Economía Mexicana. Nueva Época
ISSN: 1665-2045
ecomex@cide.edu
Centro de Investigación y Docencia Económicas, A.C.
México

Vivanco, Edgar Franco

When Change Matters. Identifying Score Gains School Determinants in Mexico: An Intra-cohort Value-added Approach

Economía Mexicana. Nueva Época, vol. II, 2013, pp. 361-405
Centro de Investigación y Docencia Económicas, A.C.
Distrito Federal, México

When Change Matters
Identifying Score Gains School Determinants in Mexico: An Intra-cohort Value-added Approach

Edgar Franco Vivanco*
Fecha de recepción: 29 de enero de 2012; fecha de aceptación: 14 de junio de 2013.

Abstract: Educational quality in Mexico is a recurrent topic in the country's public agenda. For this reason, better information about its determinants is required to guide policy makers. This discussion is particularly relevant within the context of the educational reform proposed by the federal government. Existing research for Mexican students is limited because it does not take into account the change in scores, or it is constrained by regional analysis. The aim of the present research is to fill this gap in the literature by using longitudinal data to study the impact of school characteristics on educational achievement, building a model that takes into account changes in educational achievement through time. Because existing data does not allow the tracking of students through time or their linkage to individual teachers, this article uses intra-cohort data for students that participated in the national standardized test ENLACE for primary and secondary schools between 2007 and 2010. This paper addresses endogeneity problems using fixed effects models, HLM models and spatial techniques to associate school location with census data at the neighborhood level. This research provides elements to guide public policies focused on increasing student achievement. Results show that teachers' attendance and punctuality, evaluation of teachers' knowledge of curriculum and constant evaluation of student performance have a positive relationship to student achievement growth. Additionally, there is a persistent effect of the quality of school infrastructure, teaching materials and socioeconomic level of students. These results could be useful to provide more information for the current debate on educational reform.
Keywords: education, value-added models, Mexico.

Cuando el cambio importa: La identificación de los determinantes escolares de la mejora de calificaciones en México, un enfoque de valor agregado dentro de cohortes

Resumen: La calidad educativa en México es un tema constante en la agenda pública. Por esa razón es necesario proveer de mejor información sobre sus deter-

*Edgar Franco Vivanco, edgarf1@stanford.edu, researcher, Program on Poverty and Governance, Freeman Spogli Institute for International Studies, Stanford University. Stanford, CA. USA.

minantes para la planeación de políticas públicas. Esta discusión cobra particular relevancia en la coyuntura actual en la que se propone una reforma educativa a gran escala. La investigación existente hasta la fecha para los estudiantes mexicanos tiene serias limitaciones porque no toma en cuenta el cambio en el desempeño educativo o, si lo hace, está restringida regionalmente. La presente investigación busca llenar el vacío en la literatura empleando datos longitudinales para estudiar el impacto de las características escolares en el logro educativo, por medio de un modelo que toma en cuenta el cambio en el desempeño a lo largo del tiempo. Debido a que con la información disponible no es posible seguir a los estudiantes a lo largo del tiempo o ligarlos con maestros de manera individual, este artículo utiliza información al interior de generaciones de estudiantes que participaron en la prueba ENLACE de primaria y secundaria entre 2007 y 2010. Este artículo busca reducir los problemas de edogeneidad utilizando modelos de efectos fijos, modelos multinivel y técnicas espaciales para relacionar la información de las escuelas con datos censales a nivel de sección geoelectoral. Aquí se proveen elementos para guiar la política pública enfocada a incrementar el logro escolar de los estudiantes. Los resultados muestran una relación positiva del incremento en el desempeño con el cumplimiento del tiempo de clase, con la evaluación del dominio del currículo de los maestros y con la evaluación constante del logro de los estudiantes. Asimismo, existe un efecto persistente de la infraestructura escolar, los materiales educativos y el nivel socioeconómico de los estudiantes. Estos resultados pueden servir para proveer mayor información en el debate de la actual reforma educativa en México.

Palabras clave: educación, modelos de valor agregado, México.

JEL classification: I24, I28, J33, N36.

Introduction

Poor educational achievement in Mexico is a perennial topic on the public policy agenda, with particular emphasis being paid over the past decade. This country consistently scores at the bottom of academic achievement among the OECD countries. In the 2009 PISA evaluation, 0.7 per cent of Mexican students scored in the highest level of mathematic skills; 0.4 per cent for language skills; and 0.2 per cent for science (OECD 2010, p. 35). These results rank Mexico at a lower level than poorer economies. There are many interpretations of this poor performance. For example, political economy perspectives remark the perverse incentives product of the decentralization process and the political muscle of the teachers union (Álvarez *et al.,* 2007; Ornelas, 2000). Other studies focus on the impact of particular policies or external shocks on schooling (Binder, 1999; McKenzie, 2003). Contrary to those analyses, the present research uses an educational production function approach to more accurately identify school level characteristics that determine student achievement, especially those related to teachers. Specifically, this paper addresses the question: What are the school characteristics that affect student performance change in Mexican schools? Conclusions from existing research in the same vein are limited because it is constrained to a single point in time (Álvarez *et al.,* 2007; Backhoff *et al.,* 2007; Brodziak, 2009; Valenti *et al.,* 2009; De Hoyos *et al.,* 2012; De la Vega, 2010) or to regional data (Santibáñez, 2006; Luschei, 2012; Rubio and Farías, 2013). The aim of this paper is to fill this gap in the literature developing a score gains value-added model using nationwide data to control for previous achievement and reduce selection bias.

In recent years, Mexico has taken important steps in the modernization of its educational system with the implementation of national reforms focused on educational quality, particularly the *Alianza por la Calidad de la Educación* in 2008, and *Pacto por México* in 2012; the latter led to the educational reform in 2013. The country has also been recognized by the OECD as a successful reformer after the increases presented in the PISA 2009 tests (OECD, 2010). Whereas Mexico is certainly in a better position compared to other Latin American countries regarding the quality and availability of information of the educational system, there are still important limitations. These restrictions come from the absence of long-term reliable administrative data. In Mexico, the national standardized test *Evaluación Nacional del Logro Académico en Centros Escolares* (Enlace) first began in 2006 for basic education (*primaria* and *secundaria*) and in 2008 for high school education. The information with student identifiers is not publicly available, so it is not possible to track students through time or identify their individual characteristics. There is also limited information for teachers, as there is no reliable register at the national level and it is not possible to link teachers with specific students or classrooms. School level information is also inconsistent and difficult to verify.[1] For these reasons, measuring school or teacher effects is a difficult task.

Value-added measures are a family of models built to disentangle those effects. Value-added models (VAM) recognize that educational achievement is affected by factors outside the school's control related to students' background, characteristics, funding and student mobility, among others (Harris, 2011). VAM take into account that in order to build a

[1] To avoid this limitation, some of the Mexican studies use Excale scores, which are more suitable for linking with student, teacher and school characteristics. However, a disadvantage of this test is that it is only performed in a sample with representativeness at state level.

causal link between teacher/school characteristics and student performance, it is crucial to understand the elements that determine change in student achievement across time. Additionally, VAM try to control for the effect of selection bias due to the non-random assignment of students to schools (Tiebout, 1956). This paper uses a simple definition of value, added as a model that is meant to approximate the contribution of the school on student performance (Braun *et al.,* 2010). These models have been implemented in educational research with very different approaches (McCaffrey *et al.*, 2003). However, their core assumption is that they should control for background factors and initial achievement. Although more sophisticated specifications of value-added models are interested in the difference between the predicted improvement and the actual improvement of the students, due to the characteristics of the Mexican data this research uses a score growth approach which could also be understood as a "quasi-value model", and is based on measuring the academic progress of a group of students controlling for several factors.

In order to overcome the shortcomings related to data availability, this research puts together Enlace scores from 2007 to 2010 for different cohorts. Since aggregated data at school level might be affected by confounding variables, one of the objectives of linking Enlace scores for several years is to track the same cohort from the middle point of their primary education in 3rd grade through their progression in primary school; analogously, it tracks progression of the cohort from the beginning of their secondary school in 7th grade. In addition, I complement the data with a questionnaire taken by school principals when their students take the Enlace test. These questionnaires provide a wide range of information about school conditions, teaching practices and school environment. Another source of information is the census data generated at electoral precinct level in a joint project between the *Instituto Federal Electoral* (IFE) and the *Instituto Nacional de Estadística y Geografía* (INEGI). This information allows identifying several socio-demographic characteristics at a very local geographical level which might be influencing both student performance and school variables (IFE-INEGI, 2010).

Using a fixed effects model and an HLM model, I find that schools with larger intra-cohort gains have a stronger system for monitoring teachers' content domain and student performance. Schools with larger intra-cohort gains also have teachers that are more likely to arrive on time to their classes. After controlling for several measures of student and neighborhood socioeconomic level, the impact of basic infrastructure of the school is still high, which implies that lacking adequate facilities is an important factor explaining achievement gaps (Woessmann, 2003). Although results presented in this paper do not claim causality in determining schools with large performance growths, they can be useful for pointing out the elements that are in the scope of authorities to promote long-term gains in student attainment. This research builds on the large literature studying school effects that emanates from the so called Coleman Report (Coleman *et al.,* 1966) and tries to disentangle the question of how important schools really are in increasing student achievement.

I. Determinants of Change in Student Achievement: A Literature Review

Academic achievement is understood as a "cumulative function of current and prior family, community and school experiences" (Rivkin *et al.,* 2005, p. 422). Based on this principle, recent literature interested in measuring the impact of those elements on academic achievement highlights the importance of understanding the rate of learning over time using VAM, of which the primary objective is to reduce confounding and unobserved influences. Researchers and policymakers are attracted to those models because in statistical terms they are able to separate the effects of teachers and schools from non-educational factors like family background and student characteristics.

Value-added models are used for two primary purposes: first, to hold teachers and schools accountable and to reward or punish them based on their performance; and second, to identify differences across teachers and schools in order to improve education (McCaffrey *et al.,* 2003). While there are several models that try to estimate those effects, they vary in their approach and the type of data used, and therefore differ in their results. For example, Sanders and Rivers (1996) use the Tennessee Value-Added Assessment System (TVAAS) data to study the cumulative effect of teachers in a single student cohort from grades 2^{nd} to 5^{th}, as well as the differential effects of teachers on students of different race and varying levels of achievement. Authors estimate teacher effects and effectiveness with a mixed-effects model with current-year score as dependent variable and prior year score as independent, plus a random teacher effect. They find that there are consequences of having ineffective teachers: students with three consecutive ineffective teachers score 52 to 54 percentile points behind students taught by more effective teachers. There are, however, some reasons to think that these results are biased since Sanders and Rivers do not consider the mixing of students into groups and other omitted student characteristics that might be related to teacher effects. Wright, Horn and Sanders (1997) also use the TVAAS to model gains in student test scores with a mixed model as a function of teacher and a set of student and classroom-level covariates. Then, they use standardized contributions of each variable to compare with teacher effects through a meta-analysis. They find strong and persistent teacher effects, and conclude that teacher effects are dominant to determine achievement gains. However, their results could be biased because their standardized measures are not a robust indicator of contribution to total variance in scores (McCaffrey *et al.,* 2003, pp. 20-23).

Rowan *et al.* (2002) also support the importance of teachers using residuals for classroom level variance with data from the *Prospects* study.[2] Authors use data from two cohorts to specify a different set of models to test how much classroom counts for variance in math achievement. They find that the teacher effect accounts for 72 per cent of the reliable variance[3] in growth for math scores, which supports the importance of teachers in the educational process, although severe bias of these results could stem from the nature of the data. Another relevant paper on the topic was published by Rivkin *et al.* (2005) using data from UTD Texas School Project to identify different factors behind student achievement. Rivkin *et al.* develop a strategy that uses multiple cohorts tested in multiple years, and involves multiple stages to remove stable effects of external factors on growth. Their findings are that teachers and schools do matter in increasing student achievement, and that those gains are related systematically to observable teacher and school characteristics, although this effect seems to be small and concentrated among younger students.

Lessons learned from value-added research are that teacher effects are important and cumulative, and vary according to student type. However, there is still not enough evidence to support a precise estimation of teacher and school effects; shortcomings of statistical models are also another source of bias that must be taken into account. Another important critique comes from Raudenbush's work (2004), which questions the initial objective the models are trying to measure. While teachers and schools intend

[2] *Prospects* refers to a large scale survey applied in the early 90's in a large sample of elementary schools in the United States.

[3] Reliable variance is defined as combined variance of random school and individual slopes, with the addition of variance in teacher effects.

to be used as treatments, there is no clear definition of what is intended to be used as a control. Raundenbush's critiques call for a clearer definition of the VAM in terms of a counterfactual model, which is an important challenge for future research. This challenge demands an extension of these models beyond the United States context, in particular to developing countries where many policy experiments remain unexplored.

For Mexico there are several studies attempting to measure the level of variance explained by the locality, teacher, school and individual characteristics. Those studies use different models and methods to identify determinants on student achievement using different tests to measure outcome. The first documented study was carried by Schmelkes (1997) in the southern state of Puebla to measure educational quality in 16 schools in distinct regions with varying characteristics. The research uses a test developed especially for the study, and finds a strong relationship between cognitive achievement and socioeconomic context. Recent studies take advantage of standardized tests applied at a more extended level. For example, Álvarez *et al.* (2007) use PISA 2003 results to identify state-level variations of accountability, the influence of teacher unions, and conflicts between the union and the state. Using a fixed-effects approach, they find that the presence of accountability systems at the school level has a strong and significant outcome on student achievement. At the same time, a medium level of influence has a negative effect, while a low level of conflict has a positive effect on student outcomes. The main problem with those results is that their definitions of institutional frameworks might hide within state variations. In addition, they might not be controlling for errors associated with the nested nature of the education production function.

Other studies use tests administered by national authorities to measure student outcomes. Backhoff *et al.* (2007) use Excale 2005[4] and a Hierarchical Linear Model (HLM) to identify the determinants of variance in scores. They find that differences among students explain around 66 per cent of the variance in Spanish scores, while differences across schools explain around 34 per cent. For mathematics scores, the percentages are 76 (student) and 24 (schools). De la Vega (2010) also uses Excale 2005 results and a HLM model but with a different set of explanatory variables that take into account state differences. Results confirm that student dif-

[4] *Exámenes de la calidad y el logro educativo* is a standardized test administered to a sample of students in 6th grade of primary school and 3rd grade of secondary school by the *Instituto Nacional para la Evaluación de la Educación*, an agency that collects educational data.

ferences have a bigger role in explaining score variances (86%) than school (13%) or state differences (0.32%). Another set of studies work with census data produced by the Ministry of Education with the Enlace test. Valenti *et al.* (2009) use Enlace 2008 and its contextual questionnaires to specify an HLM model for 3rd to 6th grade students in primary school, finding that family capital has the largest influence on student achievement. Somewhat different results are found by De Hoyos *et al.* (2012) which use Enlace 2008 at the high school level. Using a discriminant analysis model, they find that family resources are the factor that has the lowest explanatory power on student achievement, while student characteristics show the largest effect, followed by school resources. Authors attribute the atypical results to the fact that family characteristics might have less influence at the high school level than at earlier stages. Brodziak (2009) uses PISA 2003 and Enlace 2005 results to determine the relationship of individual and contextual socioeconomic characteristics with achievement. She finds that initial conditions are important to determine student achievement and that this relationship varies across states. Additionally, she runs a regression discontinuity to conclude that more years of schooling are related to higher achievement. Regardless of the different quality of models and data used by the studies described above, all of them suffer the major shortcoming of considering only one point in time.

A smaller set of research uses several time periods. Those specifications recognize the selection bias that exists in the school selection process. Among those studies, Santibáñez (2006) studies teacher effects on student achievement using data from *Carrera Magisterial* in Mexico City from 1996 to 2000.[5] The author finds that teacher scores have a positive —though small— effect on student achievement. Luschei (2012) also uses *Carrera Magisterial* data but for the states of Sonora and Aguascalientes. This study finds that teacher scores or training evaluations are positively correlated with student achievement, while education levels and experience are not. A major limitation of those studies is that student scores are measured with tests of doubtful quality and that there is no evidence that peer review instruments are reliable (Santibáñez *et al.,* 2006). A study that uses a different data to implement a value-added model shows how students from academic high schools have larger gains

[5] *Carrera Magisterial* is a program that rewards teachers based on credentials, seniority and test results.

than students from vocational high schools (Rubio and Farías, 2013). Although this study reduces selection bias by using propensity score matching, it only uses data for schools in Mexico City, thus reducing its external validity.

In sum, existing analyses of student achievement in Mexico are limited because they either use data from one point in time without controlling for prior performance, or because they are restricted to a particular location. This paper attempts to fill this gap in the literature using data from several years at the national level. It also looks to refocus attention on achievement changes and to set some precedent for future analysis of the Mexican educational system that could implement more robust value-added models (VAM). The next section offers more arguments and motivations for studying student achievement across time and developing the methodological framework.

II. Data and Empirical Strategy

The question this paper addresses is: What are the school characteristics that affect student performance change in Mexican schools? Given the recent efforts implemented by the federal government to increase educational quality in Mexico such as the *Alianza por la Calidad de la Educación* in 2008[6] and the *Pacto por México* in 2012,[7] it is relevant to understand which school characteristics are associated with change in student performance. From a policy perspective, having better information of which factors have a larger role would contribute to a better allocation of public monies and to improving the planning of educational policies in Mexico. To provide a guidance of these policies, this paper focuses on characteristics related to teaching practices.

The hypothesis of this paper is that school characteristics have an impact on student growth. These characteristics include the quality of the infrastructure, the quality of their materials, their teaching practices, and the internal and external environment. Additionally, a set of vari-

[6] The agreement was signed on May 15, 2008 by the Ministry of Education and the teachers' union, with the objective of "transform[ing] the educational model through public policies that enhance quality and equity in education" (Amador, 2009, p. 1).

[7] This large scope pact among political forces was promoted by the entering administration of Enrique Peña Nieto. The chapter on education sets different reform objectives to enhance educational quality, such as school autonomy, increasing school hours and improving teachers' formation, among others (Presidencia, 2013).

ables measuring dropout cases is implemented as a proxy for socioeconomic level. The main assumption of the VAM is that if family resources are stable across time and are captured by the initial score, then the change in achievement is due to school or teacher characteristics (Braun *et al.,* 2010). The final objective of these models is to control for selection bias and claim a causal link between schools, teachers and student achievement. The source of this bias is that student characteristics, family and school resources are not randomly distributed, nor is the assignment of students to certain teachers. For this reason, there is a high correlation between these elements and the initial score of the student. Then, just measuring the score at one point in time also captures elements that are not related to schools and teachers. The specification of VAM depends largely on the assumptions, the research question and the available data. The present research uses the approach of score gains at the intra-cohort level. Gain score models, as part of the VAM family, are used in the literature to explore teacher effects by Rivkin *et al.* (2005) and Wright *et al.* (1997).

Table 1 reports Pearson correlations between elements that could be a source of bias with initial and final scores of the cohort. Here we observe that if a school is located in a poor municipality, its score will be lower. Primary indigenous schools and vocational schools also have a low correlation between initial and final achievement. On the contrary, correlations with score gains are close to zero. Additionally, figure 1 shows that student achievement is highly related to the marginality level of the municipality, with a difference of about 20 percentile points between schools in richer municipalities and those in poorer municipalities. These results illustrate how using score levels instead of changes in scores could lead to biased conclusions, since the effect of schools and teachers is confounded with family, peer and community effects. However, another possible conclusion from figure 1 is that changes in scores can also be correlated with other variables, since the cohorts in poorer municipalities are presenting a larger reduction in their relative score than cohorts in other municipalities.

Since for the Mexican case it is not possible to link students with individual teachers or track students through time, to overcome this limitation this paper uses intra-cohort gains using Enlace information from students who were in 3rd grade in 2007 and 6th grade in 2010 for primary school, and those who were in 7th grade in 2008 and 9th grade in 2010 for secondary schools. The main assumption of this technique is that it com-

Table 1. Pearson correlation of school average test scores and gains with school characteristics related to selection bias

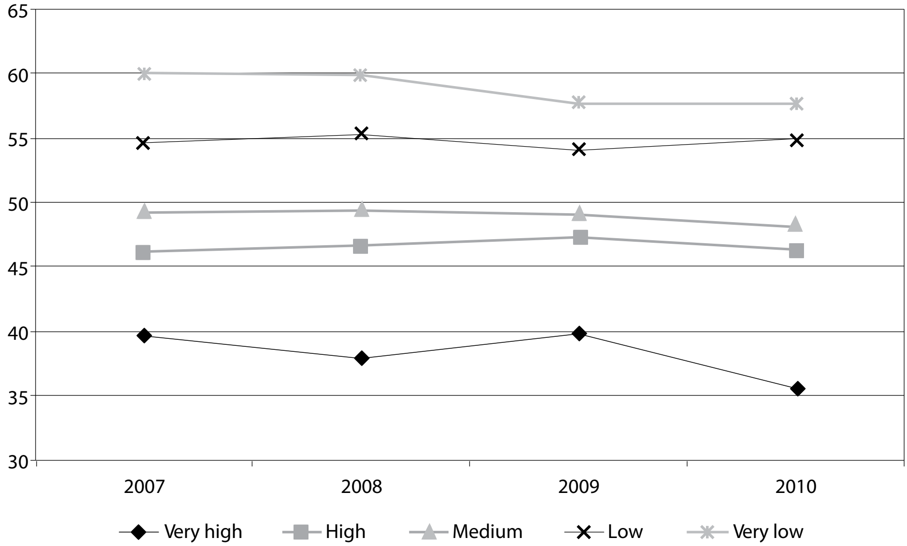
	Primary schools		
	3rd grade scores in 2007	*6th grade scores in 2010*	*Score gains 2007-2010*
Locality marginality level	-0.369 ***	-0.332 ***	0.054 ***
Precinct marginality index	-0.352 ***	-0.319 ***	0.049 ***
Regular primary school	0.188 ***	0.166 ***	-0.032 ***
Indigenous primary school	-0.188 ***	-0.166 ***	0.032 ***
Public primary school	-0.261 ***	-0.186 ***	0.085 ***
Rural primary school	-0.342 ***	-0.307 ***	0.050 ***
	Secondary schools		
	7th grade scores in 2007	*9th grade scores in 2010*	*Score gains 2007-2010*
Locality marginality level	-0.1564 ***	-0.070 ***	0.112 ***
Precinct marginality index	-0.191 ***	-0.114 ***	0.108 ***
Regular secondary school	0.1422 ***	0.054 ***	-0.111 ***
Vocational secondary school	-0.0254 ***	-0.043 ***	-0.013
Telesecundaria	-0.1171 ***	-0.023 ***	0.114 ***
Public secondary school	-0.2568 ***	-0.165 ***	0.134 ***
Rural secondary school	-0.1134 ***	-0.030 ***	0.102 ***

Source: Author's own elaboration with data from SEP, Conapo, IFE and INEGI. *Note:* Correlations are calculated with math NCE scores. ***p<0.001, **p<0.005, *p<0.1.

pares the same group of students after they experienced educational progress in the same school. The obvious shortcomings of this assumption are that some of these students could have changed or dropped out of school.[8] To better understand the role of schools in the variation of student achievement, we could use information from multiple cohorts to identify sources of variation. In this case, I included information of students that were in 3rd grade in 2008 and 5th grade in 2010, and those who were in 3rd

[8] In Mexico, drop-out rates at the primary level for the school year 2008-2009 were 1.1 per cent, while repetition rates were 3.8 per cent, compared with 6.8 and 7.5 per cent for the secondary level (SEP, 2010). For this reason, it is possible to assume that measuring cohorts at the primary level (from 3rd to 6th grade) provides a fair comparison of the same group of students. For the same reasons, results for the secondary level should be taken more cautiously.

Figure 1. Math NCE score progression for the same cohort by level of marginality of the municipality, 2007-2010



Source: Author's own elaboration with data from Enlace grade 3, 2007; Enlace grade 4, 2008; Enlace grade 5, 2009 and Enlace grade 6, 2010, from the Ministry of Education (SEP). Raw scores were transformed for each year to their NCE to be comparable year to year.

grade in 2009 and 4th grade in 2010; for an additional cohort in secondary schools I included those students that were in 7th grade in 2009. Although these cohorts have expended less time in school, they also have had less opportunities to dropout or change schools. Table 2 shows the cohorts studied in this paper. Here, it is important to consider that in Mexico primary education comprises from 1st to 6th grade, when students are between 6 and 12 years old, and secondary education comprises from 7th to 9th grade, when students are between 12 and 15 years old.

The main data source used in this research are the Enlace scores for Math and Spanish. This standardized test has been applied to all Mexican schools at basic level in 2006 and to schools at high school level since 2008. Students take the Enlace in grades 3, 4, 5, and 6 for the primary level, and grades 7, 8 and 9 at the secondary level. The objective of this test is not to provide a vertical comparison of student progression in different grades, but to establish a score within each year. For this reason, in order to compare the scores from two different years and two different grades the

Table 2. Cohorts included in the analysis

2010	*2009*	*2008*	*2007*
	Primary		
6th	5th	4th	3rd
5th	4th	3rd	
4th	3rd		
	Secondary		
9th	8th	7th	
8th	8th		

Source: Author's own elaboration. *Note:* The shaded grades are the ones included in the analysis for each cohort. The years in the first row are those in which the cohort completed the Enlace test.

scores were transformed here using a Normal Curve Equivalent (NCE). This transformation implies:

$$nce_i = zscore\ (Enlace_i) * 21.06 + 50$$

Where:

$$zscore\ (Enlace_i) = (Enlace_i - \overline{Enlace}) / (sd(Enlace))$$

The use of NCE scores is highly extended in education because it allows the comparison of relative positions within the same cohort at different points in time rather than just its absolute position (Harris, 2010). The advantage of a NCE over a z-score is that it can be interpreted similarly to percentile ranks.

The other source of information used in this paper are the contextual questionnaires applied with the Enlace test to all school principals. These questionnaires consist of 134 items about different aspects of the school, from basic infrastructure to teaching practices and internal environment. Some of these characteristics are presented in table 3. Here it is possible to observe the disparities between the two types of primary schools: indigenous and other primary schools. The first group caters to the indigenous population, who are generally bilingual and concentrated in remote rural communities. In 2010 only 46 per cent of these schools reported having running water, whereas 79 per cent of all other schools did. Although secondary schools are generally better equipped, these disparities can also be

Table 3. Descriptive statistics

	Primary (regular)	*Indigenous primary*	*Secondary (regular)*	*Vocational secondary*	*Telese-cundaria*
			Number of students		
Mean Enlace NCE math score 2010	53.0	42.2	53.4	50.2	51.7
Mean Enlace NCE math score 2007	52.4	39.5	54.5	49.1	48.2
Intra-cohort gains in NCE math scores	0.6	2.7	-1.1	1.1	3.5
Mean cohort size 2010	24.0	9.2	73.9	119.5	19.1
Mean school size 2010	106.0	44.0	271.2	370.1	63.8
			Basic infrastructure (%)		
Running water	79	46	97	92	71
Toilets	61	16	92	82	37
Electricity	93	79	99	98	94
AC	16	3	30	30	8
Garbarge collection	61	19	92	80	37
Cleaning services	69	39	94	90	45
Satellite signal	35	19	34	43	77
Internet	42	12	84	76	12
Phone	39	6	90	82	15
School transportation	6	1	18	17	2
Dining hall	14	10	20	18	4
			Educational materials (%)		
Computers for students	16	5	45	23	11
Computers for administrative activities	42	14	73	59	43
Electronic blackboard	40	21	18	13	5
Media equipment	30	12	54	38	28
Laboratory equipment	3	0	49	37	10
Furniture for students	67	59	70	56	63
			Teacher characteristics (%)		
Teachers arrive on time	72	75	37	26	78
Teachers attend classes	90	88	58	45	91

Table 3. Descriptive statistics (Cont.)

	Primary (regular)	*Indigenous primary*	*Secondary (regular)*	*Vocational secondary*	*Telese-cundaria*
Teachers fulfill class schedule	90	90	67	57	90
Teachers interrupt classes for non academic reasons	16	22	23	31	18
Teachers have control of their group	15	13	61	55	87
Teachers present low performance	18	20	34	44	15
Teachers don't have enough knowledge of new technologies	34	36	45	56	32
Teachers refuse professional development	20	19	37	46	16
Teachers disregard curriculum	9	12	18	23	7
		Evaluation (%)			
Evaluation of teachers' attendance	93	91	86	82	82
Evaluation of teachers' punctuality	93	92	84	80	79
Evaluation of content domains	73	76	48	39	35
Evaluation of class planning	91	88	70	63	59
Evaluation of student achievement	87	83	66	61	57
		Dropout causes (%)			
Low student performance is dropout cause	19	27	49	56	28
Lack of parental support is a dropout cause	54	62	65	76	62
Health problems are a dropout cause	28	37	35	38	19
Economic problems are a dropout cause	47	57	66	68	55

Table 3. Descriptive statistics (Cont.)

	Primary (regular)	*Indigenous primary*	*Secondary (regular)*	*Vocational secondary*	*Telese-cundaria*
Migration is a dropout cause	67	54	65	73	59
Adictions are a dropout cause	5	10	16	18	7
Pregnancy is a dropout cause	4	8	18	25	13
Safety in the school surroundings is a dropout cause	6	11	13	18	5
	Internal environment (%)				
Bullying in the school	47	33	52	61	31
Students disrespect teachers	6	5	9	11	13
Students steal inside the school	16	11	38	47	13
Students fight inside the school	22	10	39	51	15
Students bring weapons to the school	1	1	3	3	1
Students use alcoholic beverages	0	1	6	11	3
Students use drugs	0	1	5	7	2
	External environment (%)				
Lack of street lights	56	71	42	58	69
Use of drugs in the school surroundings	25	7	34	40	80
Robbery/assault in the school surroundings	23	9	37	37	86
Gang activity in the school surroundings	23	7	40	45	83
N	73 572	6 968	9 889	3 965	15 834

Source: Author's own elaboration with data from contextual questionnaires for Enlace 2010, from the Ministry of Education (SEP). *Note:* For basic infrastructure, educational materials and dropout indexes the percentages represent the number of schools that answered "Yes". Teachers, internal environment and external indexes environment percentages represent the number of schools that answered "Some times", "Almost always" and "Always". For the evaluation index, the percentage represents the number of schools that answered "Almost always" and "Always". N represents the maximum number of observations for each category.

observed among vocational secondary schools, *Telesecundarias* and academic secondary schools.[9]

A concern is which of these variables should be included as a control for school characteristics, particularly because there is a large correlation among them. Including too many self-correlated variables could be a problem for multicolinearity in the model. The approach used in this paper is the construction of different indexes using factor analysis and an estimation of a new variable based on varimax rotated factors. Variables used for each index are grouped in table 3. These variables include the quality of infrastructure; the educational materials; some teacher characteristics and practices; the causes of student dropout, which can be used as a proxy of the socioeconomic level of the students; and the internal and external environment. Table 4 shows the correlation among these indexes, which is low enough to avoid multicolinearity problems.

Figure 2 shows the k-density of the indexes for schools in municipalities with very low and very high levels of vulnerability, according to the *Consejo Nacional de Población* (Conapo) definition.[10] These charts show how part of the basic infrastructure index is highly dependent on the marginality level of the municipality. However, this relation seems to be weaker for the other indexes. This could have two interpretations: *1)* that some characteristics are more homogeneously distributed across schools, and *2)* that the variation of these characteristics within municipalities is high. Since the main objective of this research is to identify particular characteristics of schools impacting score gains, the two indexes analyzing teaching practices are deconstructed and integrated in the analysis as explanatory variables. The selection of the variables maintained in the regression is based on the factor loadings and its uniqueness obtained in the factor analysis. In this sense, these variables are relevant to each dimension studied without incurring in an over specification of the model.

To identify the relationship between intra-cohort gains and school characteristics, this analysis employs a school production function (Hanushek, 2007). The dependent variable is the score gains of the same group of students from the first year they take the test to 2010. This can

[9] Vocational secondary schools are focused on providing technical skills. *Telesecundaria* is a system of distance education via satellite.

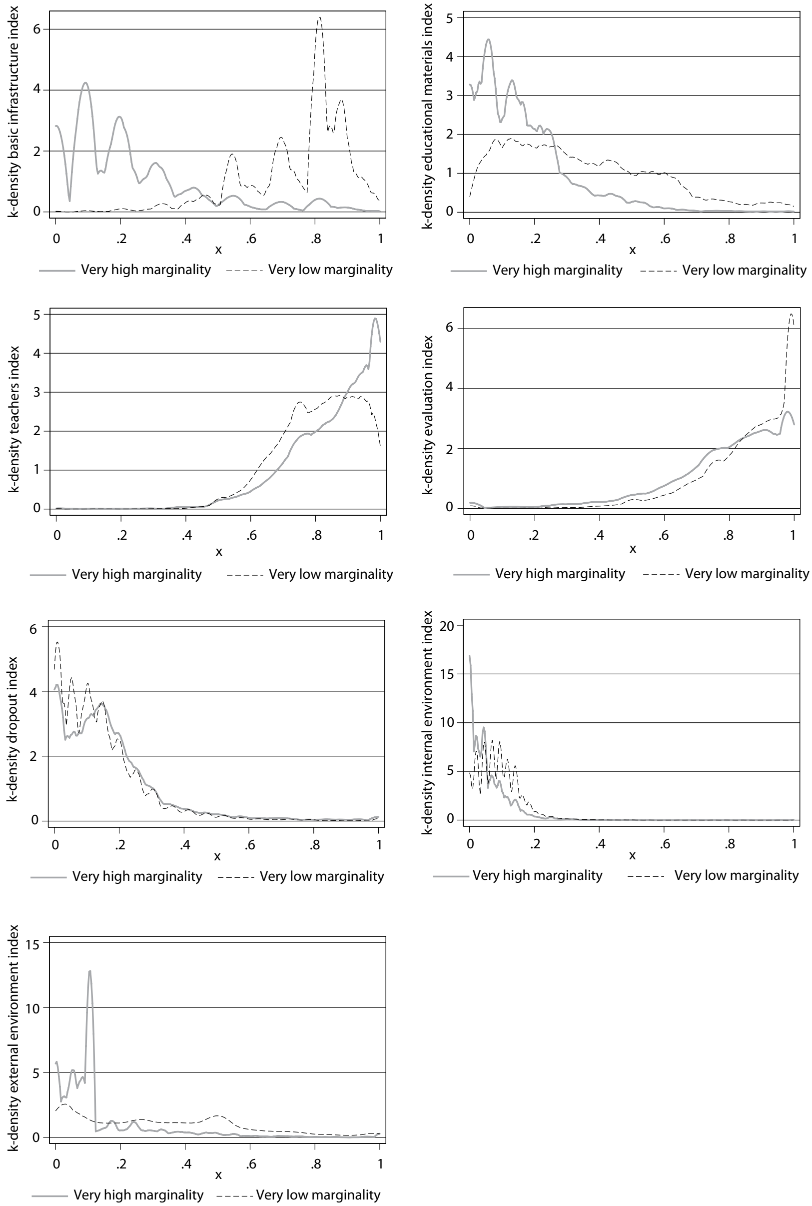
[10] Conapo builds a marginality index based on dimensions of education, household characteristics, locality characteristics and income. The index divides municipalities and localities in five categories: very low marginality, low marginality, medium marginality, high marginality and very high marginality.

Table 4. Correlation Matrix of school characteristics indexes

	Basic infrastructure index	*Educational materials index*	*Teachers index*	*Evaluation index*	*Dropout index*	*Internal environment index*	*External environment index*
			Primary schools				
Basic infrastructure index	1						
Educational materials index	-0.4424 *	1					
Teachers index	-0.1956 *	-0.0154 *	1				
Evaluation index	0.1612 *	0.1624 *	0.1986 *	1			
Dropout index	0.0013	-0.0234 *	-0.1833 *	-0.0388 *	1		
Internal environment index	0.2019 *	-0.0147 *	-0.3558 *	-0.0817 *	0.1827 *	1	
External environment index	0.2974 *	-0.0149 *	-0.240 *	0.0149 *	0.1227 *	0.3713 *	1
			Secondary schools				
Basic infrastructure index	1						
Educational materials index	0.446 *	1					
Teachers index	-0.385 *	-0.023 *	1				
Evaluation index	0.138 *	0.180 *	0.217 *	1			
Dropout index	0.238 *	0.004	-0.364 *	-0.0589 *	1		
Internal environment index	0.344 *	-0.029 *	-0.548 *	-0.115 *	0.4245 *	1	
External environment index	0.256 *	-0.081 *	-0.386 *	-0.0551 *	0.3498 *	0.5425 *	1

Source: Author's own elaboration with data from contextual questionnaires for Enlace 2010, from the Ministry of Education (SEP). *p<0.001.

Figure 2. K-densities of indexes according to vulnerability level of the municipality



Source: Author's own elaboration with data from contextual questionnaires for Enlace 2010, from the Ministry of Education (SEP).

be defined as the difference of average NCE scores of the cohort in year t and grade x, and their average NCE scores in 3rd grade in year t-1:

$$\overline{G} = \overline{(\textit{NCE score in Grade X})}_t - \overline{(\textit{NCE score in } 3^{rd} \textit{ Grade})}_{t-1}$$

This difference is calculated separately for the two subjects of the test. The model implemented here estimates score changes controlling for the initial score of the cohort and school characteristics:

$$\overline{G}_{i,l} = \alpha_{i,l} + \mu \textit{NCE score } 3^{rd} \textit{ grade}_{i,l} + \beta X_{i,l} + \lambda Z_{i,l} + \delta M_i + \gamma_l + \varepsilon_{i,l} \qquad (1)$$

Where for each school i in locality l: X represents the vector of the indexes of school characteristics. Z represents a vector of other school characteristics such as if the school is public, the size of the school, if the schools impart classes in the morning, the type of school, and the change in cohort size. Controlling for change in the cohort's size is relevant because of the possibility that some students change schools or repeat grades. M represents the marginality level of the municipality according to Conapo, and $\gamma_{l,t}$ is a term for fixed effects for the locality where the school is located. This fixed effect term controls for observable and unobservable time-invariant characteristics of each of the more than 48 000 localities in the country. The inclusion of this term allows the reduction of some endogeneity, such as different educational preferences at the locality level. The right-hand side of the equation includes the initial score in 3rd grade to estimate the impact of school characteristics while prior student achievement is maintained constant.[11] Given that some schools have two cohorts during the same school year (one attending during the morning and one during the afternoon), to avoid the artificial inflation of the standard errors they are clustered at the school level.

A second version of the model includes other cohorts in the same school that were in 3rd grade in different years. Including those cohorts not only allows having more observations, it also increases the chances that the students measured within the cohort are the same during the two years. Thus, model 2 introduces a fixed effect for each cohort c, that is

[11] Here, it is important to notice that this model is equivalent to:

$$NCE_{i,l} = \alpha_{i,l} + \textit{NCE score } 3^{rd} \textit{ grade}_{i,l} * (\mu + 1) + \beta X_{i,l} + \lambda Z_{i,l} + \delta M_i + \gamma_{l\,+}\, \varepsilon_{i,l}$$

In other words, the score of the student in the current year is controlled by a weighted effect of the initial score. The assumption here is that what students bring to the classroom in terms of achievement has a strong effect on final achievement.

4^{th}, 5^{th} or 6^{th}, the grade in which students were in 2010. This term is expressed as θ_c

$$\bar{G}_{c,i,l} = \alpha_{c,i,l} + \mu NCE\ 3^{rd}\ grade_{c,i,l} + \beta X_{i,l} + \lambda Z_{i,l} + \delta M_i + \theta_c + \gamma_l + \varepsilon_{c,i,l} \qquad (2)$$

Finally, a third version of the model introduces, individually, the variables of the evaluation index and the teacher characteristic index (τW_i). Then, the rest of the indexes are maintained as a control. This model allows identifying particular elements that could be relevant for policy analysis, since the coefficients for the indexes could be difficult to interpret and to provide enough insight for policy implementations.

$$\bar{G}_{c,i,l} = \alpha_{c,i,l} + \mu NCE\ 3^{rd}\ grade_{c,i,l} + \beta X_{i,l}' + \lambda Z_{i,l} + \delta M_i + \tau W_{i,l} + \theta_c + \gamma_l + \varepsilon_i \qquad (3)$$

- Controlling for neighborhood characteristics using spatial techniques and HLM models

One of the potential critiques of the former models is that there could be some omitted variables that might be correlated with the school characteristics and the student outcomes. As the VAM imply, obvious candidates for omitting variables are the characteristics related to the immediate surroundings of the school. In order to identify those characteristics, this paper uses a project implemented by IFE and INEGI during the national census of 2010. This project creates a national cartography of the census at electoral precinct. The country is divided in more than 66 000 of these units with an average of 2 179 inhabitants. Map 1 shows a map of the country with information on the marginality of each of its precincts. Following Díaz-Cayeros *et al.* (2012) this marginality uses the same variables and techniques employed by the *Consejo Nacional de Evaluación de la Política de Desarrollo Social* (Coneval) to create the marginality index at municipality level (Coneval, 2007). However, using data at this smaller geographical level takes into account the variation within each municipality, thus providing more accurate information about school surroundings.

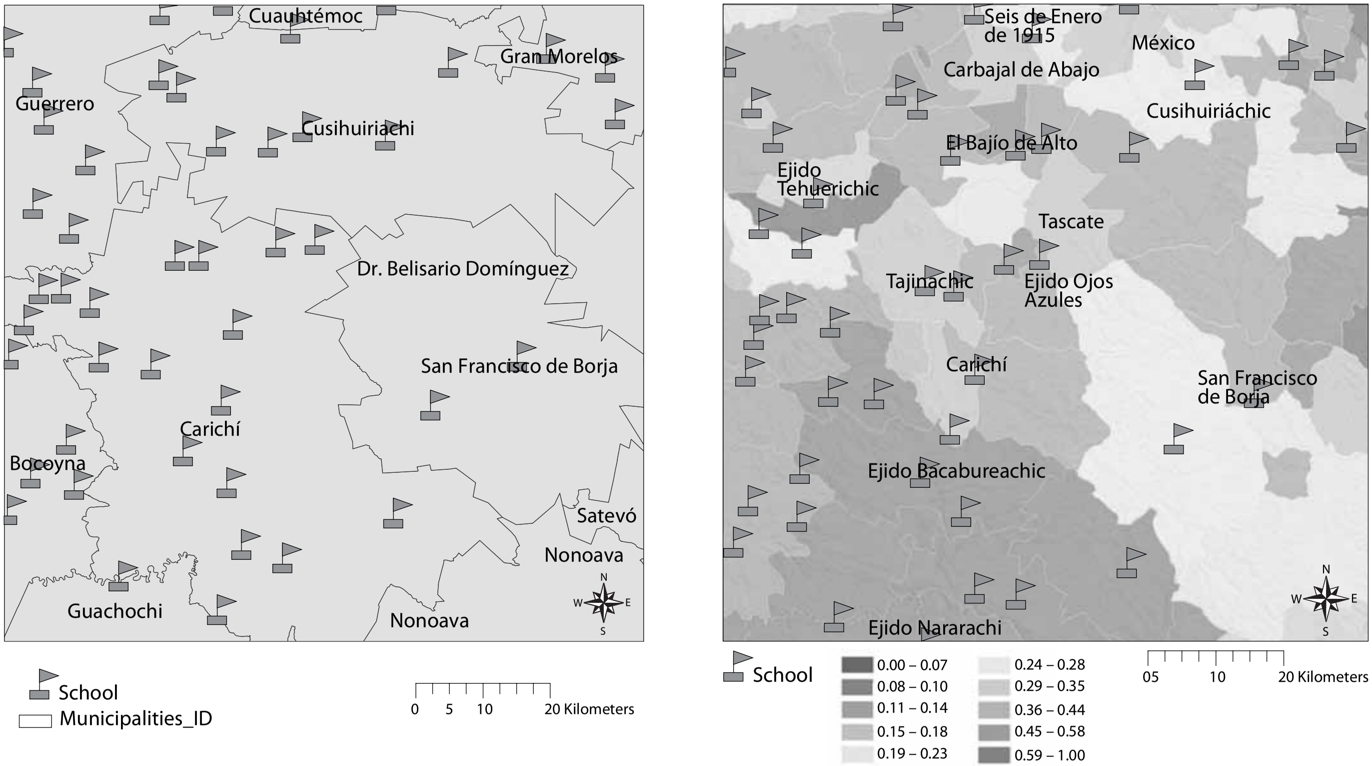
Using the latitude and longitude of each school it is possible to link its location with their corresponding precinct.[12] This provides a way to control for variables in the neighborhood of the school. Figure 3 shows an example of how schools located in the Tarahumara region of Chihuahua are

[12] This spatial joint was performed using ArcGIS.

Map 1. Marginality index at precinct level using 2010 census data

Source: Author's own elaboration with data from *Estadísticas Censales a Escalas Geoelectorales* (IFE-INEGI), based on Díaz-Cayeros *et al.* (2012).

Figure 3. Schools within municipalities and precincts in the Tarahumara region (Chihuahua)



Source: Author's own elaboration with data from *Estadísticas Censales a Escalas Geoelectorales* (IFE-INEGI) and the Ministry of Education (SEP).

linked to their precincts. In this case, it is possible to detect how aggregation at municipal level might mask some important local variation. In the case of the urban areas, this granularity can even provide information at the neighborhood level.

Given that schools are nested in neighborhoods/precincts, it is possible to create a Multilevel or Hierarchical Linear Model (HLM) in order to allow for random intercepts and random slopes for effects at each level. The key element of HLM is that these random deviations are different from those associated with the overall error term (Raudenbush and Bryk, 2002). These models provide the standard deviation of the estimate across different clusters, and are useful to understand the heterogeneous effect of each level specified. The two level model for each cohort c in school i nested in precinct p is as follows:

$$\bar{G}_{c,i,p} = \alpha_p + \mu NCE\ 3^{rd}\ grade_{c,i,p} + \beta X_{i,p} + \lambda Z_{i,p} + \theta_c + \xi_p Q_p + \varepsilon_{i,p} \qquad (4)$$

Where:

$$\alpha_p = \alpha_0 + u_p$$
$$\xi_p = \xi_0 + v_p$$

Here, Q_p is a vector of precinct/neighborhood level characteristics; α_p refers to the intercept in precinct p; α_0 refers to the overall intercept, or the grand mean of the outcome variable across all neighborhoods when all predictors are equal to zero; u_p refers to a random error component for the deviation of the intercept of a precinct to the overall intercept. The term ξ_p refers to the slope in precinct p for the relationship between precinct variables and outcome score; ξ_0 is the overall regression coefficient, and v_p refers to the error component of the slope. The assumption for the covariance is that it is unstructured.

III. Results

Results from the OLS model for primary schools (table 5) show that schools with higher initial scores have lower intra-cohort gains. This simply implies that schools with lower initial achievement have more room for improvement, and that increasing scores that are already high is a difficult task. Most of the indexes included in the analysis are significant and present the expected sign —that is, better basic infrastructure, educational

Table 5. Estimates for score gains in primary schools with locality fixed-effects

	Mathematics				*Spanish*			
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
NCE score in initial grade	-0.721*** (0.0147)	-0.722*** (0.00516)	-0.723*** (0.00515)	-0.722*** (0.00512)	-0.726*** (0.0147)	-0.722*** (0.00539)	-0.723*** (0.00537)	-0.723*** (0.00536)
Basic infrastructure index	4.684*** (0.819)	4.935*** (0.435)	5.053*** (0.431)	4.943*** (0.431)	5.850*** (0.750)	5.507*** (0.401)	5.578*** (0.397)	5.544*** (0.398)
Educational materials index	1.033* (0.497)	1.413*** (0.276)	1.429*** (0.275)	1.368*** (0.274)	1.266** (0.446)	1.563*** (0.250)	1.585*** (0.249)	1.518*** (0.248)
Teachers index	1.592* (0.791)	1.734*** (0.438)		1.616*** (0.437)	0.255 (0.708)	0.968* (0.401)		0.881* (0.400)
Evaluation index	1.433* (0.598)	1.320*** (0.324)	1.324*** (0.322)		1.322* (0.531)	1.152*** (0.292)	1.081*** (0.290)	
Droput index	-2.175** (0.702)	-2.411*** (0.380)	-2.274*** (0.377)	-2.435*** (0.378)	-1.981** (0.622)	-2.336*** (0.357)	-2.202*** (0.354)	-2.348*** (0.355)
Internal environment index	-0.586 (1.545)	-1.354 (0.825)	-1.277 (0.822)	-1.253 (0.821)	-0.507 (1.376)	-1.064 (0.746)	-0.967 (0.742)	-0.998 (0.742)
External enviornment index	0.291 (0.404)	0.0921 (0.225)	0.113 (0.224)	0.156 (0.223)	0.0860 (0.360)	-0.106 (0.204)	-0.0895 (0.203)	-0.0516 (0.203)
Public school	-3.487*** (0.386)	-5.561*** (0.204)	-5.653*** (0.205)	-5.572*** (0.202)	-4.857*** (0.370)	-7.081*** (0.193)	-7.144*** (0.193)	-7.089*** (0.191)
School size	0.0106*** (0.000874)	0.0104*** (0.000482)	0.0106*** (0.000482)	0.0104*** (0.000479)	0.0103*** (0.000818)	0.0103*** (0.000451)	0.0104*** (0.000451)	0.0103*** (0.000448)

Table 5. Estimates for score gains in primary schools with locality fixed-effects (Cont.)

	Mathematics				*Spanish*			
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Morning schedule	-0.281 (0.200)	0.155 (0.109)	0.109 (0.109)	0.154 (0.108)	0.590*** (0.176)	0.628*** (0.0990)	0.588*** (0.0994)	0.634*** (0.0984)
Indigenous school	-2.575** (0.983)	-3.008*** (0.518)	-3.031*** (0.512)	-2.791*** (0.519)	-2.727** (0.924)	-2.967*** (0.468)	-2.969*** (0.463)	-2.711*** (0.469)
Change of cohort's size	0.0341* (0.0166)	0.111*** (0.0118)	0.114*** (0.0118)	0.112*** (0.0117)	0.0453** (0.0162)	0.131*** (0.0120)	0.134*** (0.0120)	0.132*** (0.0119)
Vulnerability index of municipality	-0.457* (0.201)	-0.461*** (0.101)	-0.455*** (0.101)	-0.445*** (0.100)	-0.302 (0.191)	-0.364*** (0.0949)	-0.357*** (0.0944)	-0.348*** (0.0941)
Grade 5^{th} in 2010		-0.499*** (0.0709)	-0.500*** (0.0707)	-0.495*** (0.0705)		-0.496*** (0.0694)	-0.498*** (0.0691)	-0.500*** (0.0690)
Grade 4^{th} in 2010		-0.460*** (0.0710)	-0.466*** (0.0708)	-0.467*** (0.0706)		-0.516*** (0.0708)	-0.517*** (0.0706)	-0.529*** (0.0704)
Teacher variables								
Teachers arrive late to classes			-0.362*** (0.0824)				-0.314 *** (0.0756)	
Teachers don't attend classes			0.00338 (0.0798)				0.0865 (0.0724)	
Teachers don't complete time scheduled to class			-0.000495 (0.0784)				-0.0688 (0.0722)	
Teachers don't control of their group			-0.112 (0.0747)				-0.0580 (0.0677)	

	Mathematics				*Spanish*			
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Teachers refuse profesional development			0.0000839				0.0385	
			(0.0615)				(0.0559)	
Teachers disregard curriculum			-0.103				-0.0740	
			(0.0787)				(0.0697)	
Evaluation variables								
Evaluation of content domain				0.220**				0.188**
				(0.0686)				(0.0634)
Evaluation of class planning				0.0310				0.0576
				(0.0868)				(0.0790)
Evaluation of student achievement				0.274***				0.187**
				(0.0782)				(0.0701)
N	61454	183874	186290	186627	61454	183874	186290	186627

Source: Author's own elaboration with data from Enlace grade 3 2006, Enlace grade 6 2009 and contextual questionnaires for Enlace 2010, from the Ministry of Education (SEP). *Note:* OLS model with fixed effects at locality level; standard errrors clustered at school level. ***p<0.001, **p<0.005, *p<0.1.

materials, teaching practices and evaluation techniques are correlated with larger gains. As the causes of dropout increase, it is also more likely that score gains are lower. Since the desertion index is mostly explained by economic constraints, this can be interpreted as an effect of family income. That is, the lower the average income of the school, the lower the score gains. Indexes of internal and external environment are not significant, which could mean that the effects of elements such as internal conflict in the school, crime and gang activity are already absorbed by the locality fixed effects. Public and indigenous schools have much lower score gains than private or non-indigenous schools. Cohorts that attended classes in the morning present larger score gains in Spanish than cohorts that attend classes in the afternoon or during the night; this result is non-significant for math scores. Separating the individual variables for teacher characteristics and evaluation practices, it is possible to observe that the significant variables that explain score gains are: teachers arriving on time, evaluation of teachers' content domain, and evaluation of student performance. The inclusion of additional cohorts does not change the results as can be observed comparing columns 1 and 5 (which only include the cohort that was in 3rd grade in 2007) with the other columns.[13]

Results for secondary schools show a weaker effect of school infrastructure, school materials or teaching practices (table 6). The only effect that persists is the dropout index, which implies that elements related to economic constraints of the students have a larger effect for this educational level than school characteristics. Once again, public schools have smaller cohort gains than private schools. Vocational schools present slightly smaller gains than academic schools; this parameter also has a low significance. Interestingly, *Telesecundaria* schools have larger score gains

[13] An important concern is the amount of missing values. Whereas there are about 80 000 primary schools, due to list-wise deletion the regressions only report results for around 60 000 of them. As such, 25 per cent of the observations are missing. Results are also reported for around 24 000 of the almost 30 000 secondary schools. This could be problematic if some schools are not answering the questionnaires in a systematic manner, such as poorer schools, having a larger amount of missing values than other schools. To test this concern, first it is important to observe the number of observations for each variable. As observed in table 9, this number does not vary drastically among them. In this case, it is possible to infer that the amount of missing observations in the final regression is due to random combinations of observations with missing variables. To provide further evidence on this concern, the same table shows the t-test of the means differences by two groups of marginality. This analysis shows how, although there are some significant differences between groups regarding missing data, the percentage of missing observations is still much lower than the 25 per cent of the whole sample. This means that, although there is a difference in patterns of missing data, it is likely that the difference between the initial and final samples is driven by random combinations (table 9).

Table 6. Estimates for score gains in secondary schools with locality fixed effects

	Mathematics			
	(1)	*(2)*	*(3)*	*(4)*
NCE score in initial grade	-0.421*** (0.0267)	-0.527*** (0.0130)	-0.529*** (0.0130)	-0.527*** (0.0129)
Basic infrastructure index	1.583 (2.231)	1.630 (1.438)	1.671 (1.412)	1.805 (1.410)
Educational materials index	0.513 (0.786)	0.987* (0.502)	1.030* (0.499)	0.897 (0.494)
Teachers index	-0.0469 (1.755)	1.666 (1.063)		1.570 (1.043)
Evaluation index	1.675 (1.384)	1.295 (0.834)	1.272 (0.829)	
Droput index	-0.876 (1.094)	-2.282** (0.727)	-2.245** (0.720)	-2.256** (0.722)
Internal environment index	-2.518 (2.063)	-2.074 (1.304)	-2.056 (1.307)	-2.114 (1.302)
External enviornment index	0.879 (1.015)	0.469 (0.668)	0.443 (0.662)	0.373 (0.666)
Public school	-2.956** (0.952)	-7.843*** (0.596)	-7.629*** (0.590)	-7.732*** (0.585)
School size	0.00614*** (0.00130)	0.00915*** (0.000840)	0.00921*** (0.000837)	0.00919*** (0.000830)
Morning schedule	-0.418 (0.458)	-0.0700 (0.282)	-0.112 (0.280)	-0.0727 (0.277)
Vocational school	-0.370 (0.545)	-0.706* (0.349)	-0.687* (0.348)	-0.776* (0.346)
Telesecundaria	3.013*** (0.896)	4.347*** (0.543)	3.944*** (0.546)	4.324*** (0.538)
Change of cohort's size	0.0940*** (0.0131)	0.0728*** (0.00847)	0.0724*** (0.00842)	0.0702*** (0.00803)
Vulnerability index of municipality	0.283 (0.322)	0.518* (0.244)	0.547* (0.242)	0.533* (0.243)
Grade 8th in 2010		-1.570*** (0.120)	-1.563*** (0.120)	-1.580*** (0.119)

Table 6. Estimates for score gains in secondary schools with locality fixed effects (Cont.)

	Mathematics			
	(1)	*(2)*	*(3)*	*(4)*
Teacher variables				
Teachers arrive late to classes			-0.630* (0.266)	
Teachers don't attend classes			-0.248 (0.230)	
Teachers don't complete time scheduled to class			0.0916 (0.208)	
Teachers don't control of their group			0.0662 (0.213)	
Teachers refuse profesional development			0.0288 (0.161)	
Teachers disregard curriculum			-0.136 (0.201)	
Evaluation variables				
Evaluation of content domain				0.209 (0.187)
Evaluation of class planning				-0.161 (0.228)
Evaluation of student achievement				-0.0822 (0.209)
N	24381	48441	48878	48992

Source: Author's own elaboration with data from Enlace grade 3 2006; Enlace grade 6 2009; and contextual questionnaires for Enlace 2010 from the Ministry of Education (SEP). *Note:* OLS model with fixed effects at locality level and standard errrors clustered at school level. ***p<0.001, **p<0.005, *p<0.1.

Table 7. Estimates for score gains in primary and secondary schools, HLM model

	Primary			*Secondary*		
	(1)	*(2)*	*(3)*	*(1)*	*(2)*	*(3)*
NCE score in initial grade	-0.646***	-0.648***	-0.646***	-0.466***	-0.467***	-0.466***
	(0.00298)	(0.00297)	(0.00297)	(0.00533)	(0.00532)	(0.00531)
Basic infrastructure index	6.769***	6.849***	6.721***	2.207***	2.326***	2.197***
	(0.230)	(0.229)	(0.228)	(0.470)	(0.468)	(0.467)
Educational materials index	0.309	0.340	0.343	-0.0813	-0.0334	-0.131
	(0.217)	(0.217)	(0.216)	(0.319)	(0.318)	(0.316)
Teachers index	0.698*		0.790**	1.622**		1.812**
	(0.301)		(0.299)	(0.600)		(0.594)
Evaluation index	3.540***	3.513***		1.001*	0.911*	
	(0.215)	(0.214)		(0.438)	(0.436)	
Droput index	-4.972***	-4.773***	-4.915***	-1.995***	-1.915***	-1.976***
	(0.249)	(0.248)	(0.248)	(0.458)	(0.456)	(0.455)
Internal environment index	1.806***	2.163***	1.819**	-4.394***	-4.044***	-4.245***
	(0.582)	(0.580)	(0.578)	(0.836)	(0.832)	(0.828)
External enviornment index	0.677***	0.743***	0.698***	0.975*	0.965*	0.866*
	(0.179)	(0.179)	(0.178)	(0.384)	(0.382)	(0.380)
Public school	-4.808***	-5.036***	-4.827***	-5.646***	-5.601***	-5.709***
	(0.165)	(0.165)	(0.164)	(0.331)	(0.331)	(0.329)
School size	0.0147***	0.0152***	0.0147***	0.00745***	0.00754***	0.00750***
	(0.000423)	(0.000423)	(0.000422)	(0.000525)	(0.000523)	(0.000520)
Morning schedule	-0.474***	-0.517***	-0.463***	0.111	0.0684	0.134
	(0.105)	(0.105)	(0.104)	(0.215)	(0.214)	(0.214)
Indigenous school	-2.287***	-2.297***	-2.214***			
	(0.168)	(0.167)	(0.167)			
Vocational school				-0.459*	-0.410	-0.479*
				(0.229)	(0.228)	(0.226)

Table 7. Estimates for score gains in primary and secondary schools, HLM model (Cont.)

	Primary			*Secondary*		
	(1)	*(2)*	*(3)*	*(1)*	*(2)*	*(3)*
Telesecundaria				3.579***	3.308***	3.638***
				(0.265)	(0.268)	(0.264)
Change of cohort's size	0.164***	0.167***	0.165***	0.0722***	0.0720***	0.0716***
	(0.00568)	(0.00565)	(0.00566)	(0.00231)	(0.00231)	(0.00231)
Marginality index of precinct	-5.806***	-5.516***	-5.904***	-1.914	-1.718	-1.897
	(0.644)	(0.642)	(0.639)	(1.079)	(1.074)	(1.066)
Rural precinct	0.397**	0.369**	0.445***	2.385***	2.365***	2.307***
	(0.126)	(0.126)	(0.125)	(0.253)	(0.252)	(0.251)
Grade 5th in 2010	-0.543***	-0.550***	-0.546***			
	(0.0734)	(0.0732)	(0.0731)			
Grade 4th in 2010	-0.481***	-0.495***	-0.491***			
	(0.0740)	(0.0738)	(0.0737)			
Grade 8th in 2010				-1.619***	-1.618***	-1.628***
				(0.106)	(0.106)	(0.106)
Teacher variables						
Teachers arrive late to classes		-0.0993			-0.368**	
		(0.0529)			(0.116)	
Teachers don't attend classes		-0.418***			-0.329**	
		(0.0589)			(0.122)	
Teachers don't complete time scheduled to class		0.112			0.0619	
		(0.0575)			(0.114)	
Teachers don't control of their group		-0.133*			-0.229*	
		(0.0539)			(0.111)	
Teachers refuse profesional development		-0.0531			0.156	
		(0.0452)			(0.0918)	

	Primary			Secondary		
	(1)	*(2)*	*(3)*	*(1)*	*(2)*	*(3)*
Teachers disregard curriculum		0.106 (0.0593)			-0.0808 (0.115)	
Evaluation variables						
Evaluation of content domain			0.638*** (0.0421)			0.180* (0.0894)
Evaluation of class planning			0.141* (0.0552)			0.0887 (0.108)
Evaluation of student achievement			0.471*** (0.0514)			0.189 (0.104)
Random effect parameters						
Sd(Marginality index of precinct)	18.97*** (0.527)	18.85*** (0.525)	18.76*** (0.523)	20.39*** (2.573)	19.99*** (2.589)	17.88*** (2.796)
N	186244	188706	189067	48747	49187	49307

Source: Author's own elaboration with data from Enlace grade 3 2006; Enlace grade 6 2009; and contextual questionnaires for Enlace 2010 from the Ministry of Education (SEP). *Note:* Random coefficient and rando intercpet for precinct poverty level. Covariance is assumed to be unstructured. Only results for Math scores are reported. ***p<0.001, **p<0.005, *p<0.1.

than regular schools. This could be explained because students in this distance learning system could be more self-motivated, since this educational system requires that they develop most of the school work on their own. The only variable that stays significant once the indexes are disaggregated is teachers' punctuality to their classes.[14]

Table 7 shows the results for the HLM model using scores for mathematics in primary and secondary schools. Overall, these models corroborate the results from the OLS model. However, the educational materials index becomes non-significant, whereas the internal and external environment indexes become significant. The interpretation of these coefficients is complicated because they have the opposite expected sign, except for the internal environment index for secondary schools. Once the teachers and evaluation indexes are disaggregated, there is now a negative and significant effect of teachers not attending their classes and not having behavior management of their group. The disaggregation of the evaluation index shows the positive effect of different types of evaluation on score improvements. The coefficient for marginality level of the precinct shows how this variable has a strong effect for primary schools, but not for secondary schools. Moreover, since this parameter was calculated with a random effect, it is possible to observe its large standard deviation. This means that for some schools the direct effect of neighborhood marginality can reduce student increases by 6 NCE points, whereas in other places this effect could be as large as 25 points.

IV. Discussion

The findings of the present research are the first approach to understanding the elements that explain student achievement growth in Mexico using nation-wide data. The main question of this research focuses on school characteristics. The results show how elements such as school infrastructure and educational materials still play an important role in the changes of student scores, even after controlling for initial achievement. Structural characteristics of the school, such as whether it is public or private, also play a role in student performance. The impact of the marginality level of the community is differentiated for primary and secondary schools. Both the OLS and the HLM models show how primary school students are nega-

[14] For simplicity only Math scores for secondary schools are reported. Results for Spanish scores do not vary significantly.

tively affected by the condition of their surroundings; although there is no evidence of such an impact for secondary students. That means that initial conditions of the surroundings might play a larger role during the first years of education. The variable capturing the effect of family socioeconomic conditions is, however, significant for primary and secondary levels, with a larger role in earlier years. Moreover, the HLM model also shows a large variation of the effect of neighborhood poverty level on primary school students. This result is relevant because it implies that schools might play an important role in overcoming adverse initial conditions. That is, a student from a poor neighborhood could have a less negative effect if she attends an adequate school. However, the opposite could also be true: a student in a neighborhood with low marginality could have a negative effect due to a relatively bad school.

The results of the individual variables studied here show that teachers' punctuality and attendance have a positive effect on score gains. This result is also consistent with previous research that shows that Mexican students benefit from more school time (Aguero and Beleche, 2013). Although the majority of schools reported that teachers arrived on time or attended their classes, in 2010 about 27 per cent of primary schools reported teachers arriving late to classes was a problem and around 10 per cent of schools reported teachers not attending their classes was a problem. For secondary schools these figures are higher, with about 40 per cent of schools reporting teachers arriving late and 25 per cent reporting an attendance problem. These figures open the discussion for further analysis on the behavior of teachers in Mexico. It is important to disentangle whether teachers lack the motivation to attend their classes and arrive on time, or if the problem is related to transportation or other factors, such as teachers working more than one shift in separate locations. The different causes of this basic failure of teachers to provide a complete schedule could have very different policy implications. If it is a problem of motivation, then incentive programs to arrive on time or attend classes could be implemented. If it is a transportation problem, then the alternative could be a better scheme to locate teachers in schools closer to where they live.

Analyzing teacher punctuality and attendance across levels of marginality, urban/rural locations and school levels could be helpful to have a better understanding of the associated factors to these problems. Table 8 shows the proportion of schools that reported punctuality and/or attendance as a problem according to different marginality definitions and their location.[15] In general, these problems have a larger presence in secondary schools as compared with primary schools. Urban schools also have larger rates of teacher absence and unpunctuality as compared to rural schools. As expected, these problems are more persistent in primary schools in poorer areas; however, the opposite happens for secondary schools. Although these results should be considered as preliminary, they might be pointing that remoteness of the school is not necessarily the cause of teacher absence and unpunctuality; there might be other factors related to urban mobility, teacher motivation or work overload. In particular, for secondary schools an incentive scheme could be considered, given that motivation might explain that teachers in urban schools with low marginality have higher rates of absence and unpunctuality. Evidence about monetary incentives to teachers shows that although these instruments might be problematic when the goal is to increase student achievement (Ladd and Walsh, 2002; Springer *et al.*, 2010; Goodman and Turner, 2010; Fryer, 2011), they show to be more useful when used to promote simpler and more measureable goals, such as teacher attendance (Duflo *et al.,* 2012). In general, a deeper understanding of these problems is crucial to propose and design appropriate policies. Large scale surveys to teachers and qualitative analysis about their dynamics to arrive to the school should certainly be part of the educational reform agenda.

Evaluation related variables also have a positive impact on score gains. Schools that report a frequent evaluation of content domain or class planning have a positive impact on student achievement. These results are consistent with the work of Luschei (2012) on *Carrera Magisteral*, in which he finds that training evaluations have a positive correlation with student achievement in Aguascalientes and Sonora. The present research has the same limitation as Luschei's in regard that it is not possible to test if content domain *per se* or class planning have a positive effect, although it is possible to argue that evaluation of these fields captures something related to teachers which has an impact on student outcomes. Another variable with a positive relationship to score gains is evaluation of students. In this sense, providing a satisfactory scheme for evaluation of students constantly and communicating such results to the parents in a simple and efficient manner might be a good method to increase score gains.

[15] High marginality refers to schools within localities with a very high, high or medium marginality level as defined by Conapo. Low marginality refers to localities with a low or very low marginality level.

Table 8. Teacher punctuality and attendance across marginality levels and locations, percentage of schools reporting those problems

	Teachers arriving late is a problem				*Teachers not attending the school is a problem*		
				Primary schools			
	Rural (%)	Urban (%)	T-test		Rural (%)	Urban (%)	T-test
Higher marginality	23	37	***	Higher marginality	9	13	***
Lower marginality	22	35	***	Lower marginality	8	12	***
T-test	***	***		T-test	*	***	
				Secondary schools			
	Rural (%)	Urban (%)			Rural (%)	Urban (%)	T-test
Higher marginality	24	58	***	Higher marginality	11	38	***
Lower marginality	26	62	***	Lower marginality	13	43	***
T-test	***	***		T-test	***	***	

Source: Author's own elaboration with data from Enlace, from the Ministry of Education (SEP). *Note:* For this table it is considered that a school has this problem if in its contextual questionnaire it reports that the problem exists "some times", "almost always" or "always". These figurers only consider the questionnaires for 2010. Marginality categories are defined as follows: Higher marginality correspond to schools located in localities with a very high, high or medium level of marginality as defined by Conapo; Lower marginality corresponds to the schools in localities with low or very low level of marginality as defined by Conapo. The T-tets correspond to the difference in means across rows and columns with the following significance level: ***p<0.001, **p<0.005, *p<0.1.

Table 9. Missing data for variables in the analysis

	Observations	*% of missing data*	*% of missing data for municipalities with very low marginality*	*% of missing data for municipalities with very high marginality*	*Difference*
Math score 6^{th} grade	81 504	0.0	0.0	0.0	-
Math score 3^{rd} grade	81 504	0.0	0.0	0.0	-
Spanish score 6^{th} grade	81 504	0.0	0.0	0.0	-
Spanish score 3^{rd} grade	81 504	0.0	0.0	0.0	-
Math score gains	81 504	0.0	0.0	0.0	-
Spanish score gains	81 504	0.0	0.0	0.0	-
Basic infrastructure index	75 665	7.2	8.1	7.0	1.1 ***
Educational materials index	77 182	5.3	7.1	4.5	2.5 ***
Teachers index	77 154	5.3	9.0	3.5	5.5 ***
Evaluation index	76 793	5.8	8.7	3.9	4.8 ***
Dropout index	77 955	4.4	6.0	3.6	2.4 ***
Internal environment index	78 963	3.1	4.7	2.4	2.3 ***
External environment index	77 113	5.4	7.2	4.5	2.7 ***
Public school	81 504	0.0	0.0	0.0	-
School size	81 504	0.0	0.0	0.0	-
Morning schedule	81 504	0.0	0.0	0.0	-
Indigenous primary	80 540	1.2	2.4	0.8	1.6 ***
Change in cohort size	78 962	3.1	6.1	0.6	5.5 ***

Source: Author's own elaboration with data from contextual questionnaires for Enlace 2010, from the Ministry of Education (SEP). *Note:* Observatuions restricted for the 2007-2010 cohort in primary schols. ***$p<0.001$ in the T-test comparison of means.

Authorities could provide training to teachers and school administrators, as well as a mechanism to generate such evaluations.

The infrastructure problem of Mexican schools described here also points out the necessity to invest resources into bringing better facilities and services, since many schools might be operating in their production frontier with their available resources (Ruggiero, 1996). Here, evidence from countries where parents get involved in budget implementation could be useful. For example, a program that provides information for parents about non-wage expenditures in Uganda (Reinikka and Svensson, 2005) shows a positive impact on enrollment and learning, and a reduction in corruption. Similar steps can be implemented in Mexico taking advantage of a recently implemented program to institute school councils in primary schools.[16]

In this research there are some important caveats that should be taken into account. First, results here do not claim causality; although strategies implemented to reduce selection bias under a limited information context might provide some clues as to which school characteristics impact intra-cohort gains, selection bias might not be completely reduced. In particular, some cream-skimming processes could be assigning better students to better schools or to better teachers within the school. An important limitation is that only school-level characteristics are included in this analysis, and that these characteristics come from contextual questionnaires responded by school principals, so answers could be subject to opinion. For example, the definition of "adequate" for any of the items could vary from principal to principal and from region to region. Additionally, a strong assumption from the definition of intra-cohort gains is that dropout and repetition rates for Mexican primary schools are homogeneously distributed across schools. Those variables could differ in, for example, places with high migration or places recently affected by drug-related violence.

An important shortcoming of the present research is related to the limitations of VAM. Those limitations depend on the variation of school performance across grades and time, the amount of resources that do not depend on the school and statistical errors (Koretz, 2002; Harris, 2011). VAM are also highly sensitive to measurement and sampling errors. While sampling errors are not an issue as long as we work with a census data set

[16] The rules of the *Consejos escolares de participación social* were published on August 2010.

such as Enlace, measurement errors should be considered. Those errors depend on the quality of the test and external influences during its application. Assessing the quality of Enlace and the reliability of its application is then a relevant element to take into account. As with any other statistical model, VAM are also sensitive to omitted variables; in this case, the lack of student and teacher level data limits the prediction capacities of the model. In the same vein, Raudenbush (2004) notices that these models intend to identify causality but they fail to provide a correct definition of a treatment. The lack of a counterfactual strategy reduces its scope to descriptive and exploratory, rather than confirmatory.

Finally, it is necessary to consider the limitations of the indexes included here as control variables. The decision to include such indexes in this research is based on the necessity to control for school characteristics without incurring an over-specification of the model. These indexes provide seven different dimensions of school characteristics. However, their effect size does not have a clear interpretation because it is not possible to determine whether the distance between each two units is different than the distance between other two units. Further exploration in the construction of indexes for these variables is necessary to provide a better interpretation of the impact of each dimension.

Besides the limitations of this study, the results presented here are useful to guide educational policy in Mexico and to provide some lessons for other developing countries. First of all, they show that while the socioeconomic level of the school's location strongly determines student achievement, generalization is not always pertinent and could be hiding some characteristics and practices that schools might be implementing to increase student achievement. Among those practices the completion of class schedule and constant evaluations of teachers and students play an important role in the Mexican case.

A more general implication of the present research is that in order to fully differentiate the gains produced by the school or the teachers from those related to student characteristics, Mexico might do well in investing more in data collection at the student-level that could track students over time and link teachers with classrooms and schools. Here, the work of Kane and Staiger (2002) on the use of time series studies could be an important input for such data collection processes. More accurate public data for student and school socioeconomic level, teacher practices and characteristics, is key to better understand how to increase student achievement.

V. Conclusions

In this paper I handled the limitation to follow student outcomes through time to measure school impact on student achievement, by using intra-cohort gains in order to relate long-term student gains with school characteristics. In addition, I use information from census data to provide school location data at a very precise geographical level. This study is one of the first to analyze changes in student achievement in Mexico with national data, and might lead to future research on the topic. Its aim is also to open the field for the application of VAM to developing countries. Here, it will be particularly interesting to differentiate those elements that fall under school scope of action, those that correspond to local or federal authorities, and those that correspond to individual teachers in different contexts. As more information becomes available, more sophisticated studies could be implemented to understand which elements allow students with low initial scores to increase their achievement in Mexico. However, this analysis could be used as a first guidance to implement policies in Mexico and other developing countries.

The main results of this study point out that the relation between the socioeconomic level of the surroundings and student achievement measured through Enlace scores is strong for primary school students, but less clear for secondary school students. This relationship is also highly variable, highlighting the importance of the effect of schools on overcoming less advantaged initial conditions. School basic infrastructure has a positive relationship to score gains for primary and secondary schools, highlighting the relevance of investing in improving school facilities. Students' family constraints have a negative effect on student achievement. In addition, there are important findings related to administrative variables that are within the scope of educational authorities; schools with more frequent evaluation of teachers' knowledge and student achievement have larger gains. Schools that report that teachers complete the class schedule also have stronger gains. In this regard, understanding teachers' dynamic to arrive to the school is crucial to design adequate policies, since there are elements that point out that school remoteness is not necessarily related to teacher absence and unpunctuality; thus, policies oriented to urban mobility and incentivizing teachers to arrive to school on time might be a better approach.

This study points out several elements for future research. It also opens some questions for policy makers interested in applying VAM as accountability measures in Mexico. In this sense, it is important to take into account the lessons from the United States, where VAM have been subject to criticism in recent years. Although these models are an important tool, they are also limited and should be cautiously used. General recommendations include: adjusting the timing of the test with timing of school activities, use measures for at least two years, create measures on comparisons among teachers and schools that promote cooperation, adjust for factors outside the school control and report confidence intervals (Harris, 2011). Implemented correctly, VAM could provide better information for parents to monitor and discipline public providers and to empower citizens, particularly the poor, thereby increasing their influence in policy making and providing them with a stronger voice and "client power" (The World Bank, 2004). This research takes the first steps in understanding and analyzing changes in student achievement, as a way to improve the provision of public education in Mexico.

References

Aguero, J. M. and T. Beleche (2013), "Test-Mex: Estimating the Effects of School Year Length on Student Performance in Mexico", *Journal of Development Economics,* 103, July, pp. 353-361.

Álvarez, J., V. García and H. Patrinos (2007), "Institutional Effects as Determinants of Learning Outcomes", *Policy Research Working Papers,* The world Bank.

Amador Hernández, J. C. (2009), *La Alianza por la Calidad de la Educación: Modernización de los centros escolares y profesionalización de los maestros,* Mexico, Cámara de Diputados, Centro de Estudios Sociales y de Opinión Pública.

Backhoff, E., A.Bouzas, C. Contreras, E. Hernández and M. García (2007), *Factores escolares y aprendizaje en México: El caso de la educación básica,* Mexico, INEE.

Binder, M. (1999), "Schooling Indicators During Mexico's 'Lost Decade'", *Economics of Education Review,* 18 (2), pp. 183-199.

Braun, H., N. Chudowsky and J. Koenig (2010), *Getting Value out of Value-Added,* Committee on Value-Added Methodology for Instructional, Improvement Program Evaluation and Accountability.

Brodziak, I. (2009), *School Matters: Perspectives on Differences in Student Achievement in Mexico,* Stanford University, unpublished PhD dissertation.

Coleman, J., *et al.* (1966), *Equity of Educational Oportunity,* Washington, D. C., Government Printing Office.
Coneval (2007), *Los mapas de pobreza en México: Anexo técnico metodológico,* Mexico.
De Hoyos, R., J. Espino and V. García (2012), "Determinantes del logro escolar en México: Primeros resultados utilizando la prueba Enlace media superior", *El Trimestre Económico,* 4 (316).
De la Vega, M. (2010), *¿De qué depende la calidad de la educación en México? Un estudio jerárquico multinivel,* Mexico, ITAM, Unpublished B.A. thesis.
Díaz-Cayeros, A., E. Franco, J. Olarte and B. Magaloni (2012), *La geografía electoral del 2012,* Mexico Evalúa.
Duflo, E., R. Hanna and S. Ryan (2012), "Monitoring Works: Getting Teachers to Come to School", *American Economic Review*, 102 (4), pp. 1241-1278.
Fryer, R. (2011), "Financial Incentives and Student Achievement: Evidence from Randomized Trials", *The Quarterly Journal of Economics*, 126 (4), pp. 1755-1798.
Goodman, S. and L. Turner (2010), "Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program", Program on Education Policy and Governance Working Papers Series.
Hanushek, E. (2007), "Education Production Functions: Developed Country Evidence", *International Encyclopedia of Education,* 3rd Ed.
Harris, D. (2011), *Value-Added Measures in Education: What Every Educator Needs to Know,* Cambridge, Harvard Education Press.
IFE-INEGI (2010), *Estadísticas censales a escalas geoelectorales*, Mexico.
Kane, T. J. and D. O. Staiger (2002), "The Promise and Pitfalls of Using Imprecise School Accountability Measures", *Journal of Economic Perspectives*, 16 (4), pp. 91-114.
Koretz, D. M. (2002), "Limitations in the Use of Achievement Tests as Measures of Educators' Productivity", *Journal of Human Resources*, 37 (4), pp. 752-777.
Ladd, H. and R. Walsh (2002), "Implementing Value-Added Measures of School Effectiveness: Getting the Incentives Right", *Economics of Education Review*, 21 (1), pp. 1-17.
Luschei, T. (2012), "In Search of Good Teachers: Patterns of Teacher Quality in Two Mexican States", *Comparative Education Review,* 56 (1), pp. 69-97.
McCaffrey, D., J. Lockwood, D. Koretz and L. Hamilton (2003), *Evaluating Value-Added Models for Teacher Accountability,* Santa Monica, Rand.

McKenzie, D. J. (2003), "How Do Households Cope with Aggregate Shocks? Evidence from the Mexican Peso Crisis", *World Developement,* 31 (7), pp. 1179-99.

OECD (2010), *Strong Peformers and Succesful Reformers in Education, Lessons from PISA for Mexico,* Paris, OECD.

Ornelas, C. (2000), "The Politics of Educational Decentralization in Mexico", *Journal of Educational Administration*, 38 (5), pp. 426-441.

Presidencia de la República (2013), *Presidencia de la República,* retrieved from http://www.presidencia.gob.mx/wp-content/uploads/2012/12/Pacto-Por-M%C3%A9xico-TODOS-los-acuerdos.pdf.

Raudenbush, S. (2004), "What are Value-Added Models Estimating and What does this Imply for Statistical Practice?", *Journal of Educational and Behavioral Statistics*, 29 (1), pp. 121-129.

Raudenbush, S.W. and A.S. Bryk (2002), *Hierarchical Linear Models,* Thousand Oaks, Sage Publications.

Reinikka, R. and J. Svensson (2005), "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda", *Journal of the European Economic Association,* 3 (2-3), pp. 259-267.

Rivkin, S., E. Hanushek and J. Kain (2005), "Teachers, Schools and Academic Achievement", *Econometrica*, 73 (2), pp. 417-58.

Rowan, B., R. Correnti and R. Miller (2002), "What Large-Scale Survey Research Tells us about Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools", *Teachers College Record,* 104, pp. 1525-1567.

Rubio, D. and M. Farías (2013), "Efectos escolares en las escuelas de nivel medio superior de la Ciudad de México: Un estudio de valor agregado", *El Trimestre Económico,* 318 (2), pp. 371-399.

Ruggiero, J. (1996), "Effiency of Educational Production: An Analysis of New York School Districts", *The Review of Economic and Statistics*, pp. 499-509.

Sanders, W. and J. Rivers (1996), *Cumulative and Residual Eeffects of Teachers on Future Student Academic Achievement,* Knoxville, University of Tennessee Value-Added Research Center.

Santibáñez, L. (2006), "Why Should we Care if Teachers Get A's: Teacher Test Score and Student Achievement in Mexico", *Economics of Education Review*, 25 (5), pp. 510-520.

Santibáñez, L., J. Martínez, A. Datar, P. McEwan, C. Messan-Setodji and R. Basurto-Dávila (2006), *Haciendo camino: Análisis del sistema de evaluación y del impacto del programa de estímulos docentes Carrera Magisterial,* Santa Monica, Rand Corporation.

Schmelkes, S. (1997), *La calidad en la educación primaria: Un estudio de caso,* Mexico, FCE.

SEP (2010), *Secretaría de Educación Pública*, retrieved June 4th, 2010, from *Principales Cifras del Ciclo Escolar 2008-2009*, http://www.dgpp.sep.gob.mx/PCIFRAS2010/index.html.

Springer, M., D. Ballou, L. Hamilton, L. Vi-Nhuan, J. Lockwood, D. McCaffrey, M. Pepper and B. Stecher (2010), Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching, Nashville, The National Center on Performance Incentives.

The World Bank (2004), "Making Services Work for Poor People", *World Development Report,* Washington, D. C.

Tiebout, C. (1956), "A Pure Theory of Local Expenditures", *The Journal of Political Economy*, 64 (5), pp. 416-424.

Valenti, G., R. Salazar, N. Florez and M. Luna (2009), *Factores asociados al logro educativo: Un enfoque centrado en el estudiante,* Mexico, Flacso.

Woessmann, L. (2003), Schoolling Resources, Educational Institutions, and Student Performance: The international Evidence", *Oxford Bulletin of Economics and Statistics*, 65 (2), pp. 117-170.

Wright, S., S. Horn and W. Sanders (1997), "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation", *Journal of Personnel Evaluation in Education,* 11, pp. 57-67.