Should Students Be Allowed to Miss?

Paredes, Ricardo and Ugarte, Gabriel

April 2009

Should Students Be Allowed to Miss?

Ricardo D. Paredes

Gabriel Ugarte[1]

Abstract

In this paper we investigate the effect of class attendance on academic performance, and evaluate the existence and importance of minimum attendance requirement thresholds. We found that attendance has a relevant and statistically significant impact on performance, together with the existence of a threshold, although contrary to the expected, not associated with a decrease in performance, which questions the existence of minimum attendance requirement.

JEL Classification: I21, I28

Key Words: Attendance, Academic Performance, Hierarchical Models, Thresholds

[1]Ugarte, Escuela de Ingeniería, Universidad Católica de Chile, gaugarte@ing.puc.cl Paredes, Escuela de Ingeniería, Universidad Católica de Chile, y CEPPE, rparedes@ing.puc.cl
We are thankful for the comments and suggestions of David Fuller, Luis Beto Hernández, Paulo Volante, Ignacio Irarrázabal and especially Romulo Chumacero. We are also thankful for the support of MINEDUC, with their SIMCE data base, to SINEDUC, with their attendance data base, and for funding from Fondecyt, Project 107072.

1. Introduction

School absence is one of the biggest problems in education. In England, daily school absence rates in public schools are around 6.7%, which is equivalent to almost a half of million students absent per day (National Audit Office, 2005). In the USA, in New York and Los Angeles, the daily absence rates reach 15% and 10% respectively, and almost 30% in other cities (DeKalb, 1999). In Honduras the World Bank (1995) identified low attendance rates as one of the two most influential factors in the high school-dropout rates (Bedi & Marshall, 2002; Roby, 2004).

There is consensus that attendance has a positive effect on the quality of education and that it can be affected by school policies (Lamdin, 1996). Therefore studies regarding absenteeism are of great relevance for educational policies. There are several cases where parents and schools are given incentives and sometimes even required to reach certain attendance levels. In the USA, South Carolina and Pennsylvania have programs that compensate and recognize schools with low absentee rates (Ladd, 1996; Hoachlander, 2001). In Ohio, schools have minimum attendance rates (Roby, 2004). In England, the *Ofsted* (Office for Standards in Education) constantly inspects the absentee levels in schools and registers absences. Unjustified absences can even lead to parents going to jail (Reid, 2006; Dfes, 1996).

In Chile, the payment of a state subsidy to schools that receive a voucher, that includes almost 90% of the students, is based on individual attendance, and in order for a student to move on to the next grade they have to attend a minimum of 85% of the classes established in the Annual School Calendar (Mineduc, 1988)[2].

Policies oriented toward increasing school attendance assume that there is a clear and stable relationship between learning (or academic performance) and class attendance. However, the variety of different types of policies regarding how to encourage attendance, and the lack of research regarding the effects of these policies,

[2] For a descriptive analysis oriented toward subsidy policies, see Paredes, Ugarte and Volante (2009).

suggest that the policies are weakly supported. For example, the requirement in Chile that students must attend at least 85% of classes (i.e. if a student misses more than 27 days they must repeat the year), differs greatly from the limit of Indiana, USA of 11 absences, even though there are not any studies that validate this number. Along the same lines, in Chile attendance is a key variable to which the voucher is tied. The budget for this reaches US$3,633 million (US$1 = CL$600), which represented 61% of the budget of the Ministry of Education in 2008 (see www.dipres.cl). Many school owners oppose this system, arguing that a subsidy based on enrollment would be easier to implement and would provide more stable support. The Teachers Union claims that a subsidy based on attendance punishes the poor, promotes fraud and leads to deceit (Colegio de Profesores de Chile, 2008). They propose formulas that are easier to regulate, such as payment for enrollment or direct school subsidies. Some defenders of the actual system claim that by receiving payment for attendance, more attention is paid to this area and parents are able to choose the most adequate school (Libertad y Desarrollo, 2008). However, this discussion is also lacking empirical evaluations regarding the real effect of attendance on school performance.

This paper studies the empirical effect that attendance has on academic performance and the nature of this effect. In particular, we are interested in knowing if this effect has thresholds that justify the payment of vouchers based on attendance and the minimum requirements established levels for passing. This paper is divided into three sections in addition to this introduction. The second part describes the existing literature and in particular the relationship between attendance and academic performance and provides information regarding education in Chile. The third section presents the methodology and results and the fourth section concludes.

2. Information regarding attendance

2.1. Literature review

The literature distinguishes justified absences, those due to health reasons, death of a relative, weather or rural location; and those that are not (truancy). At a student level truancy may have to do with boredom, lack of interest, cultural factors, lack of supervision, among other things (e.g., Roderick, 1997; DeKalb, 1999; Epstein & Sheldon, 2002; Roby, 2004; Kube & Ratigan, 1992; Steward, 2002; McCarthy, 2002; Gump, 2006; and Roby, 2004). School level factors include suspension, school infrastructure and the school climate and environment (e.g., Arcia, 2006; Branham, 2004; Finn & Voelkl, 1993; Crone et al., 1993).

Epstein and Sheldon (2002) study the relationship, at a school level, between attendance and a series of variables previously identified in the literature in 12 primary schools in the USA. They find that attendance is negatively related to schools with a large number of students that receive free lunch or lunch at a reduced price, and with students who don't have a home, and that there is a positive relationship with the percentage of students that live at least a mile or more away from school.

Corville-Smith et al. (1998) find that students that are frequently absent from school have low self esteem, are less competent in their social relations, perceive less cohesion in their families, less parental acceptance and inconsistent discipline, and indicate less satisfaction regarding school characteristics and personnel.

Regarding the effect of attendance on performance, Daugherty (2008), citing Ding & Sherman (2006), indicates that if students are not attending classes, they don't have the interaction necessary for learning, and therefore the effect on their academic performance. However, despite the importance of this subject in public policy, there are few studies that analyze this effect profoundly, arguing that not only is there a lack of available data, but that it is difficult to find a significant relationship among factors due to the low variation of attendance, especially when working with cross-sectional studies and with aggregated data (Lamdin, 1996).

Strickland (1999), for a reduced sample of students at a public secondary school in Chicago, concludes that a moderate to strong positive correlation exists between attendance and average grades. Roby (2004) compares, at primary and secondary schools in the State of Ohio, the relationship between academic performance, as measured by the percentage of students that pass all of the proficiency tests in Ohio, and average annual attendance, and finds that a moderate to strong correlation exists among both factors. Daugherty (2008) realizes one of the most complete studies in the State of Delaware, over a period of three years (for students from 8th to 10th grade). The study, which controlled for variables such as gender, ethnicity and socioeconomic factors, concludes that higher rates of absenteeism translate into poor academic performance, as measured by standardized math and language tests. After 15 (16) absences the average score on the math (language) test is below the limit required by the State.

In the search for a qualitative effect, Daugherty (2008) shows that after 15 absences the average score on the math test is below the level required by the state. In turn, the existence of a non-linear effect could mean a bad specification of the relationship between attendance and performance.

Nevertheless, the poor quality of the data and in particular, the reduced size of the samples used, has led many authors to suggest that the impact of attendance on performance is greater than we think (Lamdin, 1996, 1998; Johnston, 2000). Regarding other econometric problems, Lamdin (1996) suggests that the effect of attendance can be confused with others correlated to it and that are frequently omitted, such as innate motivation of the student, concern by part of the parents or the ability of the teacher to stimulate and motivate students. These factors would over estimate the effect of attendance.

Arcia (2006) warns about the risk of bias due to the endogeneity of the attendance variable. She studies the academic performance of suspended students, controlling for socioeconomic status, gender and ethnicity, and concludes that suspensions increase the academic breach between students and schools. In addition, she finds that suspension is used mainly for students with poor performance. Those students that are absent more frequently are precisely those that should not be allowed to miss school (Murray, 2002). Daugherty (2008) suggests that the lack of effort to academically support those students that fall behind for their level, promotes absenteeism as the student passes on to the next grade, due to a loss of hope and a lack of desire to struggle throughout the school day.

2.2. Evidence for Chile

The analysis of attendance in Chile is scarce, due to a lack of available information. In this paper we use information from the databases of SINEDUC, an informational support program for schools that has complete information regarding daily attendance at a student level for almost 10% of municipal schools in Chile, from the 5^{th} to the 8^{th} regions, of which 43% are located in the Metropolitan Region.

Additionally, we use information from the standardized SIMCE[3] tests of 2005, which allowed us to know variables such as academic performance, social, economic and cultural variables at a student level, together with information regarding classroom and school levels. The website of the Ministry of Education in Chile is www.mineduc.cl and it contains descriptive data of the schools in Chile.

Some descriptive statistics of interest are presented in Table 1. It is worth mentioning that on average students miss about 9 days of school during the year (between April to November with the exception of July), although there is a large variation in this number among students. Eleven percent of students have repeated at least one year from 1^{st} to 4^{th} grade and 70% of the students have been attending the school where they took the SIMCE test since 1st grade or before. The average SIMCE score in our sample is lower than the national one, primarily because the average for municipal schools in Chile is substantially lower.

[3] In 2005 4th grade and 10th grade students were required to take the standardized SIMCE tests. Depending on the year sometimes 8th graders are required to take the test instead of 10th graders. In 1998 the average SIMCE score was fixed in 250 points at a national level with a standard deviation of 50 points.

Table 1: Average and standard deviation of the selected characteristics (year 2005).

Variable	Average	Std. Deviation
STUDENT		
SIMCE Math Score	236.3	53.8
SIMCE Language Store	243.8	52.4
Absences (days in a year)[a]	8.9	9.1
Male	0.5	-
Has Repeated a Grade Previously	0.1	-
Has Been in the School since First Grade	0.7	-
Father's Education Level (years)	10.3	-
Mother's Education Level (years)	10.2	-
Family Income (/$10.000)	18.7	17.1
Number of People in the Home	5.2	1.9
Completed Kindergarten	1	-
SCHOOL		
Urban	1	-
Number of Students in 4th Grade	838.4	410.2
Economic Vulnerability Index	31.6	13.6
Average Family Income of Students (/$10.000)	18.7	8.3
Average Educational Level of Mothers (years)	10.2	1.5
Number of Students	17,262	
Number of Schools	287	

Notes:

a. The number of absences of a student from April to November, with the exception of the month of July.

Table 2 shows daily attendance rates and suggests a pattern of absences depending on the day of the week. The percentage of students absent on Mondays (6.56%) and Fridays (6.67%) is significantly greater than the other days. Also, the same table shows that the variations by month are also relevant, being higher in the two coldest months of the year (June and July) and in the last month of the school year (December).

Table 2: The percentage of students absent during different months and days of the week in 2005.

Month	Percentage	Day of the Week	Percentage
March	3.87	Monday	6.56
April	4.47	Tuesday	5.92
May	6.62	Wednesday	6.0
June	9.25	Thursday	5.71
July	8.03	Friday	6.67
August	6.64		
September	5.71		
October	5.85		
November	5.41		
December	7.14		
Total	6.15		6.15

Note: The percentage that corresponds to Monday and Friday is significantly different from the other days (F<0,001).

Thirdly, and suggestive in relation to the goals of public policy, Table 3 shows a large difference in attendance by municipality. This suggests that municipal policies, usually oriented toward increasing income by enrolling more students and encouraging attendance, differ enormously. In effect, the general average of days absent in a year is 8.9 days, but there are municipalities that average 13.3 days like in Quilpue (5^{th} Region), or as low as 4.8 days in Macul (Metropolitan Region). Other municipalities with levels less than the average are Talca, Talcahuano, Conchalí, Melipilla and Puente Alto; while those with levels greater than the average are San Felipe, La Granja, Las Condes, San Miguel, San Ramón and Santiago. It is also worth noting that in the south students miss less days than in the districts of the Metropolitan Region, and have better daily attendance on average than the municipalities in the north (5^{th} Region).

Table 3: Average number of days absent per student by school owner.

Municipality	No. Schools	Days Absent	Std. Deviation	Minimum	Maximum
V REGION					
Los Andes	370	9.2	8.3	0	50
Quilpue	696	13.3	8.5	0	59
San Felipe	586	10.9	9.1	0	72
Total V Region	*1,652*	*11.5*	*8.8*	*0*	*72*
VI REGION					
Talca	1,320	5.2	6.5	0	69
VIII REGION					
Talcahuano	1,400	7.8	7.8	0	67
XIII REGION					
Conchalí	1,011	6.9	7.9	0	53
La Florida	1,629	9.3	9.8	0	100
La Granja	929	11.8	10.1	0	76
La Pintana	731	9.5	9.6	0	59
Las Condes	236	11.2	8.5	0	54
Macul	482	4.8	6.9	0	48
Maipú	1,931	10.3	10.3	0	93
Melipilla	636	6	8.1	0	75
Peñaflor	540	9.2	8.9	0	66
Puente Alto	2,135	7.2	8	0	62
San Joaquín	347	8.5	8.6	0	56
San Miguel	293	11	9.6	0	45
San Ramón	609	11	10.5	0	68
Santiago	1,132	11.5	9.3	0	60
Vitacura	194	10.4	7.5	0	41
Total XIII Region	*12,837*	*9.1*	*9.4*	*0*	*100*
TOTAL	17,262	8.9	9.1	0	100

Note: Only schools with more than 100 students are included.

3. Model and Results

3.1. Model

The model used follows the literature in this area, by measuring performance based on student, school and environment variables (see Mizala and Romaguera, 2000; Gallego, 2002; Sapelli and Vial, 2002). In our case, we explain academic performance in math as measured by the standardized SIMCE test, using a group of variables that are considered as exogenous in the literature (student, family, and school variables) along with a variable that indicates a student's class attendance. Specifically, we estimate a regression using a multilevel model that takes into consideration the common influences that students of the same school share, since the observations are not independent (see, Steenbergen & Bradford, 2002)[4]. In this model, academic achievement is represented by Y_{ij}, and depends on a group of factors and attendance, which is given in equation (1) that specifies the level 1 (student) of the mixed effects model:

$$Y_{ij} = \beta_{0j} + \beta_1 A_{ij} + \beta_2 A_{ij}^2 + \beta_3 S_{ij} + \varepsilon_{ij} \tag{1}$$

For $i=1,..,n_j$ students of a school j, with $j=1,..,287$. The predictor of interest, A_{ij}, represents the number of days that a student was absent during the year (between the months of April to November, with the exception of July.) The vector S_{ij} contains a series of student characteristics (sex, educational level of mother, family income and the number of people who live in the home). The error ε_{ij} is assumed to be independent and identically distributed $N(0,\sigma_e^2)$.

The level 2 (school) is represented by equation (2):

$$\beta_{0j} = \gamma_{00} + \gamma_{01} C_j + \eta_{0j} \tag{2}$$

For $j=1,..,287$. The vector C_j contains school characteristics (rural location, School Vulnerability Index, average income of the students' parents, and the average educational level of students' mothers). The error η_{0j}, that follows a distribution $N(0,\sigma_0^2)$, represents the portion of intercept that is not explained by the predictors at the school level and it is supposed to be independent of the predictors at the student level.

3.2. Effects of first order

The results of the model of academic performance as measured by the SIMCE score on the math test are shown in Table 4. By applying the Hausman test robust to

[4] The unconditional random effect model indicates that a 12.5% of the data variability is found at a school

heterocedasticity, we reject the null hypothesis that the difference in coefficients is not systematic, which indicates that the random effects model would not produce consistent estimates and the model to use is one of fixed effects ($p<0.05$). In spite of this result, in Table 4 we show the estimated coefficients for both models, although in following estimates we always use fixed effects (which do not include variables at a school level).

Apart from confirming that the signs and significance of the parameters are consistent with the literature, the number of annual absences appears to be highly significant and negative, in addition to presenting a convex effect. Therefore, a student who misses 9 days during the year (the average of the absence variable), maintaining all other variables constant, reduces performance by at least 18% of the standard deviation of the SIMCE math test. This is reflected by the standardized coefficients, and shows that the variables of greater impact are, in order, absences, mother's education and family income.

level. This suggests that an analysis using OLS would lead to mistaken results.

Table 4: Estimators of the selected parameters and robust standard error in parenthesis.

Predictor	Coefficient Random Ef.	Coefficient Fixed Effectsa	Std. Coefficient Fixed Effects
STUDENT			
Annual Absences	-1.101***	-1.15***	-0.186***
	(0.113)	(.115)	
(Annual Absences)^2	0.011***	.012***	0.076***
	(0.003)	(.003)	
Male	5.384***	5.526***	0.052***
	(0.885)	(.894)	
Mother's Education	2.454***	2.458***	0.149***
	(0.154)	(.155)	
Family Income (/10,000)	0.407***	.395***	0.126***
	(0.058)	(.058)	
(Family Income)^2 (/10,000)	-0.002***	-.002***	-0.076***
	(0.000)	(0.000)	
No. of People in the Home	-1.166***	-1.165***	-0.039***
	(0.249)	(.251)	
SCHOOL			
SVI	-0.083		
	(0.109)		
Urban	-8.816		
	(4.670)		
Average Income (/10,000)	0.684*		
	(0.294)		
Average Education of Mothers	2.381		
	(1.278)		
Teacher Experience	0.178**		
	(0.052)		
School Enrollment	0.006*		
	(0.003)		
Constant	177.064***	218.881***	
	(15.547)	2.328)	
Number of Students	12,725	12,725	
Number of Schools	285	285	
Between Variance (intercept)	124.2	333	
Within Variance	2,351.7	2,353.6	
Conditional intraclass correlation	0.05	0.124	
AJUSTE			
Pseudo- R^2_{ε} (Between R^2)	0.656	0.345	
Pseudo- $R^2{}_{e}$ (Within R^2)	0.053	0.053	
Pseudo-R^2	0.13	0.103	
Deviance	135,210		
AIC	135,279		
BIC	135,540		
Method – Degrees of freedom	ML – 35		

Notes:

*** p<0,001; **p<0,01; *p<0,05

a. Coefficients of the fixed effects model with robust errors. The Hausman test between the fixed and ramdom effects model robust to heterocedasticity indicates that we should use the fixed effects model (p<0.05)

3.3. *Endogeneity*

One problem that we eventually have in studies regarding attendance is endogeneity. The estimators β_1 and β_2 become bias if the absences are correlated with non-observed characteristics, such as effort, motivation and concern of parents; and the ability of teachers to stimulate and motivate students (Lamdin, 1996). If parents determine the attendance pattern of their children according to the expected gains in human capital and the costs of going to school (Bedi y Marshall, 2001), then a problem of endogeneity exists because of reverse causality[5].

To deal with this problem, we test the exogeneity of the attendance variable by using a test of weak exogeneity. This tests that the academic performance on the SIMCE is determined by class attendance but not vice versa. For this we use the following basic equation:

$$Y_{ij} = \beta_0 + \beta_1 A_{ij} + \beta_2 A_{ij}^2 + \beta_3 S_{ij} + \beta_4 E_j + \varepsilon_{1ij} \tag{3}$$

Where it is suspected that student attendance (and its square) may depend on performance. For this we use the following two equations:

$$A_{ij} = \gamma_0 + \gamma_1 S_{ij} + \gamma_2 E_j + \varepsilon_{2ij} \tag{4}$$

$$A^2_{ij} = \gamma_3 + \gamma_4 S_{ij} + \gamma_5 E_j + \varepsilon_{3ij} \tag{5}$$

With the goal of proving weak exogeneity of class attendance and its square, we must verify that the errors of equations (4) y (5) are not correlated. For this we use the following regressions:

[5] We should also consider as another factor of endogeneity the unobserved student characteristics such as motivation, but they aren't relevant for this study since we only work with 4th grade students.

$$\varepsilon_{1ij} = \delta_0 + \delta_1 * \varepsilon_{2ij} + \varepsilon_4 \tag{6}$$

$$\varepsilon_{1ij} = \delta_2 + \delta_3 * \varepsilon_{3ij} + \varepsilon_5 \tag{7}$$

If δ_1 and δ_3 are not significant, then we can say that the errors are not correlated and thus student class attendance (and its square) does not depend on performance.

Table 5 shows that effectively, the estimators δ_1 and δ_3 are not significant at 5%. This result indicates that we cannot reject the hypothesis that class attendance and its square are weak exogenous variables in relation to academic performance as measured by the SIMCE, which allows us to conclude that the model of fixed effects generates efficient estimators[6].

Table 5: Test of weak exogeneity for the attendance variable and its square

Predictor	Coefficient
Model $\varepsilon_{1ij} = \delta_0 + \delta_1 * \varepsilon_{2ij} + \varepsilon_4$	
Attendance Residual	8.55e-10
	(0.055)
Constant	-3.61e-09
	(0.424)
Model $\varepsilon_{1ij} = \delta_2 + \delta_3 * \varepsilon_{3ij} + \varepsilon_5$	
Attendance Residual^2	4.42e-11
	(0.001)
Constant	-3.61e-09
	(0.423)

*Notes: ***p<0.001;** p<0.01; *p<0.05*

[6] We also did tests of exogeneity using as an instrument the number of days that it rained more than 10mm during the year in the area where the student lives, which does varies at a student level. The exclusion instrument is correlated significantly with attendance and was tested using the Fisher test (Staiger and Stock, 2007), and the Kleibergen and Paap (2006) test rejected the null hypothesis that the range of the matrix wasn't complete, complying with the requirements of the matrix of instruments in the first stage. Due to the two estimates being exactly identified it was not possible to do the over-identification test of Sargan and Hansen to verify that the instruments were not correlated with the error term. After testing exogeniety, we rejected the hypothesis that class attendance is endogenous, suggesting that the estimates reported in Table 4 are consistent.

3.4. Thresholds

One question that naturally emerges and that is associated with the policy of establishing a maximum limit of absences is related to the existence of breakpoints in determined levels of absences. We are interested in knowing if there are threholds at different levels of absences which allow us, on the one hand, to analyze the consistency of the threshold determined by the Ministry of Education; and on the other, an adequate specification of the model, since it would not be lineal.

In order to determine the existence of cutoff points, we use the Hansen (2000) threshold regression method, applied to the equations (8) and (9). This allows us to identify multiple thresholds by obtaining different regression parameters depending on the number of days absent per student. The two regimes are defined as follows:

$$Y_{ij} = \beta_0 + \beta_3 S_{ij} + \beta_4 C_j + \varepsilon_{ij} \qquad \text{if } A_{ij} \leq \gamma \tag{8}$$

$$Y'_{ij} = \beta_0' + \beta_3' Sij + \beta_4' Cj + \varepsilon_{ij}' \qquad \text{if } A_{ij} > \gamma \tag{9}$$

Where A_{ij} is class attendance, γ is the critical value of attendance that divides the sample into two different groups and is not known previously, S_{ij} is the vector with the characteristics of the students, C_j is the vector with the characteristics of the schools and ε_{ij} is the regression error.

Since we don't know the threshold, we also don't know the distribution of the errors, not being able to identify the breakpoint (Hansen, 1996) and as a consequence, we were also not able to make inferences. However, based on the asymptotic distribution theory it is possible to build confidences intervals using Monte Carlo simulations (Hong et al., 2005).

The Hansen method (2000), used commonly in the analysis of cross sectional data, can be extended to panel data. This requires subtracting every variable with the average at a school level, eliminating fixed effects, which is always valid if endogeneity doesn't exist, which has already been ruled out. Therefore, by subtracting each variable from the school average we create a binary variable $d(\gamma)=\{A_{ij} \leq \gamma\}$ and defining $s_{ij}(\gamma)= s_{ij}*d(\gamma)$, the equations (8) and (9) can be written as follows:

$$y_{ij} = \beta_0' + \beta_3' s_{ij} + \delta_n*(1 + s_{ij}(\gamma)) + e_{ij} \tag{10}$$

Where the small letters express that we are using the subtraction of each variable with the school average (demeaned variables), and the sub index n applies to all possible levels of absence. The regression parameters are β0', β3', δn y γ; whose estimators $\hat{\beta}_0(\gamma), \hat{\beta}_3(\gamma), \hat{\delta}(\gamma)$ are conditional in a value γ, are obtained by the ordinary minimum squared method, minimizing the sum of the squared residuals. Later, $\hat{\gamma}$ is the value that minimizes $S_n(\gamma) = S_n(\hat{\beta}_0(\gamma), \hat{\beta}_3(\gamma), \hat{\delta}(\gamma), \gamma)$, which is the sum of the concentrated squared residuals (Ahmed & Iqbal, 2007). Following Hansen (2000), we use the Likelihood Ratio test to test the null hypothesis $\gamma = \gamma_0$. The confidence intervals robust to heterocedasticity and asymptotically correct for the LR test are obtained through bootstrap replications.

In order to test the existence of thresholds we evaluate if the estimated coefficients for one group of students (those who miss less or equal than γ days) are equal to the estimated coefficients of the other group (those who miss more than γ days). The hypothesis is that starting from γ days there is a threshold reflected in the data.

Since attendance is not endogenous, it is possible to apply the Hansen method (2000) to test the existence of thresholds.

Applying the Hansen method (2000) we find a threshold around 13 days[7]. Table 6, which shows the estimates of fixed effects for each one of the regimes found, allows us to observe that a student with average observable characteristics in the group who

[7] P-value<0.01

misses no more than 13 days in a year, obtains 11 points more than students from the other group.

We also observe that the estimator of absences is significant in both regimes, although greater in the first, which would indicate that the effect of absences is greater in the group who misses no more than 13 days in a year.

Table 6: Estimators of the selected parameters and robust standard errors in parenthesis, for each of the two regimes.

Predictor			**Standardized Coefficients**	
	Absences<=13	**Absences>13**	**Absences<=13**	**Absences>13**
Annual Absences	-1.651	-1.318	-0.13	-0.222
	(0.197)	(0.343)		
(Annual Absences)^2	0.026	0.012	0.05	0.134
	(0.008)	(0.005)		
Male	6.646	0.848	0.065	0.008
	(0.008)	(1.998)		
Mother's Education	2.521	2.206	0.148	0.132
	(0.172)	(0.342)		
Family Income (/10,000)	0.427	0.234	0.128	0.068
	(0.063)	(0.135)		
(Family Income)^2 (/10,000)	-0.002	0	-0.087	-0.017
	(0)	(0.001)		
No. of People in the Home	-1.095	-1.412	-0.038	-0.054
	(0.281)	(0.527)		
Constant	218.503	232.714		
	(2.624)	(6.898)		
Number of Observations	10,316	2,409		
Average Math Scores	241	230		

In Figure 1 we show the decline that students with average observable characteristics (for each of the two groups) experience on the SIMCE math scores as the number of absences increases. To analyze the consistency of this result with the threshold determined by the Ministry of Education, it is necessary to compare the threshold found with the minimum learning levels identified by the SIMCE[8], whose limit between intermediate and initial achievement levels is also found in Figure 1.

[8] The possible levels are: Advanced, Intermediate or Initial, where the category of each student depends on his or her score, and there is a minimum score that the student must obtain to be classified as intermediate or advanced. A group of experts define the questions that a student must answer as a minimum to receive an intermediate level therefore obtaining a cutoff point for this level. The same is true for the advanced level. See, www.simce.cl

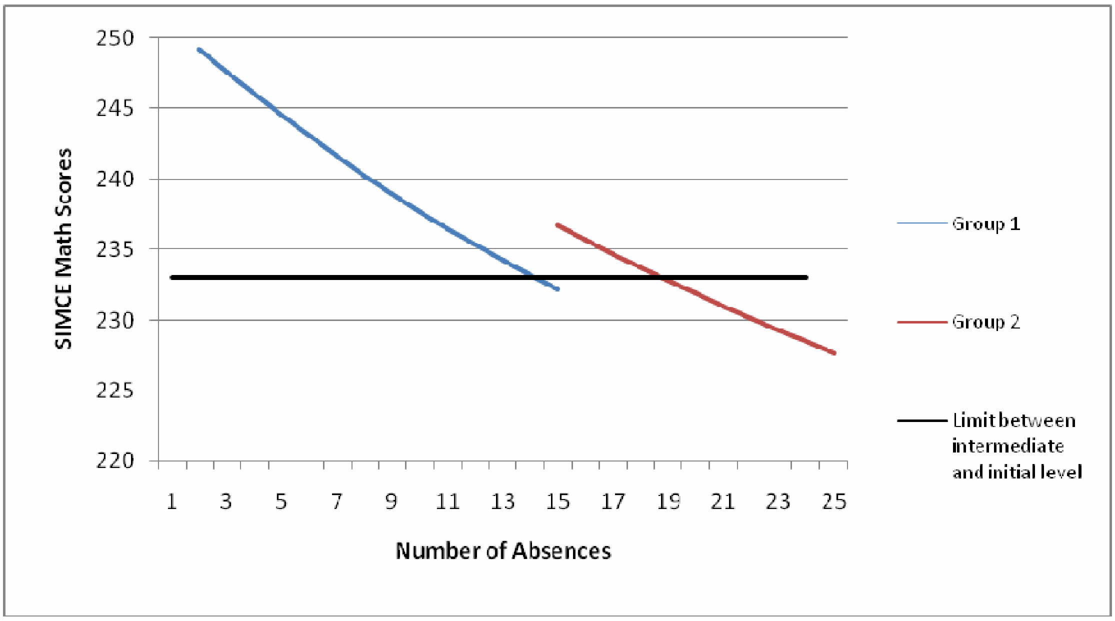

Figure 1: Effect of absences on performance.

The breakpoint that we see has an interpretation opposite of what we expected. Indeed, the population of students who miss more than 13 days (Group 2) have non-observable characteristics that are especially positive on performance. In fact, Figure 1 suggests that, on average, students that miss between 14 and 17 days, have higher scores than students who miss up to 12 days. This is quite surprising, and should not be interpreted as the existence of a positive effect of absences over a given number of days. This results seems to be the consequence of the cross section analysis, and in particular, it could be showing that that among students who miss a lot of days, those that are highly capable are over represented. One possible cause for their absences would be a lack of motivation due to the quality of education that they receive.

Of course the previous result suggests that the existence of cutoff points or thresholds as a requirement for students to pass a grade don't make sense, at least from the perspective of the existence of a relevant economic breakpoint. The decrease in performance is gradual on the population average and therefore, the existence of a breakpoint isn't logical.

Without a doubt as the number of absences increases average performance decreases, and after a certain point the level of learning achievements, on average, fall to their initial level. In fact, after 17 absences, a student with average observable characteristics from Group 2 passes from an intermediate level to an initial level, while this occurs after 12 absences for the students of Group 1. However the lack of breakpoints that reflect a drastic fall, indicate that those thresholds don't differentiate adequately between the populations. Moreover, as our results suggest, the differences in learning related to the thresholds detected, of the students who miss more and less days, is contrary to the imposition of a minimum assistance requirement.

4. Conclusions

The fact that there is consensus regarding the acknowledgement of attendance as a factor that has an impact on performance, has led to the implementation of a group of public policies that require certain minimums. Such policies, however, require concrete evidence regarding effective impacts, which don't exist because of difficulties related to how to estimate them and the availability of information.

In this paper we provide evidence that allows us to evaluate the impact of attendance on performance, address the possible problem of endogeneity, and determine the existence of certain or thresholds that allow, on the one hand, a better specification of the model, and on the other, the determination of critical points that merit policies regarding minimum attendance.

We found that indeed attendance has significant and economically important effects on educational performance. Specifically, that being absent 9 days during the school year (the sample average of absences) reduces performance by at least 23% of the standard deviation of the score on the math test.

Regarding the existence of thresholds, we found a statistically significant breakpoint at 13 absences, but contrary to expected, this is not a discontinuity that implies that above this threshold performance decreases. The absence of a breakpoint in the sense we expected questions the existence of minimum attendance requirement since this limit is not associated with a decrease in performance. Whilst we can argue that over a limit average knowledge may fall, the huge variance in knowledge for any given attendance, neither suggest minimum attendance policies.

The existence of a maximum absence requirement in Chile of close to 28 days does not have any relationship to the breakpoint found or to the level of students' learning achievements. In fact for a population with average observable characteristics of those students who miss more than 13 days, the absence of 28 days would imply that 59% of this population would have an initial achievement level, a percentage that is not so definite from a public policy perspective.

References

Ahmed, M. & Iqbal, K. (2007). "Is There Any Threshold in Mother's Education and Child Health Relationship? Evidence from Nigeria". *Carleton College Department of Economics. Working Series.*

Arcia, E. (2006). "Achievement and enrollment status of suspended students". *Education and Urban Society, Vol. 38, 3, 359.*

Bedi, A. & Marshall, J. (2002). "Primary School Attendance in Honduras". *Journal of Development Economics, 69, 1, 129.*

Corville-Smith et al. (1998). "Distinguishing Absentee Students from Regular Attenders: the Combined Influence of Personal, Family and School Factors". *Journal of Youth and Adolescence, 27, 5, 629.*

Crone, L., Franklin, B., Glascock, C., & Kockan, S. (1993). "An Examination of Attendance in Louisiana Schools". *Paper was presented at the Annual Meeting of the Mid-South Educational research Association (22nd, New Orleans, LA, on November 10-12, 1993), 1-20.*

Daugherty, M. (2008). "Attendance and Other Factors That Influence Student Achievement in a Delaware Public School District". *Dissertation submitted by the faculty of Willmington University.*

DeKalb, J. (1999). Student Truancy. *Eric Digest Number 125. ERIC Clearinghouse on Educational Management*

Dfes, U.K. Department for Education and Skills (1996). *Education Act.*

Ding C. & Sherman, H. (2006). "Teaching Effectiveness and Student Achievement: Examining the Relationship". *Educational Research Quarterly, 29, 4, 39.*

Epstein, J. & Sheldon, S. (2002). "Present and accounted for: Improving student attendance through family and community involvement". *The Journal of Educational Research, 95, 5, 308.*

Finn J. & Voelkl, K. (1993). "School Characteristics Related to Student Engagement". *The Journal of Negro Education, 62, 3, 249.*

Gallego, F. (2002). "Competencia y Resultados Educativos: Teoría y Evidencia para Chile". *Cuadernos de Economía, V.39, Nº 118.*

Gump, S. (2006). “Guess Who’s (Not) Coming to Class: Students Attitudes as Indicators of Attendance”. *Educational Studies, 32, 1, 39.*

Hansen, B. E. (1996) “Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis”. *Econometric Society, the University of Chicago.*

Hansen, B. E. (2000) “Sample Splitting and Threshold Estimation”. *Econometric Society, the University of Chicago 68, 3, 575.*

Hoachlander, G., et al. (2001). “Attending to Attendance”. *Education Week, 20, 36, 40.*

Hong Y. et al. (2005). “Identifying Threshold Effects and Typologies in Economic Growth: A Panel Approach”. *Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July 24-27, 2005*

Johnston, R. (2000). As Studies Stress Link to Scores, Districts Get Tough on Attendance. *Education week, 2000, 20, 7, 1.*

Kleibergen, F., & Paap R. (2006): “Generalized reduced rank tests using the singular-value decomposition”. *Journal of Econometrics, Vol. 133, N° 1, pp. 97-126.*

Kube B. & Ratigan G. (1992). “Putting the Attendance policy to the Test”. *Clearing House, 65, 6, 348.*

Ladd, H. (1996). Holding Schools Accountable: Performance-Based Reform in Education. *Washington, DC: The Brookings Institution.*

Lamdin, D., (1996). “Evidence of student attendance as an independent variable in education production functions”. *The Journal of educational research, 89, 3, 155, Heldref Publications, Bloomington, Ill.*

Lamdin, D., (1998). “Attendance revisited: A reply to Borland and Howsen”. *The Journal of educational research, 91, 4, 198, Heldref Publications, Bloomington, Ill.*

McCarthy, K. (2002). The effects of student activity participation, gender, ethnicity, and socio-economic level on high school student grade point averages and attendance. United States Department of Education. *Office of Educational Research and Improvement, 62, 409-424.*

Mineduc (1988). Reglamento de Evaluación y Promoción del Alumnos de Educación General Básica, de Educación Media, de Educación de Adultos, y Proceso de Titulación en Educación Técnico-Profesional. *Reglamento Ministerio de Educación Pública.*

Mizala, A. y P. Romaguera (2000). "Determinación de Factores Explicativos de los Resultados Escolares en Educación Media en Chile". *Centro de Economía Aplicada, U. de Chile.*

Murray, S. (2002). "Raising school attendance". *Education Digest, 67, 6, 54.*

National Audit Office (2005). "Improving School Attendance". *Education Journal.*

Reid K. (2006). "An Evaluation of Inspection Reports on Primary School Attendance". *Educational Research, 48, 3, 267.*

Roby, D. (2004). "Research on School Attendance and Student Achievement: A Study of Ohio Schools". *Educational Research Quarterly, Vol. 28 (1), 3-14.*

Roderick M. (1997). "Habits Hard to Break: A New Look at Truancy in Chicago's Public High Schools". *Students Life in High School Projects, http://www.consortium-chicago.org*

Sapelli, C. y B.Vial (2002). "The Performance of Private and Public Schools in the Chilean Voucher System". *Cuadernos de Economía, Año 39, Nº 118.*

Staiger D. & Stock J. (2007). "Instrumental Variables Regression with Weak Instruments". *Econometrica, 65, 3, 557*

Steward, R. (2002). "School Attendance Revisited: A Study of Urban African-American Students' GPA and Coping Strategies. *Department of Counseling and Education Psychology, Michigan State University, Lansing, Michigan. 1-20.*

Steenbergen, Macro R., and Bradford S. Jones. 2002. "Modeling Multilevel Data Structures". *American Journal of Political Science, Vol.46, No.1 January: 218-237*

Strickland V. (1998). "Attendance and Grade Point Average: A Study".