

## Explaining the indigenous test score gap in Bolivia and Chile

Patrick J. McEwan\*

Wellesley College

September 2002

---

\* Department of Economics; Wellesley College; 106 Central St.; Wellesley MA 02481; (781) 283-2987; pmcewan@wellesley.edu. I am grateful to Xavier Albó, Cristián Cox, María Luisa Talavera, Miguel Urquiola, Miguel Vera, and Donald Winkler for generously sharing their ideas. I am especially indebted to Wilson Jiménez for his previous collaboration and his assistance with the data. This research received financial support from the World Bank and from the Center for Latin American Studies at the University of Illinois at Urbana-Champaign. These individuals and institutions are not responsible for the errors or interpretations of this study.

## **Abstract**

Between 22-32% of Bolivian students in the primary grades speak Aymara or Quechua in the home. Between 5-7% of Chilean students in the primary grades have mothers that self-identify as indigenous. There is little empirical research that describes the school outcomes of these students, or that compares their outcomes to non-indigenous students. This paper explores the magnitude and explanations of the achievement test score gap between indigenous and non-indigenous students in Bolivian and Chilean primary schools. It uses Bolivia's SIMECAL data from 1997, and Chile's SIMCE data from 1997 and 1999. The test score gap varies between 0.3 and 0.5 standard deviations, always favoring non-indigenous students. It is slightly larger in Spanish than in mathematics, and it is slightly larger in higher grades. Indigenous students also live in families of lower socioeconomic status, on average, than non-indigenous students. They attend schools with less advantaged peer groups, and are less likely to attend tuition-charging private schools. A decomposition procedure suggests that 51-71% of the gap is attributable to differences in the quality of schools and classrooms that are attended by indigenous and non-indigenous students; 23-41% is attributable to differences in family variables. A small proportion of the gap is unexplained. The portion due to schools and classrooms may reflect the influences of school and classroom resources (e.g., teacher quality and textbooks) or the influences of peer-group characteristics. Each set of explanations has different implications for policies intended to reduce the test score gap.

## 1. Introduction

The four most common indigenous languages in South America are Quechua, Aymara, Mapudungun, and Guaraní (Grinevald, 1998). Bolivia and Chile are home to large numbers of each linguistic minority, particularly the first three. In Bolivia's 1992 census, 1.8 Bolivians acknowledged some facility in Quechua, 1.2 million in Aymara, and 70,000 in Guaraní or another indigenous language. In Chile's 1992 census, almost 1 million identified themselves as Mapuche, although it is not clear how many actually spoke Mapudungun (see Table 1). Another 70,000 identified themselves as Aymara or another indigenous group.

There are three reasons to expect that children of indigenous parents in Bolivia and Chile—indeed, almost anywhere in Latin America—are at greater risk of educational failure than non-indigenous children. First, indigenous adults typically have less formal schooling and lower earnings than other adults.<sup>1</sup> Both are common indicators of the quality of the educational environment in the home. Second, indigenous families are more likely to live in rural areas or marginal urban areas, where public schools may have relatively fewer and lower-quality instructional resources. Third, schools of every sort have often ignored and occasionally punished the use of indigenous languages (Comitas, 1972; Herrera Lara, 1999; Plaza & Albó, 1989). They have certainly balked at modifying their curricula and instructional approaches to accommodate linguistic and ethnic diversity, although this has slowly begun to change.<sup>2</sup> Any of these factors—alone or in

---

<sup>1</sup> See, for example, Chiswick, Patrinos, and Hurst (2000); Psacharopoulos (1993); and Psacharopoulos and Patrinos (1994).

<sup>2</sup> In recent years, particularly in Bolivia, government-led education reforms have focused upon the unique needs of indigenous children (Herrera, 1999; Hornberger & King, 1996; Hornberger, 2000).

concert—may have driven a wedge between the average achievement of indigenous and non-indigenous students.

And yet, there is a dearth of empirical research on the precise magnitude of differences in the mean achievement of indigenous and non-indigenous students, hereafter referred to as the “test score gap.”<sup>3</sup> We know even less about whether the gap is best explained by differences in family attributes, in school attributes, or some combination thereof. This knowledge would be a useful ingredient in the design and evaluation of education programs specifically targeted at indigenous students—an increasingly common feature of education reforms in Bolivia, Chile, and other countries.

This paper describes the magnitude of the indigenous test score gap in primary schools, using data from Bolivia’s SIMECAL assessment in 1997 and from Chile’s SIMCE assessments in 1997 and 1999. It then decomposes the test score gap into several components: (1) a component due to different family attributes between indigenous and non-indigenous students; (2) a component due to different peer-group and school attributes; and (3) an “unexplained” component. To do so, the paper uses a modified version of the Blinder-Oaxaca technique, commonly employed in labor economics to assess the determinants of mean wage differences between two groups.<sup>4</sup>

To preview the bottom line, the paper finds a consistent test score gap of 0.3-0.5 standard deviations in Bolivia and Chile. In each subject and grade level, more than half of the test score gap can be explained by the quality of schools and peer-groups. That is,

---

<sup>3</sup> For an early study of Paraguay, see Rivarola, Corvalán, and Zuniga (1977). For recent studies of Bolivia and Peru, respectively, see Vera (1998) and World Bank (2001). The paucity of research in Latin America contrasts with the attention lavished upon racial and ethnic test score inequalities in the U.S. (e.g., Cook & Evans, 2000; Jencks & Philips, 1998).

<sup>4</sup> Psacharopoulos and Patrinos (1994) use this technique to analyze the determinants of indigenous and non-indigenous earnings differentials in Latin America.

a substantial proportion of the gap can be explained by the fact that indigenous students attend “worse” schools, on average, with a “worse” peer group. A smaller but still important proportion of the gap—between 23% and 41%—is explained by the lower socioeconomic status of indigenous families.

The paper is organized in the following manner. Section 2 provides a brief overview of the size and distribution of indigenous populations in Bolivia and Chile, while section 3 describes the main features of recent education reforms in these countries. Sections 4 and 5 review the data and method that are used to decompose the test score gap. The results are presented in sections 6 and 7. Section 8 summarizes the results and reviews several policy implications.

## **2. Indigenous populations in Bolivia and Chile**

### *Overall population*

There is no consensus about the “proper” definition of indigenous status. The most common definitions are whether individuals report competence in an indigenous language, and whether they self-identify as belonging to an indigenous group. Regardless of the definition, there is no means of ensuring that individuals provide an honest report of either measure. The typical concern is that individuals will under-report their indigenous status, owing to the lower prestige of indigenous languages (Albó, 1995). Thus, population estimates are usually interpreted as lower bounds.

In Bolivia, indigenous status has been most commonly measured by linguistic competence. In the 1992 census, 34% report that they “know” Quechua and 23% report knowledge of Aymara; a far smaller percentage know Guaraní or another language (see

Table 1). In contrast, the Chilean census and other data sources typically measure indigenous status by self-identification. In the 1992 census, 10% of the population identify themselves as Mapuche and a very small percentage as Aymara or another indigenous group.

### *Primary school populations*

Among the students who attend primary schools, these percentages decline markedly. In 1997, 20% of Bolivian third-graders spoke Quechua in their families, and 13% spoke Aymara (see Table 2).<sup>5</sup> In sixth-grade, just 15% and 8% of sixth-graders spoke Quechua and Aymara, respectively. There are similar patterns of decline in Chile, where indigenous status is measured by the self-identification of a child's mother. In two rounds of data from the late 1990s, 6% of fourth-graders are Mapuche and just 4% of eighth-graders (see Table 3).

There are two explanations for these results. First, relatively younger cohorts in the population—including the parents of school-age children—are less likely to report themselves as indigenous than older cohorts.<sup>6</sup> This may stem from a declining grasp of indigenous languages, from generational shifts in cultural identification, or from the greater stigma that younger cohorts attach to identifying themselves as indigenous.

Second, indigenous students are more likely than non-indigenous students to enter school late, to repeat grades, and to drop out of school early (Patrinos & Psacharopoulos, 1996; Jimenez & McEwan, 2001). Hence, they will be under-represented in primary

---

<sup>5</sup> See the Appendix for exact definitions of variables used to define indigenous status in Bolivia and Chile.

<sup>6</sup> For empirical evidence of this from Bolivia, see McEwan and Jimenez (2001). In a 1997 household survey, for example, 35% of individuals aged 51-60 report knowing Quechua; among individuals aged 21-30, this declines to 25%.

school populations—relative to their population proportions—and this under-representation will tend to increase in higher grades. This provides some reason for caution in interpreting this paper’s estimates of the test score gap. The results measure the gap between indigenous and non-indigenous students who happen to attend school. If indigenous non-attendees are lower-achieving, on average, this paper’s results will understate test score gaps in the entire population of school-age children.

### *Geographic dispersion and school-based segregation*

Indigenous school-children are not evenly dispersed across either country. In some Bolivian provinces, notably Chuquisaca, Cochabamba, and Potosí, Quechua is the predominant indigenous language, while Aymara is most common in La Paz (see Table 2). In some parts of the country, like Santa Cruz, there is very little indigenous presence in primary schools. An even more pronounced pattern of geographic dispersion is evident in Chile. Mapuche populations are concentrated in the southern regions of the country, especially Region 9 where over 20% of school-children are indigenous (see Table 3). Nonetheless, the high rates of migration to Santiago (located in Region 13) have led to a substantial number of indigenous children in the capital. While only 3-5% of Region 13’s primary enrollments are Mapuche, these students account for almost 1/3 of the national total of Mapuche students.

To assess whether indigenous enrollments are evenly distributed across schools within each province or region, Tables 2 and 3 also report dissimilarity indexes that range between 0 and 1. A value of 0 would indicate an even distribution of indigenous students across schools—reflecting the overall sample proportion in each province or region; a

value of 1 would indicate perfect segregation of indigenous and non-indigenous students. It is interpreted as the percentage of indigenous students that would have to change schools in order to attain an even distribution. In both countries, the index rarely dips below 0.4, which simply confirms what is evident to a casual observer of Bolivian and Chilean primary schools. Indigenous children are often concentrated in primary schools with other indigenous children.

### **3. Education reform and indigenous students**

In the past two decades, Bolivia and Chile have each embarked upon ambitious programs of education reform that are designed to improve the quality of primary schools. Chile's reform has been widely implemented since 1990, but has placed little emphasis on the targeting of indigenous populations. The Bolivian reform, while later to start, has placed extensive emphasis on bilingual education aimed at indigenous students.

Spurred by the neo-liberal reforms of Pinochet's dictatorship, Chile's system of public schools was decentralized to municipal control in 1980.<sup>7</sup> At the same time, public and private schools were financed by equal per-student subsidies if they did not charge tuition (some private schools opted not to participate in the system, and still charge substantial tuition). Students were free to attend any public or subsidized private school, if the school admitted them. Many have interpreted this market-based approach as a "voucher" system.

With the resumption of democracy in the 1990s, Chile's Ministry of Education pursued a different tack that emphasized central government interventions. Under the

---

<sup>7</sup> See Gauri (1998), Hsieh and Urquiola (2002), and McEwan and Carnoy (2000) for a further description of Chile's education reform during the 1980s.

guise of the “P-900” program, it began distributing instructional resources to low-achieving schools, as identified by the SIMCE assessment.<sup>8</sup> In 1992, the ambitious MECE program (*Mejoramiento de la Calidad y Equidad de la Educación*) sought to endow all publicly-funded schools with infrastructure, instructional materials, and training. One sub-component of the plan—dubbed MECE-Rural—was designed to improve small schools in isolated areas. However, the MECE reform did not specifically target Mapuche populations (although many Mapuche did participate, especially if they attended rural or low-achieving schools). The Ministry of Education has subsequently provided some support for a program of bilingual education targeted at indigenous schools, but it has not been implemented on a large scale.

Bolivia began its education reform later than Chile, but it gave more prominence to components targeted at the indigenous population. In 1994, an Education Reform Law mandated new instructional materials (to accompany a new curriculum), teacher training, and increased community participation. It also emphasized the importance of bilingual education (Hornberger & King, 1996).<sup>9</sup> Under the reform, students in predominantly indigenous schools are eligible to receive instruction in both Spanish and their indigenous language (students in Spanish-speaking schools are presumably eligible to take an indigenous language as an additional subject, but this seems rare). In 1996, the reform was applied in an initial group of schools in the first grade. In subsequent years, the reform has been applied to additional schools, and to successively higher grades in participating schools. Because the Bolivian data used in this study were collected in 1997, this paper’s estimates are best considered a “pre-reform” baseline.

---

<sup>8</sup> For an evaluation of P-900, see McEwan and Urquiola (2002).

<sup>9</sup> Prior to the 1994 reform, Bolivia had experimented on several occasions with bilingual education programs, but none were implemented on a large-scale (see Plaza & Albó, 1989 for a summary).

## 4. Data

### *Bolivian data*

The Bolivia data were collected by SIMECAL (*Sistema de Medición de la Calidad de la Educación*), a unit of the Ministry of Education, which administered a national survey of achievement for the first time in 1997. A sample of primary schools was drawn, and all students in the third and sixth grades took achievement tests. In addition, students, parents, teachers, and principals completed background surveys.

A limited number of variables from this survey, described in the Appendix, are used in this paper's analyses. The Spanish and mathematics tests are each standardized to a mean of 0 and a standard deviation of 1. Hence, all subsequent test score gaps can be interpreted as percentages of a standard deviation. This will facilitate comparisons across tests and across countries.

The essential variable is INDIG, equal to 1 if a parents reports that an indigenous language is used to communicate in the home, and 0 otherwise.<sup>10</sup> Two more detailed variables, QUECHUA and AYMARA, indicate whether those languages are spoken in the home. The two variables are not always mutually exclusive, since a very small number of families report that both languages are used. About 1% of the Bolivian sample reported that Guaraní or another language was spoken in the home; these cases are excluded from the analysis.

---

<sup>10</sup> The SIMECAL data also contain a student reported variable: "what languages do you speak?" The parent-reported variable is used for two reasons. First, there are fewer missing cases. Second, the parent-reported variable may diminish the likelihood of measurement error. In any case, the later analyses were all repeated with the student-reported variable and this paper's conclusions were robust.

The remaining variables can be divided into three categories: family and student variables, peer variables, and school variables. The first category includes measures of gender, parental schooling, sewer access, electricity access, and availability of a telephone. Peer variables include the mean schooling of all mothers in each student's school. They further include the mean of INDIG—effectively the percentage of indigenous students in a given school.

School variables are limited to measures of the class size and whether a school is private or rural. In Bolivia, two kinds of private schools are considered: PRIVATE indicates the usual kind of private school that charges tuition. CONVENIO indicates that a privately-managed school receives some government subsidies.

### *Chilean data*

The Chilean data are drawn from two rounds of the SIMCE (*Sistema de Evaluación de Calidad de Educación*), administered by the Ministry of Education. The first round was collected in 1997 from eighth-graders; the second in 1999 from fourth-graders. The Chilean data differ in one important respect from the Bolivian data. They are a census rather than a sample, so they include data on the universe of fourth- and eighth-graders.<sup>11</sup> Thus, Chilean test score gaps can be calculated with greater statistical precision, despite a much smaller proportion of indigenous students.

The definitions of variables are provided in the Appendix. As with the Bolivian data, Spanish and mathematics test scores are standardized to a mean of 0 and a standard deviation of 1. The variable INDIG is equal to 1 if a student's mother identifies herself

---

<sup>11</sup> Some schools are excluded from the SIMCE measurement, but these usually account for no more than 10% of enrollments. Since they are usually located in rural areas, however, it is possible that they contain a relatively higher proportion of indigenous students.

as indigenous, and 0 otherwise. MAPUCHE and OTHINDIG separately identify students with mothers who are Mapuche or of another indigenous group.

The remaining variables are similar to those described for Bolivia. They also include measures of books in the home, a common proxy for the educational environment in the home, and family income. PRIVSUB indicates that a student attends one of the private subsidized schools, while PRIVPD indicates attendance at an elite private school where tuition is paid.

### *Descriptive statistics*

Tables 4 and 5 report descriptive statistics, dividing each country's sample by grade and by indigenous status. The discussion of test score gaps is reserved for section 6. Among the remaining variables, there are generally stark differences across indigenous and non-indigenous students. They indicate that indigenous students are of lower socioeconomic status, on average, than non-indigenous students. Furthermore, indigenous students attend school with peers who have less-educated mothers, and who are likely to be indigenous themselves. Finally, indigenous students are more likely to attend a rural school, particularly in Bolivia, and they are less likely to attend a tuition-charging private school (indigenous students are much more likely, however, to attend government-supported private schools). Any of these differences could explain an indigenous test score gap.

## **5. Method**

The analysis begins by estimating an education production function for each unique combination of country, grade level, and dependent variable:

$$A_{ij} = \beta_0 + \beta_1 INDIG_{ij} + F_{ij}\beta_2 + P_{ij}\beta_3 + S_{ij}\beta_4 + \varepsilon_{ij}$$

where  $A_{ij}$  is the achievement test score of student  $i$  in school  $j$ ,  $INDIG_{ij}$  is a dummy variable equal to 1 if the student is indigenous,  $F_{ij}$  is a vector of family variables,  $P_{ij}$  is a vector of peer-group variables, and  $S_{ij}$  is a vector of school variables. The  $\beta$ 's are coefficients to be estimated via least squares regression, and  $\varepsilon_{ij}$  is an error term.

To decompose the test score gap, note that the mean achievement of indigenous and non-indigenous students is given by:

$$\bar{A}^I = \hat{\beta}_0 + \hat{\beta}_1 + \bar{F}^I \hat{\beta}_2 + \bar{P}^I \hat{\beta}_3 + \bar{S}^I \hat{\beta}_4$$

$$\bar{A}^{NI} = \hat{\beta}_0 + \bar{F}^{NI} \hat{\beta}_2 + \bar{P}^{NI} \hat{\beta}_3 + \bar{S}^{NI} \hat{\beta}_4$$

where the  $I$  and  $NI$  superscripts indicate indigenous and non-indigenous students, respectively, and a bar indicates a mean. The mean difference can be written as

$$(\bar{A}^I - \bar{A}^{NI}) = \hat{\beta}_1 + (\bar{F}^I - \bar{F}^{NI})\hat{\beta}_2 + (\bar{P}^I - \bar{P}^{NI})\hat{\beta}_3 + (\bar{S}^I - \bar{S}^{NI})\hat{\beta}_4$$

In this formulation, the coefficient on  $INDIG$  ( $\hat{\beta}_1$ ), provides an estimate of the “unexplained” portion of the gap. Successive terms denote the portions attributable to differing endowments of family, peer, and school variables.

Given the limited number of school variables available in the Bolivian and Chilean surveys, this specification may still tell us little about the importance of school attributes. It may even cloud inferences if the included family and peer variables are highly correlated with omitted school variables, and coefficient estimates are biased. Because

there are multiple observations of indigenous and non-indigenous students within schools, however, a modified specification can be estimated:

$$A_{ij} = \beta_1 INDIG_{ij} + F_{ij} \beta_2 + \mu_j + \varepsilon_{ij}$$

where the  $\mu_j$  are a series of dummy variables, or fixed effects, denoting each school.

Because peer and school variables are constant across schools, those coefficients can no longer be estimated. However, the inclusion of fixed effects controls for *all* variables that are constant across schools, whether observed or unobserved. The new decomposition can be written as

$$(\bar{A}^I - \bar{A}^{NI}) = \hat{\beta}_1 + (\bar{F}^I - \bar{F}^{NI}) \hat{\beta}_2 + \left( \frac{1}{N^I} \sum_{j=1}^J \sum_{i=1}^{I_j} INDIG_{ij} \hat{\mu}_j - \frac{1}{N^{NI}} \sum_{j=1}^J \sum_{i=1}^{I_j} (1 - INDIG_{ij}) \hat{\mu}_j \right)$$

where  $N^I$  and  $N^{NI}$  are the total number of indigenous and non-indigenous students in the sample, respectively.  $J$  is the total number of schools and  $I_j$  is the total number of students in school  $j$ .

$\hat{\beta}_1$  is again interpreted as the “unexplained” portion of the gap, albeit with a twist. It now measures the test score gap that remains *within* each school, after controlling for parental education and other family variables. The second term is the portion of the gap attributable to differing endowments of family variables. The third term captures differences that are due to the varying fixed effects of the schools that indigenous and non-indigenous students happen to attend. It can be interpreted, roughly, as the portion of the gap attributable to differences in observed and unobserved school quality. However, it should be emphasized that this subsumes the effect of any variable that is constant across schools, including peer variables.

Because the Bolivian and Chilean data contain observations for multiple classrooms within each school, the preceding exercise can also be conducted with classroom fixed effects. Doing so permits an assessment of the added importance of classroom quality in explaining the gap. This might be the case, for example, if indigenous students are assigned—within a given school—to less able teachers than non-indigenous students or to classrooms with less privileged peers.

## **6. The magnitude of test score gaps**

Figure 1 summarizes the mean differences in test scores across countries, grade levels, and subjects. The results are obtained from achievement regressions, reported in Tables 6-9, that only control for INDIG (in every case, the coefficient on INDIG is statistically significant at 1%).

The results support three generalizations. First, the test score gap ranges between 0.3 and 0.5 standard deviations in both countries.<sup>12</sup> The size of the gap, therefore, is not overwhelming. By way of comparison, the black-white test score gap in the U.S. is currently about 1 standard deviation (Jencks & Phillips, 1998). The existing gaps are also comparable to the effect sizes of common educational interventions. For example, Urquiola (2000) finds—using the 1997 SIMECAL data—that an 8-9 student reduction in class size may raise achievement by 0.17-0.45 standard deviations.<sup>13</sup> Second, the size of

---

<sup>12</sup> In a sample of Bolivian fourth-graders from 1992, Vera (1998) found a test score gap of 0.7 standard deviations (see his Appendix 2, Model 1). The sample was drawn from the cities of La Paz and El Alto, where indigenous students are largely Aymara. World Bank (2001) analyzes Quechua and Aymara test score gaps in Peru. However, the study does not report a standard deviation of the dependent variable, making it difficult to meaningfully assess the magnitude of the gap.

<sup>13</sup> Urquiola (2000) uses a quasi-experimental method, also applied in Angrist and Lavy (1999), to address the potential correlation of the class size variable with the achievement error term.

the gaps is uniformly larger in Spanish than in mathematics, but the magnitude of these differences is not large.

Third, the gaps are relatively larger in the later grades. These differences by grade level should be interpreted cautiously. One tempting explanation is that the test score gap widens over time because families and schools have differential effects on the growth of indigenous and non-indigenous achievement (in addition to levels of achievement).

However, one must recall that the data do not follow a single cohort of students. Rather, the earlier grade's results are drawn from younger cohorts of students—particularly in Chile where the earlier grade's sample is collected at a later date. These students might have received greater exposure to incipient school reforms, for example. In this case, the “growing” test score gap could indicate that reforms are successfully diminishing the gap among younger students.<sup>14</sup> This paper will refrain from drawing conclusions about the evolution of test score gaps across time, instead focusing upon individual cross-sections.

Fourth, the magnitude of the test score gap is generally similar for specific indigenous groups within each country. In other regressions, not reported in this paper, the variable INDIG was replaced with AYMARA and QUECHUA (in the Bolivian data) and MAPUCHE and OTHINDIG (in the Chilean data). In the Bolivian regressions, one cannot reject the null hypothesis that coefficients on AYMARA and QUECHUA are equal at the 5% level of statistical significance. In the data on Chilean fourth-graders, one cannot reject the null hypothesis that coefficients on MAPUCHE and OTHINDIG are

---

<sup>14</sup> Another explanation—less consistent with the data—is that changes are driven by the drop-out behavior of students. If indigenous students, especially low-achieving ones, are more likely to drop out of school early, then one would predict a shrinking test score gap in higher grades. Of course, the opposite is observed in the data.

equal. The null hypothesis is not rejected in the eighth-grade data, but the magnitude of the two coefficients is still quite similar.

## 7. Explaining test score gaps

### *Regression results*

Tables 6-9 report estimates of achievement regressions for each country, grade level, and subject, using the specifications described in Section 5. The specifications in column (2)—including family, peer, and school covariates—are generally consistent with those of other education production functions. Parental education, especially that of mothers, is strongly related to achievement. Other proxies of the home educational environment, such as the availability of books or household services, have equally strong links. One measure of peer-group status, the school mean of EDMTH, has positive and large effects on achievement in all samples. In Chile, for example, a 1 standard deviation increase in M(EDMTH) produces a 0.3 standard deviation increase in fourth-grade Spanish scores.<sup>15</sup> The other measure, M(INDIG), has less robust links to achievement. Even when negative coefficients are statistically significant, the magnitude of the effects is not large. Class size generally has small or statistically insignificant links to achievement.<sup>16</sup> Attendance at a private school, at least in Bolivia, is generally associated with higher average achievement (though in contrast, private school effects are either small in magnitude or statistically insignificant in Chile).<sup>17</sup>

---

<sup>15</sup> Other analyses of the Chilean data have suggested that the partial correlations between peer variables and achievement are not spuriously reflecting the influences of unobserved family variables, and hence may have a causal interpretation (McEwan, in press). However, it is possible that peer variables are correlated with unobserved school variables that influence achievement.

<sup>16</sup> Using the same Bolivian data, however, Urquiola (2000) has suggested that the coefficient on class size in simple regressions is biased by omitted variables.

<sup>17</sup> This finding is consistent with other analyses of the Chilean data (McEwan, 2001).

Across all countries and dependent variables, the inclusion of family, peer, and school covariates greatly diminishes the size of coefficient on INDIG, suggesting that differing endowments of these variables can “explain” the gap. The further inclusion of school and classroom fixed effects greatly increases the amount of variance in achievement that is explained, while only reducing the magnitude of INDIG by a small amount. Without further analysis, little more can be said about which variable endowments are most responsible for explaining the test score gap.

#### *Achievement decompositions*

Following the method described in section 5, Tables 10 and 11 report the results of the decomposition exercise. To illustrate their interpretation, consider column (2) in Table 10. It decomposes the gap in third-grade Spanish scores (-0.33) into four components: (1) an “unexplained” component (which is simply the coefficient on INDIG); (2) the component due to varying endowments of family variables; (3) the component due to peer variables; and (4) the component due to school variables. In this case, -0.08, or 24%, is “unexplained.” Family, peer, and school variables explain -0.13, -0.07, and -0.05, respectively, of the gap. That is, indigenous students have smaller endowments of those variables that positively affect achievement.

And yet, the results of the decomposition could be misleading because they are based on regressions that control for a limited set of peer and school variables. Omitted variables may be correlated with achievement and with any independent variable, thus biasing coefficients. One method of reducing such bias is to base decompositions on fixed-effects specifications. They have one obvious drawback, in that they do not allow

for detailed inferences about *which* school and peer variables are most important.

However, the school fixed effects control for all variables, observed or unobserved, that are constant across schools.

Columns (3) and (4) apply the decomposition to fixed effect specifications. Focusing on the decomposition of third-grade Spanish scores, the “unexplained” gap and the portion due to family variables are little changed. Differences in school fixed effects explain -0.15 of the gap. The further addition of classroom fixed effects subsumes the observed and unobserved variables that are constant across schools *and* classrooms. The classroom fixed effects can explain -0.19 of the overall gap, a small increment.

Figure 2 provides a visual summary of the decomposition evidence, relying upon the classroom fixed-effects specifications. Perhaps the most striking result is that the differences in classroom fixed effects across indigenous and non-indigenous students can explain between 51-71% of the total test score gap. Looking within classrooms, differences in family variables can explain a further 23-41% of the gap. Even when controlling for these variables, however, between 6-19% of the gap cannot be explained.

### *Interpretations*

The empirical results in Figure 2 could have several interpretations. Consider the “unexplained” portion of the gap, which may reflect at least three (unobserved) influences. First, it may indicate the presence of unobserved family and student variables that are correlated with achievement and indigenous status. Second, it may indicate that indigenous and non-indigenous students, even *within* the same classroom, receive different amounts of classroom resources. One explanation is that individual teachers

provide instruction of relatively lower quality to their indigenous students, perhaps by maintaining lower expectations for their success.<sup>18</sup>

Third, it may indicate that indigenous and non-indigenous students reap different outcomes from *equivalent* school and classroom resources. There are a number of possibilities. One is that indigenous students benefit from being taught by an indigenous teacher, while non-indigenous students benefit from instruction by non-indigenous teachers. This has long been hypothesized about black and white students in U.S. classrooms. Relying upon experimental data from Tennessee, Dee (2001) has shown that student achievement rises when students are randomly paired with teachers of their own race.<sup>19</sup> Even without specifying the nature of this relationship, its existence in Latin American classrooms could explain some portion of the *within-class* test score gap.

It is also possible that standard instructional approaches are simply ill-suited to indigenous students, who require unique “technologies” of education production. Thus, a one-size-fits-all approach that is implemented in a diverse classroom will differentially benefit the non-indigenous students. Alternatives such as bilingual instruction have rarely been implemented as formal instructional strategies,<sup>20</sup> although Bolivia has moved decisively in recent years to institute such reforms (unfortunately, there is little empirical evidence on the impact of these reforms on the outcomes).

Even allowing for the ambiguous interpretation of the “unexplained” portion, it accounts for a relatively small proportion of the gap. In contrast, school and classroom

---

<sup>18</sup> This has been a frequently-asserted cause of the black-white test score gap in the U.S. Ferguson (1998) reviews this evidence and concludes that “teachers’ perceptions, expectations, and behaviors probably do help to sustain, and perhaps even to expand, the black-white test score gap” (p. 313).

<sup>19</sup> His analysis is based on data from the Tennessee STAR experiment, in which students and teachers were randomly assigned to classrooms.

<sup>20</sup> It may, of course, be implemented in a less formal capacity by teachers who speak an indigenous language. There is no empirical evidence of how frequently this might occur. Note that it could provide one reason to hypothesize that indigenous students will benefit from being paired with indigenous teachers.

differences account for more than half, and there might be substantial dividends to a more detailed understanding of this overly broad explanation of test score differences.

Generally, the results suggest that indigenous students have low achievement because they are more likely to attend “low-quality” schools (i.e., schools with low mean achievement). The empirical analyses provided some clues about the roots of this quality. Most prominently, indigenous students attend schools with “worse” peer groups, and they are less likely to attend private schools. But these results are hardly a conclusive explanation. The substantial importance of school quality, observed in the fixed effects decompositions, could reflect the importance of *any* school-level variable that is unobserved. There is, regrettably, no means of further assessing the relative importance of these unobserved variables (although that would be necessary for a careful assessment of policy implications, as the next section will emphasize).

The results do provide suggestive evidence on the relative importance of classroom variables, as compared with school variables. In decompositions that employ regressions with classroom rather than school fixed effects, the differences in fixed effects explain a slightly larger portion of the gap. The results indicate suggest that indigenous students are assigned to classrooms of relatively lower quality within schools (in addition to attending lower-achieving schools). For example, indigenous students might be tracked into lower-ability classes, or they may also be assigned lower-quality teachers within schools. Both explanations are speculative without further data.

## **8. Conclusions and policy implications**

This paper has provided new evidence on indigenous test score gaps in Bolivia and Chile. The magnitude of the gaps in both countries is between 0.3 and 0.5 standard deviations, depending on grade level and test. The results of a decomposition procedure suggest that between 51-71% of these gaps are attributable to differences in the quality of schools and classrooms that are attended by indigenous and non-indigenous students.<sup>21</sup> A smaller proportion (23-41%) is attributable to varying endowments of family variables like parental education. An even smaller proportion of the gap remains unexplained, though several possibilities were forwarded, ranging from unobserved family variables to unequal treatment of indigenous students within schools and classrooms.

### *Schools or peers?*

The present analysis has shown that school and classroom quality is important, but it has not indicated *which* differences in school and classroom quality are most responsible for the test score gap. The gap may be the result of an unequal distribution of school and classroom resources, such as instructional materials, teacher quality, or other variables. It may also be the result of an unequal distribution of peer-group characteristics. This paper's non-experimental analyses are somewhat consistent with both, but more research is surely needed to identify the causal impact of common school and peer variables.

Inequality in school resources can be addressed by central-government policies. Indeed, both Bolivia and Chile have made redistributive policies a hallmark of their education reforms (Bolivia has sought to directly target indigenous students, while Chile has targeted rural or poor students who happen to be indigenous). But what if the test

---

<sup>21</sup> A decomposition conducted with U.S. data suggests that school quality (proxied by school fixed effects) accounts for approximately half of the black-white test score gap (Cook & Evans, 2000).

score gap is mainly the result of differential exposure to “bad” peer groups? In that case, the optimal policy—if the primary goal is to diminish the gap—is to encourage indigenous students to mix with “better” peers. In practice, and somewhat ironically, Bolivia’s and Chile’s education systems may have encouraged exactly the opposite result.

Chile’s national voucher plan has allowed large-scale school choice across public and private schools since 1980. The empirical evidence indicates that “the first-order consequence of vouchers in Chile was a massive exodus from public schools by families from higher socioeconomic backgrounds” (Hsieh & Urquiola, 2002).<sup>22</sup> In short, school choice fundamentally altered the distribution of certain peer-group characteristics across schools, perhaps increasing segregation by socioeconomic status (there is no evidence on how sorting altered *ethnic* segregation). In Bolivia’s public schools, there is a *de facto* rather than a *de jure* system of school choice, in which students sometimes choose public schools outside their enrollment areas (Urquiola, 2000). There are also elite private schools that enroll very few indigenous students. If Chile’s experience is any guide, these institutions may have exacerbated segregation.

In light of these institutional features of school choice, how might policy be adjusted to improve the peer-group characteristics to which the average indigenous student is exposed? The U.S., of course, has a long history of “busing” and other attempts to use carrots and sticks to encourage mixing of black and white students. But consider that a legislated goal of the Bolivian reform (and, less explicitly, the Chilean one) is to aid indigenous groups in preserving their native languages and cultures. A policy that transfers indigenous students to largely non-indigenous schools—via private school

---

<sup>22</sup> If peer-effects are indeed substantial, and families consider peer-group attributes when choosing schools, then increasing segregation is a logical outcome.

scholarships or more forcible means—could have exactly the opposite effect. That is, gains in Spanish achievement might be traded off against losses in Quechua, Aymara, or Mapudungun.

### *The impact of school reform*

Given the multiple goals of the education system, the best means of reducing the test score gap may be to focus on redistributive policies that would directly improve indigenous schools. Since 1990, Chile has invested heavily in rural and low-achieving schools through its MECE and P-900 programs. Although indigenous schools were not explicitly targeted, many rural and low-achieving schools are in regions which enroll disproportionate numbers of Mapuche students. Have 12 years of reform succeeded in reducing the test score gap? In the wake of such investments, differences in school quality still explain a substantial portion of the test score gap. This could have several interpretations.

First, it possible that the investments improved the achievement of indigenous students, but that an even larger gap in test scores—and school quality—existed earlier in the decade.<sup>23</sup> That, unfortunately, cannot be assessed, because individual-level SIMCE data and measures of indigenous status are not available in the first half of the 1990's. A second interpretation, more troubling in its implications, is that the reforms did not lead to significant improvements in academic achievement of students. The empirical

---

<sup>23</sup> Another explanation, less plausible, is that the Chilean reforms contributed to lockstep improvements in the outcomes of indigenous *and* non-indigenous student populations. This would maintain the size of the test score gap, and the relative importance of school quality in explaining it. However, one must consider that the 10% of students enrolled in elite private schools—very few of them indigenous—did not participate in Ministry-led reforms. Additionally, some private subsidized schools opted not to participate in compensatory programs such as P-900.

evidence is not yet sufficient to convincingly resolve this issue.<sup>24</sup> A third interpretation, already mentioned, is that the importance of “school quality” is simply masking the importance of peer-groups effects.

In Bolivia, the 1997 SIMECAL data barely preceded the implementation of Bolivia’s large-scale education reform. Thus, it should be a priority to rigorously evaluate the causal impact of the reform, especially how its effects may vary across indigenous and non-indigenous students.

---

<sup>24</sup> It has proven difficult to evaluate the causal impact of Chile’s reforms on achievement because they were non-randomly assigned. One recent evaluation of the P-900 applied a quasi-experimental regression-discontinuity design, taking advantage of the fact that schools were more likely to receive the program if they were below a specified cut-off test score within each region (McEwan & Urquiola, 2002). It finds mixed evidence of program effects. Chile’s MECE program, including sub-components like MECE-Rural, was distributed non-randomly and it was also expanded quickly to the vast majority of publicly-funded schools. Both features have made it problematic to evaluate.

## References

- Albó, X. (1995). *Bolivia plurilingüe: Guía para planificadores y educadores* (Vol. 1). La Paz: UNICEF-CIPCA.
- Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 533-575.
- Chiswick, B. R., Patrinos, H. A., & Hurst, M. E. (2000). Indigenous language skills and the labor market in a developing economy: Bolivia. *Economic Development and Cultural Change*, 349-367.
- Comitas, L. (1972). Education and social stratification in contemporary Bolivia. In T. J. LaBelle (Ed.), *Education and development: Latin America and the Caribbean* (pp. 363-378). Los Angeles: University of California Press.
- 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.
- Dee, T. S. (2001). *Teachers, race and student achievement in a randomized experiment*. Working Paper No. 8432. Cambridge, MA: National Bureau of Economic Research.
- 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* (pp. 273-317). Washington, DC: Brookings Institution Press.
- Gauri, V. (1998). *School choice in Chile*. Pittsburgh: University of Pittsburgh Press.
- Grinevald, C. (1998). Language endangerment in South America: A programmatic approach. In L. A. Grenoble & L. J. Whaley (Eds.), *Endangered languages: Current issues and future prospects* (pp. 124-159). Cambridge: Cambridge University Press.
- 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.
- Hornberger, N. H. (2000). Bilingual education policy and practice in the Andes: Ideological paradox and intercultural possibility. *Anthropology and Education Quarterly*, 31(2), 173-201.
- Hornberger, N. H., & King, K. A. (1996). Language revitalization in the Andes: Can the schools reverse language shift? *Journal of Multilingual and Multicultural Development*, 17(6), 427-441.

- Hsieh, C.-T., & Urquiola, M. (2002). *When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program*. Unpublished manuscript, Princeton University and Cornell University.
- Instituto Nacional de Estadística (INE). (1993). *Resultados oficiales censo de población*. Santiago: INE.
- Jencks, C. & Phillips, M. (Eds.). (1998). *The black-white test score gap*. Washington, DC: Brookings Institution Press.
- Jiménez, W., & McEwan, P. J. (2001). *Estudiantes indígenas y acceso a la educación primaria en Bolivia*. Unpublished manuscript.
- McEwan, P. J. (in press). Peer effects on student achievement: Evidence from Chile. *Economics of Education Review*.
- McEwan, P. J. (2001). The effectiveness of public, Catholic, and non-religious private schools in Chile's voucher system. *Education Economics*, 9(2), 103-128.
- McEwan, P. J., & Carnoy, M. (2000). The effectiveness and efficiency of private schools in Chile's voucher system. *Educational Evaluation and Policy Analysis*, 22(3), 213-239.
- McEwan, P. J., & Jiménez, W. (2001). *Indigenous students in Bolivian primary schools: Patterns and determinants of inequities*. Unpublished manuscript.
- McEwan, P. J., & Urquiola, M. (2002). *Did Chile's P-900 program really work? A regression-discontinuity approach*. Unpublished manuscript.
- Patrinos, H. A., & Psacharopoulos, G. (1996). Socioeconomic and ethnic determinants of age-grade distortion. *International Journal of Educational Development*, 16(1), 3-14.
- Plaza, P., & Albó, X. (1989). Educación bilingüe y planificación lingüística en Bolivia. *International Journal of the Sociology of Language*, 77, 69-91.
- Psacharopoulos, G. (1993). Ethnicity, education, and earnings in Bolivia and Guatemala. *Comparative Education Review*, 37(1), 9-20.
- Psacharopoulos, G., & Patrinos, H. A. (Eds.). (1994). *Indigenous people and poverty in Latin America*. Washington, DC: The World Bank.
- Rivarola, D. M., Corvalán, G., & Zuniga, L. (1977). *Determinantes del rendimiento educativo en el Paraguay*. Asunción: Centro Paraguayo de Estudios Sociológicos.
- Urquiola, M. (2000). *Identifying class size effects in developing countries: Evidence from rural schools in Bolivia*. Unpublished manuscript, Cornell University.

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

World Bank. (2001). *Peruvian education at a crossroads: Challenges and opportunities for the 21st century*. Washington, DC: World Bank.

Table 1  
Indigenous populations in Bolivia and Chile, 1992

	Number (thousands)	Percent of population
<b>Bolivia</b>		
Knows Quechua	1,806	34%
Knows Aymara	1,238	23%
Knows another indigenous language	70	2%
Knows <i>only</i> Quechua	428	8%
Knows <i>only</i> Aymara	169	3%
Knows <i>only</i> another indigenous language	11	<1%
<b>Chile</b>		
Self-identifies as Mapuche	928	10%
Self-identifies as Aymara	48	<1%
Self-identifies as Rapanui	22	<1%

Source: The Bolivian data are from Albó's (1995) tabulations of the 1992 census. The Chilean data are from INE's (1993) tabulations of the 1992 census.

Notes: Bolivian data refer to the population ages 6 and over. Chilean data refer to the population ages 14 and over.

Table 2  
Percentages of indigenous students in Bolivian primary schools

	Indigenous	Quechua	Aymara	Dissimilarity index	Observations
<i>Third-grade, 1997</i>					
Chuquisaca	42%	40%	1%	0.61	1,232
La Paz	40%	3%	38%	0.54	3,008
Cochabamba	46%	44%	4%	0.45	1,863
Oruro	39%	21%	24%	0.42	1,477
Potosi	59%	57%	5%	0.61	1,327
Tarija	5%	3%	1%	0.48	984
Santa Cruz	9%	6%	1%	0.54	1,634
Beni	2%	1%	1%	0.65	1,280
Pando	3%	1%	1%	0.53	362
<b>National</b>	<b>32%</b>	<b>20%</b>	<b>13%</b>	<b>0.61</b>	<b>13,167</b>
<i>Sixth-grade, 1997</i>					
Chuquisaca	27%	27%	0%	0.50	1,153
La Paz	30%	4%	27%	0.50	2,703
Cochabamba	26%	25%	2%	0.37	2,092
Oruro	35%	22%	20%	0.38	1,421
Potosi	43%	43%	2%	0.53	1,424
Tarija	4%	3%	1%	0.44	1,368
Santa Cruz	5%	3%	1%	0.57	1,672
Beni	1%	0%	0%	0.70	1,308
Pando	1%	1%	1%	0.44	281
<b>National</b>	<b>22%</b>	<b>15%</b>	<b>8%</b>	<b>0.57</b>	<b>13,422</b>

Source: SIMECAL, 1997 and author's calculations

Notes: See text for definitions of indigenous

Table 3  
Percentages of indigenous students in Chilean primary schools

	Indigenous	Mapuche	Dissimilarity index	Observations
<i>Fourth-grade, 1999</i>				
Region 1	11%	2%	0.40	6,722
Region 2	7%	1%	0.42	8,444
Region 3	19%	1%	0.25	4,094
Region 4	3%	2%	0.40	9,724
Region 5	3%	2%	0.46	24,422
Region 6	4%	3%	0.46	12,717
Region 7	4%	4%	0.43	15,508
Region 8	7%	6%	0.40	30,828
Region 9	26%	25%	0.48	13,155
Region 10	13%	12%	0.39	13,064
Region 11	14%	13%	0.31	1,489
Region 12	6%	5%	0.36	2,371
Region 13	5%	5%	0.36	93,096
<b>National</b>	<b>7%</b>	<b>6%</b>	<b>0.45</b>	<b>235,634</b>
<i>Eighth-grade, 1997</i>				
Region 1	9%	1%	0.40	5,527
Region 2	5%	1%	0.47	6,474
Region 3	19%	1%	0.23	3,561
Region 4	2%	1%	0.48	7,609
Region 5	2%	1%	0.52	20,774
Region 6	2%	2%	0.51	9,955
Region 7	2%	2%	0.52	11,831
Region 8	4%	4%	0.45	24,986
Region 9	22%	22%	0.50	10,569
Region 10	9%	8%	0.40	12,084
Region 11	6%	6%	0.37	1,239
Region 12	3%	3%	0.43	2,328
Region 13	4%	3%	0.41	75,452
<b>National</b>	<b>5%</b>	<b>4%</b>	<b>0.50</b>	<b>192,389</b>

Source: SIMCE, 1997, 1999 and author's calculations

Notes: See text for definitions of indigenous. Region 13 includes Santiago.

Table 4  
Sample descriptive statistics, Bolivia

	Third-grade, 1997				Sixth-grade, 1997			
	Full sample	Non-indigenous (INDIG=0)	Indigenous (INDIG=1)	Difference	Full sample	Non-indigenous (INDIG=0)	Indigenous (INDIG=1)	Difference
SPANISH	0.00 (1.00)	0.10 (0.99)	-0.23 (0.98)	-0.33	0.00 (1.00)	0.10 (1.00)	-0.37 (0.92)	-0.48
MATH	0.00 (1.00)	0.08 (0.98)	-0.19 (1.01)	-0.27	0.00 (1.00)	0.07 (1.01)	-0.27 (0.92)	-0.35
INDIG	0.30	---	---	---	0.22	---	---	---
QUECHUA	0.19	---	0.64	---	0.15	---	0.69	---
AYMARA	0.12	---	0.41	---	0.08	---	0.37	---
FEMALE	0.52	0.54	0.49	-0.04	0.50	0.51	0.48	-0.03
EDMTH	1.79	2.08 (1.37)	1.14 (0.99)	-0.94	1.97 (1.36)	2.16 (1.36)	1.26 (1.07)	-0.90
EDFTH	2.23 (1.38)	2.47 (1.40)	1.67 (1.15)	-0.80	2.41 (1.38)	2.58 (1.38)	1.78 (1.15)	-0.80
SEWER	0.30	0.35	0.17	-0.18	0.38	0.41	0.26	-0.15
ELECT	0.62	0.70	0.44	-0.26	0.73	0.78	0.56	-0.21
PHONE	0.17	0.22	0.06	-0.16	0.22	0.27	0.07	-0.19
M(INDIG)	0.31 (0.31)	0.17 (0.19)	0.63 (0.30)	0.46	0.22 (0.25)	0.14 (0.16)	0.51 (0.30)	0.37
M(EDMTH)	1.79 (0.87)	1.99 (0.88)	1.31 (0.61)	-0.69	1.96 (0.84)	2.10 (0.85)	1.48 (0.59)	-0.62
CSIZE	32.71 (9.62)	34.23 (9.13)	29.20 (9.80)	-5.03	34.57 (9.26)	35.31 (8.90)	31.92 (10.02)	-3.38
PRIVATE	0.11	0.14	0.03	-0.11	0.15	0.18	0.03	-0.15
CONVENIO	0.06	0.07	0.05	-0.02	0.05	0.05	0.04	-0.01
RURAL	0.39	0.28	0.64	0.35	0.25	0.19	0.47	0.28
Observations	10,954	7,649	3,305		11,469	8,981	2,488	

Source: SIMECAL, 1997 and author's calculations

Note: Standard deviations are in parentheses for non-dichotomous variables.

Table 5  
Sample descriptive statistics, Chile

	Fourth-grade, 1999				Eighth-grade, 1997			
	Full sample	Non-indigenous (INDIG=0)	Indigenous (INDIG=1)	Difference	Full sample	Non-indigenous (INDIG=0)	Indigenous (INDIG=1)	Difference
SPANISH	0.00 (1.00)	0.03 (1.00)	-0.37 (0.95)	-0.39	0.00 (1.00)	0.02 (0.99)	-0.44 (1.00)	-0.47
MATH	0.00 (1.00)	0.02 (1.00)	-0.35 (0.95)	-0.37	0.00 (1.00)	0.02 (1.00)	-0.38 (0.93)	-0.40
INDIG	0.07	---	---	---	0.05	---	---	---
MAPUCHE	0.05	---	0.81	---	0.04	---	0.80	---
OTHER	0.01	---	0.19	---	0.01	---	0.20	---
FEMALE	---	---	---	---	0.52	0.52	0.51	-0.01
EDMTH	10.19 (3.57)	10.34 (3.52)	7.96 (3.54)	-2.38	9.75	9.87 (3.75)	7.50 (3.62)	-2.36
EDFTH	10.54 (3.75)	10.68 (3.72)	8.42 (3.57)	-2.26	10.16 (4.06)	10.27 (4.04)	7.94 (3.76)	-2.33
BOOKS	---	---	---	---	4.11 (2.47)	4.17 (2.47)	3.01 (2.11)	-1.16
INCOME	2.63 (3.79)	2.72 (3.87)	1.40 (1.97)	-1.32	2.99 (3.98)	3.06 (4.04)	1.66 (2.29)	-1.40
M(INDIG)	0.07 (0.10)	0.06 (0.07)	0.20 (0.23)	0.15	0.05 (0.09)	0.04 (0.06)	0.21 (0.25)	0.17
M(EDMTH)	10.09 (2.29)	10.19 (2.29)	8.71 (1.89)	-1.47	9.72 (2.41)	9.79 (2.41)	8.24 (2.06)	-1.55
CSIZE	36.41 (8.28)	36.55 (8.12)	34.41 (10.05)	-2.14	34.23 (8.53)	34.36 (8.41)	31.73 (10.22)	-2.63
PRIVSUB	0.38	0.38	0.32	-0.06	0.34	0.34	0.30	-0.04
PRIVPD	0.08	0.09	0.01	-0.08	0.09	0.09	0.01	-0.08
RURAL	0.13	0.13	0.24	0.11	0.07	0.07	0.18	0.11
Observations	196,167	183,404	12,763		163,061	155,127	7,934	

Source: SIMCE, 1997, 1999 and author's calculations

Note: Standard deviations are in parentheses for non-dichotomous variables.

Table 6  
Bolivian achievement regressions, Grade 3

	Dependent variable: SPANISH				Dependent variable: MATH			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INDIG	-0.333** (0.058)	-0.080** (0.023)	-0.080** (0.021)	-0.056** (0.020)	-0.274** (0.054)	-0.074** (0.024)	-0.080** (0.023)	-0.055* (0.022)
FEMALE	---	0.046* (0.023)	0.033* (0.016)	0.021 (0.016)	---	-0.034 (0.022)	-0.043** (0.016)	-0.050** (0.017)
EDMTH	---	0.043** (0.011)	0.044** (0.010)	0.039** (0.010)	---	0.035** (0.010)	0.033** (0.009)	0.031** (0.009)
EDFTH	---	0.012 (0.011)	0.027** (0.008)	0.027** (0.009)	---	0.008 (0.011)	0.025** (0.008)	0.023** (0.008)
SEWER	---	0.101* (0.044)	0.007 (0.023)	0.010 (0.022)	---	0.117** (0.042)	0.031 (0.023)	0.030 (0.023)
ELECT	---	0.121** (0.038)	0.096** (0.020)	0.092** (0.020)	---	0.087* (0.037)	0.079** (0.021)	0.074** (0.020)
PHONE	---	0.210** (0.039)	0.059* (0.027)	0.040 (0.026)	---	0.183** (0.037)	0.048 (0.025)	0.026 (0.024)
M(INDIG)	---	0.034 (0.124)	---	---	---	0.100 (0.124)	---	---
M(EDMTH)	---	0.120* (0.058)	---	---	---	0.118* (0.058)	---	---
CSIZE	---	-0.012** (0.004)	---	---	---	-0.013** (0.004)	---	---
PRIVATE	---	0.504** (0.131)	---	---	---	0.460** (0.126)	---	---
CONVENIO	---	0.086 (0.124)	---	---	---	0.168 (0.114)	---	---
RURAL	---	-0.154 (0.093)	---	---	---	-0.188* (0.084)	---	---
Observations	10954	10954	10954	10954	10954	10954	10954	10954
R-squared	0.02	0.16	0.51	0.56	0.02	0.13	0.47	0.52
School effects?	No	No	Yes	No	No	No	Yes	No
Classroom effects?	No	No	No	Yes	No	No	No	Yes

Source: SIMECAL, 1997 and author's calculations

Note: Huber-White standard errors are in parentheses. \*\*(\*) indicates statistical significance at 1% (5%).

Table 7  
Bolivian achievement regressions, Grade 6

	Dependent variable: SPANISH				Dependent variable: MATH			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INDIG	-0.475** (0.059)	-0.055* (0.023)	-0.057* (0.023)	-0.043 (0.022)	-0.345** (0.057)	-0.027 (0.022)	-0.022 (0.022)	-0.017 (0.021)
FEMALE	---	0.028 (0.028)	0.028 (0.017)	0.021 (0.017)	---	-0.098** (0.030)	-0.076** (0.018)	-0.087** (0.017)
EDMTH	---	0.015 (0.010)	0.021* (0.010)	0.021* (0.009)	---	0.008 (0.010)	0.012 (0.009)	0.013 (0.009)
EDFTH	---	0.034** (0.010)	0.036** (0.008)	0.035** (0.008)	---	0.013 (0.009)	0.022** (0.008)	0.021** (0.008)
SEWER	---	0.109** (0.036)	0.044 (0.023)	0.042 (0.023)	---	0.074* (0.037)	0.028 (0.021)	0.023 (0.021)
ELECT	---	0.216** (0.033)	0.234** (0.021)	0.219** (0.021)	---	0.166** (0.036)	0.185** (0.023)	0.176** (0.024)
PHONE	---	0.145** (0.029)	0.089** (0.023)	0.097** (0.023)	---	0.129** (0.033)	0.050* (0.025)	0.052* (0.025)
M(INDIG)	---	-0.177 (0.138)	---	---	---	-0.232 (0.159)	---	---
M(EDMTH)	---	0.276** (0.049)	---	---	---	0.210** (0.061)	---	---
CSIZE	---	0.000 (0.003)	---	---	---	-0.002 (0.003)	---	---
PRIVATE	---	0.373** (0.105)	---	---	---	0.413** (0.133)	---	---
CONVENIO	---	0.369** (0.085)	---	---	---	0.293** (0.091)	---	---
RURAL	---	0.026 (0.080)	---	---	---	0.150 (0.088)	---	---
Observations	11469	11469	11469	11469	11469	11469	11469	11469
R-squared	0.04	0.28	0.47	0.52	0.02	0.17	0.42	0.46
School effects?	No	No	Yes	No	No	No	Yes	No
Classroom effects?	No	No	No	Yes	No	No	No	Yes

Source: SIMECAL, 1997 and author's calculations

Note: Huber-White standard errors are in parentheses. \*\*(\*) indicates statistical significance at 1% (5%).

Table 8  
Chilean achievement regressions, Grade 4

	Dependent variable: SPANISH				Dependent variable: MATH			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INDIG	-0.392** (0.013)	-0.041** (0.009)	-0.040** (0.009)	-0.032** (0.009)	-0.374** (0.013)	-0.055** (0.009)	-0.054** (0.010)	-0.047** (0.010)
EDMTH	---	0.042** (0.001)	0.041** (0.001)	0.038** (0.001)	---	0.038** (0.001)	0.038** (0.001)	0.035** (0.001)
EDFTH	---	0.028** (0.001)	0.028** (0.001)	0.026** (0.001)	---	0.025** (0.001)	0.026** (0.001)	0.024** (0.001)
INCOME	---	0.007** (0.001)	0.008** (0.001)	0.007** (0.001)	---	0.012** (0.001)	0.010** (0.001)	0.009** (0.001)
M(INDIG)	---	-0.061 (0.048)	---	---	---	-0.049 (0.048)	---	---
M(EDMTH)	---	0.128** (0.004)	---	---	---	0.110** (0.004)	---	---
CSIZE	---	-0.001* (0.001)	---	---	---	-0.002** (0.001)	---	---
PRIVSUB	---	0.039** (0.013)	---	---	---	0.032* (0.013)	---	---
PRIVPD	---	-0.065* (0.025)	---	---	---	-0.002 (0.026)	---	---
RURAL	---	0.109** (0.016)	---	---	---	0.081** (0.016)	---	---
Observations	196167	196167	196167	196167	196167	196167	196167	196167
R-squared	0.01	0.24	0.32	0.36	0.01	0.21	0.29	0.33
School effects?	No	No	Yes	No	No	No	Yes	No
Classroom effects?	No	No	No	Yes	No	No	No	Yes

Source: SIMCE, 1999 and author's calculations

Note: Huber-White standard errors are in parentheses. \*\*(\*) indicates statistical significance at 1% (5%).

Table 9  
Chilean achievement regressions, Grade 8

	Dependent variable: SPANISH				Dependent variable: MATH			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INDIG	-0.468** (0.017)	-0.096** (0.011)	-0.097** (0.011)	-0.085** (0.011)	-0.397** (0.017)	-0.068** (0.011)	-0.068** (0.010)	-0.056** (0.010)
FEMALE	---	0.203** (0.008)	0.176** (0.005)	0.167** (0.005)	---	-0.027* (0.011)	-0.068** (0.005)	-0.076** (0.005)
EDMTH	---	0.025** (0.001)	0.025** (0.001)	0.022** (0.001)	---	0.020** (0.001)	0.019** (0.001)	0.016** (0.001)
EDFTH	---	0.013** (0.001)	0.014** (0.001)	0.012** (0.001)	---	0.007** (0.001)	0.010** (0.001)	0.008** (0.001)
BOOKS	---	0.049** (0.001)	0.044** (0.001)	0.040** (0.001)	---	0.040** (0.002)	0.036** (0.001)	0.032** (0.001)
INCOME	---	-0.007** (0.001)	-0.006** (0.001)	-0.006** (0.001)	---	-0.001 (0.002)	-0.003** (0.001)	-0.003** (0.001)
M(INDIG)	---	-0.327** (0.060)	---	---	---	-0.258** (0.072)	---	---
M(EDMTH)	---	0.120** (0.005)	---	---	---	0.131** (0.008)	---	---
CSIZE	---	0.002* (0.001)	---	---	---	0.000 (0.001)	---	---
PRIVSUB	---	0.020 (0.018)	---	---	---	-0.004 (0.025)	---	---
PRIVPD	---	0.026 (0.040)	---	---	---	0.073 (0.058)	---	---
RURAL	---	0.130** (0.021)	---	---	---	0.290** (0.026)	---	---
Observations	163061	163061	163061	163061	163061	163061	163061	163061
R-squared	0.01	0.23	0.34	0.38	0.01	0.20	0.39	0.44
School effects?	No	No	Yes	No	No	No	Yes	No
Classroom effects?	No	No	No	Yes	No	No	No	Yes

Source: SIMCE, 1997 and author's calculations

Note: Huber-White standard errors are in parentheses. \*\*(\*) indicates statistical significance at 1% (5%).

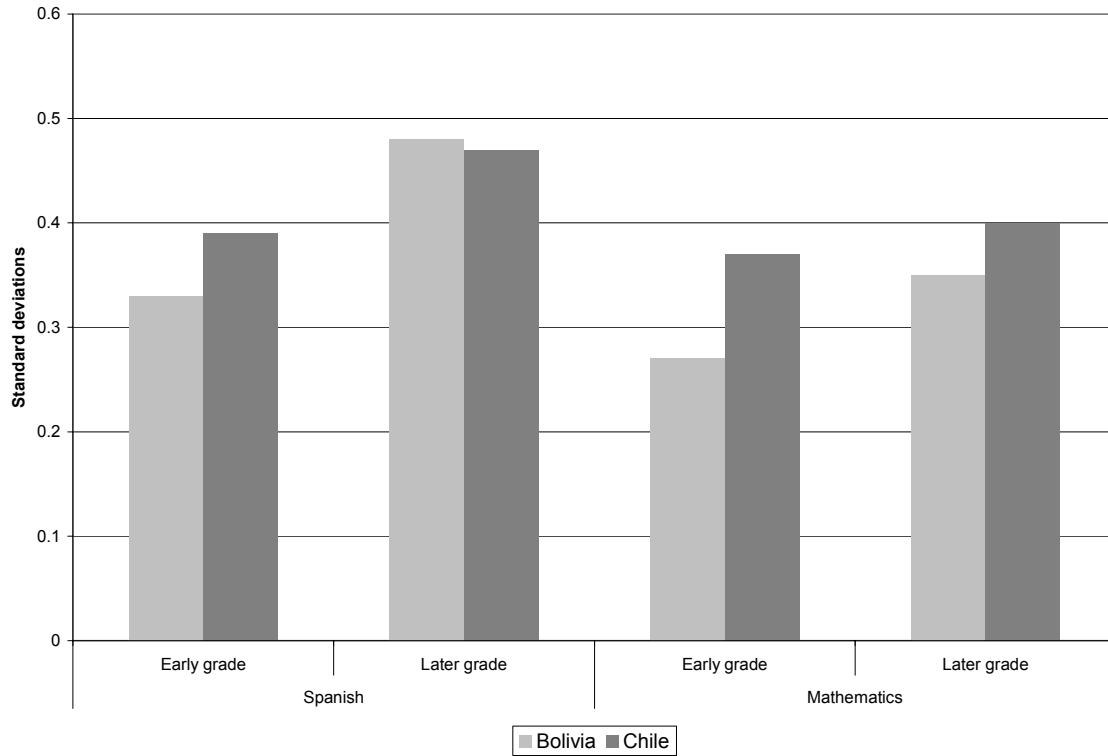
Table 10  
Achievement decompositions, Bolivia

	Dependent variable: SPANISH			Dependent variable: MATH		
	(2)	(3)	(4)	(6)	(7)	(8)
<i>Third Grade, 1997</i>						
“Unexplained”	-0.08	-0.08	-0.06	-0.07	-0.08	-0.05
Student variables	-0.13	-0.10	-0.09	-0.11	-0.08	-0.07
Peer variables	-0.07	---	---	-0.04	---	---
School variables	-0.05	---	---	-0.05	---	---
School fixed effects	---	-0.15	---	---	-0.11	---
Classroom fixed effects	---	---	-0.19	---	---	-0.14
Total	-0.33	-0.33	-0.33	-0.27	-0.27	-0.27
<i>Sixth Grade, 1997</i>						
“Unexplained”	-0.06	-0.06	-0.04	-0.03	-0.02	-0.02
Student variables	-0.13	-0.12	-0.12	-0.09	-0.08	-0.08
Peer variables	-0.24	---	---	-0.22	---	---
School variables	-0.05	---	---	-0.01	---	---
School fixed effects	---	-0.30	---	---	-0.24	---
Classroom fixed effects	---	---	-0.31	---	---	-0.25
Total	-0.48	-0.48	-0.48	-0.35	-0.35	-0.35

Table 11  
Achievement decompositions, Chile

	Dependent variable: SPANISH			Dependent variable: MATH		
	(2)	(3)	(4)	(6)	(7)	(8)
<i>Fourth-grade, 1999</i>						
“Unexplained”	-0.04	-0.04	-0.03	-0.05	-0.05	-0.05
Student variables	-0.17	-0.17	-0.16	-0.16	-0.16	-0.15
Peer variables	-0.20	---	---	-0.17	---	---
School variables	0.02	---	---	0.01	---	---
School fixed effects	---	-0.18	---	---	-0.16	---
Classroom fixed effects	---	---	-0.20	---	---	-0.18
Total	-0.39	-0.39	-0.39	-0.37	-0.37	-0.37
<i>Eighth-grade, 1997</i>						
“Unexplained”	-0.10	-0.10	-0.09	-0.07	-0.07	-0.06
Student variables	-0.14	-0.14	-0.12	-0.11	-0.10	-0.09
Peer variables	-0.24	---	---	-0.25	---	---
School variables	0.01	---	---	0.03	---	---
School fixed effects	---	-0.24	---	---	-0.22	---
Classroom fixed effects	---	---	-0.26	---	---	-0.25
Total	-0.47	-0.47	-0.47	-0.40	-0.40	-0.40

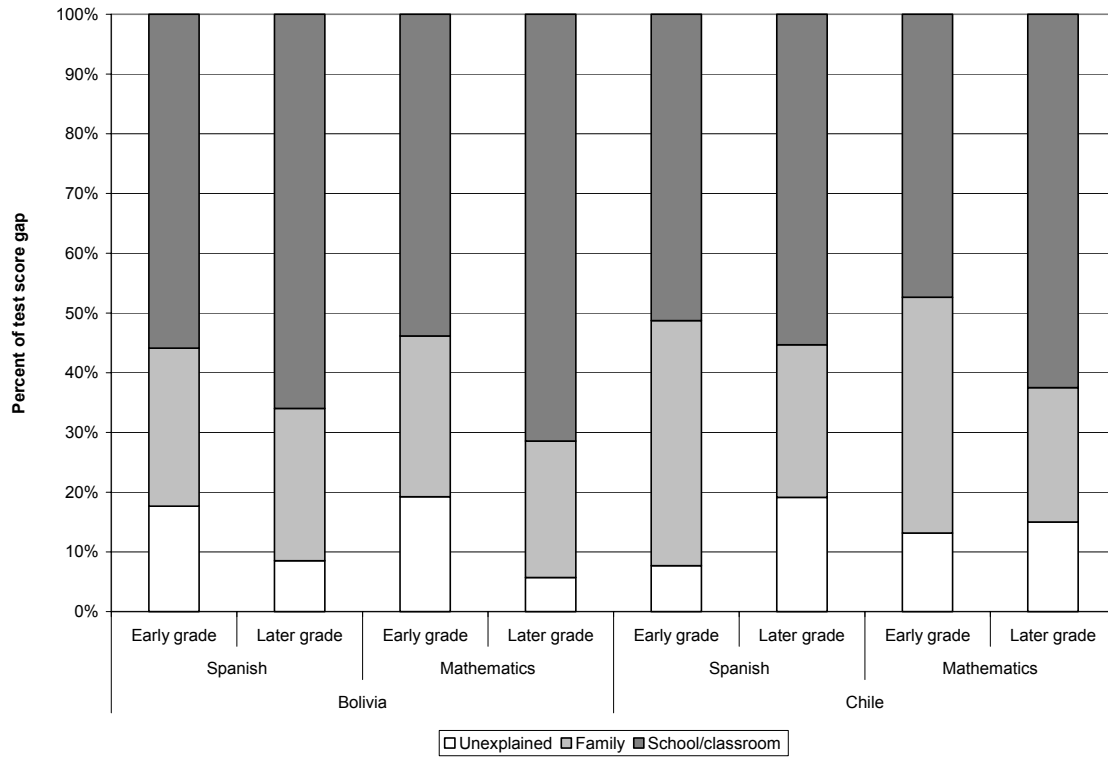
Figure 1  
 Indigenous test score gaps in Bolivia and Chile



Source: Columns 1 and 5 in Table 6-9

Note: In Bolivia and Chile, “early grade” refers to third- and fourth-grade, respectively, while “later grade” refers to sixth- and eighth-grade, respectively.

Figure 2  
 Decomposition of indigenous test score gaps in Bolivia and Chile



Source: Columns 4 and 8 in Tables 10 and 11

Note: In Bolivia and Chile, “early grade” refers to third- and fourth-grade, respectively, while “later grade” refers to sixth- and eighth-grade, respectively.

## Appendix: Variable definitions

	Bolivia	Chile
<i>Dependent variables</i>		
SPANISH	Spanish test score (standardized to a mean of 0 and a standard deviation of 1).	Spanish test score (standardized to a mean of 0 and a standard deviation of 1).
MATH	Mathematics test score (standardized to a mean of 0 and a standard deviation of 1).	Mathematics test score (standardized to a mean of 0 and a standard deviation of 1).
<i>Independent variables</i>		
INDIG	1=parent or guardian reports that indigenous languages are used to communicate in the home; 0=not	1=mother self-identifies as member of an indigenous group; 0=not
QUECHUA	1=parent or guardian reports that Quechua is used to communicate in the home; 0=not	---
AYMARA	1=parent or guardian reports that Aymara is used to communicate in the home; 0=not	---
MAPUCHE	---	1=mother self-identifies as Mapuche; 0=not
OTHINDIG	---	1=mother self-identifies as member of a non-Mapuche indigenous group; 0=not
FEMALE	1=female; 0=male	1=female; 0=male
EDMTH	Categorical variable ranging from 0 (mother did not finish primary education) to 5 (mother finished higher education)	Years of mother's schooling
EDFTH	Categorical variable ranging from 0 (father did not finish primary education) to 5 (father finished higher education)	Years of father's schooling
SEWER	1=student's family has a sewer connection; 0=not	---
ELECT	1=student's family has electricity; 0=not	---
PHONE	1=student's family has a telephone; 0=not	---
BOOKS	---	Categorical variable ranging from 1 (5 or fewer books in home) to 8 (more than 95 books in home)
INCOME	---	Family income
M(INDIG)	School-level mean of INDIG	School-level mean of INDIG
M(EDMTH)	School-level mean of EDMTH	School-level mean of EDMTH
CSIZE	Number of students in class	Number of students in class
PRIVATE	1=student attends private school; 0=not	---
CONVENIO	1=student attends "convenio" school; 0=not	---
PRIVSUB	---	1=student attends private school that receives government subsidies; 0=not
PRIVPD	---	1=student attends private school that does not receive government subsidies; 0=not
RURAL	1=rural school; 0=not	1=rural school; 0=not