FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

NATÁLIA ZANETTI MARCHI

ESSAYS ON ECONOMICS OF EDUCATION

São Paulo
2020

NATÁLIA ZANETTI MARCHI

ESSAYS ON ECONOMICS OF EDUCATION

Tese apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para a obtenção do título de Doutora em Economia. Campo de conhecimento: Microeconomia Aplicada.

Orientadora: Profa. Dra. Cristine Campos de Xavier Pinto

Coorientador: Prof. Dr. Vladimir Pinheiro Ponczek

São Paulo

2020

Marchi, Natália Zanetti.
Essays on economics of education / Natália Zanetti Marchi. - 2020.
99 f.

Orientador: Cristine Pinto.
Co-orientador: Vladimir Pinheiro Ponczek.
Tese (doutorado CDEE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Educação - Aspectos econômicos. 2. Alfabetização. 3. Interação social. 4. Redes sociais. 5. Educação integral. I. Pinto, Cristine. II. Ponczek, Vladimir Pinheiro. III. Tese (doutorado) – Escola de Economia de São Paulo. IV. Fundação Getulio Vargas. V. Título.

CDU 37.014.54

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O
Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

NATÁLIA ZANETTI MARCHI

ESSAYS ON ECONOMICS OF EDUCATION

Tese apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas como requisito para obtenção do título de Doutor em Economia de Empresas.

Campo de Conhecimento:
Microeconomia Aplicada

Data da aprovação: 28/05/2020

Banca examinadora:

Profa. Dra. Cristine Campos de Xavier Pinto
EESP - FGV

Prof. Dr. Vladimir Pinheiro Ponczek
EESP - FGV

Prof. Dr. Daniel da Mata
EESP - FGV

Prof. Dr. Daniel Domingues dos Santos
FEARP-USP

Prof. Dr. Reynaldo Fernandes
FEARP-USP

Dedico esta tese à ousadia de mudar de rota e se reinventar.

Agradecimentos

Aqui deixo meus agradecimentos, afinal, ninguém constrói nada sozinho.

Agradeço à minha família. Meus pais e mestres da minha vida, pelo amor incondicional que já descreve tudo. Minha irmã e minha parceira de vida desde sempre, pelo apoio e incentivo. Meu afilhado, por ser a razão que me leva a querer ser sempre melhor e a trabalhar por um mundo melhor para você. Meu companheiro, por toda a paciência.

Agradeço aos meus orientadores por terem confiado no meu trabalho. Tudo começou com um telefonema anos atrás que me trouxe até a defesa de um doutorado hoje.

Agradeço aos membros da banca, por suas contribuições e pela disponibilidade para participar da minha defesa.

Agradeço ao Daniel Santos. Como já disse antes, você não imagina a importância que tem para mim o trabalho para o qual me apresentou, nada que eu diga será suficiente para te agradecer.

Agradeço à Gabriela Fonseca pela amizade. Por ter dividido comigo esses anos de doutorado. Você sabe o quanto seu apoio foi importante para mim.

Agradeço a todos os professores e todos os colegas de trabalho que já fizeram parte da minha vida. Cada um certamente contribuiu para o meu aprendizado e crescimento profissional.

Agradeço a todos os colegas de turma, funcionários e professores da EESP, por terem tornado minha experiência nesses anos mais enriquecedora e mais suave nos momentos difíceis.

Agradeço aos meus amigos que me apoiaram e me incentivaram nessa trajetória e que também querem de volta a amiga que o doutorado roubou deles.

Por fim, agradeço às agências de fomento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) e Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Inicialmente este trabalho foi realizado com o apoio da CAPES - Código de Financiamento 001 - e, na sequência, minha pesquisa foi financiada pela FAPESP por meio da bolsa n° 2016/25184-5.

"When you educate one person you can change a life,
when you educate many you can change the world."
(Shai Reshef)

Resumo

Esta tese é composta por três ensaios em economia da educação.

No primeiro capítulo, investigamos os efeitos de um programa de leitura e escrita sobre o desenvolvimento do letramento de crianças na educação infantil. Para tal, conduzimos um estudo aleatorizado, incluindo 44 escolas que receberam a intervenção e 44 escolas de controle. Os resultados principais indicam que as crianças que participaram do programa tiveram desempenho superior às crianças do controle, especialmente nas medidas de leitura de uma "parlenda" e escrita de palavras. As análises de efeitos heterogêneos sugerem que a intensidade do tratamento e a infraestrutura das escolas podem intensificar os efeitos sobre o desenvolvimento das habilidades de escrita.

No segundo capítulo, nosso objetivo é entender a relação entre a influência dos pares e as competências socioemocionais. Utilizando uma base de dados com informação sobre os vínculos de amizade dos alunos, primeiro consideramos um modelo tradicional de estimação de efeitos de pares que permita a identificação dos efeitos endógeno e contextual. As evidências sugerem efeitos endógenos negativos para todas a medidas consideradas (cognitivas e socioemocionais) e a presença de efeitos contextuais da medida socioemocional Self-Management sobre a nota de português. Em nosso segundo exercício analisamos quais características mais influenciam no processo de formação das redes sociais e observamos que, entre habilidades cognitivas e socioemocionais, apenas a proximidade em termos das competências socioemocionais aumentam a probabilidade de dois alunos se tornarem amigos. Nosso terceiro exercício explora a possibilidade de efeitos heterogeneos, analisando o papel do posicionamento do aluno dentro da rede social na intensidade dos efeitos. Os resultados indicam que a centralidade do aluno é um importante preditor de seus resultados acadêmicos, mesmo quando consideramos a média da centralidade dos pares.

No terceiro capítulo, investigamos se a implementação de um programa de ensino médio integral pode afetar a estrutura da rede de amizade dos alunos. Utilizando dados de alunos de escolas públicas do Estado de Sergipe (Brasil) identificamos que o programa coloca os alunos tratados em posição mais central. Além disso, investigamos se essa variação na estrutura das redes não poderia ser um canal através do qual o programa afeta os resultados acadêmicos dos alunos. Os resultados do modelo indicam que as medidas de conectividade dos indivíduos influenciam positivamente suas notas de matemática, o que sugere que pode haver um efeito de mediação do programa sobre essas notas que deveria ser considerado quando da avaliação dos efeitos da política.

Palavras-chaves: Intervenções Escolares, Desenvolvimento da Alfabetização, Efeitos de Pares, Redes Sociais, Educação em Tempo Integral.

Abstract

This thesis consists of three essays on economics of education.

In the first chapter we investigate the effects of a reading and writing program on literacy development in early childhood education. To this end, we conducted a randomized study, including 44 intervention schools and 44 control schools. The main results indicate that the children who participated in the program outperformed the control children, especially in the measures of reading a "rhyme" and writing words. Analysis of heterogeneous effects suggest that the intensity of treatment and the infrastructure of schools may reinforce the effects on the development of writing skills.

In the second chapter our goal is to understand the relationship between peer influence and socioemotional skills. Using a database with information on students' friendship ties, we first consider a traditional peer effects estimation model to identify endogenous and contextual effects. The evidence suggests negative endogenous effects for all measures (cognitive and socioemotional) and the presence of contextual effects of the socioemotional measure Self-Management on Language score. In our second exercise, we analyzed which characteristics most influenced the networking formation process and observed that, among cognitive and socioemotional skills, only homophily in terms of socioemotional skills increases the probability of two students becoming friends. Our third exercise explores the possibility of heterogeneous effects, analysing the role of the student's position within the social network in the intensity of the effects. The results indicate that the student's centrality is an important predictor of his academic results, even when we consider the average centrality of his peers.

In the third chapter we investigate whether the implementation of a full-time high school program can affect the structure of the students' friendship network. Using data from public school students in the State of Sergipe (Brazil), we identified that the program places the treated students in a more central position. In addition, we investigate whether this variation in the networks structure could be a channel through which the program affects students' academic scores. The results of the model indicate that the individuals' connectivity measures positively influence their math test scores, which suggests that there may be a mediation effect of the program on these scores that should be considered when assessing the policy's results.

Key-words: Schooling Interventions, Literacy Development, Peer Effects, Social Networks, Full-Time Education.

List of Figures

List of Tables

Contents

1 Literacy Development in Early Childhood Education

1.1 Introduction

There is abundant literature showing the importance of acquiring certain skills during early childhood, as it is a very relevant period in terms of brain development (THOMPSON; NELSON, 2001). Thus, stimuli and experiences are especially important, considering that the accumulation of knowledge in this period supports and facilitates learning in the subsequent stages (CUNHA et al., 2006).

Regarding reading skills, studies show that it is unlikely that development lags, which can be observed at the beginning of the learning process, will be compensated. Skills levels verified in preschool and in the first years of elementary education are good predictors of future reading results (CUNNINGHAM; STANOVICH, 1997; DUNCAN et al., 2007; STORCH; WHITEHURST, 2002; COSTA et al., 2013) and go even further, as studies show a correlation between reading and writing skills and one's type of occupation around the age of 20 years (ATHANASOU, 2001).

The greatest concern accruing from the above facts is that the acquisition of cognitive skills among children from different socioeconomic levels varies considerably. In fact, these inequalities can be verified even among preschoolers (HART; RISLEY, 1995) and usually have lifelong effects (HECKMAN, 2008).

In the literacy field, for instance, the Coley report (COLEY, 2002), which resulted from a study initiated in 1998 in a representative sample of American children, shows that 85% of the children who attended preschool and belonged to the first quintile of the socioeconomic measure, were able to recognize alphabet letters, while only 39% of the last quintile presented the same ability.

In Brazil, according to data from the last National Literacy Assessment (*Avaliação Nacional de Alfabetização* - ANA 2016) , 55% of the children in the 3rd grade of public schools presented insufficient reading knowledge, while 34% presented insufficient writing skills. Even though there are no data concerning private schools[1], we can have a notion

[1] In Brazil, private elementary schools usually provide better education than public schools. According to results provided in 2017 by Basic Education Development Index (*Índice de Desenvolvimento da Educação Básica* - IDEB), calculated using data concerning pass rates obtained in School Census, and performances average obtained in the assessments conducted by the Basic Education Assessment System (*Sistema de Avaliação da Educação Básica* - SAEB), students attending the 1st to the 5th grades in private schools scored 7.1; those from the 6th to 9th grades scored 6.4; and high schoolers scored 5.8. The IDEB scores of public schools, in turn, were 5.5; 4.4; and 3.5 for students attending the same grades, respectively.

of educational inequalities when we look at ANA's desegregated data. In the north, for instance, one of the poorest Brazilian regions, 70% and 53% of the children have insufficient levels of reading and writing skills, respectively. In the southeast, in turn, one of the most developed regions in Brazil, these proportions were 44% for reading and 21% for writing skills.

The relevance of these disparities and the probability of their perpetuation has been the focus of many researchers and policymakers and the evidence points out that interventions carried out in early childhood are important mechanisms in reducing inequalities and tend to be more cost-effective (CUNHA et al., 2006). For two of the most successful and documented interventions targeting vulnerable children, such as the *Perry Preschool Program* and the *Abecedarian Program*, positive effects were found for participants through adulthood (HECKMAN; KARAPAKULA, 2019b; CAMPBELL et al., 2012). For the *Perry Preschool* program, intra and intergenerational effects (HECKMAN; KARAPAKULA, 2019a) were also observed.

This study is part of this context of research addressing interventions implemented during early childhood. Our goal is to estimate the impact of a Reading and Writing program implemented in 2016 among public preschools in Brazil[2].

This Reading and Writing program is a continuous education program intended to reach all the professionals involved in preschool education in order to promote changes in reading and writing practices. A cascade-like training takes place, through virtual and face-to-face actions, involving from technicians of the Municipal Department of Education, school principals, teaching coordinators, and teachers, up to the point in which the children are finally reached.

As far as we know, other programs have already been or are being implemented in Brazilian public early childhood schools (*Alpha e Beto Pré-Escola; Paralapracá; Pequenos Leitores; Leitura e Escrita na Educação Infantil*), however, no studies assessing the effects of these programs were found.

In other parts of the world there are also early literacy programs aimed at low-income children during childhood education and several of them had positive effects on measures such as phonological awareness, letter knowledge, word identification and some writing tests (ARAM, 2006; BAUSERMAN et al., 2005; NEUMAN, 1999; WASIK; BOND, 2001; FUCHS et al., 2001; BLACHMAN et al., 1999). Nevertheless, as far as we can tell, they keep reasonable differences in relation to our study, such as: (i) the focus of the pedagogical content, the majority focused on reading [3]; (ii) the role of those involved

[2] A randomized control trial was conducted in 2016 with 44 treatment schools and 44 control schools distributed throughout nine Brazilian cities, reaching approximately 4,000 students.

[3] Only one of the programs included in the study by Aram (2006) address alphabetic skills and "would involve games and activities promoting phonological awareness, letter knowledge, and basic writing".

and the form of implementation, in all of them the training is direct with the teacher who is responsible for developing specific and well-structured activities with the students; and (iii) the sample size, in the largest of them 500 children were assessed.

Therefore, we understand that this analysis presents a relevant contribution. Not only on the literature concerning mechanisms that may decrease inequalities in terms of skills developed by children, in both vulnerable and non-vulnerable groups, but also provides evidence that can support public policies.

This paper is organized as follows: in the next section, the program and the intervention design are presented along with an analysis of the collected data. Sections 1.3 and 1.4 present a discussion of the estimated models and main results. At the end, we present our final considerations.

1.2 Intervention Program, Evaluation Design and Data

1.2.1 The Reading and Writing Program

The primary focus of this one-year program is to strengthen early childhood education public policies by offering technical-pedagogical training to those providing education to children between 4 and 5 years and 11 months of age.

Training is organized in a cascade style, meaning that information flows from one group to another, involving all those engaged in preschool education, until it reaches the children. In the upper tier, technicians from the Municipal Department of Education attend face-to-face seminars with specialized consultants and have continuous access to material and guidance. These technicians, in turn, organize meetings with the schools' principals and teaching coordinators, while the principals become responsible for training the teachers. Teachers and principals also have access to an online course that is available on the program's platform.

This training methodology is intended to establish a collaborative network in which all actors become co-responsible for the quality of education provided to children, supporting the teachers to perform meaningful work in terms of reading and writing according to the "field of experience - listening, speaking, thinking, and imagination" of the National Common Curricular Base (*Base Nacional Comum Curricular*). Additionally, this training dynamics is intended to establish a sustainable environment in the medium and long term to ensure the continuity of this work when program ceases, considering that the goal of involving everyone is to change the teaching culture in the entire network, changing reading and writing practices.

The entire network training is based on two main axes: one linked to training strategies such as thematization of practice, and one linked to the topic of child edu-

cation. In 2016, the thematic content was the approach to reading and written culture, promoting debates on written language as a system rather than merely a code to be decoded; what and how children think about writing; how they establish hypotheses; and which interventions teachers can implement to promote the advancement of children in this process. Regarding reading, the importance of teachers' reading for children is emphasized, along with promoting conditions for children to develop reading behaviors; and also how to rethink reading practices, the quality of books, and the teachers' role as reading mediators. The methodology includes tasks that enable learning by doing, so that immediate change is observed in children's practices while concomitantly promoting the training of the different actors.

1.2.2 Evaluation Design and Measures

Before the beginning of the school year 2016, 88 early childhood schools (with 4 to 5-year-old students) distributed in nine Brazilian cities were randomly selected to participate in this study.

The sampling process excluded all rural schools. Moreover, because some principals manage more than one school and the program hinges on the principal's role, only one school per principal was selected, thus avoiding any spillover effects. Schools were paired beforehand according to the number of students and, then, 44 pairs (88 schools) were randomly drawn. In each pair, one of the schools was randomly assigned to participate in the program (i.e., training would be provided to principals and teachers) and the other assigned to the control group.

In order to access the program's effects, a field research was conducted at the end of the school year (between October and November 2016), in which four assessments were applied. The Peabody Picture Vocabulary Test, Reading Test and Writing Test were applied to all the students. The fourth assessment was determined by a draw between the Expressive Vocabulary Test and Pre-Literacy Knowledge and Skills Test. The characteristics of the instruments are presented below, and the Appendix includes examples of each test.

1. Peabody Picture Vocabulary Test (PPVT): Measures receptive language. The Hispanic American version was used in this study. It consists of 125 items and was validated and standardized for Brazil by Capovilla (1997) with children from 2 to 6 years old. The assessor presents an image card with four numbered pictures and says the name of one of the pictures, while the examinees are expected to point it out and tell the corresponding number.

2. Reading Test: Educators developed it specifically for this study. It determines the stage of the reading acquisition process and is divided into two subtests. *Reading*

a "rhyme": the assessor recites a "rhyme" (*parlenda*) (a verse with a child theme), which is written on the student's answer sheet, and asks the child to check five specific words that appear in the "rhyme", i.e., the test verifies the reading of words within a given context. *Reading a list:* the examiner reads a list with six words, which is written on the student's answer sheet and asks the child to check three specific words in the list. In each subtest, students score one point for each correct word [4].

3. Writing Test: Educators also developed this test specifically for this study in order to determine the stage of the process of writing acquisition and it is divided into two subtests. *Writing name:* the assessor asks the child to write his name on a sheet of paper. One point is assigned if the child writes his first name only and two points if he writes his full name or at least one surname. *Writing words:* the assessor dictates a set of four words, and the child is supposed to write them on a sheet. An expert corrected this subtest, and the score depended on the level of the children's writing acquisition[5,6].

4. Expressive Vocabulary Test: Measures the number of words a child can retrieve and is also related to the acquisition of reading and writing skills. The task requires children to orally name a given picture shown by the assessor (CAPOVILLA; NEGRÃO; DAMÁZIO, 2011).

5. Pre-Literacy Knowledge and Skills Test (THCP): three of the five subtests in the measure were applied (SILVA; FLORES-MENDOZA; SANTOS, 2013). *Language:* assesses children's ability to understand sentences, to identify the function of objects, to categorize elements, to recognize sounds, to segment words into syllables, and to perceive rhymes, among others. *Memory:* the items assess visual and auditory memory. The first type includes the memorization of the correct positions of stimuli after a brief interval of time. The items assessing auditory memory consider the

[4] Note that, even though the test was developed specifically for this study, it follows the same rationale as other instruments with the same purpose. The authors Fuchs et al. (2001) and Blachman et al. (1999), already mentioned in section 1.1, applied the word recognition subtest of *Woodcock Reading Mastery Test-Revised*, which consists of asking children to read words from a list aloud, assigning one point for each word spoken correctly. In the present case, the test format, which asks children to check the words, is focused on the visual recognition of words, along the lines of the National Literacy Assessment, which verifies the children's ability to read words and locate information that is explicitly provided in a text, for instance, using multiple-choice questions.

[5] There were eight possible levels: Pre-syllabic 1, 2, and 3; Syllabic 1, 2, and 3; Alphabetical Syllabic; and Alphabetical. Figure A1.1 in the Appendix A1 presents a description of each level.

[6] Similar to the Reading Test, the Writing Test is also in line with other instruments used to assess these skills. For instance, in one of the measures used in the study by Aram (2006), the assessor asked the child to write his name and two pairs of words. Blachman et al. (1999), in turn, used level 1 of the writing subtest from *The Wide Range Achievement Test-Revised* to analyze the students' ability to write dictated words. The items of the written production of the National Literacy Assessment are another example, in which students are supposed to write the names of pictures. In all the cases mentioned earlier, the score depends on the children's level of spelling.

ability to listen to a story and memorize the most important information. *Focused Attention:* degree to which a child keeps sustained attention on a task or activity for a given amount of time. This subtest is composed of a series of grouped stimuli, and the child needs to concentrate on finding targets presented throughout the groups.

The parents also completed a questionnaire addressing the children's socioeconomic status, while teachers and principals completed questionnaires addressing the activities they performed at school, as well as their socioeconomic statuses.

Professionals from the education and psychology fields, with some background involving children of the same ages as the ones addressed in this study, individually applied the instruments to children. In some schools, an external specialist applied the reading and writing tests.

1.2.3 Data

As previously noted, we collected data from 88 early childhood education schools. We asked authorization from the parents for the children's participation and 4,389 provided signed consent forms. The proportions of signed forms were very similar between the treatment and control groups. Nonetheless, not all students completed the tests, nor all the parents completed the socioeconomic questionnaire. For a more complete and accurate analysis, we decided to keep only the students who had completed all the three common tests and had provided information on their age in the sample[7]. As a result, there were no observations for two of the schools in the control group. Because randomization was carried out in pairs, we decided to exclude their respective treatment matches. Thus, a sample of 3,259 students remained.

Table 1.1 – Data Description

	Treatment		**Controls**		**Total**	**Differential Attrition**
Samples:						
Total Database	2,239		2,150		4,389	
Test Scores and Age (TS&A)	1,675	74.8%	1,584	73.7%	3,259	0.018
+ Teachers Answers (TS&A+T)	1,373	61.3%	1,319	61.3%	2,692	-0.002
+ Principals Answers (TS&A+P)	1,295	57.8%	1,273	59.2%	2,568	-0.020

Note: Differential Attrition estimated using ordinary least squares, with a set of dummies for school pairs and standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

[7] This is a piece of relevant information because the program may have a different impact on younger and older students, as will be further explored in subsection 1.4.4.

Table 1.1 presents the distribution of students between the treatment and control groups. Even though the number of students who dropped out of the initial sample was proportionally larger in the control group than in the treatment group, no significant differences were found between the groups[8].

Because we intend to make additional analyses using some information from the teachers' and principals' questionnaires, table 1.1 presents information regarding the sample size which is available to estimate these exercises and shows that there is no significant differential attrition.

We also tested for selective attrition between the sample considered in the analyses and the original sample, considering all those who provided information for each measure. Therefore, we performed a regression analysis for each of the measures on a treatment dummy, a variable indicating whether a student is in the final sample or not, and a variable of interaction between these two (which will indicate whether there is selective attrition or not).

Table 1.2 presents the results of this exercise, considering as a specific sample those presented information for the three common tests and age (TS&A), and discards the presence of selective attrition for any of the measures.

Table 1.2 – Test of Selective Attrition - TS&A Sample

	Treat x TS&A	**Treat**	**TS&A**	**N**
PPVT'	-1.986	3.464	8.900***	3,952
Reading a "Rhyme"	-0.128	0.309**	0.226**	3,942
Reading a List	0.084	-0.036	0.027	3,942
Writing Name	0.013	0.021	0.156***	3,950
Writing Words	-0.101	0.541	0.515**	3,941
THCP - Language	-0.201	0.419*	0.494***	1,969
THCP - Memory	0.058	0.037	0.372	1,960
THCP - Attention	-0.183	0.786	2.130***	1,972
THCP - Total	-0.326	1.210	3.101***	1,972
Expressive Vocabulary	-2.493	2.181	3.899***	1,976

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs and standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

We also performed this same exercise for the specific samples of principals (TS&A + P) and teachers (TS&A + T). The results are presented in tables A2.1 and A2.2 in the Appendix A2 and indicate that there may be some unbalance between the treatment and control groups for a few measures.

Finally, we verified the balance between the treatment and control groups in the observable characteristics[9]. First, as randomization was done at the school level, we com-

[8] The probability of receiving the treatment is the same for both samples. This result was obtained by estimating regression between the sample's dummy variable, equal to 1 if the student is in TS&A sample and 0 otherwise, and the treatment variable.

[9] To analyze balance, we first estimated a regression between the treatment variable and a set of dummies

pared the characteristics of the schools using data from 2015 School Census[10]. Table 1.3 shows that, together, the characteristics of the treated schools do not significantly differ from those in the control group, indicating that the schools were properly randomized.

Table 1.3 – Balance Check

		School Sample	**TS&A Sample**	
School Data	Sewer	0.047		0.123
	Teachers' room	0.074		0.030
	Computer Lab	0.110		0.012
	Library	-0.150		-0.094
	Reading Room	-0.042		0.132
	Playgroung	-0.152		-0.202
	Bathroom	0.117		0.115
	Secretary room	0.074		0.073
	Dinning Hall	0.188		0.211*
	Internet	0.006		-0.003
	N. Classrooms (prop.)	1.333		0.207
	N. Adm. Computers (prop.)	6.145		8.230
	N. Students Computers (prop.)	-0.694		0.053
	N. Employees (prop.)	-2.159**		-2.385**
	N. Students	-0.000		-0.000
Student Data	Age: 5 or younger		0.004	-0.018
	Girl		-0.004	-0.025
	Lives up to 3 people		-0.009	-0.060**
	Assessor		0.019	-0.011
	N	84	3,259	3,259
	F Test	0.257	0.997	0.00300

Note: The data source for "School Data" is 2015 School Census. Balance check estimated using ordinary least squares, with standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

We also compared the students' characteristics [11] and again, no significant differences were found between the groups. Finally, we verified balance, considering both the characteristics of the schools and students, and differences were found between the treatment and control groups. For this reason, even if such an unbalance did not indicate a positive bias in favor of the treatment group, we will consider this information in our empirical analysis.

for school pairs, and then estimated a regression between the residue of the first and all the characteristics we wanted to compare in that model.

[10] Several variables are dummies that are equal to 1 if the school has the item in question and 0 otherwise, however, other variables present the total amount of items and, in these cases, we generate variables that represent the proportion of an item in relation to the total number of students, which are indicated by (prop.).

[11] The only variables with information for the entire sample were age, gender, and number of people living in the same household. The variable age is a dummy that is equal to 1 if the child is five years old or younger and 0 otherwise. Because the program is directed to four to five-year olds, at the end of the school year, most were already five to six years old.

1.3 General Results

As the sample was drawn using the experimental method, that is, based on the randomization of the treatment/control assignment, in an ideal scenario, the group of the treated individuals would be statistically equivalent to the group of controls, both in terms of observable and non-observable characteristics. Thus, to assess the potential impact of the program on Reading and Writing skills, we would simply use a regression between the outcomes of interest and a dummy variable indicating whether the individual belongs to the treatment group or not, in addition to including a set of dummies for school pairs (as treatment/control assignment was randomized within pairs).

However, as in our balancing test, using both school and student characteristics, we obtained a statistically significant joint difference, we will also include these contextual variables and will start with the following model:

$$Y_{i,s} = \alpha + \beta Treat_s + \rho' \mathbf{W_s} + \delta' \mathbf{X_{i,s}} + \gamma' \mathbf{Pair_s} + \epsilon_{i,s} \tag{1.1}$$

where $Y_{i,s}$ is the standardized score for one of the assessments of student i from school s; $Treat_s$ is the indicator of treatment; $\mathbf{W_s}$ is the vector of the school's contextual characteristics; $\mathbf{X_{i,s}}$ is the vector of the student's contextual characteristics[12]; and $\mathbf{Pair_s}$ is the set of dummies for school pairs. All the estimations consider standard errors clustered at the school level.

Furthermore, even though the tests used to check selective attrition did not present significant unbalance between the treatment and control groups in our main sample (TS&A), to allow for a more cautious analysis of the results, we used the methodology proposed by (LEE, 2009) to include lower and upper bounds to the effects identified in our general analysis[13].

Finally, as we analyzed the effect of the program on the results measured by several assessments, we will calculate the adjusted p-values for multiple hypothesis testing, as described by Romano e Wolf (2016)[14].

Table 1.4 presents the results for our general model and considering our main sample, indicating positive and statistically significant results for virtually all the outcomes of interest.

For this general analysis, we ran some robustness tests to consolidate the results. Table A3.3 in the Appendix A3 presents the general results, considering the same main

[12] The vectors of the school's and student's contextual characteristics include all variables identified in table 1.3.

[13] Our estimations of Lee bounds do not consider the set of covariates given its large number and the characteristics of some of them.

[14] The results in bold presented in all the tables refer to an adjusted p-value at least <0.1.

Table 1.4 – Impact of the Reading and Writing Program

	ATE	Lee Bounds Lower Bound	Lee Bounds Upper bound
PPVT	**0.087***	0.052	0.142***
Reading a "Rhyme"	**0.207*****	0.066	0.164***
Reading a List	**0.105****	0.059	0.259***
Writing Name	**0.081**	0.049	0.112**
Writing Words	**0.133****	0.079	0.182***
THCP - Language	**0.107***	0.050	0.216**
THCP - Memory	**0.066***	-0.017	0.090
THCP - Attention	**0.160****	0.043	0.145
THCP - Total	**0.165****	0.089	0.207**
Expressive Vocabulary	**0.086***	-0.011	0.104
N (trimming proportion)	3,259	(0.0152)	
N (trimming proportion): THCP Total	1,628	(0.0263)	
N (trimming proportion): Vocabulary	1,624	(0.0214)	

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs, a set of variables for school and students characteristics as indicated at balance check (table 1.3) and standard errors clustered at school level. TS&A Sample. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

sample, however, including only the dummy that characterizes the assessor as a contextual characteristic, and considering a sample (TS&A + C) in which it was possible to consider a larger number of students' contextual characteristics[15]. The results that remain statistically significant in all of the exercises, even when considering the adjusted p-values[16], are the positive effects on the Reading test of a "Rhyme" and Writing Words, with not negligible magnitude, as always close to or above 0.12 standard deviations.

These results are in line with other results reported in the literature. The studies by Neuman (1999) and Blachman et al. (1999) indicate that reading programs were already effective in the acquisition of writing skills, while the study by Aram (2006) shows that children attending a program providing some writing practice performed better on the writing words test than children who only participated in a specific reading program. The studies that used the reading words measure also presented positive effects among the treated children (FUCHS et al., 2001; BLACHMAN et al., 1999), while the results obtained in the receptive language test (PPVT) did not present statistical significance in other studies (NEUMAN, 1999; ARAM, 2006).

[15] Before presenting the results of the robustness tests, we present the results of the selective attrition and balancing for this new sample, which are very close to those obtained for our main sample, in tables A3.1 and A3.2. Unbalance between the treatment and control groups was found only regarding the selective attrition test for the expressive language test, which had not been identified for the main sample.

[16] Even though the lower bounds of Lee were not statistically significant for these cases, both are above zero. Note that these bounds do not consider any covariate.

1.4 Heterogeneous Effects

In addition to assessing the extent to which the Reading and Writing program impacts the development of children's early literacy skills during childhood education, as presented in section 1.3, we also sought to verify the existence of potential heterogeneous effects using other information available.

1.4.1 Intensity of Treatment

The first question we seek to answer is whether the program's effectiveness is related to the intensity of the treatment. How frequently the program's trainer (Department of Education technician) supervises the process, according to the report of the schools' principals, will be used as a measure of intensity.

To address the intensity of the program, in addition to the regular variable indicating the treatment, we included a new dummy variable in the model presented in equation 1.1 that is equal to 1 if the treatment is intensive, that is, if the principal reports that the supervision takes one or more times a month, and 0 if supervision happened less often or if the school was not receiving the intervention. Thus, this coefficient will provide a measure of the difference between intensive treatment and regular treatment.

Table 1.5 – Heterogeneous Results - Intensity of Treatment

	Intensive	**Regular**
PPVT	**0.264*** **	0.004
Reading a "Rhyme"	0.017	**0.238*** **
Reading a List	0.160	0.066
Writing Name	0.105	0.120
Writing Words	**0.365** **	0.050
THCP - Language	0.231**	0.089
THCP - Memory	0.176*	0.047
THCP - Attention	-0.139	**0.271** **
THCP - Total	-0.006	**0.245** **
Expressive Vocabulary	0.134	0.101
N	2,568	
N: THCP Total	1,280	
N: Vocabulary	1,281	

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs, a set of variables for school and students characteristics as indicated at balance check (table 1.3) and standard errors clustered at school level. TS&A+P Sample. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

Table 1.5 shows that the students from treated schools which were more frequently supervised by local trainers performed better, especially on the PPVT and the Writing

Words test. Evidence also indicates that the program itself makes a difference in the reading results, while the intensity of the program is relevant for the development of writing skills. We believe these results correspond to the reality of this educational stage. Literacy in this stage is the subject of extensive debate and controversy. If, on the one hand, reading tasks (especially when it comes to reading to children) are already naturally encouraged and possibly implemented by schools in general, the implementation of writing tasks, however, meets greater resistance.

Thus, it is reasonable to assume that the program's regular implementation already works to improve the reading interventions carried out with children. Writing practices, however, would demand a change in the culture of teaching and greater commitment of all the actors involved in preschool education.

1.4.2 School Infrastructure

Considering that the intensity of the treatment might positively influence some of the students' results, we would like to test whether the schools' infrastructure would also be a conditioning factor to obtain the program's effects.

To access the contribution of infrastructure, we classified the schools into high infrastructure and low infrastructure based on the information provided by 2015 School Census[17], and estimated the following model:

$$Y_{i,s} = \alpha + \beta Treat_s + \phi Treat_s \cdot HInfra_s + \rho HInfra_s + \delta' \mathbf{X_{i,s}} + \gamma' \mathbf{Pair_s} + \epsilon_{i,s} \quad (1.2)$$

Unlike equation 1.1, we did not include the vector $\mathbf{W_s}$ of the school's contextual characteristics, because dummy $HInfra_s$, which identifies high infrastructure schools, was developed based on these characteristics. The results of coefficiente ϕ will indicate whether the effects of the treatment differ or not according to the school's infrastructure.

Evidence presented in table 1.6 indicates that, in general, considering the adjusted p-value, school's infrastructure does not influence the program's results. Nonetheless, the coefficients suggest that students from high infrastructure schools tend to perform better in writing tests. At the same time, no association is found for reading tests, similar to the

[17] To identify a school's infrastructure, we performed a factor analysis using the characteristics of the school identified in table 1.3 keeping only one main factor, but not all the variables were positively associated with the factor. Therefore, we performed a new analysis keeping the variables Playground, Reading Room, Bathroom, Dining Hall, N. Classrooms (prop.), N. Adm. Computers (prop.), N. Students Computers (prop.), N. Employees (prop.) in the estimation of the factor. With all this information positively associated with the factor, the higher the value, the larger a school's infrastructure. Hence, we built a dummy variable that is equal to 1 for schools with results above the factor's median factor (considered of high infrastructure) and 0 otherwise, with an equal number of treatment and control schools at the two levels.

Table 1.6 – Heterogeneous Results - School Infrastructure

	High Infrastructure	
	Interaction	**Treat**
PPVT	**0.329****	-0.069
Reading a "Rhyme"	-0.042	0.141
Reading a List	-0.040	0.099
Writing Name	0.351*	-0.097
Writing Words	0.178	0.082
THCP - Language	0.272*	-0.017
THCP - Memory	0.089	0.008
THCP - Attention	0.432**	-0.101
THCP - Total	0.410**	-0.079
Expressive Vocabulary	0.191	-0.087
N	3,259	
N: THCP Total	1,628	
N: Vocabulary	1,624	

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs, a set of variables for students characteristics as indicated at balance check (table 1.3) and standard errors clustered at school level. TS&A Sample. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

results observed for the intensity of treatment, reinforcing the assumption that writing would demand greater effort and more resources, at least in this educational stage.

1.4.3 Principals and Teachers

Another hypothesis is that the experience of principals and teachers could affect the treatment. At first, we assumed that less experienced professionals would be more open to new experiences, engaging further in training and implementation, which would enhance the effects on the students' results.

To verify this possibility, we used the principals and teachers' answers regarding how long they had been serving as principal or teacher, respectively, and generate dummies equal to 1 if professionals had up to 10 years of experience or 0 if they had more than 10 years. In the case of teachers, we also analyzed their educational level, using a variable that identifies those who held a bachelor's degree in opposition to those who also attended some graduate studies.

We estimated a model very close to equation 1.2, including the interaction between the treatment and the variable of experience(education) and the variable of experience(education) itself, but now keeping the vector of the schools' contextual characteristics, as in equation 1.1.

The results of the interaction coefficients presented in tables 1.7 and 1.8 are not

Table 1.7 – Heterogeneous Results - Principals

	Experience: Less than 10 years	
	Interaction	**Treat**
PPVT	0.369*	-0.033
Reading a "Rhyme"	-0.364*	**0.383*** **
Reading a List	-0.237	0.204*
Writing Name	0.534	-0.072
Writing Words	-0.187	0.167
THCP - Language	0.555	-0.117
THCP - Memory	-0.087	0.143
THCP - Attention	-0.486**	**0.536*** **
THCP - Total	-0.265	**0.426*** **
Expressive Vocabulary	0.436	-0.017
N	2,568	
N: THCP Total	1,280	
N: Vocabulary	1,281	

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs, a set of variables for school and students characteristics as indicated at balance check (table 1.3) and standard errors clustered at school level. TS&A +P Sample. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

statistically significant, considering the adjusted p-value. This suggests that the professionals' experience and education, as the case may be, do not affect the magnitude of the program's effects.

Table 1.8 – Heterogeneous Results - Teachers

	Experience: Less than 10 years		**Schooling: Less than Postgraduate**	
	Interaction	**Treat**	**Interaction**	**Treat**
PPVT	0.226**	0.000	-0.023	0.100
Reading a "Rhyme"	0.034	**0.171** **	-0.006	**0.187** **
Reading a List	0.124	0.048	-0.021	0.105
Writing Name	0.216	-0.000	0.058	0.064
Writing Words	0.141	0.096	0.224	0.060
THCP - Language	0.034	0.089	0.083	0.067
THCP - Memory	0.107	-0.007	0.166	-0.030
THCP - Attention	0.133	0.032	0.121	0.031
THCP - Total	0.134	0.046	0.162	0.031
Expressive Vocabulary	0.115	-0.002	0.093	0.006
N	2,692		2,692	
N: THCP Total	1,341		1,341	
N: Vocabulary	1,345		1,345	

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs, a set of variables for school and students characteristics as indicated at balance check (table 1.3) and standard errors clustered at school level. TS&A+T Sample. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

In the case of the teachers, the signs of these coefficients show a tendency that goes in the same direction as our hypothesis that less experienced teachers could positively influence the students' results. In the case of the principals, however, the tendency is not

very clear. This difference may be explained by the fact that principals also work as educators within the program's methodology, a role that may be positively affected by greater experience.

1.4.4 Students Age

Finally, we would like to test whether the program's effect would be different depending on the children's ages. As previously mentioned, the age of the program's targets ranged from four to five and 11 months old, that is, at the end of the first year of implementation, most of the students were between five and six years old, which is considered in the analysis of potential heterogeneous effects.

Thus, a model was estimated according to the terms described in subsection 1.4.3, including an interaction between the treatment and a dummy that identifies this increase in age (equal to 1 for students aged five years old or younger) and an individual dummy of age.

According to the results presented in table 1.9, there would be no significant differences between the results of younger and older children participating in the program. These children, with very similar ages, are likely in a very similar stage of development. The interaction coefficient for almost all the assessments, however, indicates a positive tendency in favor of the younger children, although some people consider that it is premature to implement certain practices of the program at this educational stage.

Table 1.9 – Heterogeneous Results - Students Age

	Age $<=$ 5 years old	
	Interaction	**Treat**
PPVT	0.129	0.009
Reading a "Rhyme"	0.005	**0.204*** **
Reading a List	0.039	0.081
Writing Name	0.046	0.053
Writing Words	0.084	0.082
THCP - Language	0.064	0.068
THCP - Memory	-0.156	**0.162** **
THCP - Attention	-0.084	**0.211** **
THCP - Total	-0.106	**0.230** **
Expressive Vocabulary	0.048	0.057
N	3,259	
N: THCP Total	1,628	
N: Vocabulary	1,624	

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs, a set of variables for school and students characteristics as indicated at balance check (table 1.3) and standard errors clustered at school level. TS&A Sample. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

In the literature analyzing the implementation of early literacy programs in childhood education, only the study by Aram (2006) assessed heterogeneous effects by age but does not report differences in the results, except in the case of the receptive language test, which was also our coefficient with the highest magnitude, though not statistically significant.

1.5 Discussion

Evidence found in the literature reveals many significant differences in the acquisition of cognitive skills among vulnerable children, which can be verified in the early years of education (HART; RISLEY, 1995; COLEY, 2002) and usually persist into adulthood (HECKMAN, 2008). Additionally, studies have shown that interventions implemented during early childhood are important mechanisms to decrease these inequalities (CUNHA et al., 2006). In this sense, this study contributes to this body of literature by investigating the possibility of a reading and writing program implemented in public preschools to impact some early literacy measures.

Our findings show that the children from the schools included in the treatment performed better, especially in the reading and writing measures, which is in line with other studies addressing literacy programs for childhood education (ARAM, 2006; BAUSERMAN et al., 2005; NEUMAN, 1999; WASIK; BOND, 2001; FUCHS et al., 2001; BLACHMAN et al., 1999). The difference of this study in comparison to other studies lies mainly in the pedagogical content and methodology of the programs analyzed. Most of the childhood education programs adopt reading tasks because the development of writing skills in this stage is controversial and seldom disseminated. Additionally, even though one of these studies has already addressed a writing-focused program (ARAM, 2006), the author reports a specific focus on the letters of the alphabet and name writing. The methodology used to implement the programs is based on training the teachers to implement a specific number of very well delimited tasks throughout the intervention. We believe that the methodology of our program is an important differential because the training dynamics seek to establish a collaborative network, promoting a sustainable environment in the medium and long term, so that the work is not interrupted after the schools' participation in the program ceases. The purpose is to involve people and promote a change in the culture of teaching practices in the entire network, so that teachers change their daily practices concerning reading and writing in the way and to the extent they consider appropriate.

This study is also expected to contribute to the implementation of evidence-based public policy. As previously mentioned, the scores obtained by children attending Brazilian public schools in literacy assessments are already poor and unequal in the 3rd grade,

despite this, new provisions were proposed by the Brazilian Common Curricular Base, approved in December 2017, establishing the focus on literacy in the first two years of the elementary education (children are now assessed in the 2nd grade, rather than in the 3r grade). Studies show that reading and phonological awareness measures assessed during preschool are important predictors of reading levels throughout elementary education (DUNCAN et al., 2007; STORCH; WHITEHURST, 2002; COSTA et al., 2013). Other studies show similar results for writing words, assessed during early childhood education (ARAM; LEVIN, 2004; ARAM, 2005). Therefore, identifying the type of program that changes the level of reading and writing skills during preschool may support the decisions of policymakers in order to obtain improved results from the children's literacy process, promoting equal opportunities for them to develop skills.

This study's results also show the importance of looking at the development of writing skills at this educational phase. Evidence indicates that the effects of interventions on writing measures may be even more significant depending on the intensity of the treatment and the school's infrastructure. In this sense, writing skills possibly depend on greater effort and more resources, which would require an actual change in the teaching culture, demanding a debate and the involvement of all those involved.

Finally, one of this study's limitations is the fact that we would like to investigate whether the program's methodology would really be sustainable in the long run and whether the effects of the intervention would continue throughout the results of children in the first years of elementary education. We do not have access to data nor have the resources that would allow such an analysis though.

2 Social networks, peer effects and socioemotional skills

2.1 Introduction

Peer effects and socioemotional skills are two subjects that have recently drawn special attention from the literature on economics of education. Various studies have attempted to identify the impact of these characteristics on the academic performance of students[1], however, few studies incorporate these two factors and, as far as we know, none of them have explored all the nuances we propose here.

Our main goal is to investigate how the influence of peers and socioemotional skills are related. We propose to answer this question through three exercises. We will use a database composed of elementary students attending municipal schools located in the city of Recife (Brazil). This database contains information that is relevant for the development of measures of socioemotional competencies, and especially, to identify the students' friendship networks[2].

In the first exercise, we will analyze how the socioemotional skills of peers affect individual outcomes, from a contextual perspective, on academic performance, and from an endogenous perspective, on the students' socioemotional skills. We will perform a peer-effect model that is similar to the empirical model most commonly used in this area (the "linear in means" model), but which addresses two of its major inconveniences, the problem of correlate effects and identification problem. Evidence indicates negative and significant endogenous effects, both in the case of socioemotional measures and cognitive measures[3], and, in some cases, positive and significant contextual peer-effects of socioemotional measures.

As far as we know, the literature is still very incipient in terms of peer-effects

1 In terms of socioemotional competencies, the literature has shown that both cognitive and socioemotional skills are equally important for an individual's development (CUNHA; HECKMAN, 2008). More specifically, Heckman, Stixrud e Urzua (2006), Borghans et al. (2008) and Almlund et al. (2011) show that socioemotional skills are as important as cognitive skills to explain not only educational outcomes but various dimensions of individuals' lives such as wage, employability and risk behavior. There is also abundant literature indicating the existence of peer effects on academic outcomes in the most varied educational phases, from elementary school to higher education (SACERDOTE, 2001; HANUSHEK et al., 2003; HOXBY, 2000; VIGDOR; NECHYBA, 2007; LAVY; PASERMAN; SCHLOSSER, 2012).

2 Peer-effect models generally assume that individuals interact in groups, that is, an individual is affected by all the individuals in his group and by no one outside it. In the context of friendship networks, potential interpersonal relationships would be formalized as direct links between agents, and the collection of all these links is called social network (DZEMSKI, 2014), which enables the existence of heterogeneous reference groups among individuals in the same network.

3 These measures are well defined in section 2.2.

of non-cognitive competencies on individual outcomes. Using Brazilian data and assuming that individuals interact in groups, Chikitani (2015) investigated the existence of endogenous peer-effects on the students' measure of Locus of Control and found negative effects, though not significant. In another study using Brazilian data Gonçalves e Raposo (2016) explores the structure of networking and two questions related to the students' self-perception[4] and observed positive contextual peer-effects of self-perception on math test scores[5], but no endogenous peer-effects for these questions. Finally, addressing data from Hong Kong, Chan e Lam (2014) used different types of friendship networks[6] to analyze peer-effects on academic outcomes and on the grade assigned by teachers to behavior. Even though they considered five factors of the Big Five inventory to analyze potential contextual effects, some significant effects are found only for conscientiousness [7]. Regarding endogenous effects, if we take into account the outcome variable closest to the socioemotional measure (grade assigned to behavior), the authors found significant (positive) spillovers only when considering the studymate network.

In the context of the peer-effect model used so far, we would observe the effects of variation on the mean of the characteristics and outcomes of the reference group of a student on his individual outcomes. It turns out that this variation may change the reference network itself, as these characteristics may be associated with the individuals' decisions to establish links.

Therefore, to more broadly identify peer-effects, it would be interesting to observe peer-effects via the direct effect of characteristics and behavior of the social network on individuals' behavior, but also, via effect of characteristics and behavior of the social network on itself.

Therefore, we will start the second round of exercises using a social network formation model such as the one designed by Goldsmith-Pinkham e Imbens (2013), to identify which characteristics most frequently influence the process of ties formation among individuals[8], that is, what increases the probability of an individual to establish interdependent relationships with another? Would it be similar socioemotional characteristics? Similar cognitive skills? Or having common friends? Evidence suggests that, between cognitive and socioemotional skills, the closer the socioemotional measures between two

[4] These are not measures of socioemotional skills, only two questions were intended to capture the perceptions of students regarding non-cognitive issues (one asks whether the students would change anything about their personality and the other asks whether the students felt excluded in the classroom).

[5] Even though the model they analyze does not incorporate the possibility of endogenous peer-effects on individual math grades.

[6] They estimated distinct effects for the peers' networks identified by the students as friends, studymates and seatmates.

[7] Contextual effect of conscientiousness on mathematics and behavioral grade, if considered the friends network, and on English grade considering both friends and studymate networks.

[8] Basically, we want to study homophily behavior (the tendency of people to relate with those who are similar to themselves) in an evolving network.

students, the greater the probability they will become friends.

The literature shows only one study that considers socioemotional competencies in a model of network formation. Although it does not provide a clear comparison between the role of cognitive and socioemotional skills in the formation of links[9], and it does not consider covariates that had been already shown to be relevant in the formation of links, such as prior existence of common links and/or friends, Chan e Lam (2014) also find that similarities in socioemotional measures, specifically conscientiousness and extroversion, influence the establishment of connections among students, when the friends network is considered.

Finally, our last objective will be to use an empirical model that will allow us to observe the possibility of heterogeneous effects. Within a social network there are heterogeneities in the friendship relations and in the complementarities that an individual can obtain in his interpersonal relationships, which also depend on the friendship established with each individual in his social network, that is, the influence that a group exerts on an individual may be particularly related to the position of this agent within the group and, therefore, vary according to this position.

Therefore, to assess the hypothesis of heterogeneous effects, we will use a model that explores the role of the student's position on the social network on his educational outcomes, using network centrality measures[10].

As far as we know, no studies investigating the impact of centrality on educational outcomes consider socioemotional skills in its specification, whether at the individual or peer level. Hahn et al. (2015) acknowledge the relevance of these skills for the model and implement exercises to show that some socioemotional measures have a positive and significant impact on centrality measures, which would corroborate the hypothesis that the effects of centrality measures would actually capture some non-cognitive measure. The fact is that even controlling for the students' and their peers' socioemotional measures, we continue to find positive and significant effects for the centrality measures, suggesting that these effects may not only be capturing the impact of socioemotional skills but also, the impact of complementarities obtained by the subjects in their friendships.

In addition to the three exercises previously mentioned, we also simulated 500 samples to investigate how peer-effects would behave if we observed social networks with links established according to the model of network formation. For the endogenous and contextual peer-effects model (first exercise) we still find evidence of negative endogenous effects for virtually all the measures of analyzed results [11], however, we found only contextual peer effects of the outcome variable itself, observed in the previous period. On

[9] Only one cognitive variable resulting from a logic test, is included in the model.
[10] We will use four measures traditionally adopted in the literature, which are defined in section 2.2.
[11] Only for one of the socioemotional measures, the effect is no longer significant.

the other hand, when we consider the model of heterogeneous effects (third exercise), the effects of individual centrality measures not only remain, but also often are even more expressive, in terms of magnitude, in relation to the results observed for the original sample. In general, the simulated results indicate the relevance of considering the mechanisms of network formation, when studying the potential influence resulting from interactions among individuals.

In summary, we understand that our study can contribute to the advancement of literature, first because it is the only study addressing the relationship between socioemotional skills and the influence of peers not only for a general peer effects model but it also expands the investigation to a peer formation model and a model that considers the complementarities an individual may obtain from his interpersonal relationships, and also because, in the case of each exercise, we explored other nuances, whether in the use of friendship networks, the use of variables that measure socioemotional skills *per se*, or in the very specification of the models.

Additionally, the results found here for the relationship between socioemotional skills and peer effects can contribute to evidence guiding public policies, for instance, supporting a more effective allocation of students, and even, supporting the work of teachers with groups of students in the same class.

2.2 Data and Measures

2.2.1 Data Set

The database used in this study is the result of an assessment of the *LEGO*® Education Program[12]. Data were collected from 3rd to 5th grade students of 30 municipal schools located in the city of Recife (State of Pernambuco, Brazil). The schools were randomly drawn from a universe of 120 schools.

Data were collected in two stages, a "baseline", which took place in July 2014, and a "follow-up", which took place in December 2014.

In both stages, 3rd to 5th grade students answered to a cognitive skills test and to a socioemotional skills test. They also completed a small questionnaire to identify their friendship network within their respective classrooms (essential for this study's objectives)[13].

In addition to the socioemotional, cognitive, and friendship network measures, the

[12] In summary, the program consists of carrying out practical activities, using *LEGO*® blocks to aid students understand a specific topic in the course curriculum.

[13] A more detailed description of these tests and questionnaires are presented in the next sections.

database provides information such as the students' birth dates (age[14]) and gender[15]. Additionally, we gathered data from two proficiency measures that resulted from standardized tests in Portuguese Language and Mathematics from the Pernambuco Educational Assessment System (*Sistema de Avaliação Educacional de Pernambuco* - SAEPE), conducted two weeks before the follow-up test were applied to assess *LEGO*® Education.

A total of 3,252 students who provided answers to the cognitive or socioemotional test in one of the two data collection stages were identified in the raw database. Of these, 2,720 students are in the follow-up and 2,419 are in the baseline (1,887 of these students are in both stages). After we generate the cognitive and socioemotional measures (as described in a later subsection), we kept in the database only the students who had (i) these three measures, (ii) information on age and gender, (iii) proficiency measures on Language and Mathematics, and (iv) who participated in both data collection stages, which resulted in a sample of 1,393 students. Finally, because of the peer-effects exercises, isolated students were excluded from the sample[16], according to data of the friendship network identified in the follow-up, thus, a final database (considering a balanced panel) composed of 1,260 students remained.

2.2.2 Socioemotional and Cognitive Measures

2.2.2.1 Exploratory Factor Analysis

The students' socioemotional skills were measured using the under-construction version of the Social and Emotional or Non-cognitive Nationwide Assessment - SENNA (SANTOS; PRIMI, 2014), developed by the Ayrton Senna Institute.

SENNA was developed as a single optimized instrument based on the selection of items from existing instruments, so that its items would represent the five factors of the Big Five[17] and a sixth additional factor[18]:

1. **Openness to new experiences**: tendency to be open to new aesthetic, cultural and intellectual experiences. Individuals with this characteristic tend to be imaginative and curious.

[14] Based on the students' birth dates, we first generate an age variable in days, which was divided by 365 to find a variable in years to be used in our exercises.
[15] To use the information on gender, we created a dummy variable that would be 1 if the student was a boy and 0 if a girl.
[16] That is, considering the definition of social network provided in subsection 2.2.3, we excluded those who nominated no one or were not nominated by no one, in line with other studies that follow this same procedure, see Goldsmith-Pinkham e Imbens (2013), for instance.
[17] Using a statistical filter, the Big Five theorists found five factors that explain most of the variation of an extensive range of personality measures(SANTOS, 2011).
[18] Examples of the items for each of the six factors that compose the test are presented in the Appendix B1.

2. **Conscientiousness**: tendency to be determined, organized, committed, and responsible. Individuals with this characteristic tend to be focused, control impulses, and postpone rewards.

3. **Extroversion**: interests and energy are driven to the external world, people and things (as opposed to the inner world of subjective experiences). Individuals with this characteristic tend to be communicative and sociable.

4. **Agreeableness**: tendency to be cooperative and generous. Individuals with this characteristic tend to have good will and empathy.

5. **Emotional Stability**: predictability and consistent emotional responses, without sudden changes in mood. Individuals with this characteristic tend to be calmer and less vulnerable to stress and anger. The reverse of emotional stability is called neuroticism.

6. **Locus of control**: it reflects the extent to which individuals attribute current experiences to past behavior and decisions (internal locus), or chance, luck or third parts' actions and decisions (external locus).

The version of the questionnaire that was applied was composed of 83 items rated on a five-point Likert scale. Using an exploratory factor analysis, we verified whether the items were highly correlated to the factor they were supposed to summarize. It is important to note that some items are designed to capture the factors' positive tendency while others to capture a negative tendency[19]. Anyway, before proceeding with the factor analysis, we reverted the negative items so all factors would present a positive interpretation.

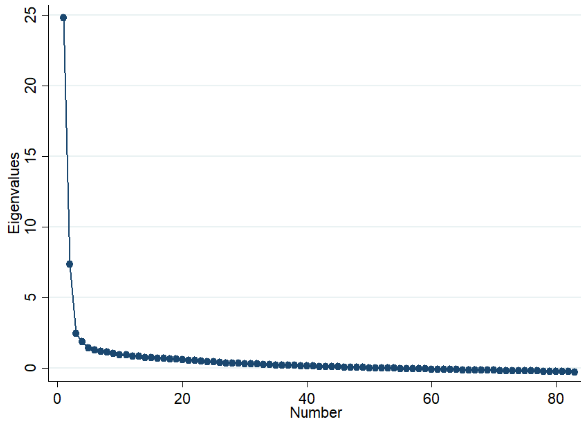
The scree plot 2.1, which was obtained after performing a factor analysis with the items from SENNA[20], shows that the point of inflection occurs in the third eigenvalue, thus the scree criterion proposed by Cattell (1966) indicates that the number of factors correspond to the number of eigenvalues above this inflection point, that is, we were able to extract two factors and not six, as previously expected.

For further estimation of these factors, we need to verify which items represent each of them. Through a common factor analysis, we estimated the factor loadings of each item, which are reported in the table B2.1 of the Appendix B2. As each SENNA's item is related to one of the six domains previously described, we observed the factor loadings

[19] For instance, higher scores assigned in the Likert scale for the item "I'm calm and control my stress well" indicate greater emotional stability (positive pole), while higher values assigned in the Likert scale for the item "I tend to lose patience" indicate lower emotional stability or higher neuroticism (negative pole).

[20] All the exploratory analysis' results in the paper consider baseline data for the students, who in this stage answered all the items presented in the SENNA instrument, in the case of the socioemotional analysis, and answered all the cognitive tests, in the case of the cognitive analysis presented in the sequence. Analyzes considering follow-up data were also performed and the results do not change.

Figure 2.1 – Scree Plot of Eigenvalues - Sociemotional Analysis



and identified which domains have greater weight in the factor we extract, which led us to call them “Reliability&Curiosity” and “Self-Management”.

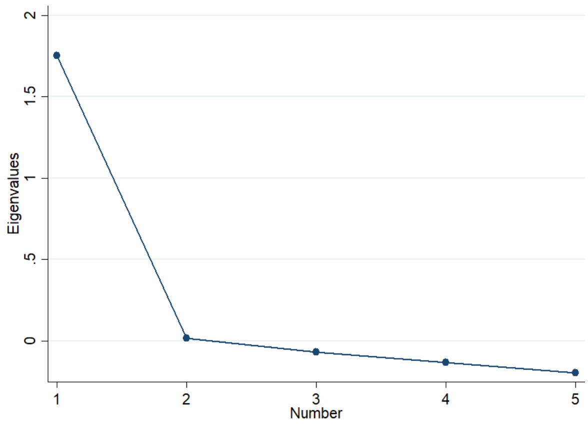
Considering the cognitive skills, in addition to specific measures of proficiency in Language and Mathematics, we have in the database measures derived from the five tests[21] of the cognitive skills assessment developed by Nakano et al. (2015) which were applied to the students, namely:

1. **Verbal Reasoning Test**: extension and depth of verbal vocabulary knowledge and ability to reason using previously learned concepts.
2. **Abstract Reasoning Test**: ability to reason in new situations, create concepts and understand implications.
3. **Spatial Reasoning Test**: ability to represent and manipulate mental images.
4. **Numerical Reasoning Test**: understanding of basic quantitative concepts, such as addition, subtraction, multiplication, and division and manipulation of numerical symbols.
5. **Logical Reasoning Test**: problems involving logical reasoning that are closed to the children’s daily life.

Each of these tests contain nine questions and result in a score of 9 points, based on the sum of the correct answers to the items of each question, or as we designed, a score that represents the proportion of correct answers. Even though in this case we have a smaller number of measures, we also performed an exploratory factor analysis with these five scores and verified that it is possible to summarize this information on a single factor, according to the scree plot 2.2, which we call “Cognitive Factor”.

[21] Examples of items of each of the five tests are found in Appendix B1.

Figure 2.2 – Scree Plot of Eigenvalues - Cognitive Analysis



2.2.2.2 Confirmatory Factor Analysis

After determining the number of factors that we can extract from the SENNA's items and from the cognitive test scores, we need to estimate the scores of factors that will represent these measures. For this analysis, we will use a factor model along the lines of the one used by Attanasio et al. (2020), assuming the following equation system:

$$\begin{aligned} m^{C}_{j,t} &= \mu^{C}_{j,t} + \beta^{C}_{j,t} lnC_t + \epsilon^{C}_{j,t} \\ m^{S1}_{j,t} &= \mu^{S1}_{j,t} + \beta^{S1}_{j,t} lnS1_t + \epsilon^{S1}_{j,t} \\ m^{S2}_{j,t} &= \mu^{S2}_{j,t} + \beta^{S2}_{j,t} lnS2_t + \epsilon^{S2}_{j,t} \end{aligned}$$

where $m^{i}_{j,t}$ is the jth measure related to factor i={C,S1,S2}, in time t={baseline, follow-up}; $\mu^{i}_{j,t}$ are the measures' expected values; $\beta^{i}_{j,t}$ are the factor loadings; and C_t, $S1_t$, $S2_t$ are the latent factors we want to estimate. The model assumes that errors $\epsilon^{i}_{j,t}$ are independent of latent factors, have a mean zero, and do not correlate over time.

Still, according to Attanasio et al. (2020), to define the scale, we established that the factor loading of the first measure, associated with each factor, would be equal to 1, that is $\beta^{C}_{1,t} = \beta^{S1}_{1,t} = \beta^{S2}_{1,t} = 1$ and normalized the factor loadings in the same measures both in the baseline and follow-up. Additionally, we know that to identify the model, we need at least three measures with a factor loading different from zero for each latent factor, which in our case would not be a problem, considering that the factor with the lowest number of measures (cognitive) has five measures.

Based on the factor model previously defined, we estimated the factors in two stages. In the first stage, we estimated the factor loadings, intercepts, and variance and covariance matrix through a minimum distance estimator. In the second stage, we used

the structures estimated in the first round to estimate the scores of the factors through the Barlett's method[22].

2.2.3 Networking Measures

The identification of the students' friendship network is extremely important for the implementation of the models we intend to estimate and the database of the *LEGO*® Education Program provides the detailed information that makes this objective possible.

During the two stages of collecting data from the Program, the students answered a small questionnaire in which they should choose (from a list provided earlier) the names of students with whom they played, studied, or talked with. The list contained the names of all the students enrolled in the same classroom. The students could choose four students at most for each activity. For the restricted number of names to be chosen in each activity did not prevent the students from representing their real friendship network, we generate, for this study's purposes, a network that included all the classmates chosen by a given student in the three activities. It implies that one student could choose up to 12 different students but the maximum number chosen was 8 students.

Based on this information, we identified the students' friendship networks and represented them in a matrix form. A social network is a social structure composed of nodes (in this case the students), who are interconnected by links, represented here by a friendship relation, reported by the students at the time of the survey.

Be R the number of networks[23], indexed by r, the relationships between students are captured by a network $\mathbf{g_r}$, with $g_{r,ij} = 1$ if "i" and "j" are friends, and $g_{r,ij} = 0$ otherwise. Since a friendship is a reciprocal relationship, we consider that there is a link between the node "i" and node "j" if "i" chose "j" or if "j" chose "i", that is, $g_{r,ij} = g_{r,ji}$ and $g_{r,ii} = 0$. Additionally, to each network $\mathbf{g_r}$ we associated an adjacent matrix $\mathbf{G_r} = [g_{r,ij}]$, which stores the relationships of $\mathbf{g_r}$. As we defined, matrix $\mathbf{G_r}$ is symmetrical and the elements of its diagonal are equal to zero.

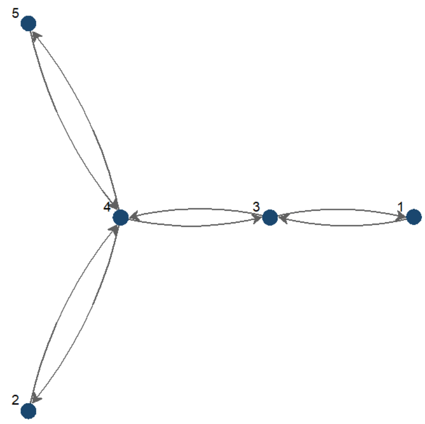
We can also design a matrix of weights for the network, which represents the proportion of the influence of each of the friends of a given student. Be $\mathbf{W_r}$ a matrix that contains elements $w_{r,ij} = \frac{g_{r,ij}}{g_{r,i}}$, with $g_{r,i} = \sum_{j=1}^{n_r} g_{r,ij}$, the number of links of "i".

Below we present a graphic and matrix example for a friendship network in our sample with five students.

[22] It obtains the scores of factors regressing the measures on the factor loadings and using the variance and covariance matrix of error as weight.

[23] In this case, R is the number of classrooms included in our sample, considering that the report of friendships was restricted to the list of students enrolled in the same class.

Figure 2.3 – Network Example



$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{W} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

The information of figure 2.3 and of the matrices complement each other and make it clear how the social network structure works according to our definition. Note the first line of matrix **G**, it shows that student 1 is friend of student 3. In the figure, an arrow connects the student 1' node to student 3' node[24]. As student 1 has only one friend, this friend exerts 100% peer-effect on student 1, as verified in the first line of matrix **W**. This example also shows a very relevant piece of information for the study's scope, that the centrality of students may present some variation within the same network, which will be explored in the analysis of heterogeneous effects. Thus, we need to define the centrality measures that will be used[25]:

1. ***Closeness Centrality***: measures how close a node is to other nodes in a given network $\mathbf{g_r}$.
   $Closeness_i(\mathbf{g_r}) = \frac{1}{\sum_{k \neq i} \Delta_{ik}} * (n_r - 1)$
   In which Δ_{ik} is the length of the shortest path between node "i" and node "k" (the length between two nodes directly linked is 1).

[24] Note the arrows are bidirectional, for instance, an arrow goes from 1 towards 3 and an arrow goes from 3 towards 1. This structure will be verified for all the students who are friends, in all the networks, as we consider friendship to be a reciprocal relationship.

[25] There are various centrality measures and we decided to use these four because they are the most frequently used and also because other studies adopted these same measures, which will facilitate making some comparisons.

2. ***Betweenness centrality***: in a given node, this measure is equal to the number of the shortest paths between each pair of nodes in the network passing through this node, that is, for a given network $\mathbf{g_r}$:
$Betweeness_i(\mathbf{g_r}) = \sum_{s \neq i} \sum_{t \neq i} \frac{\sigma_{st(i)}}{\sigma_{st}}$
In which σ_{st} is the total number of the shortest paths between node "s" and node "t" and $\sigma_{st(i)}$ is the number of these paths that passes through node "i".

3. ***Degree Centrality***: Measures the number of links of a given node.
$Degree_i(\mathbf{g_r}) = \frac{1}{(n_r-1)} * \sum_{j=1}^{n_r} g_{r,ij}$

4. ***Katz-Bonacich Centrality***: measures the importance of a given node within a social network. To assess how well a node is located, we use the weighted sum of the paths departing from this node. A certain value $\beta g_{r,i}$ is assigned to each node, a value that is proportional to its connectivity $g_{r,i} = \sum_{j=1}^{n_r} g_{r,ij}$, this value being increased with the value of the node located at a distance link of "i", two distance links, and so on, discounted by a factor that decreases as the distance increases, that is, the value of the node located "s" distance links "i" is weighted by β^{s-1}. Be $\mathbf{1}$ a vector of ones, the vector of *Katz-Bonacich* centrality can be defined as:

$$\mathbf{b}(\mathbf{g_r}, \beta) = \beta \mathbf{G_r} \mathbf{1} + \beta^2 \mathbf{G_r}^2 \mathbf{1} + \beta^3 \mathbf{G_r}^3 \mathbf{1} + ... = \sum_{s=0}^{\infty} \beta^s \mathbf{G_r}^s \cdot (\beta \mathbf{G_r} \mathbf{1})$$

And for β small enough[26], this infinite sum converges to a finite value, which gives us:

$$\mathbf{b}(\mathbf{g_r}, \beta) = (\mathbf{I_r} - \beta \mathbf{G_r})^{-1} \cdot (\beta \mathbf{G_r} \mathbf{1})$$

Given the definitions, if we go back to the network we used as an example, we see that student 3 is the most connected within the network, and in this case, is also the one with the greatest magnitude in all the centrality measures. On the other hand, even though students 1, 2 and 5 have only one friend, it does not mean they share the same results for all the measures, see the case of the Closeness measure, in which the result of student 1 is approximately 0.44, while the results of students 2 and 5 is 0.5.

2.3 Empirical Strategy

This study's first goal is to analyze whether the peers' socioemotional skills influence students, both from a contextual perspective, on academic performance, and from an endogenous perspective, on one's socioemotional skills.

[26] To ensure that $(\mathbf{I_r} - \beta \mathbf{G_r})^{-1}$ to be invertible we need $\beta < \frac{1}{\lambda_{máx}(\mathbf{G_r})}$, where $\lambda_{máx}(\mathbf{G_r})$ is the highest eigenvalue of the network $\mathbf{G_r}$, and a sufficient condition would be $\beta < \frac{1}{n_r - 1}$ (CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009).

In this investigation, we will start with a spatial autoregressive model (SAR), as the one used by Lee (2007):

$$Y_{r,i} = \beta \overline{Y}_{r,j} + \mathbf{X}_{r,i}\gamma + \overline{\mathbf{X}}_{r,j}\delta + \alpha_r + \epsilon_{r,i} \tag{2.1}$$

In the model, $Y_{r,i}$ is an outcome variable of student "i" from network "r", $\overline{Y}_{r,j} = \frac{1}{g_{r,i}} \sum_{j=1}^{n_r} g_{r,ij} Y_{r,j}$ is the mean of the outcome variable for the peers of student "i", $\mathbf{X}_{r,i}$ is a vector of characteristics of the student and of his family, $\overline{\mathbf{X}}_{r,j} = \frac{1}{g_{r,i}} \sum_{j=1}^{n_r} g_{r,ij} \mathbf{X}_{r,j}$ is a vector with the mean of the characteristics of the peers of individual "i" and of their families and α_r represents the non-observable characteristics of network r.

For the sake of convenience, we can represent equation 2.1 in matrix terms as follows:

$$\mathbf{Y_r} = \mathbf{W_r}\mathbf{Y_r}\beta + \mathbf{X_r}\gamma + \mathbf{W_r}\mathbf{X_r}\delta + \text{ł}_{\mathbf{n_r}}\alpha_{\mathbf{r}} + \epsilon_{\mathbf{r}} \tag{2.2}$$

According to the definition provided by Manski (1993), this model would cover three effects: (i) endogenous effect, verified by coefficient β, whereby an individual's behavior (outcome variable) varies with the mean of his peers' behavior; (ii) exogenous or contextual effect, verified by coefficient δ, whereby the behavior of individuals vary according to the mean of the exogenous characteristics of his peers (contextual or predetermined characteristics); and (iii) correlated effects, given that the group selects people with similar contextual characteristics, which cannot be directly observed.

Despite the advantage of simplicity, this model presents some drawbacks in its estimation, which may make it impossible to identify the effects. One of the first issues is that the model could be "contaminated" by a self-selection problem, or, as defined by Manski (1993), correlated effects. It means that the presence of an individual within a group may be due to him having certain characteristics not observed by the econometrist, which can be correlated with the individuals' other exogenous characteristics, which biases the parameters estimation. To deal with this problem, we used the strategy adopted by various studies (CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009; LEE, 2007; BRAMOULLÉ; DJEBBARI; FORTIN, 2009) that is to include fixed effects at the network level (α_r), which empirically means that we included dummies for classrooms (our unit of social network) in the specification.

The second big problem is that, in principle, the model cannot be directly estimated because of a simultaneity problem called by Manski (1993) of reflection problem. Note that our specification is actually a variation of the traditional "linear-in-means" model. In that model, the individuals interact in group (in their classroom, for instance), a case in which the outcome of any individual within the group equally and simultaneously affects the outcome of all the other individuals, not allowing us to identify whether the

behavior of one student in the class is the cause or the effect of the influence of his peers. In our case, we used information from the students' friendship network, the reference group of a student is formed by his direct links of friendship (as defined in subsection 2.2.3), therefore, within the same network, the students may have distinct reference groups, which do not overlap the network itself. It means that we can find intransitive "triads" in a social network[27], which would ensure the identification of peer effects, as shown by Bramoullé, Djebbari e Fortin (2009) and Calvó-Armengol, Patacchini e Zenou (2009). Even though some networks do not have intransitive "triads", Bramoullé, Djebbari e Fortin (2009) and Lee (2007) show that identification is possible in these cases if the groups vary in size.

Therefore, in case we have networks with intransitive "triads" or of different sizes, we can perform the first exercise to assess peer effects on socioemotional skills. Empirically, we will estimate the equation model 2.2 as a pseudo-panel, considering intra and between groups variations, using the maximum likelihood method (CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009; LEE, 2007).

Nonetheless, β and δ would measure the average effect on the outcome variable of a student given a variation in the mean of the outcome variable and mean of the characteristics of this student's reference group, keeping the social network fixed. It turns out that a change in the characteristics and/or the outcome variable of the reference network may change the network itself. Thus, to identify peer effects in a more comprehensive way, it would be interesting to observe peer effects via direct effect of the group's characteristics and behavior on the individual's behavior, but also via effect of the group's characteristics and behavior on the group itself.

Given the definition of social network, we can verify which characteristics are associated with the decision of individuals to establish links through a social network formation model close to that presented by Goldsmith-Pinkham e Imbens (2013).

The model assumes that the decision to establish links is the result of two choices, both the individuals must agree to establish a link, which will happen if the utility of a given link is positive[28]:

$$g_{r,ij} = \mathbf{1}_{U_{r,i}(j)>0} \cdot \mathbf{1}_{U_{r,j}(i)>0} \tag{2.3}$$

$U_{r,i}(j)$ is the utility of individual "i" when establishing a link with individual "j",

[27] In a set of three students ("i", "j", "k"), we may have cases in which "i" has a link with "j", "i" has a link with "k", but "j" does not have a link with "k".

[28] Additionally, we keep the condition that two students can only establish a link if they are in the same classroom "r".

which can be represented by the following equation:

$$U_{r,i}(j) = \alpha_{|x|} \mid X_{r,i} - X_{r,j} \mid + \alpha_{x_i} X_{r,i} + \alpha_{x_j} X_{r,j} + \alpha_g g_{0,r,ij} + \alpha_q Q_{0,r,ij} + \alpha_r + \epsilon_{r,ij} \quad (2.4)$$

Which depends on the distance between them in the covariates space, $\mid X_{r,i} - X_{r,j} \mid$, on their individual characteristics, $X_{r,i}$ and $X_{r,j}$, on whether they were friends on a previous period or not, $g_{0,r,ij}$, on the number of common friends in the previous period, $Q_{0,r,ij}$ and on a term α_r that comprises the non-observable characteristics of network r.

Note that the distance between the covariates are included in the function in order to reflect homophily, the utility of a link decreases as the distance in the covariates increases. Assuming that ϵ_{ij} are independent in every “i” and “j” and that they follow a logistic distribution, we will have that the probability of a link to exist between “i” and “j”, given a prior version of the social network and given covariates, will be:

$$pr(g_{r,ij} = 1 \mid \mathbf{G_0}, \mathbf{X}) = p_{r,ij} \cdot p_{r,ji}$$

Where the probability of “i” to positively rate the link with “j” is equal to the probability of “j” to positively rate the link with “i”:

$$p_{r,ij} = p_{r,ji} = \frac{exp(\alpha_{|x|} \mid X_{r,i} - X_{r,j} \mid + \alpha_{x_i} X_{r,i} + \alpha_{x_j} X_{r,j} + \alpha_g g_{0,r,ij} + \alpha_q Q_{0,r,ij} + \alpha_r)}{1 + exp(\alpha_{|x|} \mid X_{r,i} - X_{r,j} \mid + \alpha_{x_i} X_{r,i} + \alpha_{x_j} X_{r,j} + \alpha_g g_{0,r,ij} + \alpha_q Q_{0,r,ij} + \alpha_r)} \quad (2.5)$$

The estimation of this model will allow us to identify which characteristics and behaviors influence social network formation the most.

Finally, it is important to bear in mind that the gains resulting from interacting with peers within the structure of a social network may be interdependent and each individual may be able to obtain complementarities from their direct peers. Given that the complementarities would be grounded on friendship links, which are heterogeneous for each relationship between two people, these heterogeneities mean that the influence of peers vary among the group members, depending on the individual’s position within a network of interactions.

Therefore, information regarding the friendship networks included in our database should also allow for the possibility of complementing our analysis on peer effects performed so far, exploring the heterogeneity of these effects, according to the students’ positions within their friendship network.

Given centrality measures, defined in subsection 2.2.3, to identify the importance of one’s position within a social network for determining a student’s outcomes, we estimated

two models similar to those developed by Hahn et al. (2015), the first of which:

$$Y_{r,i} = \theta + \beta M_{r,i} + X_{r,i}\gamma + \overline{\mathbf{X}}_{r,j}\delta + \alpha_r + \epsilon_{r,i} \tag{2.6}$$

Similar to the model of equation 2.1, $Y_{r,i}$ is the student's outcome variable, $X_{r,i}$ is a vector of the student's characteristics, $\overline{\mathbf{X}}_{r,j}$ is a vector with the mean of the characteristics of the peers of individual "i" and α_r represents the non-observable characteristics of classroom "r". In this case, however, to capture the heterogeneous effect, we included the centrality measure M of individual "i"[29], $M_{r,i}$.

A small change is made in the second model of this exercise for the effect of the student's position on his outcomes to vary in case he is located above or below the average position of the students in his group of reference. The model follows the same structure:

$$Y_{r,i} = \theta + \beta_1 \mid M_{r,i} - \overline{M}_{r,j} \mid + \beta_2 \mathbf{1}_{\{M_{r,i} > \overline{M}_{r,j}\}} + \beta_3 \mathbf{1}_{\{M_{r,i} > \overline{M}_{r,j}\}} \cdot \mid M_{r,i} - \overline{M}_{r,j} \mid + X_{r,i}\gamma + \overline{\mathbf{X}}_{r,j}\delta + \alpha_r + \epsilon_{r,i} \tag{2.7}$$

In which $\mid M_{r,i} - \overline{M}_{r,j} \mid$ measures the distance of the centrality measure M of individual "i" to the average centrality of his peers; variable $\mathbf{1}_{\{M_{r,i} > \overline{M}_{r,j}\}}$ is a dummy equal to 1 in case individual "i" is positioned above the average of his peers for a given centrality measure M and 0 otherwise; and $\mathbf{1}_{\{M_{r,i} > \overline{M}_{r,j}\}} \cdot \mid M_{r,i} - \overline{M}_{r,j} \mid$ is a measure of interaction between two previous variables. The remaining variables have the same meaning they had in equation 2.6.

2.4 Descriptive Analysis and Results

2.4.1 General Descriptive Analysis

After presenting data and the construction of measures, we will summarize some information concerning the sample and the variables available to estimate our models. To analyze potential peer effects on students, we will consider the students' friendship networks, according to the friends they nominate in the follow-up[30]. In this sense, as we mentioned in subsection 2.2.1, we start with a raw database with 2,720 students in time "t"[31] and obtained a final database with 1,260 students in both stages. Even though we

[29] To distinguish between the importance of own student's centrality and the average centrality of his peers, we also estimated a version of this model including the average centrality of the peers of individual "i", $\overline{M}_{r,j} = \frac{1}{g_{r,i}} \sum_{j=1}^{n_r} g_{r,ij} M_{r,j}$.

[30] Except for the case of homophily exercises, in which the model also includes information from the friendship network indicated in the baseline.

[31] Let "t-1" be the initial point in time (baseline) and "t" the final point in time (follow-up).

were left with only 46% of the original sample[32], table 2.1 suggests that the number of friends (chose by the students) who remain in the final database did not change much.

Table 2.1 – Descriptive Statistics - Number of Friends (N=1260)

	Average	Standard Deviation	Min.	Max.
Number of Friends - Gross Data (t)	2.57	1.95	0	8
Number of Friends - Final Data (t)	2.07	1.52	0	7

Note: The statistics refer to the number of friends indicated by the students and who were present in the sample of the respective dataset.

Table 2.2 presents a description of the variables that can be used in the models adopted, both in terms of the students' characteristics (age and gender) and in terms of cognitive and socioemotional measures available for both stages of data collection.

Table 2.2 – Descriptive Statistics of Variables

	Average	Standard Deviation
Individual Characteristics		
Boy	0.52	0.50
Age	10.74	1.32
Measures (t-1)		
Reliability&Curiosity	1.14	0.63
Self-Management	1.31	0.87
Cognitive Factor	1.00	0.11
Measures (t)		
Reliability&Curiosity	1.09	0.7
Self-Management	1.40	1.59
Cognitive Factor	1.08	0.14
Math Test	316.02	162.15
Language Test	313.72	172.49
Observations	1260	

We have in our sample a proportion of boys slightly greater than girls and also have students aged 10.7 years old on average (an age that is closer to that of 5th grade students). Regarding information on the measures, there is a variation within the same period and between periods, while the "Self-Management" measure stands out among the factors. It is important to note that, for descriptive purposes, we included measures on the bases on which they were constructed, however, in order to compare the results, we considered the versions of the cognitive and socioemotional measures with mean zero and standard deviation equal to 1 in the models estimation.

[32] Similar proportion is verified in other studies. Calvó-Armengol, Patacchini e Zenou (2009) for instance, addressed approximately 58% of the sample in the second phase of the Add Health project and Bramoullé, Djebbari e Fortin (2009) addressed approximately 56% of the general sample of the same project.

2.4.2 Peer Effect - "Linear in Means" Approach

As mentioned in subsection 2.3, in this first exercise we will use a variation of the "linear-in-means" model to investigate how the peers' socioemotional skills affect the individuals, both from a contextual perspective, on academic performance, and from an endogenous perspective, on the individuals' socioemotional skills.

In the analysis of the endogenous effect of socioemotional skills, our outcome variables (Y) will be the two factors extracted from SENNA ("Reliability&Curiosity" and "Self-Management"), while the cognitive factor and SAEPE's Language and Mathematics scores will be the outcome variables in the investigation of the contextual effect of socioemotional skills.

Regarding the contextual variables (characteristics), we will use the students' age and gender in all the exercises because of the relationship between these and the outcome variables. In a study with data from Brazilian students (SANTOS; PRIMI, 2014), gender was a variable statistically significant in determining all the six original socioemotional competencies included in the SENNA instrument. In line with international evidence (SOTO et al., 2011; MANGER; EIKELAND, 2000), women more frequently appear as being more agreeable and conscientious than men, in addition to the fact that boys presented a more external Locus of Control. Regarding age, statistically significant evidence was found in the determination of Emotional Stability and Locus of Control with older students presenting greater Stability and internal Locus. Thus, the inclusion of age and gender was called for.

In turn, socioemotional skills measured in "t-1" will be used as contextual variables in the investigation of such effects on academic performance. Regarding the impact of the socioemotional characteristics on educational outcomes, the same Brazilian study previously mentioned (SANTOS; PRIMI, 2014) shows that in the case of performance in Language, variations in Locus of Control and Openness to New Experiences are the variations that caused the greatest impact, while Mathematics is largely affected by changes in Conscientiousness.

And since our socioemotional (cognitive) skills in one period tend to have a relevant correlation with our socioemotional (cognitive) skills in the future, we will include the socioemotional measures obtained in "t-1" as contextual variables in the specifications in which these measures obtained in "t" were outcome variables, and will also include the cognitive factor obtained in "t-1" in the specificities in which cognitive factor obtained in "t" and proficiency in Language and Mathematics was the outcome variables.

Table 2.3 presents the correlations of contextual variables with their respective outcome variables.

What we see is that, in line with the previously mentioned literature, women would

Table 2.3 – Correlation between Outcome Variables and Contextual Variables

	Reliability&Curiosity	**Self-Management**	**Cognitive Factor**	**Math Test**	**Language Test**
Boy	-0.0845*	-0.0648*	-0.0578*	-0.0548*	-0.1233*
Age	-0.0828*	-0.0616*	-0.1216*	-0.0854*	-0.1287*
Reliability&Curiosity (t-1)	0.3964*	0.0716*	0.1506*	0.1641*	0.1720*
Self-Management (t-1)	0.0393	0.2470*	0.1128*	0.0817*	0.1091*
Cognitive Factor (t-1)			0.5811*	0.4987*	0.4763*
Observations			1260		

present a positive correlation with conscientiousness (via Reliability&Curiosity factor), as would also present a positive correlation with internal locus (via Self-Management factor). The only evidence contrary to the literature would be older children, who would present a negative correlation with emotional stability and more internal locus of control[33].

Regarding cognitive outcome variables, the Reliability&Curiosity factor (in which the conscientiousness and openness to new experiences items have a great weight) presents high positive correlation with the scores obtained in Mathematics and Language, in line with what is reported in the literature. As expected, the measure that includes Locus of Control, in turn, has a higher correlation with the score obtained in Language than with the score obtained in Mathematics.

In addition to choosing the covariates that compose the model, another important point we need to address before estimating this first exercise refers to the potential (in)transitivity and variation in the size of networks, considering these are essential to identify the model. Let's recall subsection 2.3 in which the identification of peer effects would be ensured if we could find intransitive "triads" in a social network or, in the absence of intransitivity, the size of networks varied. We have 145 classes, or social networks, in our sample with a minimum of two and maximum of 20 students. A total of 85% of these networks (which represent 96% of our sample of students) have an intransitive "triad" while the size of the 22 networks that do not have intransitive "triad" range from two to four students, which meets the identification requirements.

Having made these considerations, table 2.4 presents the results of our estimations for the exercise that analyzes peer effects using the SAR model.

In the case of the endogenous effect, we found a negative and significant result for all the outcome variables (both socioemotional and cognitive). This is a somewhat unexpected result as other studies found positive and significant endogenous effects on academic performance (BRAMOULLÉ; DJEBBARI; FORTIN, 2009; GOLDSMITH-PINKHAM; IMBENS, 2013; CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009). On the other hand, even though the coefficients are not significant, Chikitani (2015) report negative endogenous results on the Locus of Control measure and Chan e Lam (2014)

33 Age, however, can capture other related information, such as school delay, for instance, which would give sense to negative correlations, including with cognitive variables.

Table 2.4 – Effects using "Linear in Means" Approach

	Reliability&Curiosity	Self-Management	Cognitive Factor	Math Test	Language Test
Endogenous Peer Effect: WY	-0.175***	-0.106***	-0.150***	-0.178***	-0.186***
Contextual Peer Effect: WX					
Boy	-0.084	0.070	0.089	0.077	-0.139*
Age	-0.005	-0.100*	-0.052	0.013	-0.155***
Reliability&Curiosity (t-1)	0.068	0.091*	0.036	0.008	0.022
Self-Management (t-1)	-0.015	0.199***	0.062	0.019	0.093**
Cognitive Factor (t-1)			0.070	0.155***	0.088*
Own Characteristics: X					
Boy	-0.083	-0.108*	0.010	-0.002	-0.068
Age	-0.062**	-0.060*	-0.070***	-0.012	-0.141***
Reliability&Curiosity (t-1)	0.428***	0.147***	0.082***	0.113***	0.131***
Self-Management (t-1)	0.117***	0.298***	0.058***	0.036	0.071***
Cognitive Factor (t-1)			0.609***	0.524***	0.450***
Observations			1260		

Note: Results estimated using maximum likelihood method, with a set of dummies for classrooms and robust standard errors. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

found negative spillovers on the grade assigned by teachers to behaviour, when multiple networks were considered (friends and studymates)[34].

From the perspective of contextual effects of non-cognitive skills, the Reliability&Curiosity factor presented a significant and positive effect only on the socioemotional factor Self-Management. The latter, in turn, showed that the friends' socioemotional skills may affect not only the socioemotional skills of individuals (positive and significant contextual effect on the Self-Management factor itself), but also their academic results (positive and significant effect on proficiency in Language). Regarding the remaining variables, the friends' age and cognitive factor are relevant to determine certain outcome variables, Self-Management and Language test scores in the case of age, and Mathematics and Language test scores in the case of the cognitive factor. Note that the contextual effect of the Self-Management factor (which represents domains such as internal locus of control and emotional stability) appeared slightly more relevant (in terms of magnitude) in determining the Language test score than the cognitive factor itself. Chan e Lam (2014) also report a more robust contextual factor of the socioemotional measure (conscientiousness) on the students' English grade[35].

The effect of non-cognitive variables at the individual level is not the focus of this study; however, it is worth noting that they are important predictors of all outcome variables, especially the Reliability&Curiosity factor (which represents domains like Conscientiousness and Openness to New Experiences).

Even though a negative endogenous effect is not expected, the interpretation of models with spatial dependence go beyond the coefficients reported when we estimate the

[34] Chan e Lam (2014) also found negative spillovers, though not significant, on the Math test score when they considered the network of seatmates.

[35] In the case of the English grade, the friends' conscientiousness has a positive and significant effect on individual results, both considering the network of friends and the network of studymates, while the cognitive measure is significant only when the studymates network was considered.

general model presented in table 2.4. To briefly explain this analysis, consider, for simplicity purposes, a model that has only spatial dependence on the dependent variable. If we look at the model coefficients, we would say that the coefficient we call endogenous effect would be, itself, the spillover effect, while the coefficients of the individual characteristics would be the direct effects of each exogenous variable. However, in this case, if an exogenous measure "x" varies, it impacts the endogenous variable "y", in the magnitude of the coefficient "x", but this impact on "y" spills over to produce an additional variation on "y", in the magnitude of the coefficient of the endogenous effect, weighted by the weight matrix, and this variation spills over again to produce another impact on "y", and so on, successively. Therefore, in the interpretation of results we should take this recursive process into account. We should also consider that in a model with spatial dependence on exogenous characteristics, as in our case, a variation on "x" also causes a variation on "y" in the magnitude of coefficients of the contextual effect, weighting by the weight matrix, which will also spill over within the recursive process previously mentioned.

This entire process was considered when we presented the results in table 2.5 and, basically, what we estimated are the effects of individual characteristics and the contextual effects on the reduced-form mean of the outcomes variables.

Table 2.5 – Effects on the Reduced-Form using "Linear in Means" Approach

	Reliability&Curiosity	**Self-Management**	**Cognitive Factor**	**Math Test**	**Language Test**
Contextual Peer Effect					
Boy	-0.064	0.076	0.08	0.071	-0.114
Age	0.005	-0.088*	-0.038	0.013	-0.117***
Reliability&Curiosity (t-1)	-0.007	0.071	0.022	-0.011	-0.002
Self-Management (t-1)	-0.032	0.158***	0.049	0.011	0.072*
Cognitive Factor (t-1)			-0.02	0.056	0.003
Own Characteristics					
Boy	-0.078	-0.111*	0.006	-0.007	-0.06
Age	-0.062**	-0.056*	-0.068***	-0.013	-0.133***
Reliability&Curiosity (t-1)	0.428***	0.144***	0.081***	0.114***	0.131***
Self-Management (t-1)	0.119***	0.292***	0.055**	0.035	0.066***
Cognitive Factor (t-1)			0.610***	0.520***	0.45***
Observations			1260		

Note: Average Effects of independent variables (X) on the reduced-form mean of the outcome variable (Y) - post estimation results of the model reported in the table 2.4. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

The result that stands out the most in table 2.5, is the fact that we still found the contextual effect of the socioemotional factor Self-Management on the Language test score, while the effect of the cognitive factor disappears.

2.4.3 Networking Formation - Homophily Results

As already mentioned, variations in the group's characteristics and outcomes may affect individual outcomes, but may also change the reference network itself, given that these characteristics may be associated with the individuals' decisions when forming their links.

In the previous exercise, we observed that the results and characteristics of a student's reference network are actually potential predictors of an individual's own outcome. In particular, we found evidence that the mean of socioemotional skills of one's group may affect not only one's socioemotional skills, but also, some of one's academic results. Now, we want to verify whether there is a relationship between these characteristics and results and friendship network formation, being our main interest to identify the role of socioemotional skills in the establishment of these links.

To achieve this objective, we used a social network formation model (as defined in subsection 2.3) to identify the characteristics that are associated with the students' decisions to establish friendship ties, that is, to establish links among them[36].

The variables we include in the model determine which aspects will be captured as predictors of link formation. We estimated three specifications, a complete one, including the distance between each pair of students, both in their cognitive measures and socioemotional measures, while other specifications sometimes include only the distance for the cognitive measures and sometimes include only the distance for socioemotional measures. According to the case of each specification, we also included cognitive and socioemotional measures at the students level. Additionally, in all the specifications we included a variable that considers the number of common friends at baseline[37] and a dummy that identifies whether there was a link between two friends in "t-1".

Before addressing the results, we performed an analysis of the behavior of friendship relations (existence of links) between the dyads in our sample in the two periods.

Table 2.6 – Friendship Links (both periods)

		Link (t-1)		
		0	**1**	**Total**
Link (t)	**0**	7146	1504	8650
	1	2088	2042	4130
	Total	9234	3546	12780

Table 2.6 shows there is reasonable variation in the students friendship relations between the two periods. The individuals in approximately 32% of the 12,780 potential dyads verified in our sample identified themselves as friends in period "t", while 28% of them represent friendship relations in period "t-1". Additionally, only 50% of the pairs

[36] The estimation of this model requires the generation of all possible links between two students within a social network. A link in a social network, or pair between two nodes, is called dyad. In the case of this study, as we considered it a reciprocal relationship, the dyad between student "i" and student "j" (D_{ij}) and between student "j" with student "i" (D_{ji}) behave the same, that is, both represented connected dyads (with a link between two students) or both represent disconnected dyads (with no links between students).

[37] Similar to the previous exercise, here we considered the baseline as the initial point in time or "t-1" and the follow-up to be the final point in time or "t".

who were friends in period "t-1" remained connected in period "t", while 22% of those who were not friends in "t-1" became connected in "t". It shows that the friendship relations are dynamic and that a prior existence of a link between friends is certainly an important predictor in the formation of a link in the future, but it is not deterministic.

Table 2.7 presents the results for this second exercise. It is worth noting that the distance between the covariates are included in the function in order to reflect homophily, that is, the utility of a link decreases as the distance in the covariates space increases.

Table 2.7 – Homophily Results

	Complete	**Cognitive**	**Socioemotional**
Link (t-1)	0.202***	0.203***	0.204***
Number of common friends (t-1)	0.041***	0.042***	0.043***
|Boy(i) - Boy(j)|	-0.206***	-0.206***	-0.204***
|Age(i) - Age(j)|	-0.001	-0.001	-0.002
|Reliability&Curiosity(i) -Reliability&Curiosity(j)|	-0.016**		-0.019**
|Self-Management(i) - Self-Management(j)|	-0.026***		-0.027***
|Cognitive Factor(i) - Cognitive Factor(j)|	-0.008	-0.008	
|Math(i) - Math(j)|	-0.011	-0.012	
|Language(i) - Language(j)|	0.002	0.000	
Observations		12702	

Note: This table reports variables' marginal effects. Results estimated using logit, with a set of dummies for classrooms, a set of exogenous variables of the differences $\mid X_i - X_j \mid$ on the level of student "i" and "j". Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

In addition to the relevance of gender, of the previous existence of a link between individuals, and the number of common friends in the previous period in the formation of current links, the influence that also remains robust is the distance between socioemotional measures, that is, the closer the Reliability&Curiosity or Self-Management measures between two children, the greater the probability of them to become friends.

It is noteworthy that the distance between the cognitive measures is not statistically significant, not even in the specification which does not include socioemotional measures. Even if they were significant, the magnitude of the Self-Management factor is greater than the sum of all the cognitive measures, showing that similarity between socioemotional skills would be much more relevant than similarities in cognitive skills in the students' decision to establish friendship links. Chan e Lam (2014) also observe that proximity in terms of characteristics such as conscientiousness and extroversion increase the probability of two students to become friends. On the other hand, Goldsmith-Pinkham e Imbens (2013) found a significant influence in favor of homophily in cognitive terms[38], which we did not.

[38] But they do not include homophily in terms of gender, which is relevant both in this study and in the study by Chan e Lam (2014), and do not control for the cognitive variable at the student level, only include the difference.

2.4.4 Peer Effect - Centrality Approach

Finally, our last goal was to extrapolate the possibilities of the peer effects model previously adopted, investigating whether the heterogeneities between the reference groups of students in the same social network could make the peer effect on a group's members to vary, according to the position of the student within the friendship network.

Anyway, for this investigation, we will take into account the results found in the previous exercises. First, the evidence presented in table 2.7 indicates socioemotional skills as predictors of the formation of friendship links between the students, and possibly, of the social networks geometry. Additionally, the results in table 2.4 indicate these same skills as predictors of the individuals' outcomes, both from an individual point of view and, in some cases, from the perspective of peers.

In this sense, for this analysis, (i) we consider that socioemotional skills affect centrality measures, not the opposite, which led us to consider only cognitive measures as outcome variables[39], and (ii) we included the socioemotional factors in the model's covariates vector both at the individual and network levels.

Before presenting the results of the estimation of models 2.6 and 2.7, we will briefly analyze the four centrality measures considered in these estimations.

Table 2.8 – Descriptive Statistics - Centrality Measures

	Average	Standard Deviation	Min.	Max.
Closeness	0.543	0.164	0.169	1
Betweeness	0.103	0.149	0.000	1
Degree	0.376	0.220	0.053	1
Katz-Bonacich	0.274	0.170	0.042	1.4
Observations		1260		

When we present the example of one of our social networks in subsection 2.2.3, we pointed out that we could have a variation in the students' centrality, which is corroborated by table 2.8 and supports our analysis of heterogeneous effects. The parameters of the distribution of our measures are even similar to those reported in other studies that consider such measures (CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009; HAHN et al., 2015), especially in the case of the Closeness and Degree centralities.

After observing the behavior of centrality measures, we proceeded with our investigation. Table 2.9 presents the results of the model 2.6 estimation considering only the student's centrality measure.

[39] Only to confirm our assumption, we estimated the results of the model considering the socioemotional factors as outcome variables and, as presented in the Appendix B3, we did not find significant effects of centrality measures on the socioemotional measures.

Table 2.9 – Effects using Centrality Approach (own centrality)

	Cognitive Factor	**Math Test**	**Language Test**
$\mathbf{M_{r,i}}$			
Closeness	0.273	0.813***	0.725***
Betweeness	0.170	0.415**	0.341**
Degree	0.122	0.548***	0.499***
Katz-bonacich	0.087	0.665***	0.517***
Observations		1260	

Note: Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age, socioemotional and cognitive measures), both on student and group level. Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

The results seem to corroborate our hypothesis that individuals would be able to obtain complementarities from their direct peers and, as the friendship relations are heterogeneous within a group, a student's position within the network is relevant for determining some of his individual outcomes, that is, the more central a student is within a social network, the greater his standardized scores in Language and Mathematics. When we consider only the student's centrality measures, these results are positive and significant for any of the centrality measures considered. This model is very close to the alternative model estimated by Calvó-Armengol, Patacchini e Zenou (2009), though, they found statistically significant results for the American sample only for the Degree centrality (the Closeness and Betweenness coefficient are positive, but not significant).

Like Hahn et al. (2015), in the sequence we estimated the same model, only this time we included in the specification the average centrality of the students' reference group. Table 2.10 presents the results of this exercise.

When we decompose the effects of the individuals' centrality from the effects of the peers' average centrality, we continued to verify positive and significant effects on the Language and Mathematics test scores for three of the four measures considered. Now, like Calvó-Armengol, Patacchini e Zenou (2009) and Hahn et al. (2015) we did not find significant effects for the Betweenness centrality (note that this was the measure with the lowest mean and smallest variation).

On the other hand, our results diverge from those reported by Hahn et al. (2015) regarding the influence of peers' centrality. While they did not find significant results for any peers' centrality coefficient, we found, in some cases, a negative and significant result (even if, in the case of peers, the result does not seem as robust as the individual results - only the Closeness measure is significant both for Language and Mathematics). It means that belonging to a reference network in which, on average, your friends have a high centrality level, would harm your academic results.

Table 2.10 – Effects using Centrality Approach (own and peers centrality)

	Cognitive Factor	Math Test	Language Test
$\mathbf{M_{r,i}}$			
Closeness	0.247	0.713***	0.622**
Betweeness	0.167	0.298	0.250
Degree	0.079	0.452***	0.391**
Katz-bonacich	0.096	0.686***	0.545***
$\overline{\mathbf{M}}_{\mathbf{r,j}}$			
Closeness	-0.348	-1.321**	-1.358**
Betweeness	-0.010	-0.345	-0.266
Degree	-0.240	-0.549*	-0.614
Katz-bonacich	-0.269	-0.645	-0.858*
Observations		1260	

Note: Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age, socioemotional and cognitive measures), both on student and group level. Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Observing the results of this second specification, which suggest that the effect of centrality itself could be different from the peers' centrality effects, we have even more reasons to investigate the last specification involving centrality measures, in which we explore the possibility of asymmetrical effects, depending whether the student is located above or below the average centrality of his reference group. Table 2.11 presents the estimation results for the model 2.7.

Table 2.11 – Effects using Centrality Approach (heterogeneous centrality)

	Cognitive Factor	Math Test	Language Test
$\mid \mathbf{M_{r,i}} - \overline{\mathbf{M}}_{\mathbf{r,j}} \mid$			
Closeness	-0.502	-1.117**	-0.555
Betweeness	-0.103	-0.171	-0.010
Degree	-0.384	-0.414	-0.361
Katz-bonacich	-0.548*	-0.646	-0.291
$\mathbf{1}_{\{\mathbf{M_{r,i}} > \overline{\mathbf{M}}_{\mathbf{r,j}}\}}$			
Closeness	-0.063	-0.045	-0.032
Betweeness	-0.005	0.146*	0.137*
Degree	-0.166*	0.055	-0.002
Katz-bonacich	-0.154	0.012	-0.002
$\mathbf{1}_{\{\mathbf{M_{r,i}} > \overline{\mathbf{M}}_{\mathbf{r,j}}\}} \cdot \mid \mathbf{M_{r,i}} - \overline{\mathbf{M}}_{\mathbf{r,j}} \mid$			
Closeness	0.954	1.973**	2.047**
Betweeness	0.238	0.164	0.100
Degree	0.939	0.730	0.995*
Katz-bonacich	1.170	1.281	1.475*
Observations		1260	

Note: Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age, socioemotional and cognitive measures), both on student and group level. Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

In this case, we understand that more robust evidence is concentrated on the model's coefficient β_3, that is, on the results of variable $\mathbf{1}_{\{M_{r,i} > \overline{M}_{r,j}\}} \cdot \mid M_{r,i} - \overline{M}_{r,j} \mid$, which are positive and significant for three of the four centrality measures, when we analyze the results on the Language score [40]. As only this coefficient is significant for the Language standardized score, it means that, if a student's centrality measure is above his peers' average centrality, the more central he is, the higher his score[41].

Here it is important to note that none of the two studies with which we compare our results include in their specifications controls for the students' and/or their peers' socioemotional skills. As noted in the beginning of this subsection, these variables proved to be very relevant within this context of social networks and peer effects analysis. In fact, Hahn et al. (2015) recognize this relevance and even consider that the effect of centrality measures would be actually capturing some non-cognitive measure of students. To test this hypothesis, they implement exercises to show that some of the socioemotional measures they have available positively and significantly impact centrality measures, which would not be true for the cognitive measures[42]. The fact is that even controlling for the students' and their peers' socioemotional measures, we keep finding positive and significant effects for centrality measures, suggesting that these effects may not only capture the impact of socioemotional skills but also the impact of complementarities obtained by the individuals from their friendship relations.

2.5 Simulating Social Networks

We performed some additional exercises to investigate how peer effects would behave if we observed social networks with links established according to our network formation model.

In this stage, we first analyzed the results of the network formation model if we kept in our sample the isolated individuals[43], that is, those individuals with no links with the remaining students of their respective classrooms, as reported in the follow-up.

[40] Like in our case, the main results found by Hahn et al. (2015) are also for the coefficient β_3 and they did not find significant effects for all the measures they considered either.

[41] we have two analyses in the case of the Closeness results on Mathematics score, the negative result of β_1 means that, if a student's centrality measure is below his peers' average centrality, the less central he is and the lower his score. Considering that for the students whose centrality is above the average, we need to sum the coefficients β_1 and β_3, both significant, and, since the result is positive, the more central a student is, the greater his score. Finally, in the case of the centrality measure Betweenness, in which only the coefficient β_2 is positive and significant, both on the Language and Mathematics test scores, we have that a centrality measure above the reference group's average has a positive effect on a student's scores.

[42] In a way, these results reported by Hahn et al. (2015) are in line with the results we found in subsection 2.4.3.

[43] It implies that we used the database with 1,393 students in both stages of data collection, which was obtained after excluding the observations with no information for some of the variables considered, as we explained in subsection2.2.1.

Table 2.12 presents the results of this exercise, which are very similar to the results found when we estimated the model for the sample without isolated individuals (table 2.7). Therefore, we will start with the coefficients estimated in the current model (with isolated students) to simulate the social networks that will be used in the peer effects estimation.

Table 2.12 – Homophily Results - Sample with isolated students

	Complete	**Cognitive**	**Socioemotional**
Link (t-1)	0.192***	0.193***	0.195***
Number of common friends (t-1)	0.039***	0.039***	0.041***
|Boy(i) - Boy(j)|	-0.189***	-0.189***	-0.187***
|Age(i) - Age(j)|	-0.003	-0.003	-0.004
|Reliability&Curiosity(i) -Reliability&Curiosity(j)|	-0.016**		-0.020***
|Self-Management(i) - Self-Management(j)|	-0.024***		-0.026***
|Cognitive Factor(i) - Cognitive Factor(j)|	-0.007	-0.007	
|Math(i) - Math(j)|	-0.010	-0.011	
|Language(i) - Language(j)|	-0.003	-0.004	
Observations		14376	

Note: This table reports variables' marginal effects. Results estimated using logit, with a set of dummies for classrooms, a set of exogenous variables of the differences $\mid X_i - X_j \mid$ on the level of student "i" and "j". Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Considering the value predicted for the "complete" specification of table 2.12, we constructed 500 samples with simulated networks according to the following procedure:

1. We assigned a random value[44] for each of the 14,376 potential dyads observed in our sample.

2. We generated a dummy variable to identify the existence of links between the students of each dyad, which is equal to 1 if the predicted value in the peer formation model is greater or equal to the random value assigned to the respective dyad, and 0 otherwise.

3. As we are assuming that friendships is a reciprocal relationship, we consider there is a link between node "i" and node "j" if, in the previous procedure, we identified a link between these nodes, be in the dyad between "i" and "j" ($D_{i,j}$) be in the dyad between "j" and "i" ($D_{j,i}$).

4. Finally, since we did not include the students without a friend in the peer effects estimation, we excluded from the sample those who remained isolated after we concluded the previous stages.

[44] Random variables uniformly distributed over interval [0,1].

The final product are 500 samples, with an average of 1,340 students[45]. For each of these samples, we estimated the endogenous peer effects model (the variation of the "linear in means" model defined in equation 2.2) and the models that consider the heterogeneous peer effects, based on the student's centrality (equations 2.6 and 2.7).

The results reported in each case are the means of the coefficients obtained in each of the 500 estimations. For the test of hypothesis, we considered a t test built from the distribution of the 500 observed coefficients.

2.5.1 Simulation Results: Peer Effect - "Linear in Means" Approach

Table 2.13 presents the results for the model of equation 2.2.

Table 2.13 – Simulation Results: Effects using "Linear in Means" Approach

	Reliability&Curiosity	**Self-Management**	**Cognitive Factor**	**Math Test**	**Language Test**
Endogenous Peer Effect: WY	**-0.241**	-0.097	**-0.215**	**-0.287**	**-0.275**
Contextual Peer Effect: WX					
Boy	-0.032	0.049	0.133	0.073	-0.049
Age	-0.045	-0.067	-0.043	-0.037	-0.085
Reliability&Curiosity (t-1)	**0.193**	0.039	0.01	0.042	0.038
Self-Management (t-1)	0.02	**0.128**	0.036	0.023	0.04
Cognitive Factor (t-1)			**0.177**	**0.228**	**0.208**
Own Characteristics: X					
Boy	**-0.098**	**-0.109**	-0.007	0.009	**-0.105**
Age	**-0.054**	**-0.056**	**-0.057**	**-0.027**	**-0.131**
Reliability&Curiosity (t-1)	**0.443**	**0.137**	**0.079**	**0.124**	**0.138**
Self-Management (t-1)	**0.117**	**0.297**	**0.056**	**0.041**	**0.069**
Cognitive Factor (t-1)			**0.616**	**0.531**	**0.471**
Observations			1340		

Note: Average results for the 500 simulated networks. Results estimated using maximum likelihood method, with a set of dummies for classrooms and robust standard errors. Results in bold are significant at 10% (or smaller) level.

As observed in table 2.4 for our original sample, the endogenous effects remain negative and significant (except in the case of the socioemotional factor Self-Management). On the other hand, for the peers contextual effects, now we only observe impacts of the outcomes variables in "t-1", while here the results also seem to be positive and significant when we consider the socioemotional factor Reliability&Curiosity and the cognitive factor. The effects observed at the individual level are very close to those found in the original sample.

Consider the links established according to the network formation model, to estimate peer effects, may be capturing some endogeneity problem related to the fact that the original network is self-reported, which could justify the results found for the simulated samples.

[45] Our smaller sample has 1,320 observations and the largest sample has 1,362 observations, that is, the number of observations in our simulated samples is between the number of observations of our real sample with and without isolated students, that is, 1,260 and 1,393 students, respectively.

2.5.2 Simulation Results: Peer Effect - Centrality Approach

Tables 2.14 and 2.15 present the simulated results of the first model that consider the centrality measures and corroborate the results for the individual centrality measures we had previously found.

Table 2.14 – Simulation Results: Effects using Centrality Approach (own centrality)

	Cognitive Factor	**Math Test**	**Language Test**
$\mathbf{M_{r,i}}$			
Closeness	0.174	**1.061**	**0.771**
Betweeness	0.082	**0.566**	**0.373**
Degree	0.118	**0.666**	**0.489**
Katz-bonacich	0.142	**1.041**	**0.809**
Observations		1340	

Note: Average results for the 500 simulated networks. Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age, socioemotional and cognitive measures), both on student and group level. Standard errors clustered at school level. Results in bold are significant at 10% (or smaller) level.

When we consider only the individual centralities (2.14), the coefficients of all measures continue to indicate positive and significant impacts on both proficiency scores. In terms of magnitude, we found even more expressive results, especially when the outcome variable is the Mathematics test score.

Table 2.15 – Simulation Results: Effects using Centrality Approach (own and peers centrality)

	Cognitive Factor	**Math Test**	**Language Test**
$\mathbf{M_{r,i}}$			
Closeness	0.148	**1.032**	**0.8**
Betweeness	0.08	**0.644**	0.468
Degree	0.101	**0.649**	**0.507**
Katz-bonacich	0.113	**0.972**	**0.778**
$\overline{\mathbf{M}}_{\mathbf{r,j}}$			
Closeness	-0.118	-0.134	0.119
Betweeness	-0.006	0.133	0.165
Degree	-0.074	-0.076	0.07
Katz-bonacich	-0.321	-0.768	-0.339
Observations		1340	

Note: Average results for the 500 simulated networks. Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age, socioemotional and cognitive measures), both on student and group level. Standard errors clustered in school level. Results in bold are significant at 10% (or smaller) level.

When we consider the specification that includes both individual measures and the mean of the group's centrality measures, the individual results in the original sample

remain positive and significant (except for the Betweenness measure), but they decline for almost all measures in comparison to the first specification. What we see for the simulated results is that many individual coefficients become more expressive or almost do not change, if compared to the first specification. Additionally, we had previously found some negative and significant results for the peers' centrality measures, which we no longer found. No result of the group's centrality is significant and various coefficients are positive.

Finally, the results for the last model, summarized in table 2.16, do not show considerable robustness for the potential asymmetrical effects, just as it happened in the original sample.

Table 2.16 – Simulation Results: Effects using Centrality Approach (heterogeneous centrality)

	Cognitive Factor	**Math Test**	**Language Test**
$\mid \mathbf{M_{r,i}} - \overline{\mathbf{M}}_{\mathbf{r,j}} \mid$			
Closeness	-0.133	-0.6	-0.443
Betweeness	-0.048	-0.258	-0.12
Degree	-0.103	**-0.395**	-0.297
Katz-bonacich	-0.138	**-0.83**	**-0.695**
$\mathbf{1}_{\{\mathbf{M_{r,i}} > \overline{\mathbf{M}}_{\mathbf{r,j}}\}}$			
Closeness	0.011	0.089	0.069
Betweeness	0.024	**0.135**	**0.105**
Degree	0.005	0.073	0.053
Katz-bonacich	0.01	0.04	0.011
$\mathbf{1}_{\{\mathbf{M_{r,i}} > \overline{\mathbf{M}}_{\mathbf{r,j}}\}} \cdot \mid \mathbf{M_{r,i}} - \overline{\mathbf{M}}_{\mathbf{r,j}} \mid$			
Closeness	0.197	0.955	0.601
Betweeness	0.029	0.252	0.109
Degree	0.157	0.68	0.461
Katz-bonacich	0.221	**1.574**	**1.303**
Observations		1340	

Note: Average results for the 500 simulated networks. Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age, socioemotional and cognitive measures), both on student and group level. Standard errors clustered at school level. Results in bold are significant at 10% (or smaller) level.

Anyway, the results presented in tables 2.14 and 2.15 suggest that if we could form social networks according to the standards observed in the network formation model, the influence of a student's position within his group could be an even more relevant predictor of his school performance.

We also understand that observing the different results of the simulated networks shows how important it is to consider the mechanisms of network formation when studying the potential influence resulting from the interaction between individuals.

2.6 Final Considerations

This study investigates the relationship between peer effects and the students' socioemotional skills. Using a database rich in information on the students' interpersonal relationships, which enabled us to develop what we call friendship networks, we set out to answer this question through three exercises.

First, we used a peer effects model similar to the empirical model most commonly used in this area (the "linear in means" model), to identify potential endogenous and contextual peer effects. Evidence shows negative endogenous effects for all the outcome measures (both cognitive and socioemotional measures). Even though some studies had already found negative endogenous results, non-significant though, for socioemotional measures (CHIKITANI, 2015; CHAN; LAM, 2014), these results were not expected because other studies found positive and significant endogenous effects on academic results (BRAMOULLÉ; DJEBBARI; FORTIN, 2009; GOLDSMITH-PINKHAM; IMBENS, 2013; CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009). On the other hand, in the scope of contextual effects, for the Language score outcome, we verified that the socioemotional measure Self-Management has a positive and slightly greater influence than the cognitive factor, in line with what was found by Chan e Lam (2014) for the English score.

In the second stage, using a social network formation model, we identified that proximity in terms of socioemotional competencies would be more relevant than proximity in terms of cognitive results in the establishment of links between students. This result is somewhat ambiguous in other studies. Even though they do not consider socioemotional measures in their model, Goldsmith-Pinkham e Imbens (2013) observe homophily, in cognitive terms, a significant predictor for the establishment of links. On the other hand, Chan e Lam (2014) obtain some significant evidence in favor of homophily for measures such as Conscientiousness and Extroversion.

In the third exercise, we used an empirical model that allowed us to observe the possibility of heterogeneous effects, assessing whether the influence a group exerts on an individual would be particularly related to the position of this agent within the group and, therefore, vary according to this position. Evidence found was robust and pointed out that an individual's centrality is a good predictor of his academic performance. As we controlled for the socioemotional measures, both at the student and the group level, we understand that the effects of centrality may not only be capturing the impact of socioemotional skills (as suggested by Hahn et al. (2015)), but also the impacts of complementarities obtained by the subjects in their friendships.

Finally, we performed some additional exercises, simulating new samples, to investigate how peer effects would behave if we observed social networks with links established according to our network formation model. The simulated results still presented negative

endogenous effects. On the other hand, in the case of contextual peer effects, we began to verify only impacts of the outcome variable itself, observed in the previous period. For the heterogeneous effects model, the simulated results continue to corroborate the relevance of individual centrality on academic results, with magnitudes more expressive than those found for the original sample.

We understand that despite our database being a rich source of information for this study, with data that enabled us to construct the students' friendship networks and socioemotional measures, we are aware that it has also some limitations such as a lack of more complete information regarding the students' exogenous characteristics (both at the personal level such as ethnicity information, and at the level of family environment, such as the parents' educational level and information of some socioeconomic measure of these students) and lack of adherence of the socioemotional test with the original competencies that it was intended to measure (possibly because of our students' age and due to the fact that this applied questionnaire was still under development and could be considered long for the age group addressed). Because, as we have seen, the variables that compose the model could be as important as dealing with identification problems, as they determine which aspects will be captured as predictors of the results we assessed.

Anyway, not only because of the limitations, the empirical results obtained here, which sometimes are in line with evidence of some studies and sometimes are in line with evidence of other studies, show how relevant the relationship between the topics we address here is relevant and how much it is necessary to explore it within this literature to obtain clearer and even more robust results to effectively support public policies.

3 Impact of Full-Time High School on Network Connectivity Measures

3.1 Introduction

The need to improve the quality of education offered to children and young people, especially in the most vulnerable environments, has led to the implementation of several programs focused on expanding the school time. The increase of time would permit exploring new content and activities, extending the students' time of study and expanding the school integration, promoting more comprehensive education and reflecting in increased academic performance. Nevertheless, there is no consensus in the literature on the effectiveness of this type of program.

Studies around the world have sought to evaluate the relationship between increased time and school performance, with varying results, both in developed countries and developing countries. In the United States, Dynarski et al. (2004) investigated the impact of community programs that offered additional educational activities in the extracurricular period and did not observe significant effects on the grades of elementary or high school students. In turn, Zimmer, Hamilton e Christina (2010) assessed the impact of a program implemented in public schools in Pittsburgh and verified significant effects on the math scores of elementary and high school students. Also in the context of developed countries, Meyer e Klaveren (2013) found no effect on the mathematics or language tests of elementary school students in a Dutch program and Battistin e Meroni (2016) registered that the expansion of school time entailed positive results for the mathematics grades of students from Southern Italy.

Looking at emerging countries, Hincapie (2016) observed positive effects for mathematics and language in a full-time education program implemented in Colombian elementary education, although the effects on language tests were smaller in magnitude and less robust. In Chile, using a method of differences in differences, Pires e Urzua (2010) have identified positive effects on cognitive tests and reduction in school dropout of students who opted for full-time high school, although these results vary in magnitude and significance depending on the subsample analyzed. In Brazil, much of the studies evaluate full-time teaching programs for elementary school students and find little or no effect on the test scores (AQUINO; KASSOUF, 2011), even identifying negative effects on mathematics grades (ALMEIDA et al., 2016; GANDRA, 2017).

In view of such heterogeneous results, our study fits into this context of evaluation of full-time educational programs, but seeking to look from another angle, after all, if

the prospects for gaining from these programs are so promising, why do the results seem unsatisfactory? One type of literature that has grown a lot in recent times and that could shed some light on this inquiry is the literature that studies how the effects of peers can influence the individuals' educational outcomes. Beyond demonstrating what students can gain from their friendship ties (BRAMOULLÉ; DJEBBARI; FORTIN, 2009; SACERDOTE, 2001; CARRELL; FULLERTON; WEST, 2009), some studies suggest that the formation of these links is dynamic,[1] and may vary if there are changes in the parameters the individuals are subject to (GOLDSMITH-PINKHAM; IMBENS, 2013; CHRISTAKIS et al., 2010; CARRELL; SACERDOTE; WEST, 2013).

It seems reasonable to assume that moving from a regular education model to a full-time model would have great potential to change the dynamics of the relationships between students. The coexistence between them increases, new activities are shared and they spend much less useful time in other environments, therefore, with less possibility of forming external links. The schoolmates, and particularly the classmates increasingly become their references.

That brings us to the question this study intends to investigate. Does the implementation of full-time high school change the structure of the students' friendship network? As far as we know, there is no other study that has explored the possibility of educational interventions affecting the structure of the networks. In the educational sphere, what comes closest to what we want[2] is the study by Carrell, Sacerdote e West (2013). They show that an intervention that separated the students into groups according to their academic outcomes affected the type of link created among the students and, therefore, the reference group of each student within the network.

In the area that investigates the exposure to formal credit markets, there are some studies with questions close to the one we seek to answer here, but sometimes with different results. Banerjee et al. (2018) investigated the impact of the introduction of microcredit institutions in Indian villages on the structure of their social networks and identified a significant reduction in these network ties. On the other hand, Comola e Prina (2014) analyzed how access to savings accounts in Nepalese villages affected the structure of informal financial transaction networks. In addition to reporting increased links within these networks, the same authors conducted some exercises to demonstrate that, if this change is not taken into account, the effects of the policy of expanding the access to savings accounts would be underestimated[3].

[1] The individuals permanently have the opportunity to build a relationship with the other person. This decision will take into account all the benefits this relationship offers, depending on factors such as the characteristics of the individual, a possible friend, and other parameters of his network.

[2] Unlike us, the intervention they analyze directly affects the construction of the students' network.

[3] Other studies investigating the peer effect have already pointed to the importance of taking into account the network formation when estimating the effects. Carrell, Sacerdote e West (2013), for example, show that, in the investigated intervention, the groups were constructed according to the premise that students

In our case, we used data from 2018 for high school students from public schools in the state of Sergipe (Brazil) and observed that the implementation of full-time high school led to an increase in the centrality of these students in their friendship networks. These results, together with the evidence from the literature on the importance of considering such changes in the evaluation of the effects of the policy itself, have led us to a new question. Would the absence of robust effects in the studies that investigated the impacts of school time extension policies hide a possible channel through which the program would affect the students' academic outcomes, that is, the social network structure?

To try and answer this question, we first verified that, in the case of our sample, we would not find a significant effect either if we simply investigated the impact of full-time high school on the students' grades. Then, we explored a model that includes the possibility of test scores being affected by students' centrality measures and observed positive and significant effects on Mathematics scores. This result suggests that, if the full-time education policy affected students' connectivity measures and if some of these measures affect the students' academic outcomes, then the network structure seems to be a possible channel through which the program would affect the students' scores.

The sequence of this work is divided as follows: in the next section, we will present the context for the implementation of the evaluated program, as well as an analysis of the sample construction and the collected data. In section 3.3, we briefly analyze the structure of the observed social networks and investigate the impact of the program on the network structure. In section 3.4, we discuss a possible mediation effect of the intervention on the students' cognitive test scores through the social network structure. Finally, we present our final considerations about the study.

3.2 Context and Data

3.2.1 The Context of Full Time High School

In 2014, the National Education Plan (*Plano Nacional de Educação* - PNE), which was approved by Federal Law 13,005 and remains in force until 2024, has set, among other goals: (i) the universalization of school attendance for youth 15 to 17 years of age by 2016; (ii) the promotion of the quality of basic education, in order to improve the flow of school and learning, in accordance with the biannual goals; and (iii) the provision of full-time education in at least 50% of the public schools, in order to attend to at least 25% of the students in basic education. However, what the data from 2016 revealed was a very different reality.

---

with lower grades would benefit from the presence of students with higher grades in their networks, but when students were separated into groups with these characteristics, there was a change in the pattern of links formed within the heterogeneous networks, which led to very different results than expected, generating a negative effect for students in the lower end of the distribution of academic skills.

Data from the National Continuous Household Sample Survey (*Pesquisa Nacional por Amostra de Domicílios Contínua* - Pnad) showed that 12.8% of the population in this age group were still out of school. The Basic Education Development Index[4] (IDEB) also indicated a distance of this quality parameter from the planned goals[5]. And data from the School Census showed that only 6.7% of public school enrollments in high school were in the full-time modality.

It is in this context that the Program to Promote the Implementation of Full-Time Schools[6] (*Programa de Fomento à Implementação de Escolas em Tempo Integral* - EMTI) was established at the national level, aiming at the improvement of high schools and the permanence of young people in this school stage. The program provides for support[7] for the implementation of a pedagogical proposal of full-time high schools in the public education system, based on the extension of the school day and on the integral and integrated student education, considering both cognitive and socioemotional aspects.

It is also in this context, and given that adherence to the EMTI is given to the state department of education, that the "School Educates More" (*Escola Educa Mais*) program was created in the State of Sergipe. Sergipe is the smallest Brazilian state and, despite having presented the third highest Gross Domestic Product (GDP) per capita among the nine states of the Northeast in 2016[8], its education rates went in the opposite sense. In 2015, the IDEB of Sergipe, for state-owned high schools[9], amounted to only 2.6, the worst indicator in the Northeast, and much lower than the national average. Besides worrying school performance rates, 17% of state high school students failed and 16.3% dropped out of school in 2015, the learning rates were also very critical, as only 0.52% and 0.46% of the students in state high schools presented proper knowledge on the Portuguese and Mathematics tests of Saeb.

With only 4 full-time high schools until 2016, the program "School Educates More" combines the understanding that the state public high schools panorama requires the construction of more effective public policies and the opportunity for federal financial support for the implementation of a full-time education policy, to set the goal of offering full-time education in 38 high schools of the state education system, between the years 2017 and 2018.

In fact, Sergipe reached the established goal and implemented the new pedagogical

---

[4] The IDEB (*Índice de Desenvolvimento da Educação Básica*) is calculated based on the students' Portuguese and Mathematics test scores (*Sistema de Avaliação da Educação Básica* - Saeb and *Prova Brasil*) and on the school flow-pass rate (School Census).

[5] The High School IDEB of State schools was 3.5 in 2015 (the IDEB is biannual and we have no coefficients for 2016), while the goal of the State schools for that year was 3.9.

[6] Ministry of Education (MEC) Decree 1,145 from 2016.

[7] Through the transfer of financial resources to the department of education of the participating states.

[8] According to data from the Brazilian Institute of Geography and Statistics.

[9] Like for Brazil, more than 97% of enrollments in public high schools in Sergipe come from the state education system.

model in 13 schools in the year 2017 and in 24 schools in the year 2018. The database constructed from the assessment of the implementation of the program in 2018 will allow us to develop the analyses we propose in this article.

3.2.2 Data

As mentioned, for the analyses in this article, we will use the data collected in 2018 to assess the implementation of full-time high school education in Sergipe.

For the year 2018, the State Department of Education submitted for evaluation of the MEC 24 schools that manifested their interest in the adoption of the full-time high school education model and obtained the official approval for all of them. It is important to highlight that the appointment of a school to participate in the EMTI only occurs if the school itself takes interest, which involves the school community's approval of the new model. In addition, if on the one hand the MEC recommends ranking the submitted schools according to some infrastructure requirements[10] and physical capacity, on the other hand, it also recommends that education departments give priority to the choice of schools in regions of social vulnerability.

To permit the evaluation, in addition to the data from the 24 schools that have adopted the new teaching model, the Department of Education further ranked 24 regular high schools, which had characteristics as close as possible to the schools in the program, to serve as a comparison group. Thus, our original sample would consist of 48 schools, 24 for treatment and 24 controls, distributed among 28 cities in the state, in addition to the capital.

As Sergipe adopted a gradual implementation format, with the conversion of one grade per year, in 2018, the data collection only involved students in the 1st grade of high school (or the 12th grade of basic education), being the only students the program directly affects at the treated schools. In addition, for budgetary reasons, the data collection was done in a single class of each school[11].

Data were collected at two points in time. In March 2018, at the beginning of the school year and, hence, at the start of the adoption of the full-time high school education model, a baseline survey was conducted through the application of assessments such as the revised version of the SENNA (Social and Emotional or Non-cognitive Nationwide Assessment), an instrument developed by the Ayrton Senna Institute to measure five non-cognitive skills[12], and standardized tests of Portuguese Language and Mathematics.

[10] Related to spaces such as the principal's office, teachers' room, secretary room, computer lab, library, bathrooms, indoor courtyard, among others.

[11] To avoid arbitrariness, the data were always collected in Class A of the 1st grade of high school at each school.

[12] The measures obtained by the SENNA are: Agreeableness, Self-Management, Engagement with Others, Emotional Resilience, and Openness. Agreeableness includes traits such as empathy, respect, and

In November 2018, that is, at the end of the school year, the follow-up research took place, involving the application of the same baseline assessments. In addition, the students answered a socioeconomic questionnaire, with information on personal characteristics such as age, gender and race, and information on the family environment, such as the mother's level of education and items regarding infrastructure and consumer durables present at home.

The essential data for our research were collected only in the follow-up research and allow us to identify the students' friendship network and generate the connectivity measures explored in our empirical exercises. The friendship information was obtained through a questionnaire that asked each student to identify up to five best friends, within the list of classmates. Assuming that friendship relations are reciprocal, that is, there is a link between two students if at least one of them indicated the other as a friend, the structure of each friendship network is captured by a symmetric matrix $\mathbf{G_r}$ of $n_r \times n_r$ dimensions, where n_r is the number of students in class r. The elements of the matrix $\mathbf{G_r}$ will be equal to $g_{r,ij} = g_{r,ji} = 1$ if we observe a link between students i and j and 0 otherwise, while all the elements of its diagonal, $g_{r,ii}$, are equal to zero.

Our sample is based on the collection of data from the baseline and follow-up assessments, the socioeconomic questionnaire and the friendship questionnaire. At first, we departed from a sample of 48 schools, however, observing the baseline data, we found that we had no information for 1 Control School, so that we could build a sample with a maximum of 47 schools. In addition, when we collected the friendship information, we identified the absence of information for three more control schools, which left us with a final sample of 44 schools, being 24 treatment, and 20 control units. In terms of students, we started with a sample with 1203 students, being 641 treated students and 562 control students. After excluding the students who did not have information from the instruments in the follow-up or information for the variables in the socioeconomic questionnaire and who did not have any friendship link, we got a final base with 717 students, being 392 treatment students and 325 controls.

We know that our program was not implemented randomly and our final sample did not continue with the same number of schools with and without the presence of full-time education, so before defining our empirical approach we did some balancing and attrition tests.

As the program is adopted at the school level, we first compared the characteristics of the schools that adopted full-time high school education and the schools in the compar-

modesty; Self-Management includes organizational skills, determination, focus, persistence, and responsibility; Engagement with Others involves traits of assertiveness, enthusiasm, and ease of communication; Emotional Resilience refers to aspects such as tolerance to stress, self-confidence, and self-control; and Openness includes traits such as curiosity, creativity, and interest in the arts.

ison group, using data from the 2017 School Census[13]. Table 3.1 presents the balancing result, considering the schools in the original sample and in the sample that considers the school that was lost due to the absence of baseline information, which we will call the initial sample.

Table 3.1 – Balance Check - School Sample

		Original Sample	**Initial Sample**
School Data	Sewer	0.226	0.201
	Teachers' room	0.104	0.068
	Science Lab	-0.052	-0.035
	Library	0.313	0.294
	Reading Room	-0.132	-0.175
	Bathroom with shower	-0.072	-0.066
	Dinning Hall	0.272	0.295*
	Indoor Courtyard	0.194	0.186
	Number of Classrooms	0.021	0.029
	Number of Students Computers	-0.007	-0.003
	Number of Employees	-0.006	-0.007
	State's Capital	-0.030	0.032
	N	48	47
	F Test	2.849	3.089

Note: The data source for "School Data" is 2017 School Census. Balance check estimated using ordinary least squares, with standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

In the case of the original sample, although we cannot rule out the unbalance, the results suggest that the Department of Education actually indicated a group of comparison schools very similar to the group of schools treated, in terms of infrastructure and location in the State capital. When we look at the initial sample we see that the loss of one of the control schools does not substantially change the balancing results.

After observing the patterns of the preliminary samples, we proceed to investigate the behavior of the final sample. In this case, besides comparing the schools' characteristics, we also analyzed possible differences in the students' characteristics, considering the variables of the socioeconomic questionnaire and the scores of the socioemotional instrument and the Language and Mathematics tests obtained at baseline.

The results are presented in table 3.2 and, at the school level, they indicate that the loss of the other three control schools, because of the friendship data, does not greatly affect the balancing of the groups when considering what was observed for the previous samples.

[13] One of the treatment schools was new and had no information for the 2017 School Census. Thus, in order not to impair our analysis, we used, only for this school, the information present in the 2018 School Census. It should be noted that, for the variables considered, the variations from 2017 to 2018 are not expressive. On average, the difference between the two years is always greater in favor of the treated schools. Therefore, the complementation of the data of the treated school, at the limit, tends to go against and not in favor of the balance.

Table 3.2 – Balance Check - Final Sample

		School Sample	**Student Sample**	
School Data	Sewer	0.175		0.150
	Teachers' room	-0.517**		-0.533***
	Science Lab	-0.052		-0.132
	Library	0.336		0.453**
	Reading Room	-0.151		-0.152
	Bathroom with shower	0.033		0.150
	Dinning Hall	0.318*		0.329**
	Indoor Courtyard	0.205		0.132
	Number of Classrooms	0.026		0.040
	Number of Students Computers	-0.002		-0.002
	Number of Employees	-0.005		-0.005
	State's Capital	-0.068		-0.104
Student Data	Age		0.039	0.019
	Race: White		0.017	0.010
	Race: Brown		-0.017	-0.017
	Woman		0.046	0.018
	Retention		-0.060*	-0.025
	Early Childhood Education		-0.104*	-0.111**
	Mother's Schooling: Elementary School		-0.093	0.007
	Mother's Schooling: High School		-0.062	-0.023
	Mother's Schooling: College or more		-0.073	-0.012
	Socioeconomic Status		-0.035	-0.059
	Agreeableness (t-1)		0.024	0.002
	Self-Management (t-1)		0.053*	0.024
	Engagement with Others (t-1)		0.031	0.024
	Emotional Resilience (t-1)		-0.001	0.002
	Openness (t-1)		-0.020	-0.004
	Language Test (t-1)		0.018	-0.004
	Math Test (t-1)		-0.008	0.008
	N	44	717	717
	F Test	14.65	3.855	131.6

Note: The data source for "School Data" is 2017 School Census. Balance check estimated using ordinary least squares, with standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

When we look at the comparison of the students' characteristics, in combination with the school characteristics, we observe that we cannot reject the hypothesis of an unbalance, but there is no indication of a clear difference between the groups; the differences do not clearly appoint a possible bias in favor of the treated units. In terms of the school characteristics, the negative coefficient of teacher's room is opposed to the positive coefficients of library and dinning hall, leaving no clear indication if the school infrastructure of the treated schools, before the program, is better or worse when compared to the control schools. In the case of the students' characteristics, the negative coefficient of the variable that indicates whether the student attended early childhood education would be an unfavorable indication for the full-time high school sample, as studies indicate that early childhood education would be a predictor of positive educational outcomes throughout life (HECKMAN; KARAPAKULA, 2019b; CAMPBELL et al., 2012). Anyway, as we cannot rule out the presence of some difference between the groups, we will consider all

these characteristics in our empirical analysis.

We have also tested whether there would be selective attrition between the final sample and the initial sample, that is, the sample of students who had some information for the baseline instruments, considering all students with information for each score of the socioemotional assessment, cognitive tests and variables of the socioeconomic questionnaire. For the sake of this analysis, we regressed each of the measures in a treatment dummy, a variable that indicates if the student is part of the final sample or not and an interaction variable between both (which will indicate the presence of selective bias or not).

Table 3.3 presents the results of this exercise and presents evidence of some attrition in the case of the socioemotional measure scores for Self-Management and Emotional Resilience. This underlines the need to take these measures into account in our empirical exercises.

Table 3.3 – Test of Selective Attrition - Final Sample

	Treat x Final Sample	**Treat**	**Final Sample**	**N**
Age	-0.270	0.291*	-0.247**	1,198
Race: White	0.059	-0.036	-0.052	1,187
Race: Brown	-0.052	0.047	0.062	1,187
Woman	-0.009	0.049	0.050	1,202
Retention	-0.196	0.085	-0.105	1,186
Early Childhood Education	0.039	-0.090***	0.029*	1,189
Mother's Schooling: Elementary School	0.009	-0.016	0.011	1,187
Mother's Schooling: High School	-0.053	0.032	0.083*	1,187
Mother's Schooling: College or more	0.022	-0.017	-0.057*	1,187
Socioeconomic Status	-0.016	-0.017	0.019	1,118
Agreeableness (t-1)	0.223	-0.032	0.037	1,175
Self-Management (t-1)	0.319***	-0.105	-0.061	1,175
Engagement with Others (t-1)	0.200	0.019	-0.078	1,175
Emotional Resilience (t-1)	0.179*	-0.076	-0.068	1,175
Openness (t-1)	0.120	0.055	-0.048	1,175
Language Test (t-1)	0.122	0.016	0.164	1,180
Math Test (t-1)	0.125	-0.115	-0.055	1,180

Note: Results estimated using ordinary least squares, with a set of variables for school characteristics and standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

3.3 Descriptive Analysis and Results

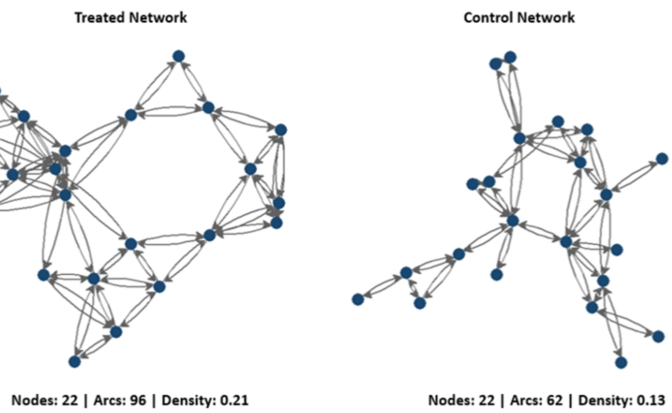
3.3.1 Descriptive Analysis

The main goal in our analysis is to investigate whether the implementation of the full-time high school education policy can change the structure of the friendship network of the attended students. We will use four measures of students centrality, which are: "Closeness"," Betweenness", "Degree" and "Katz-Bonacich". All these measures are properly defined in Appendix C1.

Although a descriptive analysis of all possible variations observed in the social network structures between the pre-and post-treatment periods is not possible, as we do not have information on the friendship links for the previous period, if we analyze the final structure of our 44 networks[14], we can extract some preliminary information about the behavior of the networks that joined and did not join the program.

In our sample, overall, the control networks contain a slightly higher number of students than the treated networks[15], but the data show that this is not reflected in friendship links among their students. Let us look at an example of the picture of friendship relations for one school in each group:

Figure 3.1 – Network Example



Note that both networks contain the same number of students, but the similarity stops there. In the school that adopted full-time high school education, the students built a much larger number of friendship links, consequently resulting in a denser network[16], but that is not all. Figure 3.1 also shows that, in general, the students at the treated school are in a more central position within their network when compared with the students at the control school.

The pattern observed in this example, to some extent, seems to extend to the other cases in our sample. Table 3.4 suggests that, on average, the treated networks are denser and their students more central in three out of four measures considered.

This is only descriptive evidence, but indicates that the program can affect the network structure. Nevertheless, due to the lack of information for the initial period, we could raise the question if the sampling and the friendship questionnaire applied could not

14 In our case, as the relationships are restricted to students in the same classroom, with only one class per school, each social network represents one school.
15 The mean and median of the number of students in the control schools are, respectively, 20 and 22, against a mean and median of approximately 18 for the treated schools.
16 "Arcs" and "Density" definitions are also available in Appendix C1.

Table 3.4 – Descriptive Statistics - Connectivity Measures

	Treated Network		**Control Network**	
	Mean	Std. Dev.	Mean	Std. Dev.
Number of Students	18.33	5.71	19.93	6.82
Arcs	78.75	34.16	78.54	41.7
Density	0.26	0.07	0.21	0.08
Closeness	0.44	0.11	0.39	0.13
Betweeness	0.08	0.12	0.08	0.13
Outdegree	0.26	0.13	0.21	0.12
Katz-bonacich	0.39	0.25	0.3	0.21

"impair" the control students' network measures. Therefore, we will present some statistics showing that this should not be the case.

First, according to table 3.4, although the average number of students is higher in the case of the control schools, the average links among the students (captured by the Arcs variable) is very close for the two groups. This would indicate that the students of the control schools actually form less links (as captured by the Density variable) or that, by limiting the number of indications to 5 friends, the questionnaire applied to students would represent a restriction on the true network of students, especially the network of controls.

If we look at the number of friends the students originally indicated in the final sample, 63.3% of the treated students indicated five friends, while 62.5% of the controls indicated the maximum number. In other words, this limit does not seem to be a substantial restriction for both types of schools. If considered a restriction, the figures indicate that it could affect the treated networks even a little more than the control networks. Furthermore, a comprehensive analysis of the number of indications shows that, on the one hand, approximately 81.3% of the treated students indicated between four and five close friends, while 73.8% of the control students indicated in the same interval. On the other hand, 17.2% of the control students in the sample did not indicate anyone (and only continued in the database because other students indicated them), while only 6.6% of the treated students did not appoint any tie, suggesting a lesser number of links for the controls as a matter of choice, reflecting the actual network structure.

But, beyond the questionnaire, we could consider that, when constructing the sample, the observations we lost would be affecting the links the control students appointed more than the treated students' links. To test this hypothesis, we have developed three analyses: (i) calculate the number of lost links when comparing the number of friends the students indicated and the number of friends in the initial sample; (ii) calculate the number of lost links when comparing the number of friends indicated and the number of friends in the final database; and (iii) calculate the number of students lost in each network when comparing the final and initial samples. In all three cases, we tested whether

there would be a significant difference between the treated and control students. Note that, in the first two analyses, we are testing whether the observations lost reduce the control students' links of friendship more than those of the treated students. In the final analysis, we are testing whether the number of links of the treated students could seem higher because the treated networks decreased more than the control networks.

The results are presented in table 3.5 and refer to the number of the analysis defined in the above paragraph.

Table 3.5 – Balanced Tests - Missed data

	Treatment		**Control**		**Means Difference**
	Mean	Std. Dev.	Mean	Std. Dev.	
Analysis 1	-0.20	0.48	-0.19	0.51	-0.01
Analysis 2	-1.14	1.08	-1.18	1.25	0.04
Analysis 3	-7.07	3.06	-10.99	5.24	3.92***

*** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Note that there is no significant difference in the number of lost links, but there is a significant difference in the number of lost students. It turns out that, on average, the control networks lose more students and, as the number of lost links is approximately the same between the two groups, this difference should favor the control networks in terms of density, instead of the treated networks.

Finally, we have tested whether some school characteristics could be interfering in this proportionally smaller number of links in the regular high schools. First, we have observed that the number of schools offering the final years of elementary education in the previous year (2017) was equal between the two groups[17], which indicates that the possibility of longer links of friendship is proportionally greater for the control students, as the total number of treated schools in the sample is greater. And we have investigated if the number of high school students, considering both the three years of this stage and only the first year, could be significantly higher in the regular schools, which could indicate that the students have more friendship options beyond their class, reducing the indications of friendship within the class.

Table 3.6 shows that there is no significant difference between the number of enrollments in full-time and regular high schools.

Therefore, as far as we were able to investigate, we found no case that could indicate the presence of an artificial limitation to the structure of the control networks to the detriment of the treated networks.

[17] 17 treated and 17 control schools would also attend the second stage of elementary school, according to data from the 2017 School Census.

Table 3.6 – Balanced Tests - High School Enrollment

	Treatment		Control		Means Difference
	Mean	Std. Dev.	Mean	Std. Dev.	
1st year	158.33	95.08	175.45	112.31	-17.12
1st, 2nd and 3rd years	465.46	318.24	395.6	252.34	79.86

Note: The data source is 2018 School Census. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

3.3.2 Impact of Full Time High School on Network

To investigate our hypothesis on the impact of full-time high school education on the structure of friendship networks, we will follow an empirical strategy close to those used by Banerjee et al. (2018) and Comola e Prina (2014) and we will estimate the following model:

$$M_{ir,t} = \alpha + \beta Treat_r + \gamma' \mathbf{Y_{ir,t-1}} + \delta' \mathbf{X_{ir,t-1}} + \rho' \mathbf{S_{r,t-1}} + \epsilon_{ir,t} \tag{3.1}$$

where $M_{ir,t}$ represents one of the measures of connectivity for student i from network r at time t (follow-up); $Treat_r$ is the treatment dummy; $\mathbf{Y_{ir,t-1}}$ is the vector of student i's socioemotional and cognitive scores measured at baseline (t-1); $\mathbf{X_{ir,t-1}}$ is the vector of the student's socioeconomic characteristics; $\mathbf{S_{r,t-1}}$ is the vector of the school's characteristics[18]; and $\epsilon_{ir,t}$ is the error term of the model. All estimations consider the standard errors clustered at the school level.

It is important to note that we are dealing with an important identification problem. As the treatment was not randomly assigned, the selection for the program may be correlated with observed or non-observed characteristics of the treated units and we need to address this self-selection problem. If the characteristics involving the decision to join the program or not are totally unknown, they will be incorporated into the model error and our estimated β will be biased. On the other hand, if we can include in the model all the characteristics that represent the relevant differences between treated and controlled units, then we can get consistent estimates for our parameter of interest (DUFLO; GLENNERSTER; KREMER, 2007).

In this regard, we believe that, in this study, we have reasons to believe that the available information should be sufficient to assume the hypothesis of selection on observables, which is equivalent to saying that we have an error term with a conditional mean equal to 0 and that we will be able to properly identify the relationship between the treatment and the variation in the structure of the friendship networks (RAVALLION, 2007; TODD, 2007).

[18] The vector of the school's characteristics and the vectors of the students' characteristics and scores include the variables identified in table 3.2.

As mentioned in subsection 3.2.2, the decree of the MEC that establishes the EMTI recommends ranking the participating schools according to some infrastructure and physical capacity requirements, in addition to giving priority to schools located in regions of social vulnerability. To capture these aspects, we are considering in our model several variables that characterize the infrastructure and availability of resources at the schools, including items expressly mentioned in that Decree. We have also included a range of characteristics of students and their family environments, which capture the socioeconomic conditions of the school community. In addition, although we do not have data on the structure of the friendship network for the period prior to the treatment, our specification considers both the students' socioemotional and cognitive scores measured at baseline, under the assumption that past performance measures would be a sufficient statistic for the students' history of non-observable information and innate ability (TODD; WOLPIN, 2003).

Given these considerations, we can proceed to the analysis of the results of the model 3.1.

Table 3.7 – Impact of Full Time High School on Network

	Closeness	**Betweeness**	**Degree**	**Katz-bonacich**
Treatment	0.074***	0.012	0.059***	0.124**
Observations	717	717	717	717

Note: Results estimated using ordinary least squares, with a set of students variables (socioeconomic, socioemotional and cognitive measures at baseline) and a set of variables for school characteristics. Standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

The results suggests that the implementation of the full-time high school can generate an important change in the structure of the friendship networks of the attending students. We observe statistically significant variations in three of the four measures that capture the students' level of centrality within their networks. Although the estimated coefficients may seem very small, we describe the change as relevant because they correspond to effects ranging from 0.46 (Degree) to 0.60 (Closeness) standard deviations.

But if the treatment can influence a reasonable change in the friendship networks, should we not consider the influence of such effects when investigating the possible impacts of the program on the students' academic outcomes? Studies investigating the effects of peers on the students' cognitive scores demonstrate the importance of considering the formation of networks in the calculation of these effects (CARRELL; SACERDOTE; WEST, 2013; GOLDSMITH-PINKHAM; IMBENS, 2013; CHAN; LAM, 2014). Furthermore, Comola e Prina (2014) show that an intervention that provided access to the formal financial market modified the structure of the informal financial transaction network and

that ignoring such changes underestimates the general impact of the intervention.

Thus, in the next subsection, we will investigate a possible mediation effect of the intervention on the students' cognitive test scores, through the social network structure.

3.4 Mediation Effects

As mentioned in the subsection 3.1, different studies have investigated the possible effects of full-time high school education on the students' cognitive test scores and the observed effects show to be rather heterogeneous, with several results indicating no significant effects.

But given the results we have found regarding the influence of the program on the behavior of the students' friendship network, the question that arises is: in some cases, the absence of effects verified by these studies would not be hiding a possible channel through which the program would impact the students' academic outcomes, that is, the social networks' structure?

To investigate the hypothesis of a possible mediation effect through the friendship network, we will estimate a model that explores the influence of network connectivity measures on students' academic outcomes, according to the following equation:

$$Y_{ir,t} = \alpha + \beta Treat_r + \theta M_{ir,t} + \gamma' \mathbf{Y_{ir,t-1}} + \delta' \mathbf{X_{ir,t-1}} + \rho' \mathbf{S_{r,t-1}} + \epsilon_{ir,t} \tag{3.2}$$

Where $Y_{ir,t}$ is the standardized scores on the Language or Mathematics test of student i from class r at time t (follow-up); $Treat_r$ is the treatment dummy; $M_{ir,t}$ represents one of the four centrality measures of student i from network r at time t; $\mathbf{Y_{ir,t-1}}$ is the vector of student i's socioemotional and cognitive scores, measured at baseline (t-1); $\mathbf{X_{ir,t-1}}$ is the vector of the student's socioeconomic characteristics; $\mathbf{S_{r,t-1}}$ is the vector of the school's characteristics; and $\epsilon_{ir,t}$ is the error term of the model. All estimations consider the standard errors clustered at the school level.

For the mere sake of comparison, before presenting the results of equation 3.2, let us estimate a model like equation 3.1 to know the direct result of the treatment for the students' cognitive outcomes[19].

The results of this estimation are presented in table 3.8 and show that, without considering the networks' influence, we did not find any significant effects either for the implementation of full-time high school education, in line with other studies.

[19] Basically we estimate the model of the equation 3.1, only considering students' standardized scores on the Language or Mathematics test, as a dependent variable.

Table 3.8 – Impact of Full Time High School on Cognitive Scores

	Language Test	**Math Test**
Treatment	0.077	0.137
Observations	717	717

Note: Results estimated using ordinary least squares, with a set of variables for student characteristcs (socioeconomic, socioemotional and cognitive measures at baseline) and a set of variables for school characteristics. Standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

We will now analyze the results of the equation model 3.2 presented in tables 3.9 and 3.10. For the Language test score, we did not observe significant results, suggesting that, in this case, centrality measures would not be a channel for mediating the effects of the intervention[20].

Table 3.9 – Impact of Centrality Measures on Language Test Score

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
				Language Test				
Treatment		0.046		0.074		0.057		0.051
Closeness	0.463	0.423						
Betweeness			0.261	0.252				
Degree					0.383	0.350		
Katz-bonacich							0.233	0.214
Observations	717	717	717	717	717	717	717	717

Note: Results estimated using ordinary least squares, with a set of students variables (socioeconomic, socioemotional and cognitive measures at baseline) and a set of variables for school characteristics. Standard errors clustered in school level. The results presented in the even columns consider exactly the model of the equation 3.2, whereas the results of the odd columns follow the equation 3.2, excluding the treatment dummy. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

On the other hand, in the table 3.10, if we observe the odd columns, in which the model does not consider the treatment variable, we find positive and significant results for three of the four centrality measures, indicating that the network structure can be an important predictor of students' math scores. When we move to the even columns, which includes the treatment dummy, we continue to observe positive and significant results for the same measures of centrality, but with some reduction in magnitude, suggesting that such measures would be capturing some effect of the intervention on this score.

Therefore, what is relevant in this analysis is to verify that the network structure seems to be an actual channel through which the program would affect at least the

[20] The results presented in the even columns consider exactly the model of the equation 3.2, whereas the results of the odd columns follow the equation 3.2, excluding the treatment dummy.

Table 3.10 – Impact of Centrality Measures on Math Test Score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Math Test				
Treatment		0.090		0.133		0.103		0.098
Closeness	0.710**	0.632*						
Betweeness			0.381	0.365				
Degree					0.639*	0.579*		
Katz-bonacich							0.348*	0.311*
Observations	717	717	717	717	717	717	717	717

Note: Results estimated using ordinary least squares, with a set of students variables (socioeconomic, socioemotional and cognitive measures at baseline) and a set of variables for school characteristics. Standard errors clustered in school level. The results presented in the even columns consider exactly the model of the equation 3.2, whereas the results of the odd columns follow the equation 3.2, excluding the treatment dummy. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

students' math scores and if we did not merely consider and observe only the results of table 3.8, we might say that full-time high school education does not have any effect whatsoever.

3.5 Conclusions

Using data from high school students at 44 public schools in the State of Sergipe (Brazil), we have investigated if the implementation of a full-time education program can alter the structure of the students' social networks.

Although the program does not have an experimental design, our sample includes students from 24 schools that adhered to the program and from 20 regular high schools who served as a comparison group and which showed to be very similar to the treated schools in terms of school characteristics as well as student characteristics. In any case, to deal with the problem of self-selection, we included these characteristics in the control vector of our econometric model and we believe that they capture very well the factors that influenced the schools' adherence to the new educational model. In addition, the vector of the students' characteristics includes their cognitive and socioemotional scores measured in the period before treatment, which would capture the history of the students' non-observable characteristics and together with the other controls allowed us to assume the hypothesis of selection on observables and obtain consistent estimators for the treatment parameter.

The results of our analysis for our first hypothesis suggest that the policy of expanding school time increases the density of the network and puts the students of these schools in a more central position within their friendship networks, considering three of the four measures of centrality that we analyzed.

Therefore, as we have evidence that the program can affect the networks' structure,

we proceed to investigate our second hypothesis that the variation in the network structure could be a possible channel through which the program would affect the student's cognitive test scores and not considering this possible spillover effect could lead us to underestimated effects in the evaluation of the educational policy.

Thus, we first verified what would be the impact of the implementation of full-time high school on the students' Language and Mathematics test scores, without considering any influence of the network, and we found no significant results for this analysis, as was the case in several other studies on this matter. Next, we estimated a model that explores the possible influence of the students centrality measures on their academic outcomes and observed positive effects for some of these measures in the case of math test score. This evidence appoints that, if the program can affect the network connectivity measures and if these measures can affect the students' academic outcomes, then the network structure would be a mechanism through which the program affects these results so that it cannot be neglected in the assessment of this type of policy.

We consider that our study results offer a valuable contribution to the literature that investigates the effects of school time expansion policies. This type of policy has been increasingly used to offer high-quality teaching and enhance the student's educational performance, without there being a consensus in the literature on the effectiveness of its results. It would also be an excellent contribution to the literature that assesses other educational policies in a broad sense as, as far as we know, no study existed in this area that demonstrated the possible spillover effect of such interventions through the students' social network.

We know that, despite the contribution, the sample and the implementation design of the program are some of the limitations in this study, as we are not dealing with an experiment and our sample is restricted to students from a single Brazilian state. Anyway, the results justify future studies that aim to corroborate the results found and to strengthen the internal and external validity of these results.

Bibliography

ALMEIDA, R. et al. *Assessing the impacts of Mais Educacao on educational outcomes: evidence between 2007 and 2011.* [S.l.]: The World Bank, 2016.

ALMLUND, M. et al. Personality psychology and economics. In: *Handbook of the Economics of Education*. [S.l.]: Elsevier, 2011. v. 4, p. 1–181.

AQUINO, J. M.; KASSOUF, A. L. A ampliação da jornada escolar melhora o desempenho acadêmico dos estudantes? uma avaliação do programa escola de tempo integral da rede pública do estado de são paulo. *Rede de economia aplicada*, 2011.

ARAM, D. Continuity in children's literacy achievements: A longitudinal perspective from kindergarten to school. *First Language*, Sage Publications Sage CA: Thousand Oaks, CA, v. 25, n. 3, p. 259–289, 2005.

ARAM, D. Early literacy interventions: The relative roles of storybook reading, alphabetic activities, and their combination. *Reading and writing*, Springer, v. 19, n. 5, p. 489–515, 2006.

ARAM, D.; LEVIN, I. The role of maternal mediation of writing to kindergartners in promoting literacy in school: A longitudinal perspective. *Reading and Writing*, Springer, v. 17, n. 4, p. 387–409, 2004.

ATHANASOU, J. A. Young people in transition: factors influencing the educational-vocational pathways of australian school-leavers. *Education+ Training*, MCB UP Ltd, v. 43, n. 3, p. 132–138, 2001.

ATTANASIO, O. et al. Estimating the production function for human capital: results from a randomized controlled trial in colombia. *American Economic Review*, v. 110, n. 1, p. 48–85, 2020.

BANERJEE, A. V. et al. Changes in social network structure in response to exposure to formal credit markets. *Available at SSRN 3245656*, 2018.

BATTISTIN, E.; MERONI, E. C. Should we increase instruction time in low achieving schools? evidence from southern italy. *Economics of Education Review*, Elsevier, v. 55, p. 39–56, 2016.

BAUSERMAN, K. L. et al. Kindergarten literacy achievement: The effects of the plato integrated learning system. *Literacy Research and Instruction*, Taylor & Francis, v. 44, n. 4, p. 49–60, 2005.

BLACHMAN, B. A. et al. Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing*, Springer, v. 11, n. 3, p. 239–273, 1999.

BORGHANS, L. et al. The economics and psychology of personality traits. *Journal of human Resources*, University of Wisconsin Press, v. 43, n. 4, p. 972–1059, 2008.

BRAMOULLÉ, Y.; DJEBBARI, H.; FORTIN, B. Identification of peer effects through social networks. *Journal of econometrics*, Elsevier, v. 150, n. 1, p. 41–55, 2009.

CALVÓ-ARMENGOL, A.; PATACCHINI, E.; ZENOU, Y. Peer effects and social networks in education. *The Review of Economic Studies*, Oxford University Press, v. 76, n. 4, p. 1239–1267, 2009.

CAMPBELL, F. A. et al. Adult outcomes as a function of an early childhood educational program: an abecedarian project follow-up. *Developmental psychology*, American Psychological Association, v. 48, n. 4, p. 1033, 2012.

CAPOVILLA, F. Desenvolvimento linguistico na crianca brasileira dos dois aos seis anos: traducao e estandardizacao do peabody picture vocabulary test and language development survey de rescorla. *pesquisa e aplicacao*, v. 1, p. 353–380, 1997.

CAPOVILLA, F.; NEGRÃO, V.; DAMÁZIO, M. Teste de vocabulário auditivo e teste de vocabulário expressivo. *São Paulo: Editora Memnon Edições Cientificas*, 2011.

CARRELL, S. E.; FULLERTON, R. L.; WEST, J. E. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, The University of Chicago Press, v. 27, n. 3, p. 439–464, 2009.

CARRELL, S. E.; SACERDOTE, B. I.; WEST, J. E. From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, Wiley Online Library, v. 81, n. 3, p. 855–882, 2013.

CATTELL, R. B. The scree test for the number of factors. *Multivariate behavioral research*, Taylor & Francis, v. 1, n. 2, p. 245–276, 1966.

CHAN, T. J.; LAM, C. T. Type of peers matters: A study of peer effects of friends, studymates and seatmates on academic performance. *The Society of Labor Economists*, 2014.

CHIKITANI, M. *Peer effects on locus of control*. Dissertação (Mestrado) — Escola de Economia de São Paulo da Fundação Getúlio Vargas, 2015.

CHRISTAKIS, N. A. et al. *An empirical model for strategic network formation*. [S.l.], 2010.

COLEY, R. J. An uneven start: Indicators of inequality in school readiness. policy information report. ERIC, 2002.

COMOLA, M.; PRINA, S. Do interventions change the network? a panel peer-effect model accounting for endogenous network changes. IZA Discussion Paper, 2014.

COSTA, H. C. et al. Emergent literacy skills, behavior problems and familial antecedents of reading difficulties: a follow-up study of reading achievement from kindergarten to fifth grade. *Research in developmental disabilities*, Elsevier, v. 34, n. 3, p. 1018–1035, 2013.

CUNHA, F.; HECKMAN, J. J. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of human resources*, University of Wisconsin Press, v. 43, n. 4, p. 738–782, 2008.

CUNHA, F. et al. Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, Elsevier, v. 1, p. 697–812, 2006.

CUNNINGHAM, A. E.; STANOVICH, K. E. Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental psychology*, American Psychological Association, v. 33, n. 6, p. 934, 1997.

DUFLO, E.; GLENNERSTER, R.; KREMER, M. Using randomization in development economics research: A toolkit. *Handbook of development economics*, Elsevier, v. 4, p. 3895–3962, 2007.

DUNCAN, G. J. et al. School readiness and later achievement. *Developmental psychology*, American Psychological Association, v. 43, n. 6, p. 1428, 2007.

DYNARSKI, M. et al. When schools stay open late: The national evaluation of the 21st century community learning centers program–new findings. *US Department of Education*, ERIC, 2004.

DZEMSKI, A. *An empirical model of dyadic link formation in a network with unobserved heterogeneity*. [S.l.], 2014.

FUCHS, D. et al. Is reading important in reading-readiness programs? a randomized field trial with teachers as program implementers. *Journal of Educational Psychology*, American Psychological Association, v. 93, n. 2, p. 251, 2001.

GANDRA, J. M. F. V. *O impacto da educação em tempo integral no desempenho escolar: uma avaliação do programa mais educação*. Dissertação (Mestrado) — Federal University of Viçosa, 2017.

GOLDSMITH-PINKHAM, P.; IMBENS, G. W. Social networks and the identification of peer effects. *Journal of Business & Economic Statistics*, Taylor & Francis Group, v. 31, n. 3, p. 253–264, 2013.

GONÇALVES, M. B. C.; RAPOSO, I. P. d. A. *Peer Non-Cognitive Self-Perception And School Performance: Estimates Using The Student Friendship Network At Classroom Level*. [S.l.], 2016.

HAHN, Y. et al. Network structure and education outcomes: Evidence from a field experiment in bangladesh. IZA Discussion Paper, 2015.

HANUSHEK, E. A. et al. Does peer ability affect student achievement? *Journal of applied econometrics*, Wiley Online Library, v. 18, n. 5, p. 527–544, 2003.

HART, B.; RISLEY, T. R. *Meaningful differences in the everyday experience of young American children*. [S.l.]: Paul H Brookes Publishing, 1995.

HECKMAN, J. J. Schools, skills, and synapses. *Economic inquiry*, Wiley Online Library, v. 46, n. 3, p. 289–324, 2008.

HECKMAN, J. J.; KARAPAKULA, G. *Intergenerational and Intragenerational Externalities of the Perry Preschool Project*. [S.l.], 2019.

HECKMAN, J. J.; KARAPAKULA, G. *The Perry Preschoolers at late midlife: A study in design-specific inference*. [S.l.], 2019.

HECKMAN, J. J.; STIXRUD, J.; URZUA, S. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, The University of Chicago Press, v. 24, n. 3, p. 411–482, 2006.

HINCAPIE, D. *Do longer school days improve student achievement? Evidence from Colombia.* [S.l.], 2016.

HOXBY, C. *Peer effects in the classroom: Learning from gender and race variation.* [S.l.], 2000.

LAVY, V.; PASERMAN, M. D.; SCHLOSSER, A. Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, Wiley Online Library, v. 122, n. 559, p. 208–237, 2012.

LEE, D. S. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, Wiley-Blackwell, v. 76, n. 3, p. 1071–1102, 2009.

LEE, L. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, Elsevier, v. 140, n. 2, p. 333–374, 2007.

MANGER, T.; EIKELAND, O.-J. On the relationship between locus of control, level of ability and gender. *Scandinavian Journal of Psychology*, Wiley Online Library, v. 41, n. 3, p. 225–229, 2000.

MANSKI, C. F. Identification of endogenous social effects: The reflection problem. *The review of economic studies*, Oxford University Press, v. 60, n. 3, p. 531–542, 1993.

MEYER, E.; KLAVEREN, C. V. The effectiveness of extended day programs: Evidence from a randomized field experiment in the netherlands. *Economics of Education Review*, Elsevier, v. 36, p. 1–11, 2013.

NAKANO, T. C. et al. Battery for assessment of high abilities/giftedness: Analysis conducted using item response theory. *Estudos de Psicologia (Campinas)*, SciELO Brasil, v. 32, n. 4, p. 729–741, 2015.

NEUMAN, S. B. Books make a difference: A study of access to literacy. *Reading Research Quarterly*, Wiley Online Library, v. 34, n. 3, p. 286–311, 1999.

PIRES, T.; URZUA, S. Longer school days, better outcomes. *Northwestern University*, p. 1–55, 2010.

RAVALLION, M. Evaluating anti-poverty programs. *Handbook of development economics*, Elsevier, v. 4, p. 3787–3846, 2007.

ROMANO, J. P.; WOLF, M. Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, Elsevier, v. 113, p. 38–40, 2016.

SACERDOTE, B. Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly journal of economics*, MIT Press, v. 116, n. 2, p. 681–704, 2001.

SANTOS, D.; PRIMI, R. Social and emotional development and school learning: a measurement proposal in support of public policy. *São Paulo, Brazil: Ayrton Senna Institute*, 2014.

SANTOS, D. D. *A importância socioeconômica das características de Personalidade*. [S.l.], 2011.

SILVA, R.; FLORES-MENDOZA, C.; SANTOS, M. Teste de habilidades e conhecimento pré-alfabetização (thcp). *São Paulo: Vetor*, 2013.

SOTO, C. J. et al. Age differences in personality traits from 10 to 65: Big five domains and facets in a large cross-sectional sample. *Journal of personality and social psychology*, American Psychological Association, v. 100, n. 2, p. 330, 2011.

STORCH, S. A.; WHITEHURST, G. J. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental psychology*, American Psychological Association, v. 38, n. 6, p. 934, 2002.

THOMPSON, R. A.; NELSON, C. A. Developmental science and the media: Early brain development. *American Psychologist*, American Psychological Association, v. 56, n. 1, p. 5, 2001.

TODD, P. E. Evaluating social programs with endogenous program placement and selection of the treated. *Handbook of development economics*, Elsevier, v. 4, p. 3847–3894, 2007.

TODD, P. E.; WOLPIN, K. I. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, Wiley Online Library, v. 113, n. 485, p. F3–F33, 2003.

VIGDOR, J.; NECHYBA, T. Peer effects in north carolina public schools. *Schools and the equal opportunity problem*, MIT Press, p. 73–101, 2007.

WASIK, B. A.; BOND, M. A. Beyond the pages of a book: interactive book reading and language development in preschool classrooms. *Journal of educational psychology*, American Psychological Association, v. 93, n. 2, p. 243, 2001.

ZIMMER, R.; HAMILTON, L.; CHRISTINA, R. After-school tutoring in the context of no child left behind: Effectiveness of two programs in the pittsburgh public schools. *Economics of education Review*, Elsevier, v. 29, n. 1, p. 18–28, 2010.

Appendix

APPENDIX A - Appendix from Chapter 1

A1 Measures

Figure A1.1 – Description of Writing Levels

Níveis de conceitualização da escrita

Nível pré-silábico 1

- ✓ as garatujas - linhas onduladas, contínuas ou fragmentadas, bolinhas...;
- ✓ as escritas unigráficas - escrevem somente uma letra para cada palavra;
- ✓ e as escritas que são consideradas completas apenas quando alcançam o limite do papel.

Nível pré-silábico 2

- ✓ escritas com diferenciação intrafigural (variação qualitativa das letras dentro de cada item da lista).

Nível pré-silábico 3

- ✓ escritas com diferenciação interfigural (produzindo diferenças de tipo quantitativo e/ou qualitativo entre os itens da lista).

Nível silábico 1

- ✓ escreve uma letra para cada sílaba sem valor sonoro convencional.

Nível silábico 2

- ✓ escreve uma letra para cada sílaba com valor sonoro convencional e com predominância das vogais.

Nível silábico 3

- ✓ escreve uma letra para cada sílaba com valor sonoro convencional e com presença de vogais e consoantes pertinentes.

Nível silábico-alfabético

- ✓ escreve usando ora uma letra para representar cada sílaba, ora mais de uma letra para essa representação, com predominância de letras com valor sonoro convencional.

Nível alfabético

- ✓ escreve alfabeticamente, mas com falhas na ortografia.

Nível alfabético ortográfico

- ✓ escreve alfabeticamente e com poucas falhas na ortografia.

Figure A1.2 – PPVT Measure - Page Example

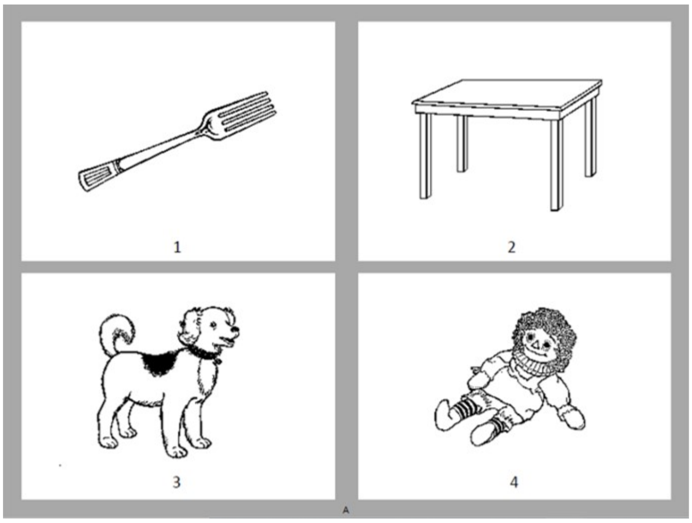


Figure A1.3 – Reading Measure - Examples of Subtests

MARCAREM AS PALAVRAS: PIANO, VENENO, BEBEU, DO, QUEM.

QUESTÃO 1: LEITURA DE UMA PARLENDA

LÁ EM CIMA DO PIANO

TEM UM COPO DE VENENO

QUEM BEBEU MORREU

O AZAR FOI SEU

marquem as palavras, **MACACO e GIRAFA**.

QUESTÃO 2: LEITURA DE UMA LISTA

CACHORRO

MACACO

GATO

PAPAGAIO

LEÃO

GIRAFA

Figure A1.4 – Writing Measure - Examples of Subtests

Instruções do Teste [Questão 1]

1. Escreva seu nome na primeira página da prova

[mostre para a criança em sua folha de prova o lugar indicado]

Instruções do Teste [Questão 2]

Esta questão pretende avaliar o nível de conhecimento dos alunos sobre o sistema de escrita, ou seja, procura verificar como cada um compreendeu até este momento o funcionamento e as regras de geração da escrita.

Instruções de aplicação:

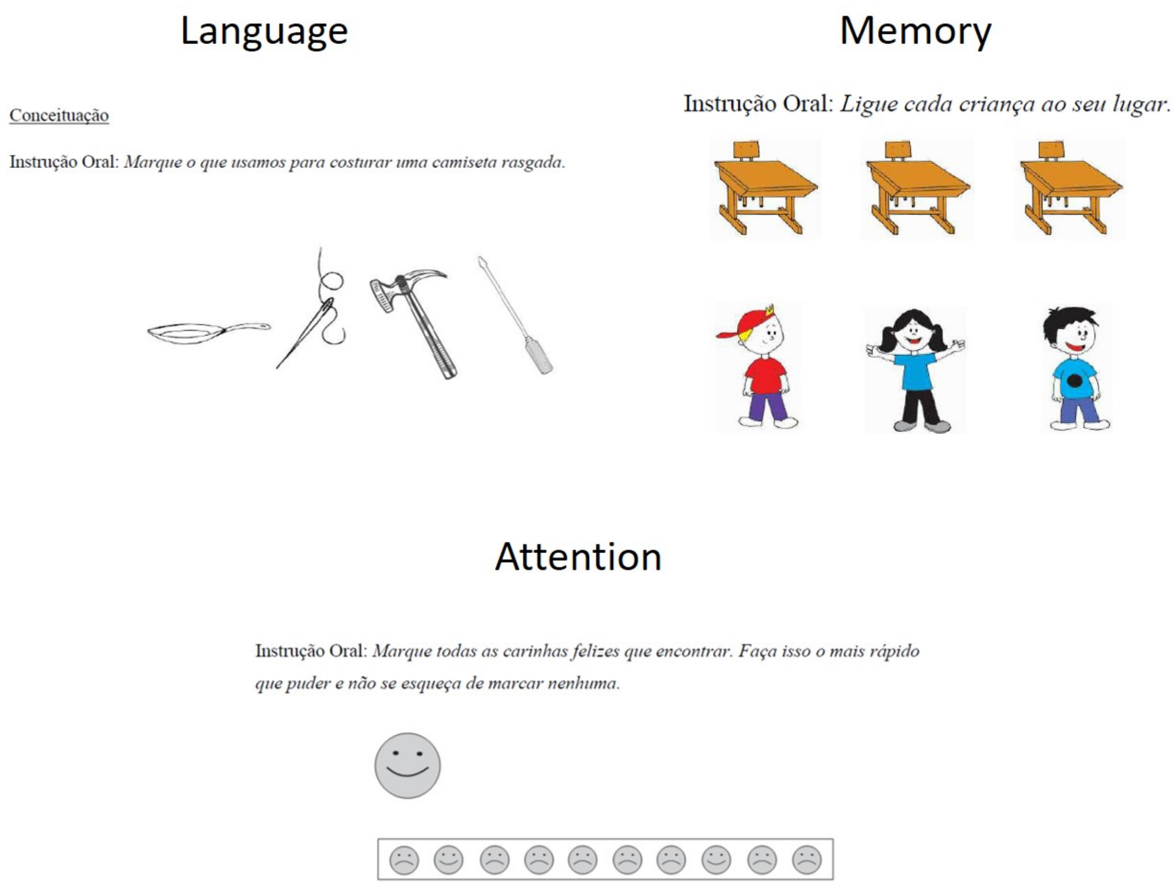
A professora Helena escreveu com seus alunos uma lista dos brinquedos preferidos do grupo.

Leia a comanda e as palavras, em seguida, dite uma de cada vez, dando tempo para a criança escrever.

PATINETE	BONECA	BOLA	PIPA
PALAVRA 1	PALAVRA 2	PALAVRA 3	PALAVRA 4

Figure A1.5 – Expressive Vocabulary Measure - Examples of Figures

Figure A1.6 – THCP Measure - Examples of Subtests



A2 Testing Selective Attrition

Table A2.1 – Test of Selective Attrition - TS&A + P Sample

	Treat x TS&A+P	**Treat**	**TS&A+P**	**N**
PPVT	2.873	0.186	3.343	3,952
Reading a “Parlenda”	0.168	0.093	-0.123	3,942
Reading a List	0.144*	-0.061	-0.098	3,942
Writing the name	0.109	-0.037	-0.016	3,950
Writing words	0.314	0.253	-0.136	3,941
THCP - Language	0.141	0.174	0.195	1,969
THCP - Memory	0.619**	-0.301	0.071	1,960
THCP - Attention	1.043	-0.004	0.180	1,972
THCP - Total	1.838	-0.180	0.535	1,972
Expressive Vocabulary	1.574	-0.812	0.679	1,976

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs and standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Table A2.2 – Test of Selective Attrition - TS&A + T Sample

	Treat x TS&A+T	**Treat**	**TS&A+T**	**N**
PPVT	-0.143	2.049	6.452***	3,952
Reading a “Parlenda”	-0.049	0.241**	0.260***	3,942
Reading a List	0.069	-0.011	0.088**	3,942
Writing the name	0.066	-0.011	0.076*	3,950
Writing words	0.226	0.313	0.468	3,941
THCP - Language	0.081	0.205	0.292	1,969
THCP - Memory	0.297	-0.107	0.403*	1,960
THCP - Attention	-0.710	1.154	1.958***	1,972
THCP - Total	-0.300	1.201	2.734***	1,972
Expressive Vocabulary	-3.015**	2.209*	3.398***	1,976

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs and standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

A3 Robustness Test

Table A3.1 – Test of Selective Attrition - TS&A + C Sample

	Treat x TS&A+C	**Treat**	**TS&A+C**	**N**
PPVT	0.108	1.649	4.401***	3,952
Reading a "Parlenda"	-0.115	0.274**	0.129	3,942
Reading a List	0.053	0.000	-0.034	3,942
Writing the name	-0.008	0.036	0.059**	3,950
Writing words	0.148	0.357	0.154	3,941
THCP - Language	0.243	0.097	0.099	1,969
THCP - Memory	0.321	-0.123	0.172	1,960
THCP - Attention	0.432	0.350	0.767	1,972
THCP - Total	1.043	0.263	1.069	1,972
Expressive Vocabulary	-1.709*	1.184	1.750***	1,976

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs and standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Table A3.2 – Balance Check

		TS&A + C Sample	
School Data	Sewer		0.127
	Teachers' room		0.048
	Computer Lab		0.013
	Library		-0.099
	Reading Room		0.108
	Playgroung		-0.176
	Bathroom		0.090
	Secretary room		0.127
	Dinning Hall		0.225*
	Internet		0.059
	N. Classrooms (prop.)		-0.369
	N. Adm. Computers (prop.)		7.806
	N. Students Computers (prop.)		0.053
	N. Employees (prop.)		-2.080**
	N. Students		-0.000
Student Data	Age: 5 or younger	-0.009	-0.015
	Girl	-0.015	-0.033**
	Lives up to 3 people	-0.014	-0.058**
	Assessor	0.031	-0.003
	White	-0.010	-0.040
	Age attended school: 3 or younger	-0.056	-0.070*
	Live with both parents	-0.060**	-0.109***
	Mother works out	-0.027	-0.032
	Mother's schooling: at least high school	0.028	0.038
	Socioeconomic status	0.132**	0.021
	N	2,484	2,484
	F Test	0.205	0

Note: The data source for "School Data" is Censo Escolar 2015. Balance check estimated using ordinary least squares, with standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Table A3.3 – Impact of the Reading and Writing Program: Robustness test

	TS&A Sample	**TS&A + C Sample**
PPVT	0.062	0.039
Reading a "Parlenda"	**0.130****	**0.138*****
Reading a List	0.072	0.057
Writing the name	0.065	0.014
Writing words	**0.161****	**0.119****
THCP - Language	0.096	0.107*
THCP - Memory	0.037	0.052
THCP - Attention	0.087	0.119*
THCP - Total	0.096	0.128*
Expressive Vocabulary	-0.023	0.006
N	3,259	2,484
N: THCP Total	1,628	1,231
N: Vocabulary	1,624	1,248

Note: Results estimated using ordinary least squares, with a set of dummies for school pairs. Column "TS&A Sample" also includes a dummy indicating the kind of assessor and Column "TS&A + C Sample" also includes a set of variables for school and students characteristics as indicated at balance check. Standard errors clustered in school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1. Results in bold means adjusted p-values at least <0.1.

APPENDIX B - Appendix from Chapter 2

B1 Socioemotional and Cognitive Measures - Examples

Figure B1.1 – SENNA Examples

Fator	**Abaixo, mostramos algumas características pessoais que podem ou não ter a ver com você. Para responder às perguntas, pense em como você é/se sente/se comporta na maioria das situações.**	**1** **Nada** Não tem nada a ver comigo	**2** **Pouco** Tem um pouco a ver comigo	**3** **Mais ou menos** Às vezes tem, às vezes não tem a ver comigo	**4** **Muito** Tem muito a ver comigo	**5** **Totalmente** Tem tudo a ver comigo
Conscienciosidade	Sou um(a) aluno(a) que se esforça.	☐	☐	☐	☐	☐
Estabilidade Emocional	Perco a cabeça com facilidade.	☐	☐	☐	☐	☐
Amabilidade	Sou amável e legal com quase todo mundo.	☐	☐	☐	☐	☐
Abertura à novas experiências	Tenho ideias novas e originais.	☐	☐	☐	☐	☐
Locus de Controle	Sinto que é inútil me esforçar na escola porque a maioria dos alunos é mais inteligente do que eu.	☐	☐	☐	☐	☐
Extroversão	Gosto de conversar.	☐	☐	☐	☐	☐

Figure B1.2 – Verbal Test Example

Instrução: Essa prova é constituída por frases nas quais a última palavra está faltando. Sua tarefa é encontrar a palavra correta que completa a frase. Veja esse exemplo:

Ex. A: Dia está para **noite** como **pequeno** está para ___________

A. **Luz** B. **Grande** C. **Forte** D. **Criança** E. **Escuro**

Figure B1.3 – Abstract Test Example

Instrução: Nos próximos exercícios a primeira figura (1) sofre uma transformação passando à figura (2). Você deverá aplicar esta mesma transformação à figura (3) e descobrir em qual dos desenhos (A, B, C, D ou E) essa figura se transformaria. Na sua resposta você deve indicar qual figura ficaria em 4. Veja o exemplo:

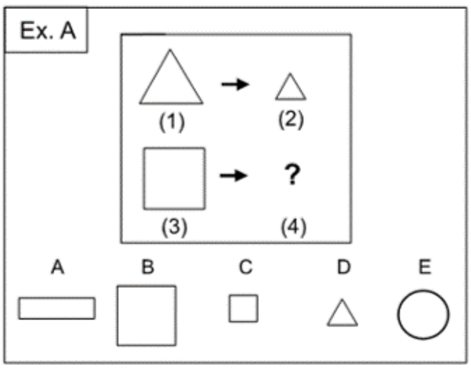


Figure B1.4 – Spatial Test Example

Instrução: Agora você terá que resolver problemas com um cubo. Em cada problema você encontrará um cubo que está girando. Você verá o mesmo cubo em três posições diferentes. Conforme ele vai rodando os seus lados vão mudando. Você deve olhar os lados que estão aparecendo e descobrir para que direção está girando. Você deve descobrir qual seria a próxima figura que apareceria se ele continuasse a girar da mesma forma, e escolher a alternativa (A,B,C,D ou E) te o desenho correto.

Veja o Exemplo 1:

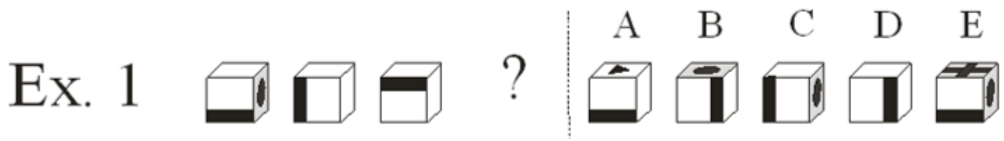


Figure B1.5 – Numeric Test Example

Instrução: Em cada série, os números aparecem de acordo com uma determinada ordem. Nessa atividade, sua tarefa é descobrir os dois próximos números que completam a série apresentada. Veja o exemplo abaixo. Analise a série de números e descubra quais os dois números que viriam a seguir nos locais marcados com pontos de interrogação (? ?):

Exemplo A: 1 3 5 7 9 ? ?

Figure B1.6 – Logic Test Example

Instrução: Esta prova é constituída por problemas de raciocínio lógico. Nela você deverá escrever a sua resposta nas linhas da folha de resposta. Veja os exemplos abaixo:

Exemplo A

Joana e a Paula são amigas. Uma possui um cão e a outra possui um gato. A Paula tem no seu cão um grande amigo. A quem pertence cada um dos animais?

B2 Socioemotional and Cognitive Measures - Factor Analysis

As previously mentioned, we estimated the factor loadings of each of the SENNA's items using a common factor analysis with two factors. After estimating the factor loadings we applied an oblique "promax" rotation, which admits the correlation between the factors, to identify which of the factors would be associated with each item[21]. Below we present the results obtained for the factor loadings, with the items already separated by factor:

[21] Note that it was not necessary to perform this stage of the exploratory analysis in the case of the cognitive measures as we identified only one factor composed of five measures.

Table B2.1 – Factor Loadings

Items loaded on Factor 1			
SENNA's domains	**SENNA's item**	**Factor 1**	**Factor 2**
Agreeableness	ise_27	0.707	0.094
Agreeableness	ise_32	0.604	-0.027
Agreeableness	ise_37	0.543	0.027
Agreeableness	ise_41	0.802	0.159
Agreeableness	ise_46	0.313	-0.229
Agreeableness	ise_50	0.749	0.116
Agreeableness	ise_55	0.392	-0.271
Agreeableness	ise_59	0.657	-0.063
Agreeableness	ise_68	0.639	-0.03
Agreeableness	ise_75	0.637	-0.115
Conscientiousness	ise_9	0.722	0.185
Conscientiousness	ise_12	0.495	0.024
Conscientiousness	ise_15	0.789	0.308
Conscientiousness	ise_18	0.79	0.295
Conscientiousness	ise_21	0.773	0.212
Conscientiousness	ise_23	0.562	0.039
Conscientiousness	ise_25	0.89	0.276
Conscientiousness	ise_29	0.709	0.103
Conscientiousness	ise_34	0.767	0.127
Conscientiousness	ise_43	0.6	-0.024
Conscientiousness	ise_47	0.735	0.117
Extroversion	ise_10	0.547	0.01
Extroversion	ise_13	0.327	-0.132
Extroversion	ise_16	0.652	0.047
Extroversion	ise_22	0.251	-0.223
Extroversion	ise_24	0.371	-0.063
Extroversion	ise_73	0.75	0.063
Extroversion	ise_77	-0.487	0.109
Extroversion	ise_80	0.614	-0.058
Extroversion	ise_83	-0.518	0.091
Extroversion	ise_86	0.445	-0.176
Extroversion	ise_90	0.666	-0.118
Emotional Stability	ise_11	0.307	-0.034
Emotional Stability	ise_14	0.367	-0.014
Emotional Stability	ise_17	0.42	-0.059
Emotional Stability	ise_20	0.394	0.002
Emotional Stability	ise_49	0.63	-0.058
Emotional Stability	ise_58	0.56	-0.09
Internal Locus of Control	ise_54	-0.352	0.287
Internal Locus of Control	ise_74	-0.456	0.245
Internal Locus of Control	ise_78	0.66	-0.007
Internal Locus of Control	ise_89	-0.371	0.252
Openness to Experience	ise_28	0.621	-0.002
Openness to Experience	ise_33	0.724	0.156
Openness to Experience	ise_42	0.474	-0.123
Openness to Experience	ise_45	0.402	-0.207
Openness to Experience	ise_51	0.625	-0.072
Openness to Experience	ise_60	0.621	-0.031
Openness to Experience	ise_64	0.713	0.055
Openness to Experience	ise_72	0.636	-0.105
Openness to Experience	ise_76	0.563	-0.134
Openness to Experience	ise_79	0.456	-0.093
Openness to Experience	ise_82	0.504	-0.164
Openness to Experience	ise_85	0.419	-0.24
Openness to Experience	ise_87	0.614	-0.1

Items loaded on Factor 2			
SENNA's domains	**SENNA's item**	**Factor 1**	**Factor 2**
Agreeableness	ise_63	0.145	0.683
Agreeableness	ise_71	-0.031	0.613
Conscientiousness	ise_38	0.193	0.689
Conscientiousness	ise_52	0.169	0.675
Conscientiousness	ise_56	-0.167	0.479
Conscientiousness	ise_57	-0.004	0.673
Conscientiousness	ise_61	0.072	0.598
Conscientiousness	ise_65	0.127	0.708
Conscientiousness	ise_69	0.159	0.762
Extroversion	ise_19	-0.053	-0.44
Extroversion	ise_88	-0.11	0.482
Extroversion	ise_91	0.241	-0.362
Emotional Stability	ise_26	0.095	0.591
Emotional Stability	ise_30	0.054	0.573
Emotional Stability	ise_35	0.044	0.634
Emotional Stability	ise_39	0.085	0.684
Emotional Stability	ise_40	0.119	0.59
Emotional Stability	ise_44	0.135	0.607
Emotional Stability	ise_48	0.103	0.657
Emotional Stability	ise_53	0.131	0.804
Emotional Stability	ise_62	0.063	0.696
Emotional Stability	ise_66	-0.31	0.372
Internal Locus of Control	ise_31	0.131	0.624
Internal Locus of Control	ise_36	0.182	0.537
Internal Locus of Control	ise_67	-0.08	0.594
Internal Locus of Control	ise_70	0.106	0.719
Internal Locus of Control	ise_81	-0.229	0.427
Internal Locus of Control	ise_84	-0.235	0.455

To name the factors we extracted, we considered the proportion of the absolute values of the items' loadings of a given domain in the SENNA instrument on the total of loading in our factor to which they are related. Thus, in the case of Factor 1, the most representative SENNA's domains are Conscientiousness (25%) and Openness to New Experiences (23%), followed by Agreeableness (19%) and Extroversion (18%), which is why we call it "Reliability&Curiosity". In the case of Factor 2, the SENNA's domain proportionally most relevant to explain the factor is Emotional Stability (37%), followed by Conscientiousness (27%) and Internal Locus of Control (20%), which led us to call it "Self-Management"[22]

B3 Peer Effect - Centrality Approach

Table B3.1 – Effects using Centrality Approach (own centrality)

	Reliability&Curiosity	**Self-Management**
$\mathbf{M_{r,i}}$		
Closeness	0.238	-0.306
Betweeness	0.149	-0.292
Degree	0.098	-0.121
Katz-bonacich	0.137	-0.082
Observations	1260	

Note: Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age and socioemotional measures), both on student and group level. Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

[22] According to our analysis, the items' loadings related to the SENNA's domains that were not mentioned, in the two cases, do not represent even 10% of the respective factor.

Table B3.2 – Effects using Centrality Approach (own and peers centrality)

	Reliability&Curiosity	Self-Management
$\mathbf{M_{r,i}}$		
Closeness	0.264	-0.306
Betweeness	0.271	-0.144
Degree	0.110	-0.139
Katz-bonacich	0.135	-0.072
$\mathbf{\overline{M}_{r,j}}$		
Closeness	0.360	0.001
Betweeness	0.354	0.432*
Degree	0.070	-0.105
Katz-bonacich	0.039	-0.248
Observations	1260	

Note: Note: Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age and socioemotional measures), both on student and group level. Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

Table B3.3 – Effects using Centrality Approach (heterogeneous centrality)

	Reliability&Curiosity	Self-Management
$\mid \mathbf{M_{r,i}} - \mathbf{\overline{M}_{r,j}} \mid$		
Closeness	-0.025	0.624
Betweeness	0.117	0.819*
Degree	-0.148	0.320
Katz-bonacich	-0.030	0.272
$\mathbf{1}_{\{\mathbf{M_{r,i}} > \mathbf{\overline{M}_{r,j}}\}}$		
Closeness	0.027	0.030
Betweeness	0.099	0.108*
Degree	0.019	0.056
Katz-bonacich	0.093	-0.013
$\mathbf{1}_{\{\mathbf{M_{r,i}} > \mathbf{\overline{M}_{r,j}}\}} \cdot \mid \mathbf{M_{r,i}} - \mathbf{\overline{M}_{r,j}} \mid$		
Closeness	0.028	-0.490
Betweeness	-0.226	-0.743*
Degree	-0.044	-0.248
Katz-bonacich	-0.479	0.291
Observations	1260	

Note: Note: Each centrality coefficient is obtained from a different regression, estimated using ordinary least squares, with a set of dummies for classrooms and a set of exogenous variables (gender, age and socioemotional measures), both on student and group level. Standard errors clustered at school level. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

APPENDIX C - Appendix from Chapter 3

C1 Network Connectivity Measures

In this appendix, we give the definition of the connectivity measures used in the paper.

Be R the number of networks in our sample, indexed by r, the relationships between students are captured by an adjacent matrix $\mathbf{G_r}$, with $g_{r,ij} = 1$ if i and j are friends, and $g_{r,ij} = 0$ otherwise. Since a friendship is a reciprocal relationship, we consider that there is a link between the node i and node j if i chose j as one of his best friends or if j chose i, that is, $g_{r,ij} = g_{r,ji}$ and $g_{r,ii} = 0$. As we defined, matrix $\mathbf{G_r}$ is symmetrical and the elements of its diagonal are equal to zero.

1. ***Arcs***: Total number of links between two students. Note that, as we consider friendship to be a reciprocal relationship, the arcs are bidirectional, that is, if students i and j form a friendship tie this means that we have two links, one from i with j and one from j with i.

2. ***Density***: Ratio between the Arc and the total number of possible links in a network $(n_r * (n_r - 1))$.

3. ***Closeness Centrality***: measures how close a node is to other nodes in a given network $\mathbf{g_r}$.
   $Closeness_i(\mathbf{g_r}) = \frac{1}{\sum_{k \neq i} \Delta_{ik}} * (n_r - 1)$
   In which Δ_{ik} is the length of the shortest path between node i and node k (the length between two nodes directly linked is 1).

4. ***Betweenness centrality***: in a given node, this measure is equal to the number of the shortest paths between each pair of nodes in the network passing through this node, that is, for a given network $\mathbf{g_r}$:
   $Betweeness_i(\mathbf{g_r}) = \sum_{s \neq i} \sum_{t \neq i} \frac{\sigma_{st(i)}}{\sigma_{st}}$
   In which σ_{st} is the total number of the shortest paths between node s and node t and $\sigma_{st(i)}$ is the number of these paths that passes through node i.

5. ***Degree Centrality***: Measures the number of links of a given node.
   $Degree_i(\mathbf{g_r}) = \frac{1}{(n_r - 1)} * \sum_{j=1}^{n_r} g_{r,ij}$

6. ***Katz-Bonacich Centrality***: measures the importance of a given node within a social network. To assess how well a node is located, we use the weighted sum of

the paths departing from this node. A certain value $\beta g_{r,i}$ is assigned to each node, a value that is proportional to its connectivity $g_{r,i} = \sum_{j=1}^{n_r} g_{r,ij}$, this value being increased with the value of the node located at a distance link of i, two distance links, and so on, discounted by a factor that decreases as the distance increases, that is, the value of the node located s distance links i is weighted by β^{s-1}. Be $\mathbf{1}$ a vector of ones, the vector of *Katz-Bonacich* centrality can be defined as:

$$\mathbf{b}(\mathbf{g_r}, \beta) = \beta \mathbf{G_r} \mathbf{1} + \beta^2 \mathbf{G_r}^2 \mathbf{1} + \beta^3 \mathbf{G_r}^3 \mathbf{1} + ... = \sum_{s=0}^{\infty} \beta^s \mathbf{G_r}^s \cdot (\beta \mathbf{G_r} \mathbf{1})$$

And for β small enough[23], this infinite sum converges to a finite value, which gives us:

$$\mathbf{b}(\mathbf{g_r}, \beta) = (\mathbf{I_r} - \beta \mathbf{G_r})^{-1} \cdot (\beta \mathbf{G_r} \mathbf{1})$$

[23] To ensure that $(\mathbf{I_r} - \beta \mathbf{G_r})^{-1}$ to be invertible we need $\beta < \frac{1}{\lambda_{máx}(\mathbf{G_r})}$, where $\lambda_{máx}(\mathbf{G_r})$ is the highest eigenvalue of the network $\mathbf{G_r}$, and a sufficient condition would be $\beta < \frac{1}{n_r - 1}$ (CALVÓ-ARMENGOL; PATACCHINI; ZENOU, 2009).