UNIVERSITAT
ROVIRA I VIRGILI

DEPARTAMENT D'ECONOMIA

(creip)

CENTRE DE RECERCA EN ECONOMIA
INDUSTRIAL I ECONOMIA PÚBLICA

# WORKING PAPERS

## Col·lecció "DOCUMENTS DE TREBALL DEL DEPARTAMENT D'ECONOMIA - CREIP"

The Impact of Public Libraries on School
Achievement: The Case of Medellin

Gustavo Canavire-Bacarreza
Luis Diaz-Serrano
Alejandro Corrales-Espinosa

Document de treball n.25 - 2018

**DEPARTAMENT D'ECONOMIA – CREIP**
**Facultat d'Economia i Empresa**

# The Impact of Public Libraries on School Achievement:

# The Case of Medellin

Gustavo Canavire-Bacarreza[1]
Luis Diaz-Serrano[2]
Alejandro Corrales-Espinosa[3]

## Abstract

This paper explores the relationship between public libraries and school achievement. Medellin counts on a system of public libraries. However, public library parks are part of a separate system. The library parks are understood as cultural centers for social development that seek to encourage the meeting of citizens and to develop educational activities that involve the digital culture. The project began in 2004 and was consolidated until 2011 with the introduction of 9 different library parks. Using an approach of differences-in-differences and matching techniques, we explore the effect of the library parks in Medellin on school performance. We found a significant effect on the performance of the language test in the students treated, especially in the long-term specification.

**Key Words:** Public libraries, school achievement, impact evaluation, public investment.
**JEL Classification:** C33, H54, H76, I2, I26.

_____

[1] Universidad EAFIT, School of Economics and Finance, Medellin Colombia. Email: gcanavir@eafit.edu.co
[2] Universitat Rovira i Virgili, Departament d'Economia, Av. de la Universitat 1. 43204 Reus, Spain. E-mail: luis.diaz@urv.cat.
[3] Universidad EAFIT, School of Economics and Finance, Medellin Colombia. Email: acorral5@eafit.edu.co1

## 1. Introduction

The implementation of libraries in the different "Comunas"[1] of Medellin can have a potentially positive effect on the quality of education due to an increase in the proximity and availability of resources for students. Additionally, reading development is of great use for cognitive processes, so these programs create an incentive for the students to develop the ability to increase comprehension and retention. In turn, these factors can positively impact both the academic performance of students in the short term and their potential long-term job market outcomes.

Medellin counts on a System of public Libraries, and a separate system of public library parks are detached. In turn, this system of public libraries is part of the Library Network of Medellin and its Metropolitan Area. The library parks are understood as Cultural Centers for social development that seek to encourage social interaction and to develop educational activities that involve digital culture (Empresa de Desarrollo Urbano). In this context, these libraries seek to improve the quality of life of individuals and improve their education. The project began in 2004 and was consolidated until 2011, with the introduction of 9 different library parks. However, little is known formally about the impacts of these parks on the quality of education. Thus, the project seeks to study the impact of library parks on the quality of education in Medellin. More specifically, through the use of econometric techniques, the project aims to answer the question: What would be the results of the Saber 11 tests in the regions affected by the library parks in Medellin if the library parks had not been implemented? Additionally, the project examines whether there is any kind of heterogeneity in effects. Since the implementation of library parks has not been random, it is necessary to implement quasi-experimental techniques to reduce the effect of selection in the estimation.

Libraries play an important role in the cognitive development of the students who use them. However, much of the analysis of the impact of libraries worldwide has been concentrated on the use of libraries by the number of users, number of books, etc., leaving aside the impacts they may

---

[1] Comunas is the name that districts receive in Medellin

have on the quality of education (Celano and Neuman, 2001). This project seeks to analyze the impact of the implementation of Library Parks on the educational performance of students who are in the vicinity of the parks and who have a high probability of attending them. According to the Empresa de Desarrollo Urbano, "Library Parks are Cultural Centers for social development that promote citizen meeting, educational and recreational activities, collective construction, approach to new challenges in digital culture. And they are also spaces for the provision of cultural services that allow the cultural creation and the strengthening of existing neighborhood organizations. This project promotes neighborhood, zonal and city improvement. It transforms mentalities and becomes an engine of change for communities. For its quality, beauty and contribution to the development, the Library Parks will become urban and architectural references. The Library Parks are generators of opportunities for children, youth and adults, inclusive places that dignifies its citizens." (Empresa de Desarrollo Urbano).

Given these conditions it is expected that Library Parks will have different effects on society; one of the most important effects is to improve the conditions and capacities of those individuals who are in the area of influence. This should be reflected in higher levels of academic performance, given that there is scientific evidence that children and adolescents require high exposure to a wide range of activities and high-quality books in different subjects, categories and perspectives to increase their levels of understanding and reasoning. Additionally, there is evidence of an increase in the academic performance of students who are confronted with diverse and multicultural literature so that children can identify themselves (Neuman, 2000).

Thus, the main objective of this study is to analyze whether the implementation of Library Parks in Medellin has influenced the school performance of the students potentially affected by them, due to the area of influence in which they are developed. A priori, it is possible to hypothesize that the greater the student access to these Parks, the greater the possibility that the students use them and the greater the direct effects on the academic performance. It is also important to mention that there are multiplicative effects (spillover effects) on the neighborhood or the commune and

society as a whole. Finally, it is possible to examine whether there are any heterogeneous effects that allow for the identification of the robustness of this type of interventions

**2. Library Parks**

Library parks are part of the system of public libraries in Medellin and the network of libraries in the metropolitan area. These library parks are the most representative project of the development plan: "Medellin la más educada". This plan was implemented in the period 2004-2007 by the administration of Mayor Sergio Fajardo, a period in which five library parks were constructed. During the period of 2008-2011, Mayor Alonso Salazar Jaramillo continued the program and was responsible for strengthening the network of libraries with the construction of four more parks.

The library parks were intended as comprehensive centers of cultural and social development, by means of which they provide educational, cultural and recreational opportunities to the population. With the implementation of this project it was intended that the community would perceive that the library parks are more than just places to store books, aside from promoting equity and citizen participation, because the public libraries are spaces enabling inclusion since the egalitarian offer of technological. The Impact of Public Libraries on School Achievement: The Case of Medellin computer, cultural and social services and the access to knowledge (Cuadros-Rodríguez, Valencia, & Valencia-Arias, 2013). However, there are no studies that relate the impact of these libraries on the school performance of the people who benefit from them, so this work can be taken as a starting point for evaluations of this type of facilities that go more beyond qualitative assessments.

The company of urban development of Medellin defines the library parks as cultural centers of social development that encourage the meeting of citizens, educational and fun activities, the formation of groups and the approach to new challenges of digital culture. The majority of the parks are located in the peripheral areas of the city, which are characterized as being low income areas with high levels of violence. However, the parks are situated in close proximity to various modes

of public transport which enables easy access, an aspect which positively influences the probability of use of the library. (Japzon & Gong, 2005) The library parks have an approximate extension of 15,000 m2 and are comprised of the following spaces:

- Centers of regional business development: they are aimed at the development and consolidation of entrepreneurship projects.

- Reading rooms: they are aimed at children and adults. The rooms are adapted to different lifecycles.

- Virtual navigation room: classrooms with computers that feature free connection to the internet.

- Auditorium: a space equipped for the performance of concerts, musical presentations, plays and other artistic works.

- Mi Barrio room: a room allocated to the customs and traditions of the neighborhood.

- Toy library: a space for children between the ages of 0 to 10 years. Activities which impact your imagination and motor function are carried out.

- Exhibition room: a room intended for artistic expression.

- Services: cafeteria, stationer's, commercial premises among others.


## 3. Where do we stand? A Brief Review of the Existing Literature

The investment of public resources with the objective of increasing human capital or reducing levels of poverty are topics that have interested both economists and policymakers in recent decades (Rawlings & Rubio, 2003). Multiple studies have approached this issue from different perspectives, including different forms of intervention, different target populations and varied techniques to face the methodological challenges presented when it is necessary to evaluate these types of policies (Lee, 2009; Carneiro, Heckman, & Vytlacil, 2010; Attanasio, Guarín, Medina, & Meghir, 2015; Horowitz & Manski, 2013).

One of the most interesting lines of research is related to the returns generated by investment in education. Some work has focused on showing how government programs can help improve the educational performance of children and young people through conditional cash transfers or through investment in infrastructure, which may include building educational institutions or improving of existing facilities. This pair of policies is consistent with what Thapa, Cohen, Guffey, & Higgins-D'alessandro (2013) propose in their analysis of the school environment, in which student motivation and adequate facilities are determinants of good school performance.

Angrist, Bettinger and Kremer (2006), for example, concentrated on studying the long-term effect of the PACES program in Colombia. This program was responsible for subsidizing secondary education for low income young people. Program subsidies were made permanently conditional on good school 3 Canavire et al. performance during the year. The researchers applied parametric and non-parametric methods to correct for possible biases due to self-selection. They concluded that the treatment group had higher graduation rates, better school performance, and higher scores on state tests than the control group.

On the other hand, Kremer, Miguel, & Thornton (2009) evaluated a policy of merit scholarships for girls in Kenya that rewarded girls who achieved good academic tests by giving them cash and paying their school fees. This policy was successful, as it increased the grades in the examinations of the girls who were eligible for the scholarship. The main characteristic of this work was the positive externality generated by this intervention in the male students, who also had considerable increases in their school exams despite the fact that they could not participate in the scholarships.

Analyzing investment in infrastructure and its relationship with school results provides evidence in favor of this type of policy. Belskaya, Posso, & Peter (2014) evaluated the expansion of higher education in Russia through the increase of campuses and university graduates concentrating mainly on the marginal individual, who is directly affected by campus expansion. To evaluate the impact of the policy, they used the marginal treatment effect method in both versions,

parametric and semi-parametric. According to the researchers, university expansion attracts individuals with lower returns to college. They found that individuals in small cities and places that previously did not have a university campus are the ones that receive the most benefits from this policy.

Paxson & Schady (2002) focused on evaluating the impact of investment in education, specifically on infrastructure that FONCODES carried out in Peru, answering two main questions: Who benefited from this investment, and did the investment improve educational outcomes? The research showed that spending on school infrastructure improved the attendance rate of the youngest children because of the good policy, which focused on serving the poorest households and districts in Peru.

On the other hand, Duflo (2001) contributed to the literature related to the question of whether investment in infrastructure increases human capital and reduces poverty. Her work evaluated the impact on education and wages of the construction of 61,000 primary schools in Indonesia. The goal of the government's policy was to increase the percentage of children between the ages of 7 and 12 enrolled in schools. This research concluded that the construction of primary schools generated increases in educational achievements in Indonesia, as it encouraged an even higher percentage of the population to complete their primary education, which at the same time implies percentage increases in the wages of the beneficiaries by the government program.

As was previously discussed, infrastructure investment positively impacts enrollment rates and school performance, and it even has longer-term effects that are represented by increases in wages. However, these works focused on a particular type of infrastructure, such as schools and university campuses, while this research focuses on another kind of infrastructure, such as public libraries. It is necessary to know what other investigations related to this structure have found.

Public Libraries have been designed as a space where individuals can meet and interact with each other. It is also a place to obtain information, learn and teach. (Francis, Lance, & Lietzau, 2010; Gong, Japzon, & Chen, 2008). However, the literature review has shown two possible ways

of analyzing the impact of the public libraries on society. On one side, there exist articles that study the probability of usage of public libraries from a spatial perspective, including social, economic, academic and cultural factors. The second is through the relationship between school libraries (public and private) and academic outcomes. There are some traditional variables that influence the probability of using a public library; some of those are location, level of education, income, and race. Libraries with a closer influence area and better accessibility increase their probability of usage. For example, Koontz (1992) highlights the location of libraries as a long-term determinant of their use; she suggests that the decision of where the library should be located must be based on the geographic range of the library's market area and the characteristics of the local inhabitants.

The Impact of Public Libraries on School Achievement: The Case of Medellin On the other hand, Japzon & Gong (2005) recognized the importance of these traditional variables, but they also added another two variables with spatial and social components that could help to explain the use of public libraries in the neighborhoods of New York City. First, they applied a bivariate correlation method with an expected result of a positive correlation between individuals with the following characteristics: high levels of education, being White or Asian, being a middle-class individual and the use of public libraries. In addition to this, Japzon & Gong made a multivariate regression and found a causal relation between the use of public libraries and traditional variables, confirming the behavior shown in the previous correlation method. With respect to the new variables that represent the spatial and social components, this research proved that the use of public libraries could be higher in territories with better spatial accessibility, stronger social connections and cultural and racial diversity.

Almost all of the studies and researchers consider that school libraries have a positive impact on the skills of students and their scholastic achievement. According to Chan (2008), this trend began with a series of studies by Lance, Welborn and Hamilton-Pennell in Colorado, where they showed that libraries with adequate staff and equipment were the second-best predictor of academic performance. This study was replicated in Pennsylvania and Alaska with similar results.

Nevertheless, Chan warned about the low impact of libraries on school achievement, attributed mainly due to the lack of communication and coordination between teachers and library staff. In the same line, Haycock (2011) supported the idea that school libraries positively impact scholastic achievement, although he was focused on the possible higher scores on standardized tests of the students that attend schools that had more resources assigned to their school libraries. Haycock established through a correlation method that library usage has a positive impact on the skills of the students, as his work concluded that schools with libraries that are well-stocked, well-equipped and have adequate staff produce students with better scores on standardized tests.

Moreover, Bhatt (2010) used an IV approach to investigate the impact of library use on the amount of time spent reading, homework completion rates, average grades, behavior problems at school and grade repetition. The main reason to use IV regression was that OLS estimates were biases toward zero, because the students that enjoy reading the most were less likely to use a library. She found that living close to a library increases the probability of using it. Furthermore, library use positively impacts student performance through an increase in homework completion rates, reduces misbehavior, increases the amount of time an individual dedicates to reading, and decreases the time spent watching television.

**4. Methods**

The implementation of libraries in the different communes can have a potential positive effect on the quality of education due to the increase in the availability and proximity of resources for students. Additionally, the development of reading is of great use in developing cognitive processes, so libraries create an incentive for the students to increase comprehension and retention. In turn, these factors can positively impact both the academic performance of students in the short term and their potential long-term job market outcomes.

Medellin counts on a System of Public Libraries, which is separate from the public library parks. In turn, this system of public libraries is part of the Library Network of Medellin and its

Metropolitan Area. The library parks are understood as Cultural Centers for social development that seek to encourage the civic interaction and to develop educational activities dealing with digital culture (Empresa de Desarrollo Urbano). In this context, these libraries seek to improve the quality of life of individuals and improve their education. The project started in 2004 and was consolidated until the introduction of 9 different library parks in 2011. However, little is known formally about the impacts of these parks on the quality of education. Thus, this project seeks to study the impact of library parks on the quality of education in Medellin. More specifically, using econometric techniques, the project aims to answer the Canavire et al. question: What would be the results of the Saber 11 tests in the regions affected by the library parks in Medellin if the library parks had not been implemented? Additionally, this project examines whether there is any kind of heterogeneity in effects. Since the implementation of library parks has not been random, it is necessary to implement quasi-experimental techniques to reduce the effect of selection in the estimation. For this reason, we will use parametric and non-parametric estimates based on an estimator of differences in spatial differences, such as those developed in Canavire, Duque and Urrego (2016) and Chagas (2014).

Libraries play an important role in the cognitive development of the students who use them. However, much of the analysis of the impact of libraries worldwide has been concentrated on the use of libraries by the number of users, number of books, etc., leaving aside the impacts they may have on the quality of education (Celano and Neuman, 2001). This project seeks to analyze the impact of the implementation of library parks on the educational performance of students who are in the vicinity of the parks and who have a high probability of attending them.

According to the Empresa de Desarrollo Urbano, "Library parks are Cultural Centers for social development that promote citizen meeting, educational and recreational activities, collective construction, approach to new challenges in digital culture. And they are also spaces for the provision of cultural services that allow the cultural creation and the strengthening of existing neighborhood organizations. This project promotes neighborhood, zonal and city improvement. It

transforms mentalities and becomes an engine of change for communities. For its quality, beauty and contribution to the development, the library parks will become urban and architectural references. The library parks are generators of opportunities for children, youth and adults, inclusive places that dignifies its citizens." (Empresa de Desarrollo Urbano).

Given these conditions it is expected that library parks will have different effects on society. One of the most important effects is to improve the conditions and capacities of those individuals who are under its influence. This should be reflected in higher levels of academic performance given that there is scientific evidence that children and adolescents require high exposure to a wide range of activities and high quality books in different subjects, categories and perspectives to increase their levels of understanding and reasoning. Additionally, there is evidence of an increase in the academic performance of students who are confronted with diverse and multicultural literature so that children can identify themselves (Neuman, 2000).

Thus, the main objective of this study is to analyze if the implementation of library parks in Medellin has had an effect on the school performance of the students in the surrounding areas. A priori, it is possible to hypothesize that the greater the access to these parks, the greater the possibility that the students use them and the higher their academic performance. It is also important to mention that there are multiplicative effects (spillover effects) on the neighborhood or the commune and society as a whole. Finally, it is possible to examine whether there are any heterogeneous effects that allow the identification of the robustness of this type of interventions.

To estimate the effect of library parks on students' academic performance, we designed an evaluation model that considers the spatial distributions of students, the geographical location of the library parks and the location of the schools they serve the students. Initially, we started with a difference-in-differences approach that allows us to establish the effect of libraries on students' school performance, conditional on factors, such as distance to school, migration parameters, and parental education.

The differences-in-differences approach, due to Ashenfelter and Card (1985), has become very popular and is the basis for our study. In this context, the results for the two groups of two time periods are observed. One of the groups contains students who have been exposed to the existence of libraries in a nearby territory in the second period, but not in the first period. The second group consists of students 6 The Impact of Public Libraries on School Achievement: The Case of Medellin who are not exposed to libraries during any of the periods. For the case that concerns us, we will use neighborhoods that allow us to make the estimation of the model and to contrast the robustness at the individual level. In general, the difference-in-differences estimator is presented as follows:

$$y_{it}^b = \beta_0 + \beta_1 dB + \delta_0 d2 + \delta_1 d2 * dB + \mu_{it} \tag{1}$$

where Y is the result of the tests, d2 is a dummy variable for a period after the implementation of the libraries and dB captures the effects of possible differences between the groups. Additionally, the interaction of d2 and dB captures the desired effect. We will use robust estimators to control for heteroscedasticity.

Although the difference-in-differences estimator is adequate, we believe that there is a possibility of spatial correlation. Following the approach of Chagas et al. (2014), a spatial model of differences in differences can be estimated to quantify the effect of treatment in the treated area (area affected by libraries), as well as the effect on neighboring untreated areas. This model divides the common treatment coefficient into the direct treatment effect and the spatial side effects of the untreated units.

We define the result associated with treated and untreated units prior to treatment. It is important to note that spatial models consider that there is a relationship between geographical units, enabling the spread of a particular effect:

$$y_{it,0}^b = \beta X + \mu_{it} \tag{2}$$

$$y_{it,1}^b = y_{it,0}^b \tag{3}$$

where ybit;0 is the pretreatment outcome variable for the untreated units and ybit;1 is the result before treatment of the treated units; X represents the set of independent variables. Then, assuming the existence of a spatial impact on the untreated units caused by the treatment, it must depend on the proximity of each unit to the treated units. This proximity is measured and defined using a contiguity matrix, W. For the matrix W, different definitions can be used. Therefore, the outcome for each group after treatment will be as follows:

$$y_{it,0}^a = \beta X + w_i \rho d_{it} + \mu_{it} \tag{4}$$

$$y_{it,1}^a = y_{it,0}^a + \alpha \tag{5}$$

The parameter _ represents the direct impact of the treatment in the treated areas, while represents the spatial impact of the treatment. The treatment variable is denoted as dit, and wi represents the column i of the contiguity matrix W. Then, defining an indicator, Dit, which takes 0 for the untreated and 1 for the treated, we have the following:

$$y_{it} = (1 - D)y_{it,0} + D_{it}y_{it} \tag{6}$$

After this set of definitions, we can calculate the corresponding effects we are looking for, the treatment effects on the mean and the average treatment effect on untreated units (spatial effect of treatment).

$$ATE = E[y_{it,1}^a - y_{it,1}^b] - E[y_{it,0}^a - y_{it,0}^b] = \alpha \tag{7}$$

$$ATENT = E[y_{it,0}^a - y_{it,0}^b] = w_i \rho d_{it} \tag{8}$$

In matrix structure, we have:

$$Y = \Delta WY + \beta X + (\alpha + I \otimes W\rho)D + U \tag{9}$$

where Y is the result variable of the analyses, D identifies the dummy variable of treatment, I is the identity matrix, and U is the error vector. Then, using particular strategies of spatial econometrics, the necessary parameters can be estimated. This model is just one particular case of the general spatial model proposed in Elhorst (2014). The advantage of this strategy is that we can measure the indirect impact of the treatment and then calculate the total impact, which we expect should be much higher than the impact calculated without the spatial component analysis.

## 5. Data

The main source of data for this study was provided by ICFES. This database contains the socioeconomic information of the students of the city of Medellin that presented the tests of the state Saber 11. This test is directed to the students in the last year of High School. Information is available from 2006 to 2014. However, in this research we are interested in specifically analyzing some years that allow us to find short and long-term effects. For this we will take the years 2009, 2010 and 2013 as reference years. The most important data contained in this database is the result that the student obtained in the different branches of knowledge on the test, which are Mathematics, language, social sciences, philosophy, biology, physics, chemistry and English. The academic performance information in the test is accompanied by a characterization of the students. The database contains information such as age, socioeconomic status, household size, number of siblings, occupation and education of the parents. It also contains features of the home such as internet access, computers, and television, among other appliances. Finally, the database presents information about the school, such as the study day, type of institution, and other relevant characteristics of the educational establishment.

Table 2 present the summary statistics of our dataset. In this table, it is possible to identify the characteristics of the students who took the test Saber 11. The data show that on average, the students who took the test are 18 years old, live in urban areas and their socioeconomic stratum is between level two and level three. Parents of these students, on average, have a professional

education level and work as independent professionals, while mothers work as workers and their level of education is on average technical. The average household of a student who takes the test has a wage that is between two and three legal minimum salaries and has a computer, but not all computers have an internet connection. On the other hand, the average student attends in the morning or evening a public school of academic and technical character located in an urban area. These characteristics are similar for all reference years showing only small variations in some cases. For example, for the year 2013, the level of education of parents differs from the average behavior mentioned above, since the average parents go from having a professional level of education to having a high school education. In the same way, the average parents in 2013 are characterized as workers but are characterized as independent workers in other years.

In addition, to completing the information of the people who took the test, coordinates were generated that allowed us to locate the students' homes, educational institutions attended and park libraries. For each student, we calculated the distance between their home and the educational institution they attended at the time of the test, the distance between the student's home and the nearest Library Park, and the distance between the nearest Park library and educational institution. According to Table 3 the different measures of distance have a similar behavior for all reference years, on average the distance between a home and the school is 13 km, while the distance between the school and the nearest library is 12 km; finally, the shortest distance is the one between homes and libraries; on average a home is 1.3 km from the nearest library. These results reinforce the decision to take as a unit of measurement the individuals and their homes through the Coordinates to know the location of their home. As mentioned at the beginning of this article, being exposed to this type of structures can improve the skills of the students.

**Table 1**

Saber 11 Average score test by topic and year

| Variable | Year 2006 | | | | Year 2009 | | | | Year 2010 | | | | Year 2013 | |
| | Obs | Mean | Std. Dev. | Obs | Mean | Std. Dev. | Obs | Mean | Std. Dev. | Obs | Mean | Std. Dev. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Language | 24,670 | 49.44132 | 6.935455 | 27,109 | 46.85855 | 6.678676 | 30,355 | 51.11759 | 10.21235 | 26,795 | 48.46438 | 7.975968 |
| Mathematics | 24,670 | 46.01445 | 8.73295 | 27,109 | 44.44924 | 9.831213 | 30,355 | 49.89941 | 10.43567 | 26,795 | 45.33682 | 10.78135 |
| Social science | 24,670 | 45.99042 | 8.177727 | 27,109 | 45.20994 | 8.881729 | 30,355 | 50.40109 | 10.40337 | 26,795 | 44.9663 | 8.377853 |
| Philosophy | 24,670 | 47.82988 | 7.845662 | 27,109 | 41.38722 | 7.990526 | 30,355 | 49.67155 | 10.16298 | 26,795 | 40.49804 | 9.604928 |
| Biology | 24,670 | 47.07625 | 7.499611 | 27,109 | 45.66867 | 6.516779 | 30,355 | 49.89881 | 10.17166 | 26,795 | 45.10435 | 7.759346 |
| Chemistry | 24,670 | 44.87879 | 5.749865 | 27,109 | 44.82478 | 6.725775 | 30,355 | 48.63513 | 10.26798 | 26,795 | 44.63754 | 8.942339 |
| physics | 24,670 | 45.4568 | 7.513835 | 27,109 | 43.46385 | 7.546302 | 30,355 | 48.61098 | 10.30586 | 26,795 | 43.77809 | 11.21244 |
| English | 24,670 | 44.21779 | 10.28478 | 27,109 | 45.42879 | 11.08764 | 30,355 | 50.26948 | 12.25206 | 26,795 | 45.71644 | 12.21116 |

**Table 2**

Summary statistics of the covariates

| Variable | 2006 | | 2009 | | 2010 | | 2013 | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. | Mean Std. | Dev. |
| Father education level | 10.60226 | 6.493786 | 19.01335 | 25.39343 | 19.21231 | 26.10397 | 5.015425 | 2.878807 |
| Mather education level | 11.75882 | 8.114972 | 14.91622 | 17.25386 | 14.35067 | 16.54227 | 5.001498 | 2.514565 |
| father occupation | 20.50353 | 4.889057 | 20.50338 | 3.710909 | 20.68397 | 3.598818 | 8.643018 | 3.223217 |
| Mather occupation | 21.02821 | 3.076114 | 20.97466 | 2.702027 | 21.06982 | 2.663029 | 9.147922 | 2.886398 |
| Internet | 0.8031161 | 0.3979257 | 0.4551592 | 0.4979944 | 0.4898955 | 0.4999061 | 0.7163447 | 0.45078 |
| Computer | 2.592645 | 1.024268 | 1.777696 | 1.473973 | 1.837899 | 1.461335 | 0.8135295 | 0.4969372 |
| Age | 18.00927 | 4.52524 | 18.08158 | 4.578051 | 18.53977 | 5.347984 | 17.84568 | 4.570468 |
| School location | 2.968808 | 0.1813757 | 2.959328 | 0.2086301 | 2.962017 | 0.2025664 | 2.944209 | 0.2575252 |
| School character | 1.542643 | 0.9024458 | 1.495334 | 0.8933596 | 1.488618 | 0.8918011 | 1.500877 | 0.9045003 |
| School day | 2.910012 | 1.450191 | 3.020805 | 1.442932 | 3.017328 | 1.402317 | 2.960403 | 1.431043 |
| School type | 1.623389 | 0.4845459 | 1.631377 | 0.4824403 | 1.609982 | 0.4877621 | 1.62911 | 0.4830521 |
| Socioeconomic status | 2.97067 | 1.367139 | 2.568047 | 1.08458 | 2.487467 | 1.070986 | 2.262849 | 0.8260021 |

16

**Table 3**

Average distances

| Variable | 2006 | | 2009 | | 2010 | | 2013 | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| School-Library | 15.99688 | 242.478 | 13.55501 | 290.3568 | 8.387 | 208.7739 | 12.65567 | 273.6039 |
| House-Library | 1.401241 | 0.8498126 | 1.372697 | 0.8699943 | 1.391179 | 0.851872 | 1.390578 | 0.836929 |
| House-School | 17.24255 | 243.1098 | 14.86932 | 290.6483 | 9.812222 | 209.0016 | 13.95036 | 273.8814 |

## 6. Results

To evaluate the impact of libraries on school results, it is pertinent to consider that library parks opened in different years. Between the years 2006 and 2008, 5 libraries entered operation, while the remaining four began work between the years 2011 and 2013. For this reason, the estimates for the Saber 11 tests are divided into groups and show the progressive effect of the library parks on the results.

$$Y_i = \beta_0 + \beta_1 * T_i + \beta_2 * t_i + X_i \beta_3 + \mu_i \qquad (10)$$

The variable Ti takes the value of 0 if the individual belongs to the control group and takes the value of 1 if it belongs to the treatment group. On the other hand, ti is a variable indicating the period in which the test was performed, where t = 0 refers to the base year and t = 1 refers to the final year used for the comparison. Finally, Xi is a set of covariates that contains characteristics of the student and the school that can influence the result of the test Saber 11, which is the dependent variable Yi.

Table 4 presents the results of the estimates for the first set of libraries, that is, those that opened between 2007 and 2008. The table presents the coefficients associated with the variable that measures the effect of the intervention and its level of significance. As seen in the table, two periods of time in 2009and 2010 were analyzed to show the differences that may exist when the library has been operating in a territory for a longer time. As expected, most of the effects found in 2009 were not significant except for the math test, where for the 200 and 300-meter specifications, it was found that being treated increased the math score by 1.2 and 0.9 points.

On the other hand, in Table 4, when the specification is made for the year 2010, the math, language and English scores have a positive and statistically significant effect. It is important to note that now there are more types of tests that are positively impacted. This result could be related to some of the purposes of the library parks, in particular encouraging reading and writing, skills that may be on this type of test. Until now, significant and positive effects have been found in the short term, with the tests of language, mathematics and English being the ones that have been most

impacted by the proximity of students to the park libraries. As mentioned at the beginning of this article, one of the main objectives of the construction of the library parks was that these places were perceived by the community as places where they could meet and do cultural, social, pedagogical and academic activities, such as those that help improve the skills necessary to do reading comprehension exercises.

When we move on to the long-term analysis, it is important to mention that another library has been included in the analysis, which implies a larger treatment group. Additionally, those that were already treated by the proximity to the libraries already built have had more time exposed to the treatment, so it is expected that in this longer time horizon, more evidence in favor of an impact greater will be found than that found in the short-term situation. When modifying the restriction of the opening year of the libraries and included in the analysis those that opened in 2011 the results suffer variation. Table 5 shows how the coefficient that measures the effect of the intervention is statistically significant for the Language test in all possible distance specifications that condition on the formation of the control and treatment groups. As seen, in the long term the effect of the library parks is concentrated on the scores in language, which is the only statistically significant difference among all the subtopics that made up the test. In contrast to what occurred in the short term, where the difference between the scores of the individuals who belonged to the group of treaties and control was between 0.8 and 1.2 units, in the long term this difference increased between 3 and 5 points depending on the reference distance.

The fact that the results are statistically significant for the language test is not only striking but also coherent with the objectives of the library parks, since these places have as main objectives the use of free time and the meeting of citizens as mentioned by the urban development company of Medellin. The urban development company of Medellin defines the library parks as cultural centers of social development that encourage the meeting of citizens, educational and fun activities, the formation of groups and the approach to new challenges of digital culture through multiple spaces that libraries have, such as reading rooms.

| | 2009 | | | | 2010 | | | |
| | Language | Mathematics | Social science | Philosophy | Language | Mathematics | Social science | Philosophy |
|---|---|---|---|---|---|---|---|---|
| 200 mts | 0.174 | -0.003 | 0.4 | -0.144 | 1.006 | 1.157 | 0.114 | 0.075 |
| p-value | 0.7322 | 0.9971 | 0.556 | 0.8325 | 0.1323 | 0.0744 | 0.8731 | 0.9144 |
| 300 mts | 0.201 | 1.265 | 0.289 | -0.09 | 1.109 | 0.835 | 0.641 | 0.761 |
| p-value | 0.5158 | 0.0061 | 0.4752 | 0.8124 | 0.006 | 0.0362 | 0.1126 | 0.0629 |
| 400 mts | 0.39 | 0.977 | 0.599 | 0.546 | 0.998 | 0.631 | 1.264 | 0.886 |
| p-value | 0.0651 | 0.0021 | 0.0334 | 0.0296 | 0.0006 | 0.03 | 0 | 0.0031 |
| | Biology | Chemistry | physics | English | Biology | Chemistry | physics | English |
| 200 mts | 0.445 | 0.066 | 0.45 | 0.439 | 0.496 | 0.188 | -0.561 | -0.295 |
| p-value | 0.4103 | 0.9007 | 0.3574 | 0.5482 | 0.501 | 0.7752 | 0.4123 | 0.674 |
| 300 mts | -0.164 | 0.369 | 0.634 | 0.786 | 0.493 | 0.48 | 0.26 | 1.295 |
| p-value | 0.59 | 0.2367 | 0.0472 | 0.0804 | 0.2546 | 0.2591 | 0.5282 | 0.0037 |
| 400 mts | 0.268 | 0.376 | 0.753 | 1.187 | 0.905 | 0.632 | 0.896 | 1.446 |
| p-value | 0.2023 | 0.0781 | 0.001 | 0.0005 | 0.0026 | 0.0335 | 0.0029 | 0 |

**Table 5**

Effect of Libraries Parks on Saber 11 test. Libraries build 2007-2011

|  | 2013 | | | |
|---|---|---|---|---|
|  | **Language** | **Mathematics** | **Social science** | **Philosophy** |
| 200 mts | 5.931 | 2.866 | 0.623 | -0.781 |
| p-value | 0.0557 | 0.202 | 0.7898 | 0.6984 |
| 300 mts | 5.055 | 0.39 | -0.101 | 1.84 |
| p-value | 0.0105 | 0.8617 | 0.9527 | 0.2539 |
| 400 mts | 3.406 | -2.002 | -2.02 | -0.91 |
| p-value | 0.0095 | 0.1811 | 0.1599 | 0.6165 |
|  | **Biology** | **Chemistry** | **physics** | **English** |
| 200 mts | 0.35 | -0.669 | 1.206 | -2.122 |
| p-value | 0.8495 | 0.6681 | 0.5081 | 0.3397 |
| 300 mts | -0.032 | -0.232 | -1.408 | 1.395 |
| p-value | 0.9831 | 0.8344 | 0.5799 | 0.6663 |
| 400 mts | -0.552 | 0.745 | -0.325 | -0.281 |
| p-value | 0.7237 | 0.6031 | 0.8497 | 0.8784 |

## 7. Robustness: Placebo Test and Matching.

To identify the impact of the library parks of the city of Medellin on school performance, this article implements a placebo test. So far, the positive effect of the library parks on the Saber 11 tests for the language test under the long-term specification has been shown. The objective of implementing this test is to show that the satisfactory results that have been found are not given by patterns of increases in the scores of the language test across the city of Medellin.

To generate the placebo test, we must create a false treatment group. For this we will extend the distance that defines the treatment to 600 meters, 900 meters and 1.2 km. This means that the groups to be compared are groups that in the previous specifications belonged to the control group. What should we expect? If the library parks really have a significant effect on the language section of the Saber 11 Tests, we should not find an impact under this new specification. Put the opposite way, if we find a significant impact, that means increases in language test scores come from sources other than the library parks. According to the results shown in Table 6, there is no evidence of significant differences in the scores in the

Spanish subject between the new control and treatment groups, thus reinforcing the results obtained previously where it was found that being at a distance from the libraries parks had a significant and positive effect on school performance in language. In addition to the placebo test, matching methods were used to compare between individuals who have similar characteristics and have presented the test in 2013, since this was the period in which the most significant results were obtained.

**Table 6**
Placebo test

| | Language | | |
| | 600 mts | 900 mts | 1200 mts |
| --- | --- | --- | --- |
| Coefficient | 0.9777105 | 0.6426371 | 0.1069384 |
| p-value | 0.327 | 0.375 | 0.876 |

The Matching strategy consisted in estimating the likelihood of belonging to the control group or the treatment group controlling for characteristics of the student, their home and their school. That is, we wanted to find individuals who were very similar in their characteristics to compare their average score in the language test, where the main difference was that one of them belonged to the treatment group and the other to the control group. According to Table 7, there is a significant difference between the language scores of those students who belong to the treatment group and those who belong to the control group. As the specification of the treatment depends on the distance from the Library Park, the difference in the scores decreases, confirming the behavior shown in the estimates presented above.

**Table 7**
Radius Matching with the maximum common support and calipers

| | Language | | |
| | 200 mts | 300 mts | 400 mts |
| --- | --- | --- | --- |
| Difference | 9.22313596 | 5.48614861 | 3.11828992 |
| p-value | 0.00782199 | 0.0098662 | 0.0408644 |

**8. Discussion.**

The main objective of this study was to analyze whether the implementation of library parks in Medellin has influenced the school performance of the students potentially affected by them, due to their proximity to the library. A priori, it was possible to hypothesize that the greater the access to these parks, the greater the possibility that the students use them and the greater the direct effects on academic performance.

In this study, we showed that the library parks of Medellin, places that were thought of as cultural centers of social development, that encourage the meeting of citizens, provide educational and fun activities, form groups, and approach the new challenges of culture through digital media, have a positive impact on school performance as measured by state tests Saber 11. There are statistically significant differences between students who are closer to the library and those who are further away, specifically for language tests.

**References**

Andrea C . Japzon, & Gong, H. (2005). A Neighborhood Analysis of Public Library Use in New York City. The Library Quarterly, 75(4), 446–463.

Angrist, J., Bettinger, E. and Kremer, M. (2006). Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. The American Economic Review, 96(3), 847–862.

Ashenfelter, O & Card, D Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs (1985) The Review of Economics and Statistics, 67(4).

Attanasio, O., Guarín, A., Medina, C., & Meghir, C. (2015). Long Term Impacts of Vouchers for Vocational Training: Experimental Evidence for Colombia. NBER WP.

Bhatt, R. (2010). The impact of public library use on reading, television, and academic outcomes. Journal of Urban Economics, 68(2), 148–166.

Belskaya, O., Posso, C., & Peter, K.S. (2014). Education: Evidence from Russia College Expansion and the Marginal Returns to Education: Evidence from Russia. IZA Discussion Paper #8735.

Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2010). Estimating Marginal Returns to Education Estimating Marginal Returns to Education, 101(5275), 2754–2781.

Celano, Donna, Neuman, Susan B. (2001). The Role of Public Libraries in Children's Literacy Development: An Evaluation Report. Pennsylvania Library Association Technical Report, Pennsylvania Department of Education. (http://www.ifpl.org/ Junior/ studies/RoleofLibraries.pdf) (03.03.10).

Castris, D., & Pellegrini, G. (2012). Does Spatial Agglomeration decrease Regional Unemployment? Some evidence from Europe. In 24th Annual Conference of the European Association for Evolutionary Political Economy.

Chagas, A.; Almeida, A. & Azzoni, C. (2014). Sugar Cane Burning and Human Health: A Spatial Difference-in-Difference analysis. FEA-USP Working Paper Series, Universidade de São Paulo, 2014-20.

Chan, C. (2008). The impact of school library services on student achievement and the implications for advocacy: A review of the literature. Acces, 22(4), 15–20.

Francis, B. H., Lance, K. C., & Lietzau, Z. (2010). Continue to help students achieve standards: the third Colorado study (2010), (November), 20.

Gong, H., Japzon, A. C. A., & Chen, C. (2008). Public Libraries and Social Capital in Three New York City Neighbourhoods. Tijdschrift Voor Economische En Sociale Geografie (Journal of Economic & Social Geography), 99(1), 65–83.

Haycock, K. (2011). Connecting British Columbia (Canada) School Libraries and Student Achievement: A Comparison of Higher and Lower Performing Schools with Similar Overall Funding. School Libraries Worldwide, 17(1), 37–50.

Horowitz, J. L., & Manski, C. F. (2013). Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data, 95(449), 77–84.

Koontz, C. (1992). Public Library Site Evaluation and Location: Past and Present Market-Based Modelling Tools for the Future. Library and Information Science Research, 14(4), 379–409.

Kremer, M., Miguel, E., & Thornton, R. (2009). Incentives to learn. The Review of Economics and Statistics, 91(3) (February), 437–456.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. Review of Economic Studies, 76(3), 1071–1102.

Neuman, S. B. (2000) Every child a reader. Paxson, C., & Schady, N. R. (2002). The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru. The World Bank Economic Review, 16(2), 297–319.

Rawlings, L. B., & Rubio, G. M. (2003). Evaluación del impacto de los programas de

    transferencias condicionadas en efectivo.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D 'alessandro, A. (2013). A Reviewof School

    Climate Research. Review of Educational Research, 83(3), 357–385.