

# A longitudinal analysis of the effect of bilingual schooling on the educational outcomes of indigenous children in Mexico

Erica Soler-Hampejsek and Kelly Hallman

Indigenous populations in Latin American countries face persistent disadvantages. Previous studies have shown that, on average and in different countries of the region, non-indigenous children are more likely to be enrolled in school and attain higher levels of education than indigenous children (Hall and Patrinos 2005; Hallman et al 2007). In Mexico, according to the 2000 census, about seven percent of individuals over the age of five speak an indigenous language and are, therefore, considered to be indigenous. Mexico's public educational system offers three types of primary schools: general; communitarian (which operate in rural localities with less than 100 inhabitants), and indigenous or bilingual (which operate in a number of indigenous communities and use bilingual teachers and textbooks in the indigenous language of the community). Parker, Rubalcava, and Teruel (2003), who look at the determinants of educational outcomes of indigenous children in Mexico using cross-sectional data, find that being monolingual (that is, only speaking an indigenous language) has very large negative effects on school attendance and schooling attainment relative to other indigenous children that do speak Spanish. These authors also present some initial evidence that primary schools that practice bilingual education improve the educational performance of monolingual children relative to bilingual children.

This paper will use longitudinal data to explore the educational disadvantages that indigenous children face in Mexico and whether bilingual schools help mitigate these disadvantages. In particular, we will go beyond standard measures of education—that is, school participation and highest grade attained—and also look at differences in grade repetition, grade attainment for age, and cognitive skills between monolingual, bilingual, and non-indigenous children. We will use longitudinal data from the Mexican Family Life Survey (MXFLS) conducted in 2002 and 2005 and from the country's Secretary of Public Education (SEP). The longitudinal nature of the MXFLS allows to compare children's progress in school and in cognitive skills between the two rounds as well as to explore transitions from being monolingual to being bilingual, while addressing unobserved heterogeneity.

## Data

The MXFLS contains detailed information on a wide array of social, economic, demographic and health behaviors of individuals and their families. It also includes comprehensive information at the community and facility levels, including data on the availability of bilingual instruction in schools that is not readily available in most data sets in the region. Furthermore, the MXFLS asked the name and location of the schools that children in the sample younger than age 15 attend. This data can be linked to school-level information from the Secretary of Public Education (SEP) to determine the type of school that children attend (general, communitarian, or bilingual).

As of now, two waves of the survey have been conducted. The first round, conducted in 2002, covered over 8,400 households in 150 communities across Mexico. The second wave, conducted during 2005-2006, had a 90% re-contact rate at the household level. According to summary data from the first wave, slightly over 11% of children ages 5-14 reported speaking an indigenous language and of these around 14% do not to speak any Spanish.

## Analysis

First, to explore the educational disadvantages that monolingual and bilingual indigenous children face versus non-indigenous children, for each round of data we will estimate a model of the type:

$$s = a_0 + a_1\text{Monoling} + a_2\text{Biling} + a_3X + a_4F + a_5C + u + e \quad (1)$$

where:

s: educational outcome, such as school enrollment, grade repetition, cognitive skills

Monoling and Biling: indicator variables for monolingual and bilingual children respectively (reference category is non-indigenous children)

X: vector of individual characteristics such as age and sex

F: vector of family-level socioeconomic characteristics

C: vector of community-level characteristics, including characteristics of schools

u: other unobserved heterogeneity

e: random error

Second, to estimate the effect of bilingual education on the schooling outcomes of indigenous children, we will focus only on the sample of indigenous children and compare the performance of monolingual versus bilingual children and explore the effect of indigenous schools as compared to other types of school.

$$s = b_0 + b_1\text{Monoling} + b_2\text{IndigSch} + b_3\text{Monoling}*\text{IndigSch} + b_4X + b_5F + b_6C + u + e \quad (2)$$

where:

IndigSch: indicator variable for attending indigenous school

A problem with equations (1) and (2) is that we cannot directly control for the unobserved heterogeneity term, which in both equations might as well be correlated with the schooling outcome and therefore produce biased estimates. For example, children who have higher innate ability might be more likely to be both bilingual (if they come from an indigenous background) and to perform better in school. Equation (2) also faces a potential problem of endogenous placement of indigenous schools. To address concerns of potential endogeneity, we will attempt to re-estimate both equations using a two-step instrumental variable approach.

Finally, to further look into the effects of bilingual education and take advantage of the longitudinal nature of the data, we will estimate models to explore the effect of having

attended a bilingual school on changes in educational outcomes and transitions from being monolingual to being bilingual:

$$y = c_0 + c_1\text{IndigSch} + c_2X + c_3F + c_4C + u + e \quad (3)$$

where:

y: change in language ability; educational outcome

IndigSch: indicator variable for attending indigenous school

X: vector of individual characteristics such as age and sex

F: vector of family-level socioeconomic characteristics

C: vector of community-level characteristics, including characteristics of schools

u: other unobserved heterogeneity

e: random error

#### Reference List

Hall, Gillette and Harry Anthony Patrinos, eds. 2005. *Indigenous Peoples, Poverty and Human Development in Latin America 1994-2004*. Washington, DC: World Bank.

Hallman, Kelly and Sara Peracca, with Jen Catino and Marta Julia Ruiz. 2007. "Indigenous Girls in Guatemala: Poverty and Location," in M. Lewis and M. Lockheed, eds., *Exclusion, Gender and Schooling: Case Studies from the Developing World*. Washington, DC: Center for Global Development.

Parker, Susan W., Luis Ruvalcaba, and Graciela Teruel. 2003. "Language Barriers and Schooling Inequality of the Indigenous in Mexico". Pp. 145-177 in *Who's In and Who's Out: Social Exclusion in Latin America*. Jere R. Behrman, Alejandro Gaviria, and Miguel Szekely, eds. Inter-american Development Bank. Washington, DC: The John Hopkins University Press.