

Language Barrier to Human Capital Development

(Thesis abstract)

Fidel Pérez Macal
University of Western Ontario

The root of Guatemalan students' language barrier: households' preferences for school attributes or spatial segregation. (Job Market Paper)

In countries where more than one language is spoken, not being proficient in the language of instruction at schools may affect human capital development. In Guatemala, non-Spanish mother tongue (non-SMT) students' performance on national tests is much lower, on average, than SMT students. Pérez-Macal (2018) shows that linguistic segregation across schools is partially to blame for non-SMT students' lack of Spanish comprehension. This paper identifies how school attributes, children's non-Spanish language use, and spatial segregation determine parents' school choice, and hence shape students' linguistic sorting across schools. I estimate a model of demand for junior high schools in Guatemala in which parents consider schools as differentiated products. In particular, I allow for the degree of differentiation across schools to depend on characteristics of other parents who select the school in equilibrium. The results show that non-SMT parents value schools in which their child is likely to speak/learn Spanish. However, non-SMT parents prefer to sort into schools where other parents have a similar mother tongue. This latter preference dominates the former as we move away from the Guatemalan capital city, which leads to both spatial and linguistic segregation at schools.

Are non-Spanish mother tongue students facing a language barrier in the Guatemalan education system?

Non-Spanish mother tongue (non-SMT) students' achievement on the Guatemalan national tests is poor. On average, non-SMT students' reading and math test scores are 0.6 and 0.4 standard deviations lower, respectively, than SMT students' test scores. This paper analyzes how students' linguistic sorting across schools affects non-SMT students' Spanish comprehension given their non-Spanish language use, and therefore their educational achievement. To explain test score gaps between SMT and non-SMT students, I employ a longitudinal dataset which is representative of grade six students in 2010. Furthermore, to account for the endogeneity of students' non-Spanish language use, I use different estimation approaches such as two stage least squares, first differences, and local instrumental variables. The results show that, first, students' linguistic sorting across schools is a prominent factor for students' non-Spanish language use. Second, non-SMT students who frequently speak their mother tongue are not yet proficient in Spanish by the time they attend secondary schools. Last, students' non-Spanish language use explains about 70 percent of the test score gaps between SMT and non-SMT students.