# Impact of Full Time High School on Network Connectivity Measures

Natália Zanetti Marchi [*] Cristine Pinto[†] Vladimir Ponczek[‡]

São Paulo School of Economics - FGV

**Abstract**

The main goal of the present study is to 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.

**Keywords:** full-time education, changes on network structure, mediation effects

[*] natalia.marchi@fgv.br
[†] cristine.pinto@fgv.br
[‡] vladimir.ponczek@fgv.br

# 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 extra-curricular period and did not observe significant effects on the grades of elementary or high school students. In turn, Zimmer et al. (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 and Van Klaveren (2013) found no effect on the mathematics or language tests of elementary school students in a Dutch program and Battistin and 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 and 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 and 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 et al., 2009; Sacerdote, 2001; Carrell et al., 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 and Imbens, 2013; Christakis et al., 2010; Carrell et al., 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 et al. (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 and 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].

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

[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 et al. (2013), for example, show that, in the investigated intervention, the groups were constructed according to the premise that students 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.

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, we briefly analyze the structure of the observed social networks and investigate the impact of the program on the network structure. In section 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.

# 2 Context and Data

## 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.

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 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.

[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.

## 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. 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 informa-

[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 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.

tion 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 comparison group, using data from the 2017 School Census[13]. Table 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

[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.

will call the initial sample.

Table 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 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.

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 and Karapakula, 2019; 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.

Table 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.

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 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: 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 Descriptive Analysis and Results

## 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.

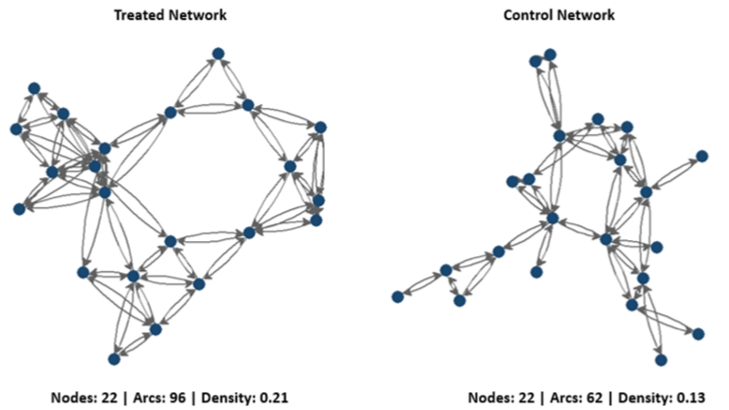
Although a descriptive analysis of all possible variations observed in the social network structures between the pre-and post-treatment periods is not possible, as we do not have information on the friendship links for the previous period, if we analyze the final structure of our 44 networks[14], we 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

[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.

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 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 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 4 suggests that, on average, the treated networks are denser and their students more central in three out of four measures considered.

Table 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 |

This is only descriptive evidence, but indicates that the program can affect the network structure. Nev-

[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.

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

First, according to table 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 5 and refer to the number of the analysis defined in the above paragraph.

Note that there is no significant difference in the number of lost links, but there is a significant difference

Table 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.

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 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.

Table 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.

### 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 and Prina (2014) and we will estimate the following model:

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

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 $\beta$ 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 et al., 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).

As mentioned in subsection 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

[18]The vector of the school's characteristics and the vectors of the students' characteristics and scores include the variables identified in table 2.

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 and Wolpin, 2003).

Given these considerations, we can proceed to the analysis of the results of the model 1.

Table 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 et al., 2013; Goldsmith-Pinkham and Imbens, 2013; Chan and Lam, 2014). Furthermore, Comola and 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.

# 4 Mediation Effects

As mentioned in the subsection 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:

$$Y_{ir,t} = \alpha + \beta Treat_r + \theta M_{ir,t} + \gamma' \mathbf{Y_{ir,t-1}} + \delta' \mathbf{X_{ir,t-1}} + \rho' \mathbf{S_{r,t-1}} + \epsilon_{ir,t} \tag{2}$$

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 2, let us estimate a model like equation 1 to know the direct result of the treatment for the students' cognitive outcomes[19].

The results of this estimation are presented in table 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.

Table 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.

[19]Basically we estimate the model of the equation 1, only considering students' standardized scores on the Language or Mathematics test, as a dependent variable.

We will now analyze the results of the equation model 2 presented in tables 9 and 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 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 2, whereas the results of the odd columns follow the equation 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 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.

Table 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 2, whereas the results of the odd columns follow the equation 2, excluding the treatment dummy. *** means p-value <0.01, **p-value <0.05 and * p-value <0.1.

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 students' math scores and if we did not merely

[20] The results presented in the even columns consider exactly the model of the equation 2, whereas the results of the odd columns follow the equation 2, excluding the treatment dummy.

consider and observe only the results of table 8, we might say that full-time high school education does not have any effect whatsoever.

# 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.

## References


**Almeida, Rita, Antonio Bresolin, Bruna Borges, Karen Mendes, and Naercio Menezes-Filho**, *Assessing the impacts of Mais Educacao on educational outcomes: evidence between 2007 and 2011*, The World Bank, 2016.

**Aquino, Juliana Maria and Ana Lúcia Kassouf**, "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.

**Banerjee, Abhijit V, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson**, "Changes in social network structure in response to exposure to formal credit markets," *Available at SSRN 3245656*, 2018.

**Battistin, Erich and Elena Claudia Meroni**, "Should we increase instruction time in low achieving schools? Evidence from Southern Italy," *Economics of Education Review*, 2016, *55*, 39–56.

**Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin**, "Identification of peer effects through social networks," *Journal of econometrics*, 2009, *150* (1), 41–55.

**Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou**, "Peer effects and social networks in education," *The Review of Economic Studies*, 2009, *76* (4), 1239–1267.

**Campbell, Frances A, Elizabeth P Pungello, Margaret Burchinal, Kirsten Kainz, Yi Pan, Barbara H Wasik, Oscar A Barbarin, Joseph J Sparling, and Craig T Ramey**, "Adult outcomes as a function of an early childhood educational program: an Abecedarian Project follow-up.," *Developmental psychology*, 2012, *48* (4), 1033.

**Carrell, Scott E, Bruce I Sacerdote, and James E West**, "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 2013, *81* (3), 855–882.

**_ , Richard L Fullerton, and James E West**, "Does your cohort matter? Measuring peer effects in college achievement," *Journal of Labor Economics*, 2009, *27* (3), 439–464.

**Chan, Tszkin Julian and Chungsang Tom Lam**, "Type of Peers Matters: A Study of Peer Effects of Friends, Studymates and Seatmates on Academic Performance," *The Society of Labor Economists*, 2014.

**Christakis, Nicholas A, James H Fowler, Guido W Imbens, and Karthik Kalyanaraman**, "An empirical model for strategic network formation," Technical Report, National Bureau of Economic Research 2010.

**Comola, Margherita and Silvia Prina**, "Do interventions change the network? a panel peer-effect model accounting for endogenous network changes," 2014.

**Duflo, Esther, Rachel Glennerster, and Michael Kremer**, "Using randomization in development economics research: A toolkit," *Handbook of development economics*, 2007, *4*, 3895–3962.

**Dynarski, Mark, Susanne James-Burdumy, Mary Moore, Linda Rosenberg, John Deke, and Wendy Mansfield**, "When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program–New Findings.," *US Department of Education*, 2004.

**Gandra, Juliana Mara Fátima Viana**, "O impacto da educação em tempo integral no desempenho escolar: uma avaliação do programa mais educação," Master's thesis, Federal University of Viçosa 2017.

**Goldsmith-Pinkham, Paul and Guido W Imbens**, "Social networks and the identification of peer effects," *Journal of Business & Economic Statistics*, 2013, *31* (3), 253–264.

**Heckman, James J and Ganesh Karapakula**, "The Perry Preschoolers at late midlife: A study in design-specific inference," Technical Report, National Bureau of Economic Research 2019.

**Hincapie, Diana**, "Do longer school days improve student achievement? Evidence from Colombia," Technical Report, IDB Working paper series 2016.

**Meyer, Erik and Chris Van Klaveren**, "The effectiveness of extended day programs: Evidence from a randomized field experiment in the Netherlands," *Economics of Education Review*, 2013, *36*, 1–11.

**Pires, Tiago and Sergio Urzua**, "Longer school days, better outcomes," *Northwestern University*, 2010, pp. 1–55.

**Ravallion, Martin**, "Evaluating anti-poverty programs," *Handbook of development economics*, 2007, *4*, 3787–3846.

**Sacerdote, Bruce**, "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly journal of economics*, 2001, *116* (2), 681–704.

**Todd, Petra E**, “Evaluating social programs with endogenous program placement and selection of the treated,” *Handbook of development economics*, 2007, *4*, 3847–3894.

_ **and Kenneth I Wolpin**, “On the specification and estimation of the production function for cognitive achievement,” *The Economic Journal*, 2003, *113* (485), F3–F33.

**Zimmer, Ron, Laura Hamilton, and Rachel Christina**, “After-school tutoring in the context of no child left behind: Effectiveness of two programs in the Pittsburgh public schools,” *Economics of education Review*, 2010, *29* (1), 18–28.

## 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 $\Delta_{ik}$ 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 $\sigma_{st}$ 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 $\beta^{s-1}$. 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 $\beta$ small enough[21], this infinite sum converges to a finite value, which gives us:

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

[21] 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 et al., 2009).