0.07**
Number of students	5,321	3,361	1,960	
Number of schools	510	365	228	

Source: 2001 PRONERE survey and authors' calculations.
Note: Standard deviations (in parentheses) are not reported for dummy variables. ** (*) indicates statistical significance at 1% (5%). The standard errors used in hypothesis tests are adjusted for school-level clustering.

Table 6
Differences between nonindigenous and indigenous achievement in PRONERE surveys

	Grade 3		Grade 6	
	Spanish	Math	Spanish	Math
Panel A: Rural schools				
1997	--	--	-0.75	-0.60
1999	--	--	-0.90	-0.40
2000	-1.00	-0.50	-1.17	-0.45
2001	**-1.03**	**-0.55**	**-0.83**	**-0.39**
Panel B: Urban schools				
1997	--	--	-0.73	-0.17
1999	--	--	-0.80	-0.21
2000	-1.07	-0.54	-0.80	-0.34
2001	--	--	--	--

Source: PRONERE (1997, 1999, 2000, 2001) and authors' calculations.
Note: -- indicates that an estimate is not available. Each cell reports the mean difference between indigenous and nonindigenous students' test scores, divided by the standard deviation in the pooled sample. Bold numbers refer to the data used in this paper.

Table 7
Indigenous test score gaps in multiple countries

Study	Country	Grade	Gap
Panel A: Language test scores			
Hernandez-Zavala et al. (2006)	Guatemala	3, 4	-1.11
This study	Guatemala (rural)	3	-1.03
This study	Guatemala (rural)	6	-0.83
Hernandez-Zavala et al. (2006)	Peru	3, 4	-0.83
Freeman and Fox (2005)	U.S.	8	-0.74
Hernandez-Zavala et al. (2006)	Mexico	5	-0.73
Freeman and Fox (2005)	U.S.	4	-0.73
Vera (1998)	Bolivia (La Paz)	4	-0.70
McEwan (2004)	Bolivia	6	-0.48
McEwan (2004)	Chile	8	-0.47
Garcia Aracil and Winkler (2004)	Ecuador	5	-0.43
McEwan (2004)	Chile	4	-0.39
Ma and Klinger (2000)	Canada (New Brunswick)	6	-0.36
McEwan (2004)	Bolivia	3	-0.33
Panel B: Math test scores			
Hernandez-Zavala et al. (2006)	Guatemala	3, 4	-0.90
Freeman and Fox (2005)	U.S.	4	-0.71
Hernandez-Zavala et al. (2006)	Mexico	5	-0.69
Freeman and Fox (2005)	U.S.	8	-0.69
This study	Guatemala (rural)	3	-0.55
Hernandez-Zavala et al. (2006)	Peru	3, 4	-0.58
McEwan (2004)	Chile	8	-0.40
This study	Guatemala (rural)	6	-0.39
McEwan (2004)	Chile	4	-0.37
McEwan (2004)	Bolivia	6	-0.35
McEwan (2004)	Bolivia	3	-0.27
Ma and Klinger (2000)	Canada (New Brunswick)	6	-0.25
Garcia Aracil and Winkler (2004)	Ecuador	5	-0.20

Note: The final column reports the unadjusted difference between the means of indigenous and nonindigenous students' test scores, divided by the standard deviation in the pooled sample. The U.S. estimates are from 2003 National Assessment of Educational Progress (NAEP) data, reflecting the difference in the mean achievement of white and Native-American/Alaskan natives (Freeman & Fox, 2005).

Table 8
Achievement regressions, grade 3

	Dependent variable: MATH			Dependent variable: SPANISH		
	(1)	(2)	(3)	(4)	(5)	(6)
MAYA	-0.553**	-0.456**	-0.130*	-1.025**	-0.888**	-0.302**
	(0.051)	(0.049)	(0.064)	(0.050)	(0.046)	(0.063)
FEMALE		-0.267**	-0.281**		-0.121**	-0.129**
		(0.028)	(0.028)		(0.024)	(0.024)
MOTHSCH2		0.052	0.016		0.056	-0.001
		(0.037)	(0.035)		(0.034)	(0.033)
MOTHSCH3		0.146**	0.052		0.179**	0.049
		(0.054)	(0.052)		(0.049)	(0.047)
MOTHSCH4		0.300*	0.183		0.332**	0.104
		(0.117)	(0.119)		(0.096)	(0.097)
MOTHSCH5		-0.053	-0.063		-0.057	-0.076
		(0.057)	(0.053)		(0.051)	(0.050)
FATHSCH2		0.019	0.026		0.046	0.058
		(0.041)	(0.039)		(0.037)	(0.034)
FATHSCH3		0.073	0.090		0.128**	0.082
		(0.051)	(0.048)		(0.046)	(0.045)
FATHSCH4		0.343**	0.282**		0.432**	0.341**
		(0.081)	(0.087)		(0.067)	(0.066)
FATHSCH5		0.066	0.056		0.116**	0.084
		(0.050)	(0.048)		(0.044)	(0.044)
BOOKS2		0.059	0.039		0.077*	0.015
		(0.039)	(0.038)		(0.035)	(0.033)
BOOKS3		0.080	0.026		0.099	0.066
		(0.247)	(0.249)		(0.211)	(0.198)
TV2		0.247**	0.062		0.385**	0.151**
		(0.041)	(0.035)		(0.036)	(0.034)
TV3		0.334*	-0.086		0.220	-0.107
		(0.162)	(0.155)		(0.150)	(0.158)
Constant	0.171**	-0.027	0.044	0.314**	-0.110*	-0.035
	(0.032)	(0.060)	(0.050)	(0.028)	(0.051)	(0.049)
Observations	4,504	4,504	4,504	4,504	4,504	4,504
R-squared	0.065	0.112	0.430	0.224	0.286	0.539
School fixed effects?	No	No	Yes	No	No	Yes

Source: 2001 PRONERE survey and authors' calculations.
Note: ** (*) indicates statistical significance at 1% (5%). Robust standard errors, adjusted for school-level clustering, are in parentheses.

Table 9
Achievement regressions, grade 6

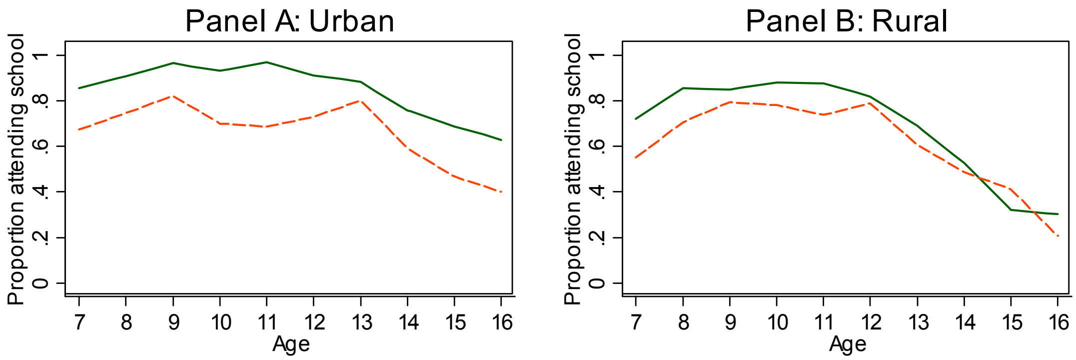
	Dependent variable: MATH			Dependent variable: SPANISH		
	(1)	(2)	(3)	(4)	(5)	(6)
MAYA	-0.394**	-0.353**	-0.124	-0.833**	-0.754**	-0.375**
	(0.054)	(0.055)	(0.070)	(0.049)	(0.049)	(0.079)
FEMALE		-0.251**	-0.252**		-0.151**	-0.166**
		(0.031)	(0.029)		(0.028)	(0.028)
MOTHSCH2		0.102**	0.047		0.124**	0.036
		(0.039)	(0.036)		(0.036)	(0.033)
MOTHSCH3		0.234**	0.150**		0.346**	0.232**
		(0.057)	(0.050)		(0.050)	(0.042)
MOTHSCH4		0.152	-0.006		0.270**	0.162*
		(0.086)	(0.077)		(0.085)	(0.082)
MOTHSCH5		0.005	-0.041		0.067	-0.031
		(0.051)	(0.045)		(0.046)	(0.045)
FATHSCH2		0.064	0.007		0.109**	0.054
		(0.045)	(0.038)		(0.039)	(0.039)
FATHSCH3		0.058	0.003		0.183**	0.097*
		(0.058)	(0.051)		(0.049)	(0.046)
FATHSCH4		0.173*	0.032		0.336**	0.204**
		(0.069)	(0.059)		(0.066)	(0.061)
FATHSCH5		0.146**	0.029		0.235**	0.115*
		(0.055)	(0.048)		(0.047)	(0.046)
Constant	0.144**	0.093	0.116*	0.304**	0.093	0.094*
	(0.035)	(0.060)	(0.046)	(0.029)	(0.049)	(0.045)
Observations	5,321	5,321	5,321	5,321	5,321	5,321
R-squared	0.036	0.059	0.365	0.162	0.190	0.391
School fixed effects?	No	No	Yes	No	No	Yes

Source: 2001 PRONERE survey and authors' calculations.
Note: ** (*) indicates statistical significance at 1% (5%). Robust standard errors, adjusted for school-level clustering, are in parentheses.

Table 10
Decomposition results

	Specification (1)		Specification (2)	
MATH, Grade 3				
Unexplained	-0.456	(82%)	-0.130	(24%)
Family variables	-0.096	(17%)	-0.042	(8%)
School fixed effects	--		-0.381	(69%)
Total	-0.553		-0.553	
MATH, Grade 6				
Unexplained	-0.353	(90%)	-0.124	(31%)
Family variables	-0.042	(11%)	-0.010	(3%)
School fixed effects	--		-0.260	(66%)
Total	-0.394		-0.394	
SPANISH, Grade 3				
Unexplained	-0.888	(87%)	-0.302	(29%)
Family variables	-0.137	(13%)	-0.057	(6%)
School fixed effects	--		-0.666	(65%)
Total	-1.025		-1.025	
SPANISH, Grade 6				
Unexplained	-0.754	(91%)	-0.375	(45%)
Family variables	-0.080	(10%)	-0.040	(5%)
School fixed effects	--	--	-0.418	(50%)
Total	-0.833		-0.833	

Figure 1
School enrollment by age, 2000



Source: *Encuesta de Condiciones de Vida* (ENCOVI), 2000 and authors' calculations.
Notes: The graph plots the proportion of children within each age group—applying the ENCOVI sampling weights—that attend school. Solid lines indicate nonindigenous children, and dotted lines indicate indigenous children. Indigenous status is based upon self-reported knowledge of an indigenous language.

Table A1
Variable definitions

Variable	Definition
SPANISH	Spanish test score (standardized to mean 0, s.d. 1)
MATH	Mathematics test score (standardized to mean 0, s.d. 1)
MAYA	1=child speaks a Mayan language; 0=otherwise
FEMALE	1=female child; 0=male child
MOTHSCH1	1=mother did not attend school; 0=otherwise
MOTHSCH2	1=mother completed 1-5 years of school; 0=otherwise
MOTHSCH3	1= mother completed 6 years of school; 0=otherwise
MOTHSCH4	1= mother completed great than 6 years of school; 0=otherwise
MOTHSCH5	1=missing data on mother's schooling; 0=otherwise
FATHSCH1	1=father did not attend school; 0=otherwise
FATHSCH2	1=father completed 1-5 years of school; 0=otherwise
FATHSCH3	1= father completed 6 years of school; 0=otherwise
FATHSCH4	1= father completed great than 6 years of school; 0=otherwise
FATHSCH5	1=missing data on father's schooling; 0=otherwise
BOOKS1	1=books not present in home; 0=otherwise
BOOKS2	1=books present in home; 0=otherwise
BOOKS3	1=missing data on books in home; 0=otherwise
TV1	1=child does not watch television; 0=otherwise
TV2	1=child watches television; 0=otherwise
TV3	1=missing data on television; 0=otherwise

Source: 2001 PRONERE survey.