UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

# SOCIAL NETWORKS AND ACADEMIC ACHIEVEMENT: PEER-EFFECTS WITHIN SAO PAULO'S PUBLIC SCHOOL SYSTEM.

**Renata Mayer Gukovas**

**Prof. Dr. Marcos de Almeida Rangel**

SÃO PAULO
2013

Prof. Dr. João Grandino Rodas
Reitor da Universidade de São Paulo

Prof. Dr. Reinaldo Guerreiro
Diretor da Faculdade de Economia, Administração e Contabilidade

Profa. Dra. Elizabeth Maria Mercier Querido Farina
Chefe do Departamento de Economia

Prof. Dr. Pedro Garcia Duarte
Coordenador do Programa de Pós-Graduação em Economia

**RENATA MAYER GUKOVAS**

# SOCIAL NETWORKS AND ACADEMIC ACHIEVEMENT: PEER-EFFECTS WITHIN SAO PAULO'S PUBLIC SCHOOL SYSTEM.

Dissertação apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito para a obtenção do título de Mestre em Ciências.

**Orientador: Prof. Dr. Marcos de Almeida Rangel**

Versão Corrigida

SÃO PAULO

2013

**FICHA CATALOGRÁFICA**
Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Gukovas, Renata Mayer
Social networks and academic achievement: analyzing the peer-effect on São Paulo Public School System / Renata Mayer Gukovas. – São Paulo, 2013.
98 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2013.
Orientador: Marcos de Almeida Rangel.

1. Redes sociais 2. Educação 3. Microeconomia 4. Segregação social 5. Políticas públicas I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 302.3

**Aos meus pais.**

## AGRADECIMENTOS

Gostaria primeiramente de agradecer aos meus pais. Sem as oportunidades e o apoio que eles me deram durante toda a vida, o caminho até aqui teria sido extremamente mais dificil. Ao meu irmão Gustavo também, futuro economista.com que sempre foi um ótimo companheiro.

Gostaria também de agradecer ao meu orientador pelo apoio durante esse processo. Também aos professores Ricardo Madeira, Fernando Botelho e Juan Bonilla pela ajuda com as bases de dados e outras sugestões essencias para este trabalho.

Tenho muito a agradecer à minha turma de mestrado. À Laura pela paciência quase interminável nas matérias que cursamos e fomos monitoras juntas e pelo socorro nos muitos momentos de tensão. À Eliane pela voz da responsabilidade quando necessária. Ao Rafael também pela paciência e companhia em várias matérias, monitorias e no final de curso do mestrado. Mas também a todos os outros colegas, Vivia, Leonardo, Fernando, Lígia, Lilian, Otávio, Bruno, Tarik, Thiago, André, Ana e nossos colegas doutorandos Joelson e Moisés. Foram poucas as vezes que tive a oportunidade de contar com uma turma que desse tanto apoio uns aos outros, fazendo com que os momentos exigentes não fossem assim tão ruins e tornando os bons melhores ainda. Muito obrigada a cada um de vocês!

Não posso deixar de citar também os colegas de outros anos e amigos que apoiaram muito nesse períoro. Rafael Ferreira Neves pelos socorros urgentes em qualquer assunto de programação. Eduardo Sanches Astorino pela ajuda a qualquer hora. Arturo Jose Villanueva pelo apoio na revisão do texto. Aya e Danielly pela ajuda com a base de dados. Raymundo, Fernando, Bruna, Thamara, Paula, Stella pela amizade e compreensão durante esses 2 anos e 8 meses. E por fim meus novos colegas de apartamento, Welton, Luciana e a honorária Marê pela força na reta final.

Finalmente, agradeço à FIPE, CNPQ e FAPESP pelo apoio financeiro.

**“I knew who I WAS when I got up this morning,**
**but I think I must have been changed**
**several times since then."**
***Lewis Carroll***

## RESUMO

O ser humano é um animal social e a relevância das redes nas quais cada indivíduo se insere, em diversos aspectos de sua vida já foi comprovada em vários estudos. Neste trabalho descreve-se intensa porém não exaustivamente as redes de amizades nas escolas estaduais de São Paulo. São detalhadas diversas características, algumas que confirmam e outras que vão contra o que diz o senso comum. Em seguida, é avaliado o possível impacto que o programa *" TEM +Matemática"* provocou sobre as estruturas de redes de amizades nos colégios. Esse programa é composto de aulas extras de matemática, e a aleatorização das escolas interessadas na sua implementação permitiu sua avaliação. Observou-se que nas escolas onde o programa foi realizado, as redes passaram a ser menos coesas de maneira geral, com menos conexões e clusters. Entre os alunos dessas escolas, aqueles que eram elegíveis e não demonstraram interesse em participar das aulas formam os mais afetados, com menos amigos e um grau de centralidade na rede menor. Esses resultados apontam para uma possível segregação de alunos com desempenho ruim que não demonstram motivação para melhorar.

## ABSTRACT

Human beings are social animals, and the relevance of the networks an individual participates throughout his or her life has already been proved by several studies. In this dissertation, the social networks of 107 schools in São Paulo's public System are described intensively but not exhaustively. Several characteristics are detailed. Some of the characteristics observed go in the same direction as the common sense, while others go against it. Then, it is evaluated if the program *"TEM+Matemática"* has had an impact over the structure of friendship in some of these schools. This program consists of extra mathematics classes, and during its implementation, interested schools were randomized to participate, allowing this evaluation. It was observed that the schools that participated in the program had, in general, less cohesive networks, with fewer bonds and clusters. Among these schools students, the most affected were those who were eligible but did not show interest in participating on the classes. These students became less connected and less central in the networks. These results indicate that low performing students that do not show motivation suffer some sort of segregation.

## SUMÁRIO

# LIST OF TABLES

# LIST OF FIGURES

## 1 INTRODUCTION

Most sociologists and economists agree that a given individual's behaviour is partly determined by the opinion and attitudes of the people he relates to. According to Jackson (2005), the depth in which social networks affect human behaviour in many dimensions makes the understanding of their functioning vitally important. How and why are the bonds formed, are they kept, how strong these relations are, how many connections an individual has, and how many of them are necessary for this individual to reach other members of the network are just a few characteristics that may be of interest when studying networks.

On the situations where there is interaction, people influence each other in several aspects. This influence is known by the literature as *peer effect*, the impact that friends and colleagues have over each other. Even though researchers understand its influence, there is little evidence about the magnitude of several forms of peer effect, or even how they operate. One connection can propitiate positive externalities between peers, can facilitate job intermediation, a student's motivation can positively affect the rest of the group or other good or bad behaviours can be disseminated through a friendship network. The existence of these effects itself can drive an individual to try to integrate a given group, or, on the other hand, can be turned into a public police mechanism with the objective of motivating a certain behaviour or fight the inertia of social inequality.

At the school level, it is believed that students tend to perform better when they study together with high achieving classmates, but also can be harmed by undisciplined children either by their behaviour in classroom, or by the dissemination of opinions and motivation. In the United States, there are several empirical studies about the friends' and classmates' influence on students' academic achievements, decisions and behaviour, mostly focusing on university students. Little is known about peer effects on adolescence, a stage of life typically characterized by insecurity and the necessity to reassure oneself among peers. It is possibly at this age that friends' opinions matter the most when a adolescent makes his own decisions. Giorgi, Pellizzari e Redaelli (2010) suggests that more studies are made within this age group, and preferentially in an environment more heterogeneous then university.

There are several evidences that suggest that a few low performing students could benefit from being in the same class as high performing students without provoking negative externalities.

Together with the idea that bad connections in life tend to worsen existing gaps, arbitrarily forming peer groups could be seen as one form of public policy.

In Brazil, little is known about social networks in general. There are very few quantitative studies about peer effects and networks are rarely described with more details. With this in sight, the first part of this dissertation takes advantage of an unpublished data set with detailed information about Sao Paulo state's public school students to describe intensively but not exhaustively friends' networks among adolescents. Section 2 resumes part of the networks' theory that supports the analysis and delineate the existing literature on peer effects, segregation and interventions related to network structures. While section 3 describes the data set formation. And in section 4 several aspects of theses schools' networks are discussed.

The second part tries to assess the impact on the networks' formation of an exogenous intervention made to a randomly assigned set of schools. At those schools, low performing students were given the opportunity to join extra mathematics classes in groups of less than five. This study uses the detailed information available for the end of the school year when the intervention was made for both the treatment and control schools. It tries to identify effects on the structure of the networks on the whole school, among the eligible and ineligible students and finally among those who showed interest in taking part of the extra classes and those who did not. In order to do this, section 5 explain the design and implementation of the extra classes program, the *"TEM+Matematica"*. Section 6 describes the results obtained. Finally, section 8 concludes.

# 2 ABOUT NETWORKS

## 2.1 How to measure networks

By definition, a network can be defined as something that resembles an openwork fabric or structure in form or concept. A network, for instance, can be a group of rail roads, an interconnected system of computers or a group of people with similar interests. In network' theory, a network is defined by a group of two or more nodes and by the edges formed between them. Every member in a network, a student, an enterprise, a computer or a country is referred to as a node, while the connection built between any two nodes is called an edge. Every node has characteristics of its own: a student can be a boy or a girl, a company can operate in different sectors or a country can adopt a given government system, similarly edges can also differ. For instance, a bond can go both ways when we refering to marriages, but when considering the internet, a website can have a link to another website that does not necessarily have a link back, in this case, the edge is directed. Edges can also have different weights, as the value of a transaction between two banks, the number of publications co-authored by two researchers or even the amount of trust between two friends.

A network can be represented in several ways. By definition, a graph $(N, g)$ consists of a set of nodes $N = \{1, \ldots, n\}$ and of an *adjacency matrix* (JACKSON, 2010). An *adjacency matrix* is a $n \times n$ matrix $G$, on which each element $g_{ij}$ is 1 if the elements $i$ and $j$ are connected (or the weight given to the connection between $i$ and $j$) or 0 if they are not. It is easy to see that on an undirected network, the *adjacency matrix* is symmetric, since if $i$ is connected to $j$, then $j$ is connected to $i$. That is not necessarily true on a directed network. The graph can also be represented on a diagram in order to allow a better visualization of the network. Typically, on a diagram, each node is represented by a dot or a circle and each edge is represented by a segment between two nodes. In the case of a directed network, arrows are used to represent edges. The characteristics of nodes and edges can be represented by the use of different colors, sizes or shapes for the circles and segments and arrows. There are several algorithms that can be used to determine the position of each node on a diagram. The one used in this dissertation is *Force Atlas 2*, that use the concept of magnetism in physics to simulate a repulsion force on each node and an attraction force associated with the edges. It allows also to determine a gravitational force on the center of the diagram to prevent that disconnected nodes are pushed to far away, and to create a repulsion force between the edges to avoid overlapping and consequently facilitate the visualisation.

Not only can nodes and edges can be described, but so can the structure of the network, and the characteristics of these networks can, for instance, show how well connected a given country is, how long it would take for a disease or a piece of information given to one individual to reach the entire network or if it would reach everyone, if a group of students is segregated by one characteristic and even if the bonds are made randomly or following a given rule. Those characteristics can be analyzed through a set of indicators, which include, among others: degree, indegree, outdegree, number and size of the components, shortest path, diameter, cohesiveness, centrality, existence of positive assortitiviness, clustering.

The first three characteristics refer to the number of connections a given node has. While indegree and outdegree are the number of edges that arrive or leave from a given node, respectively, degree is the total number of edges that connect a given node to the rest of the network. Indegree and outdegree are usually only mentioned for a directed network, and for the purposes of this analysis, they refer to the number of found citations that any student receives and makes. On the other hand, the degree of a node, or the average degree of a network, can be analyzed both on directed and undirected networks. In the case of a directed network, the degree of a node is not necessarily the sum of its indegree and outdegree. In a directed network where two nodes have edges going in both directions between them, they are counted as one citation made and one received, but in an undirected network, both of them correspond to only one connection.

A component is a subset of nodes that are all connected, either directly or indirectly via their edges. Suppose a network is composed of people and that its edges indicate with whom you would share knowledge. If a piece of information is given to any member of a component and passed forward to his/her connections, eventually all the members of this component would be aware of it. In an undirected network, it is easy to separate the components on a diagram; it is only necessary to identify the groups that are not connected by any edge. For instance, by observing School 1, represented by figure 3, it is easy to see that there are only two components if the direction of the connections is not considered. Interestingly one of them is composed only by girls. But if the direction of the edges is considered, this group of girls is divided into three components: the two small circles are two isolated components because no connection reaches them, while all the others are one component, because from any of them, it is possible to reach all the others by following the arrows. Apart from the number of components on a given network, this dissertation also shows statistics related to components. The average size of the components, the average number of single student components (isolated nodes), the size of the largest component on each school and the percentage of students on that component are shown in the next session.

Another characteristic that is analyzed in several ways is the distance between two nodes. In network theory, the distance between two nodes is the lowest number of edges that are crossed to connect them, also called the shortest path. For instance, the distance between two friends is 1, and between this individual and the friend's friend is 2, unless this is a friend in common. In an undirected network, the distance between $i$ and $j$ is the same distance between $j$ and $i$, but that is not true in a directed network, where $i$ can be connected directly to $j$ and $j$ only connects to $i$ through other nodes or does not connect at all. When two nodes are in separate components in an undirected network, they cannot reach one another. As a result, the distance between them is set to infinity. A network's diameter is defined as the maximum distance between any two nodes.

One way of calculating the distance between two nodes in an undirected network is to calculate the lowest $k$ for which the element $h_{ij}$ of the matrix $H_k = (G)^k$ is different from zero, where $G$ is the network's *adjacency matrix*. This happens because when calculating the *adjacency matrix*'s $k^{th}$ power, the element $h_{ij}$ represents the number of $k$ long paths between $i$ and $j$, as shown in Jackson (2010). When two nodes are in separate components, the $h_{ij}$ element of the matrix $H_k$ will never differ from zero.

The average distance between one node and the rest of the network, the node's average path length, can tell how central the node is, and the average of that measure across the network can tell how tied together it is. But, when the nodes are not all connected, the nodes' or the network's average path length and the diameter will always be infinity. An alternative to circumventing this problem is to calculate the average path length and the diameter only inside the components, and, for the whole network, report the statistics of the largest component. A second way, as suggested by Newman (2003), is to calculate one statistic similar to the average path length using the formula below, where $\ell(i, j, g)$ is the length of the shortest path between $i$ and $j$ and $g$, which equals infinity if the nodes are not connected.

$$\frac{n(n+1)}{2\sum_{ij}\frac{1}{\ell(i,j,g)}} \tag{1}$$

As mentioned before, it can be said that the lower is the average path length of a node, the more central the node in the network is. So, a possible measure of centrality would be the inverse of the average path length. A second way to evaluate the node's centrality is to consider the total number of connections it has compared to the total number of connections it could have. This is called the *degree centrality* and is calculated by simply dividing the node's degree by the number of nodes in the network minus one, as demonstrated by equation (2) below. A third way of measuring the centrality is to weight the nodes that are closer and further away differently. This is known as the *closeness centrality*. This can be done by calculating the *decay centrality* as it is written in equation (3), by using a *decay parameter* $\delta$ that varies between 0 and 1. The closer $\delta$ is to 1, this measure show how large is the component the node is, while when it approaches zero, it gives less weight to nodes that are further away.

$$\frac{d_i(g)}{n-1} \tag{2}$$

$$\sum_{i \neq j} \delta^{\ell(i,j)} \tag{3}$$

The transitiveness of bonds can also indicate how tightly clustered they are. One measure of this closeness is the numbers of cliques the network has, being a clique a group of nodes that fully connected, that have links going between all of them. But not only it is harder to measure, but also a clique is very sensitive to small changes on the network: removing one single edge already dismantles a clique. The clustering measure was created to solve this problem. It consists of the number of totally connected trios of nodes over the number of different trios that have at least two edges between them on the network as shown by equation 4 below, where $g_{ik}$ is extracted from the *adjacency matrix*.

$$Cl(g) = \frac{\sum_{i;j\neq i;k\neq j;k\neq i} g_{ij}g_{ik}g_{jk}}{\sum_{i;j\neq i;k\neq j;k\neq i} g_{ij}g_{ik}} \tag{4}$$

Similarly to the *overall clustering* explained above, an individual measure of clustering can be calculated using the equation 5 and then averaged over the network. The difference between these two measures is that the second method gives more weight to low degree nodes then the first one.

$$Cl_i(g) = \frac{\sum_{j\neq i;k\neq j;k\neq i} g_{ij}g_{ik}g_{jk}}{\sum_{j\neq i;k\neq j;k\neq i} g_{ij}g_{ik}} \tag{5}$$

$$Cl^{Avg}(g) = \frac{\sum_i Cl_i(g)}{n} \tag{6}$$

Finally, this dissertation also analyses one measure of correlation between the characteristics of connected nodes, the positive assortitiveness. It is calculated using the formula below, where $d_i$ is the node's characteristic of interest (usually the nodes degree), $m$ is the average of that measure across the network, and $i$ and $j$ are connected. This number varies from $-1$ to 1, when it is close to 1, it means that popular nodes (with high degree, for instance), tend to tie with other popular nodes while the least popular tend to tie between them. When it is close to -1, it means that bonds tend to be formed between individuals with opposite characteristics, and when it is close to 0, the links are formed randomly acording to that characteristic.

$$\frac{\sum_{i,j\in g}(d_i-m)\left(d_j-m\right)}{\sum_{i,j\in N}(d_i-m)^2} \tag{7}$$

## 2.2 Literature on Peer Effects and Segregation

Jackson (2010) states that networks can be formed either randomly, where each connection has a probability of being formed, or strategically, when the individuals chose voluntarily to form a bond with a certain objective or for a certain purpose. In the second case, the reasons and results of a network formation can concern policy makers and social analysts. Since networks play an important role in the life of an individual in several aspects, not being well connected may be detrimental. For instance, when a person is searching for a new job, or trying to learn in groups at school, if he/she is not well connected, his/her success can be limited. The high relevance of networks therefore explains the existence of many studies on segregation and peer-group formation.

In Linden (2009) they perform an experiment in a poor communities in India to assess the existence of peer-effects in enrolment and attendance patterns of current out-of-school children. They randomly select a group of children to have positive and repeated incentives to go to especial classes that prepare out-of-school children to go back to the formal system. Then, through a series of questionnaires they build the social network of that community and assess whether non-influenced children changed their behaviour. They found out that having one treated friend increased the probability of going back to school by 20% of the treated effect in the treated student. Having additional treated friends did not show any significant difference, but when the friendship was mutual (both friends named each other in the questionnaire) the effect was larger.

Echenique e Fryer (2007) developed a segregation measure, the SSI (Spectral Segregation Index), based on two premises: this measure should be disaggregated at the level of the individuals, and each individual should be considered more segregated the more segregated the people with whom he or she connects become. By using this measure on networks of students in grades 7 and 12, they found that the percentage of students that belong to a certain minority does not explain the real segregation level they face.

Both in Echenique e Jr (2005) and in Austen-Smith e Fryer (2005), a similar methodology is used to estimate an effect known as *acting white*, which refers to the segregation suffered by afro-American students among their peers when they have good grades. By studying, it is believed that they are trying to act as white students. They estimate one model in which one afro-American student has to decide to allocate his time to study, and that would be a signal for employers and then bring better wages in the future, or to social activities, that would lead to being accepted by the group. The researchers built a popularity measure based on the number of friends a given individual has and on these friends' own popularity, and they observed that this measure is positively correlated to the student's grades for all races except for Hispanics and afro-Americans. For these two races there is an inflexion and studying and popularity become negatively correlated after a certain grade. This result supports the existence of the *acting white* phenomenon.

Jr (2010), on the other hand, studies the relation between segregation, discrimination and the friendship dynamics with the evolution of the grades observed among 9 and 13 year old white and afro-Americans in the United States. His studies reveal out that this differential rises on a period of life that the returns to education are high. According to his estimates, this movement could not be explained by an increase in segregation, once it lowers in the United States since the 40s, not even by discrimination models. Two explanations that would be more plausible are the identity or the friendships' dynamics, being the first related to a possible change in preferences between a group of friends and the second that correlates negatively good schooling outcomes and social mobility (acting white). Those hypotheses cannot be separated, and the data did not allow reaching a clearer conclusion about which would better explain the rise in the differential between white and afro-American students.

Using the idea of peer-effect in classrooms, there are some studies that state that the outcome of students can be optimized by assorting them in classes according to their prior achievements, in order to have a few low performing students in top performing classes (GRAHAM; IMBENS; RIDDER, 2009) (BHATTACHARYA, 2009). These studies, among others, use the existing formation of several classes to build a production function that would allow policy makers to build an optimal division of students. Using those results, Carrell, Sacerdote e West (2013) conduct an experiment at the Unites States Air Force Academy, building the classes in order to improve academic performance of the lowest ability students. Their results contradict the prior studies, showing an actual negative impact in the targeted students. What they suggest as an explanation is that the other studies did not take into account the process of group formation inside the classrooms. This exogenous intervention might have caused a form of segregation between the students, worsening the situation of these low achieving students.

# 3 DATA SET CONSTRUCTION

For the analysis conducted on this dissertation, three different datasets were combined. The first dataset was the result of the second questionnaire applied to the year 9 [1] public school students, the *Pesquisa C-IDEAS/LEIA-ME/BID de Caracterizacao Multidimensional do Alunado Paulista*, in November 2011, as part of the program *TEM+Matematica*. This program, and the database resultant of the questionnaires related to it, was conducted by the research centres C-IDEAS and LEIA-ME and financed by the *Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira* (INEP), the Interamerican Development Bank (IDB) and non-governmental organizations. The questionnaires were read with the program *Abby Flexy Capture 10* and the data set is under the responsibility of Prof. Dr. Marcos de Almeida Rangel. The questionnaires were marked with the same student id used by the Secretariat of Education, making possible the use of the two other datasets. The first was the the administrative records from SARESP, which contained some individual information of the students along with their names and score on the state standardized test. The other database contained all the grades given to the student in mathematics and language inside the school. These were the ones used to determine whether a student was eligible to join the program and also the only grades students are aware of.

The table below shows the size of the sample in terms of number of schools, classes, students. It shows that out of the 7,989 students who answered the questionnaire, 7,696 were found in the Secretariat database, corresponding to $96,33\%$ . Some of the students were not found because their id could not be identified. The questionnaires were previously marked with the students' id both in numbers and on a bar code according to the list of enrolled students, and when a student did not have a questionnaire, he or she was given a blank one and was supposed to fill in his or her id. But the id is a sequence of numbers containing the school code, class code and student code and it is usually not known by the students, so these questionnaires were lost. Another possible cause of attrition is the transference of students between schools. Due to the creation of the code system, students cannot be traced once they move to a different school.

The average size of the schools was over 1000 students, of which about 152 were in their year 9. Each school had on average four different year 9 classes, with about 36 students each. Of the 15,601 students in year 9 in the 107 schools, an average of 49.33% per school answered the questionnaire (7,696), or about 71 students per school. This low average can be attributed to the number of schools in which almost no students filled out the questionnaires while, on the other

[1] Last grade on Middle School in Brazilian's school system

hand, several schools had a high response rate. The schools with a response rate close to zero were, therefore, not analyzed. Among the others, the response rate was around $78,00\%$. Of the responses $50.53\%$ were boys, and their average percentage of correct answers in language was $45,66\%$ and in mathematics $37.19\%$.

Table 1: General description of the database

| **Sample description** | |
|---|---|
| Total number of questionnaires | 7989 |
| Total number of questionnaires matched with the administrative data | 7696 |
| Percentage of questionnaires matched with the administrative data | 96.33% |
| Total number of schools | 107 |
| Average number of students per school | 1192.93 |
| Average number of classes per school | 32.15 |
| Average number of students per class | 36.33 |
| Average number of $9^{th}$ year students per school | 152.60 |
| Average number of $9^{th}$ year classes per school | 4.19 |
| Average number of students per $9^{th}$ year class | 36.36 |
| **Among the $9^{th}$ year students** | |
| Average percentage of boys | 50.53% |
| Average percentage of right answers in language | 45.46% |
| Average percentage of right answers in mathematics | 37.19% |
| Average number of questionnaires answered per school | 71.55 |
| Average percentage of $9^{th}$ year students that answered the questionnaire | 49.33% |

The tables and diagrams in this dissertation were built upon the data bases from the questionnaires, the Secretariat and SARESP, together with the student's school report information for 2010 and 2011. The information about gender and race was taken from the questionnaires and is self-declared. SARESP grades go from 125 to 500 and are comparable across school years, so the older the student is, the higher the grade he is expected to achieve. Students from each school year are expected to reach a certain level and they are classified into four categories according to these requirements. Table 2 shows how the year 9's classification is divided. After

Table 2: SAEB Scale for outcomes in language and mathematics

| **Classification** | **Outcome in Mathematics** | **Outcome in Language** |
|---|---|---|
| Below basic | $< 225$ | $< 200$ |
| Basic | Between 225 and 300 | Between 200 and 275 |
| Adequate | Between 300 and 350 | Between 275 and 325 |
| Advanced | $> 350$ | $> 325$ |

merging the two databases it was necessary to match the friends named by the students with their registry. Apart from the expected issues with misspelling, which is very common when dealing with names, the process also had to overcome poor handwriting, misunderstanding of the questions and, since most of the students are between 14 and 15 years old, sometimes nicknames and drolleries. The students were asked to fill in slots with the first and then last names of four of their closest friends in that year. Expecting to encounter problems when trying to find those friends at the school, they were also asked to tell their friends' gender, race and class. Not every student filled all the slots. Out of the 31,972 possible citations, only 19,496 (61%) were reported. There were also some students who named two different friends on the same slot. In order to avoid distortions, only the first friends were considered. As for the other characteristics, class was rarely filled in, and the race could not be used because even for the perfect matches, the self-declared race was different than the one friends designated. In total, as it can be seen in table 3, 77% of the friends of self-declared white students gave the same answer, 60% for brown students and only 54% for afro-descendents. For the few indigenous and Asian students, the percentage was even lower, close to 10%.

This matching process could have been done directly using the record linkage methodology, but in order to better check the quality of the match, it was done in four different stages. Firstly, the number of perfect matches was verified by considering the students' full name on the year 9 list, the entire citation (first and last name slots) and the school code (both students, the individual and the possible friend were required to be at the same school in all cases). Only 2,511 friends were found this way (12.9% of the total number of citations). Then, only the very first and very last names were considered on both the list and on the citation [2]. In this case, another 2.154 (11.0%) were found. Before using the record linkage, the first name and first last name were matched, excluding particles such as "de", "da", "dos" and others that precede some surnames. This time, there were 6,979 positive matches (35.8% of the total number of citations). It is important to note that all homonyms in any one school were dropped from the list in order to avoid double matches.

Finally, the remaining names were matched using record linkage, a methodology that enables the merging imperfect string variables. As explained in Blasnik (2010), record linkage allows identifiers, such as names and addresses, to have spelling and formatting variations. *Formal record linkage methods often employ a combination of approximate string comparators and*

[2] In Brazil, people traditionally carry both their mother's and father's last name. As a result, it is common for people to have two last names, and very often even more than that. In those cases, usually only their very last and sometimes only their first surname is known. Also it is not rare to have two first names as well, and be known by both of them.

*probabilistic matching algorithms to identify the best matches and assess their reliability*. In this case, it *employs a modified bigram string comparator and allows user-specified match and non-match weights*. The record linkage used the first and last names separately and required that the students be enrolled at the same school. By setting the minimum match probability to 0.6, a total of 5,084 positive matches were found (26,0% of the total). The average merge probability was 80.03% and not many noticeable false matches were recorded. Table 4 shows examples of names that were matched using record linkage.

Table 5 summarizes the outcomes, number of friends named and found according to some of the characteristics of the sample. AS it can be seen, girls perform better in language while boys in mathematics. White students and the ones that are on the right school year for their age have better outcomes in both subjects. Figures 1 and 2 show the distribution of outcomes according to race. The vertical lines represent the division between classifications for this age. In both cases, there is a large mass of students of all races that fall below the basic expected for them.

## 4 ANALYSIS OF THE SCHOOLS NETWORKS

Figures from 3 to 23 represent some examples of diagrams of the school networks analyzed and tables from 6 to 9 and graphs from 24 to 45 show a summary of the whole sample characteristics. On the diagrams, each node represents a student and the arrow represents a citation as a friend, being the source of the citation the beginning of the arrow and the friend named its direction. On every diagram, the size of the nodes is directly proportional to its degree and the colors of the arrows represent the answer to the question "Would you consider leaving the school if this friend decided to drop out?", being red "no", yellow "may be", green "yes" and black when the student did not answer. Each school is represented by four diagrams that show through the nodes' colors a different set of characteristics.

The first diagram for each school shows the gender of the students, where the girls are represented by red nodes, boys are blue and the ones that did not answer the questionnaire are represented by the color green. The information about gender was collected from the questionnaire and, even though it was also available at the Secretariat data base (and so also to the nodes that did not answer the questionnaires) it appeared to be important to always identify the nodes that did not answer the questionnaire in order not to take them for a student that has no found friends.

On the gender diagrams it can be clearly seen that bond tend to be formed between students of the same gender, a typical situation among children and young adolescents. In most of the schools groups composed mainly by girls or boys can be identified. At the first school, there is a separate large component made of 9 girls, one fourth out of the total 36 students. School 8 also show a large division between boys and girls. When considering popularity in terms of degree of a node, it can be seen that in some schools there are both boys and girls that are popular, but in general if there are outstanding nodes they tend to be girls. On the other hand, isolated nodes are mostly boys.

The second diagram for each school shows the students' race, where the self declared white students are represented by the color yellow, the brown by brown, black by black, Asian students by light blue, indigenous by red, and finally, green again represents students who did not answer the questionnaire. When concerning race, it is harder to identify a pattern just by looking at the diagrams. Most students are white or brown so it is more probable to see a white or brown student as the most popular, but at school 7 there are many black popular kids. Unlike it happens with gender, groups composed mainly by one race can not be identified. Isolated nodes are also race mixed.

The third and the forth diagrams for each school show the students' outcome in language and mathematics. A lighter shade of green or red represent a student with lower outcome, while a darker shade represents one with better grades. The students who did not answer the questionnaire are represented by blue on both diagrams. Again, a clear pattern cannot be identified. Popular students and isolated nodes had every possible outcome and homogeneous groups are not apparent. At school 3 the bottom right area of the diagram seems to present darker colors, but it appears to be more an exception then a rule.

Another interesting aspect that can be noticed on the diagrams is the recurring appearance of pentagon shaped groups. Several groups of five students all connected by edges that often go in both directions, and that have few links connecting them to other students, can be identified. This indicate a high level of clustering and that the student' own impression about his peer group is shared among his friends. The five student group then is expected because each had to name four friends.

Figure 24 describes the distribution of the number of friends each student named on his questionnaires. It is interesting to see that in most of the cases either a student did not fill any of the spaces where he was supposed to put his friends' names or he filled all the four spaces. Only few students answered between 1 and 3 names of friends. But not all of them were legible, real names or could be found among the students of that school year. Table 25 shows the distribution of the number of friends who were actually found, or what was considered the outdegree. The number of students with zero friends rises slightly, while the number of students that had all their four citations found fall. But even after losing some citations, the number of students with at least one found friend is almost twice as large as the number of students with no friend. The following table shows the number of citations that were not found. For very few students all four citations, or even more than one, were lost. In most cases, either no or only one citation was lost.

Table 6 summarizes the main individual network characteristics analyzed according to gender, race, age and outcomes. As it could be seen on the diagrams, girls have on average higher degree, indegree and outdegree, with a difference of around 0.5 friends. Brown students are also more popular, but the difference between them and white or black students is much smaller than the difference between genders: it varies from less than 0.1 on the indegee until around 0.2 on the degree and outdegree. There are not enough Asian and indigenous students to draw conclusions, so they are not analysed. Students that are on the right school year for their age, aged 14 and 15 on this case, also have the most friends. The three measures are also positively related to the students' outcome in language, while only the outdegree is positively related to outcomes in mathematics. The indegree and total degree grow from low performing students to those with adequate outcomes, but it is lower for advanced students. If it is assumed that better performing students understand better the questionnaire, or that they fill it in more responsibly, it is expected that the outdegree would be higher for them, and that is confirmed, with the total difference reaching over 0.6 friends. But both the indegree and the total degree are more influenced by the potential friends. But unlike it is commonly said about Brazilian schools, bad performing students don't seem to be the most popular when the degree is used as a popularity measure.

The table also shows the individual measures for clustering, both directed and undirected. On both measures, girls are on average more closely knit then boys. When considering different races, the difference is smaller (except on the first measure for Asian students, that appear to have the tightest groups, but there are only 24 Asian students, so no conclusions can be drawn), with white students having a higher clustering average. Students on the right age are more knit together than older students, so do students with an adequate outcome in language.

Table 7 shows the average measures average shortest path, centrality and decay centrality with $\delta$ equals to 0.25, 0.5 and 0.75. Boys have a lower average shortest path, ...
The following table describes the average indegree according to the students's self declared race, but also according to his/her gender, age and outcomes in language and mathematics. Both among girls and boys brown students are the most popular in terms of number of nominations with black students close behind. But the difference between brown and black girls is much smaller than among the boys. The averages according to age follow the same pattern as before, brown students and those on the right school year have higher indegree. This measure appears as positively correlated both to outcomes in language and mathematics for all races, with the exception of advanced level students in mathematics, which are not numerous.

Figures from 29 to 32 show the relation between indegree and outcomes in mathematics and language according to the students self declared race with and without confidence intervals. The vertical dotted lines show the division of the outcomes' categories (below basic, basic, adequate and advanced). As it could be seen on the tables, brown students tend to have larger indegree then white and brown. But except for a small interval in language where the confidence interval for black students falls below the confidence interval for brown students, it cannot be said that the averages are statistically different. Again, there were not enough Asian and indigenous students to calculate the fitted polynomials. The graphs also show a positive relation between outcome and indegree, only falling for the top performing, where the number of observations is already low to draw conclusions. Only on the case of black students, the average indegree does not increase for better performing students in language.

Table 9 show a possible evidence of homophily, segregation of different students based on personal characteristics. According to (JACKSON, 2010), a basic measure of segregation is to compare the proportion of friends with a given characteristic a student has and the proportion of possible friends with that same characteristic. If the proportions are different, there are evidence to support the existence of segregation. In table 9, it can be clearly seen that boys and girls tend to form friendship bonds between themselves. When race is concerned, the pattern is not so evident, but it is still possible to identify that students connect slightly more with others of the same race.

Figure 33 show the average shortest path length between students according to the school size. As it is usually seen in networks, they are positively correlated, but the average shortest path does not grow in the same proportion of the school size. The same happens to the diameter of the largest component and the number of components, as it can be seen in figure 34 and 35. Figure 36 show the number of students that are on the largest component, showing that the probability of a new student to enter this giant component is relatively large. Combining this with the average number of connections a given student has, the hypothesis that these networks are formed randomly can be rejected.

The level of clustering is not positively or negatively correlated with the size of the school, but small schools have a much larger variance in this measure of cohesiveness. Figures fro 40 to 45 try to identify any correlation between positive assortativity in terms of several characteristics, but they appear not to be correlated with school size.

# 5 PROGRAM *"TEM+MATEMATICA"*

In the last decade, the Brazilian government implemented a cycle system in public schools. In this new system students can fail a subject, but can only repeat the school year if it is the final year of a cycle [3]. This policy has had a positive impact on school attendance, but as a result, in addition to the already low quality of the school system, many children move forward without knowing even basic material. Typically, school grades vary between zero and ten, with ten being the best outcome, and with five or seven serving as the minimum passin score. In elementary and middle school [4] 7 is usually the standard.

The program *"TEM+Matematica"* consists of extra Mathematics classes given to small groups of students who have had grades below seven on the previous year's final school exam on the subject. The classes were given to 7th and 9th year' students. Prior to the implementation of the program, the students were asked about their interest in taking part of those classes, and, according to the supply of instructors and the demand, cohorts of schools were built for each year in order to have schools where students had the same probability of being selected for the classes. Schools were then randomized inside the cohorts to receive the program. In theory, interested students were also supposed to be randomized, but because of the high levels of absence, new students were called to replace the *drop outs*. As a result, the information about participation is imprecise. A preliminary evaluation was conducted by the research centre CIDEAS to see the program's impact on grades. The results showed little or no impact on students outcomes, which was probably consequence of the irregularity of the classes.

As in Carrell, Sacerdote e West (2013), these extra classes can also be seen as an exogenous intervention on the existing networks. To assess if the program had an impact on the structure of these networks, the second part of this dissertation tries to identify these effects by comparing the students from treated and control schools, the eligibles and the students that enrolled or not for the classes. The results are presented in the next session.

[3] The cycles may vary between states and whether the school is administrated by the state or the municipality, but it usually lasts between 3 and 4 grades.

[4] Elementary School, or the Fundamental School's first stage, starts at year 1, when the student is about six years old, and continues until the year 5. Middle School, or the Fundamental School's second stage, comprehends years 6 through 9.

## 6 IMPACT OF THE EXTRA CLASSES ON GROUP FORMATION

Among the schools that showed interest in participating in the program, some were selected to have the classes and other schools were left for the control group. They were always compared within the cohort in which they were randomized. Unfortunately, there were some schools that were lost because of the low response rate, and other schools had to be left out because in the end they were in a cohort with representatives of only one of the groups (treatment or control). Finally there were some of the 107 schools analyzed in the previous sessions that were not selected to participate in the randomization because they were not comparable with any other. After that, there were 56 schools remaining in 26 different cohorts. 29 of the schools were in the treatment group, with a total of 2002 students, and 27 in the control group, with 2221 students. The average response rate was near 80%

In the first group of analysis, the informations about the interest of individual students in taking part of the extra classes was not used. The analyzed variable was regresses against a dummy that indicated if the student was in a school selected for the program, and controls were included in three stages. On the second part, the the treatment dummy was interacted with the eligibility dummy, that equals 1 for students with previous year's final score in mathematics below 7. Thirdly, it was investigated whether the impact was different between different range of grades, including dummies for the groups between 4 and 5, between 6 and 7 and above 7, with the group below 4 as the base line. In all of the cases the control variables were introduced at times. Then this whole exercise was repeated for students that showed interest in participating of the extra classes both in the treatment and control schools, and finally for those who did not.

Ten different characteristics of the networks were compared between the students from schools that were selected to participate in the program and students of those left for the control group. The first group of characteristics involved the probability of a node to be connected with the rest of the network, divided into three variables: a dummy that indicated whether a student was an isolated node, a dummy that indicated if he or she had named at least one friend, and finally a dummy that indicated if the student was named as a friend by any other student of the network. Then it was evaluated if there was a change in the indegree or the total degree of the students in participating schools. The cohesiveness of the network was analyzed through the decay centrality with the three different $\delta$ calculated before, and through the clustering level calculated for the individuals, both considering the direction of the nodes and not taking them into consideration. The last groups of regressions analyzed if there was a change in the composition of groups of friends in terms of outcomes in mathematics: first observing the correlation between the student's own grade with his or her friends, an then by verifying the average grade among friends for different groups of students.

Tables from 10 until 43 show the results of the regressions. As it can be seen on table 10, the probability of a given node being an isolated node increases with the treatment, specially for those with grades below 4. Mostly because it lowers the probability of being named as someone's friend then because it lowers the probability of naming someone as friend. As a result of that, the coefficient for the treatment dummy is also negative when regressed against the indegree. The table 14 shows that the indegree is significantly lower for the students with grades of the previous year below 7 in the treatment schools, being lower the worse the grade was. The same results are observed for the total degree.

Not surprisingly, they also became on average less central in their respective networks, once the decay centrality calculated with all the $\delta$ falls for the students on the schools with extra classes. Interestingly it falls more for students with grades below 4 or above 7. Also the clustering level, considering both directed and undirected measures, fall for the students in treated schools, except for those with grades higher than 7. The correlation between students' own grades on the previous year and their friends did not change, it was already higher for students with grades above 7. But the students with grades between 6 and 7 became friends with students that had worse grades on the previous year.

If having more connections and being more central is considered a good outcome, the treatment appeared to have a negative impact on the structure of these schools. In order to identify better who were the students that were more affected, the exercise above was repeated for students that showed interest in participating of the extra classes and those who did not. Remembering that enrolling for the classes did not necessarily mean attending them. Tables from 21 to 31 show the results for the interested students and tables from 32 to 43 for the not interested students. If the impact was caused by the classes themselves, the coefficients are expected to be higher in absolute terms.

The effect of the treatment on the probability of being an isolated node was indeed stronger for students who enrolled on the extra classes, except for students with grades below 4 on the previous year. On the other hand, the probability of having at least one person naming the student as a friend was lower on the treatment groups both for the interested and the not interested students. But it was much lower for this second group, specially for those with really low grades. The change in probability of naming at least one friend was also due to a reduction in the group of students that did note enrol for the extra classes.

The reduction on the indegree of students in treated schools was also mostly due to the reduction for the students who did not enrol. The case of the indegree, was very interesting: the reduction in the average indegree was higher for students who enrolled, and who, on the previous year, had grades in math above 4. On the other hand, the very bottom students were more damaged when they did not enrol. The same was observed for the decay centrality with $\delta$ equals 0.25 and 0.50. For $\delta$ equals 0.75, the effect was divided in two groups: students with grade below 4 who did not enrol and students with grades above 7 who did enrol. The impact on the clustering level, both directed and undirected, fell over not interested students. The positive assortiviness in terms of grades in mathematics on the previous year was not affected in any of the groups. But the reduction on the average grade of the friends happened among non interested students.

This results suggests that when the lack of interest of certain students became clear to the rest of the classmates, this group became more isolated. Because of the connectivity of the network, this affects the measures of the other students as well. It cannot be stated that the change was due to the motivation itself because not interested students were compared to students who also did not enrol for the classes in the control group (even though they would not receive the extra classes, they were asked to manifest their interest because, as it was said in the previous sessions, the relation between the supply and demand of classes was used in the randomization process.). The impact came when it was perceived by the other students. This result is interesting because it goes against the common sense in Brazil that more motivated and interested students would be segregated by their friends.

## 7 CONCLUSION

The results of the analysis of the school networks and the impact of the extra mathematics' classes in group formation both supported some ideas of the common sense and proved others wrong. In the first place, there was a strong division between girls and boys on the network structure. Girls also shown to be more connected to the rest of the network and also more clustered and central in the networks. In terms of race, the evidence suggested the existence of some discrimination, but not as strong as it appears on the United States and other countries. On the other hand, unlike it is commonly believed, better performing students are not segregated. The evidence suggests exactly the contrary that students with low grades seem to have fewer friends and be less central in their networks.

Moreover, the evaluation of the program *"TEM+Matematica"* allowed to see that the segregation could be worse for students known to have low performance, if we consider that participating on extra classes could signal that to other students. One assumption that could be drawn from these results is that the extra classes could be effective in improving the grades of the students if this effect was measured separately. But these classes also affect the structure of the network the students are part of, leaving the worse students less connected and also connected to other with lower grades. If a great part of what is learned depends on one?s interaction with his or her peers, this could have a negative impact on the students? grades. Finally, this could cancel a possible positive impact of the classes, explaining the lack of significance found on the traditional impact evaluation.

# 8 BIBLIOGRAPHY

AUSTEN-SMITH, D.; FRYER, R. An economic analysis of "acting white". **The Quarterly Journal of Economics**, Oxford University Press, v. 120, n. 2, p. 551–583, 2005.

BHATTACHARYA, D. Inferring optimal peer assignment from experimental data. **Journal of the American Statistical Association**, Taylor & Francis, v. 104, n. 486, p. 486–500, 2009.

BLASNIK, M. **RECLINK: Stata module to probabilistically match records**. 2010. Disponível em: <http://EconPapers.repec.org/RePEc:boc:bocode:s456876>.

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.

ECHENIQUE, F.; FRYER, R. A measure of segregation based on social interactions. **The Quarterly Journal of Economics**, Oxford University Press, v. 122, n. 2, p. 441–485, 2007.

ECHENIQUE, F.; JR, R. F. **On the measurement of segregation**. [S.l.], 2005.

GIORGI, G. D.; PELLIZZARI, M.; REDAELLI, S. Identification of social interactions through partially overlapping peer groups. **American Economic Journal: Applied Economics**, American Economic Association, v. 2, n. 2, p. 241–275, 2010.

GRAHAM, B. S.; IMBENS, G. W.; RIDDER, G. **Complementarity and aggregate implications of assortative matching: a nonparametric analysis**. [S.l.], 2009.

JACKSON, M. The economics of social networks. California Institute of Technology, 2005.

JACKSON, M. **Social and economic networks**. [S.l.]: Princeton University Press, 2010.

JR, R. F. **The Importance of Segregation, Discrimination, Peer Dynamics, and Identity in Explaining Trends in the Racial Achievement Gap**. [S.l.], 2010.

LINDEN, L. Bridge classes and peer networks among out-of-school children in india. **Essays on Household Decision Making in Developing Countries**, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, p. 59, 2009.

NEWMAN, M. E. The structure and function of complex networks. **SIAM review**, SIAM, v. 45, n. 2, p. 167–256, 2003.

Table 3: Proportion of friends that say a student is of a given race according to the self-declared race

| | | Friends' opinion | | | | |
|---|---|---|---|---|---|---|
| | N | White | Brow | Black | Asian | Indigenous |
| Self declared race | | | | | | |
| White | 2792 | 0.77 | 0.18 | 0.03 | 0.01 | 0.01 |
| Brown | 3872 | 0.26 | 0.6 | 0.12 | 0.01 | 0.02 |
| Black | 1018 | 0.1 | 0.33 | 0.54 | 0 | 0.02 |
| Asian | 29 | 0.4 | 0.27 | 0.1 | 0.15 | 0.08 |
| Indigenous | 121 | 0.19 | 0.45 | 0.18 | 0.03 | 0.15 |
| NA | 161 | 0.32 | 0.52 | 0.11 | 0.01 | 0.04 |
| Total | 7993 | 0.41 | 0.42 | 0.14 | 0.01 | 0.02 |

Table 4: Examples of created names that would be matched using record linkage

| | Names as if filled in the questionnaire | |
|---|---|---|
| Name as if on the School List | First Name | Last Name |
| MICHELE CAMPINAS CORREA DOS SANTOS | FRANCIELE | CORREIA |
| ALEX PAIVA DE JESUS | ALEX | PAUIRA |
| PAULA BELMIRO SAES | PAULA | BEOMIRO |
| TIAGO VASCO DIAS | IIAGO | VASCO DIAZ |
| VIVIANE DE SOUZA HEDEGUSCH | VIVIANE | HEDEGUSN |
| FABIO TOL FORTUNATO SILVA | FABIO | FORTUNATO |
| VITOR SANTOS COSTA | VICTOR | DOS SANTOS DA COSTA |
| INGRID PRADO SALGUEIRO | INGRIDI | PRADO SALGUEIRO |
| PAULO HENRIQUE DE JESUS | PAULO | HENRIQUE |
| ALEF RODRIGO MACIEL RAMOS | ALEF RODRIGUES | |
| VITORIA BORSATTI DE LIMA | VITORIA | BORSATE |
| ELEN DA MATA | HELLEN | MATA |
| JENIFER FLAVIA MAFRA | JENEFER | FLAVIA MAFRA |
| RICARDO JESUS DA SILVA | RICARDO | JESUS |
| FLAVIO EZIQUIEL PREREIRA SOUZA | TIAGO | EZAQUIEL |
| VINICIUS VICTOR PASSOS DA SILVA | VINICIUS | VICTOR PASSOS DA S |
| EDSON PRADO CAJAIBA | SHREK PRADO CAJAIBA | |
| JULIANA SOUZA FRANCA | JULIANA FRANA | |
| KELSON JARBAS DA SIVLA VAZ | KELSON | VAS |

| | | |
|---|---|---|
| WALDISNEY JACINTO DAS OLIVEIRAS | WALDISNEY | JACITO DA OLIVEIRA |
| JESSICA BRUNA RAMOS | JECCIKA | BRUNA |

Table 5: Outcomes and number of friends according to other characteristics

| | Total sample | Proficiency in language | Proficiency in Mathematics | Number of friends cited | Outdegree |
|---|---|---|---|---|---|
| Gender | | | | | |
| F | 3835 | 226.07 | 232.79 | 2.68 | 2.36 |
| M | 3865 | 210.43 | 234.57 | 2.37 | 1.98 |
| Total | 7700 | 218.28 | 233.68 | 2.53 | 2.17 |
| Race | | | | | |
| White | 2792 | 224.85 | 238.42 | 2.38 | 2 |
| Brown | 3872 | 216.45 | 232.72 | 2.66 | 2.23 |
| Black | 1018 | 209.37 | 226.76 | 2.51 | 2.02 |
| Asian | 29 | 211.43 | 229.84 | 2.03 | 1.52 |
| Indigenous | 121 | 207.71 | 227.88 | 2.54 | 2.04 |
| Did not answer | 161 | 207.81 | 222.96 | 1.56 | 1.12 |
| Total | 7993 | 218.2 | 233.72 | 2.52 | 2.09 |
| Age | | | | | |
| 12 | 6 | 199.4 | 220.94 | 4 | 1.67 |
| 13 | 23 | 233.43 | 249.17 | 2.48 | 2.26 |
| 14 | 4726 | 223.24 | 237.56 | 2.63 | 2.31 |
| 15 | 2054 | 213.72 | 229.97 | 2.42 | 2.05 |
| 16 | 640 | 201.42 | 221.65 | 2.26 | 1.82 |
| 17 | 210 | 200.68 | 216.97 | 1.98 | 1.53 |
| 18 and above | 41 | 188.54 | 213.49 | 2.07 | 1.71 |
| Total | 7700 | 218.28 | 233.68 | 2.53 | 2.17 |

Figure 1: Outcome distribution in Language according to self declared race

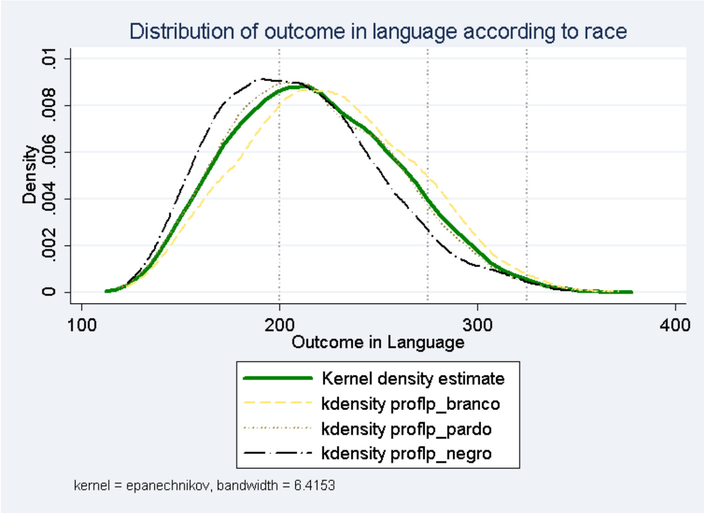
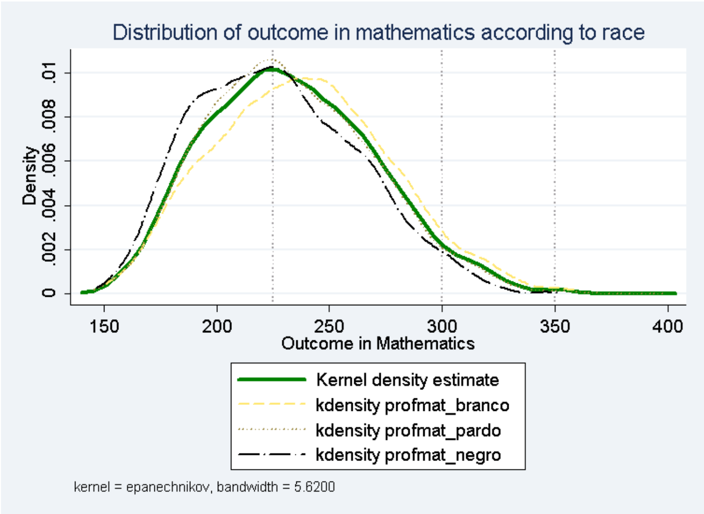



Figure 2: Outcome distribution in Mathematics according to self declared race



# 1 DIAGRAMS

## School 1

Figure 3: Diagram: school 1 - according to gender

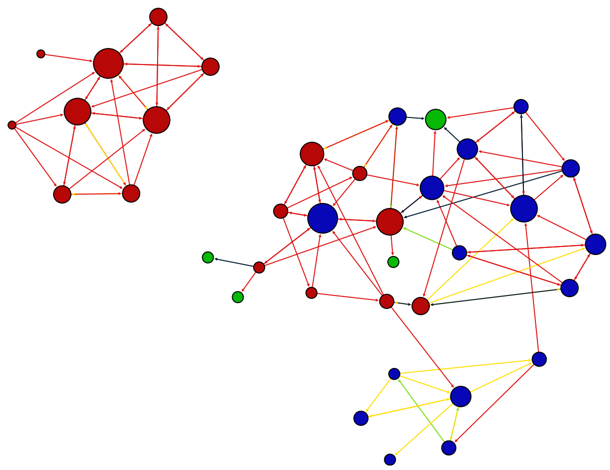


Figure 4: Diagram: school 1 - according to self declared race

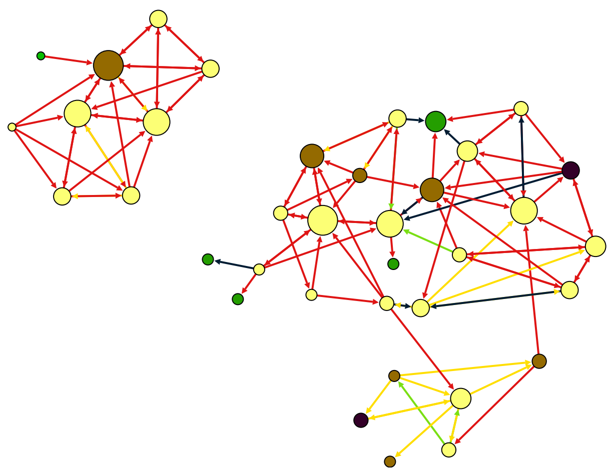
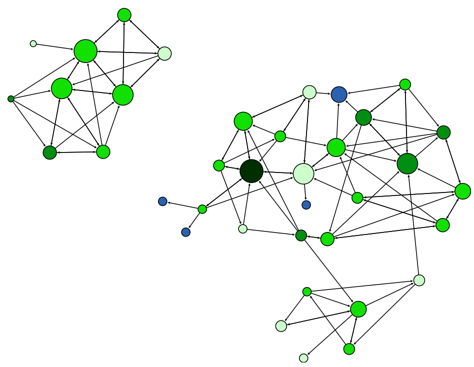
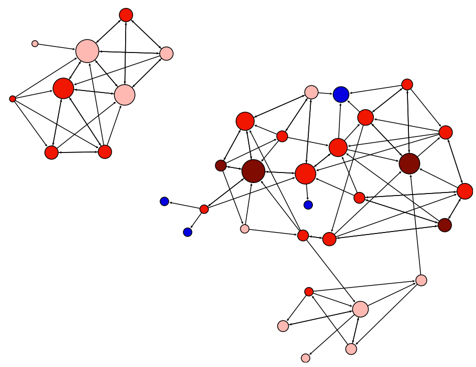


Figure 5: Diagram: school 1 - according to outcomes in Mathematics and Language



(a) Outcomes in Language



(b) Outcomes in Mathematics

**School 2**

Figure 6: Diagram: school 2 - according to gender

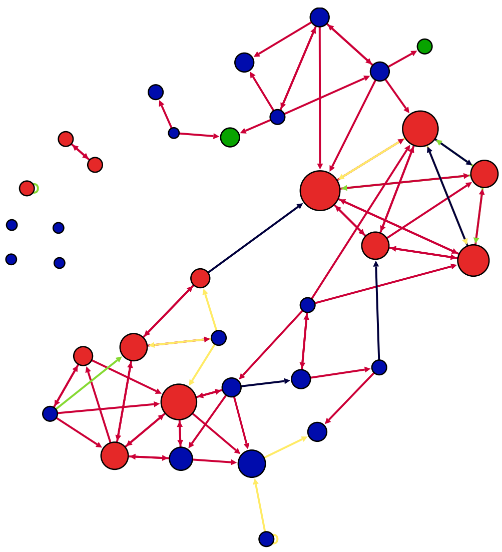


Figure 7: Diagram: school 2 - according to self declared race

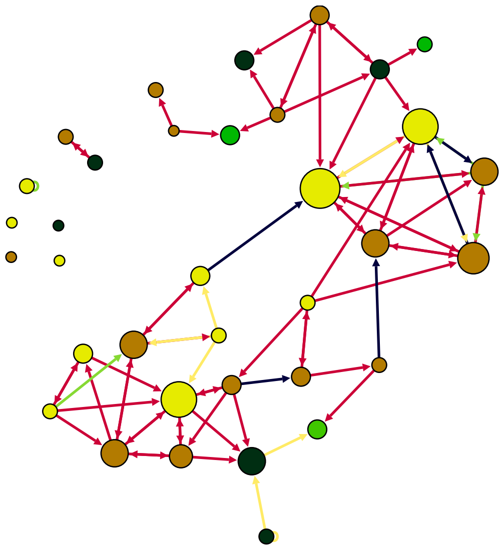
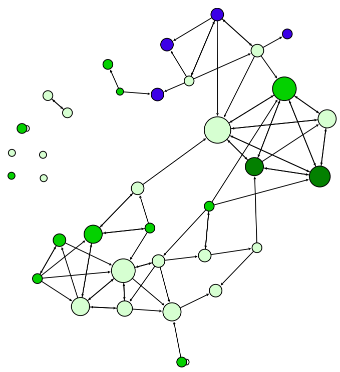
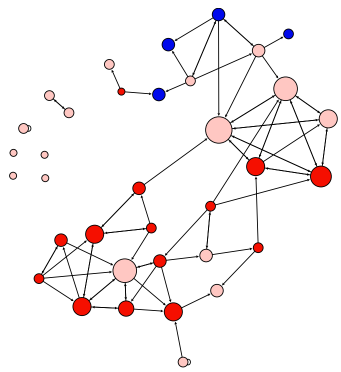


Figure 8: Diagram: school 2 - according to outcomes in Mathematics and Language



(a) Outcomes in Language



(b) Outcomes in Mathematics

**School 3**

Figure 9: Diagram: school 3 - according to gender

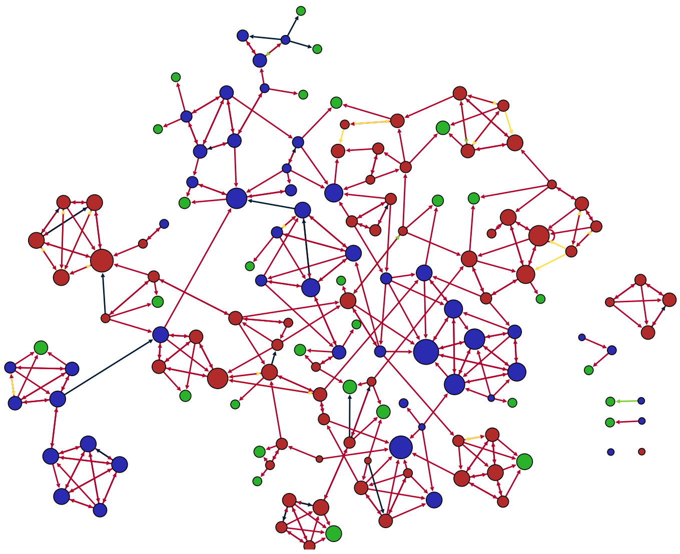


Figure 10: Diagram: school 3 - according to self declared race

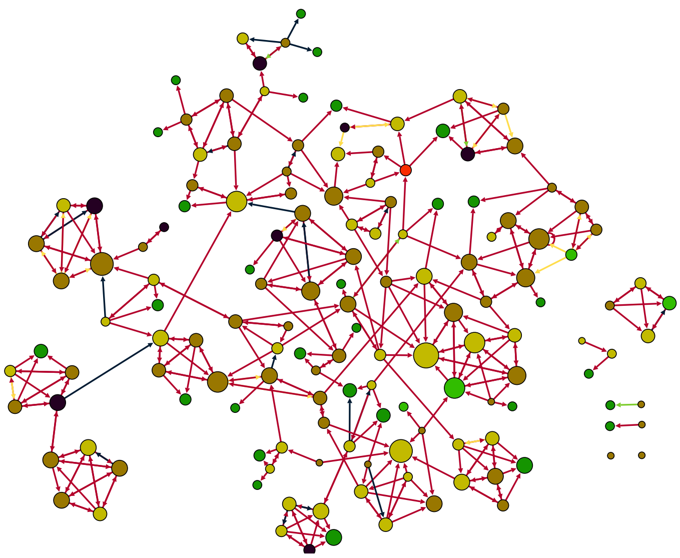
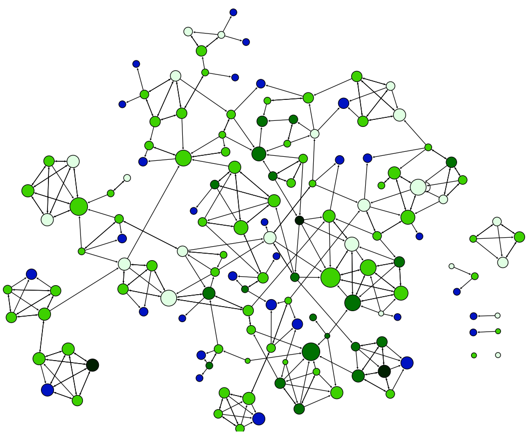
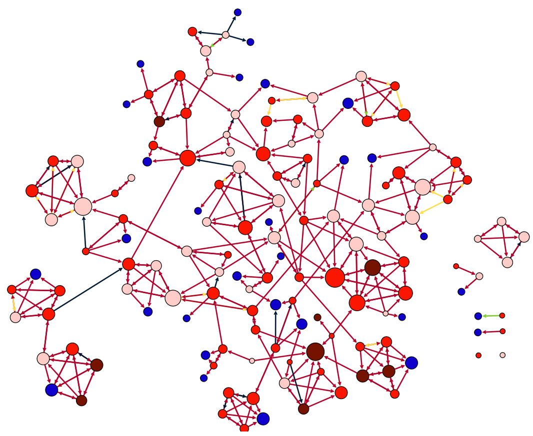


Figure 11: Diagram: school 3 - according to outcomes in Mathematics and Language



(a) Outcomes in Language



(b) Outcomes in Mathematics

**School 4**

Figure 12: Diagram: school 4 - according to gender

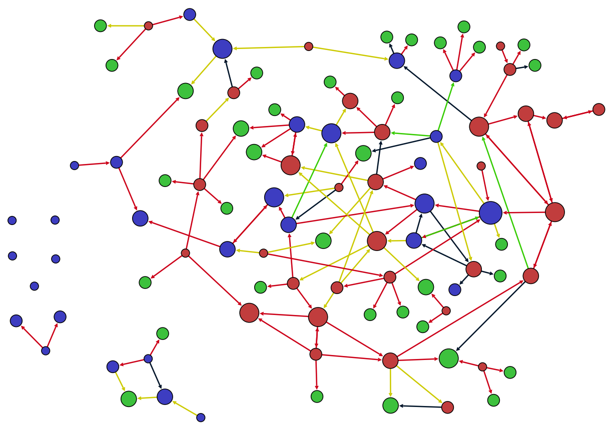


Figure 13: Diagram: school 4 - according to self declared race

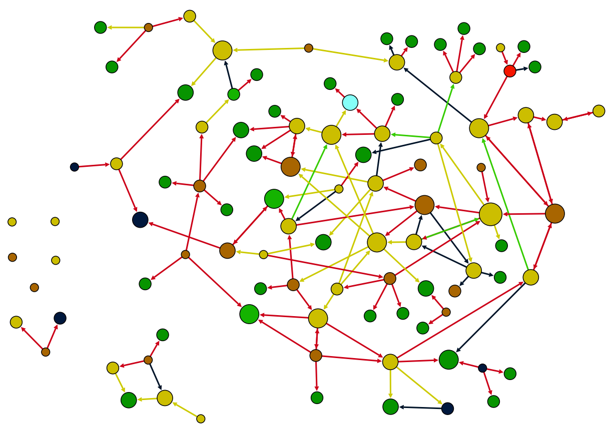
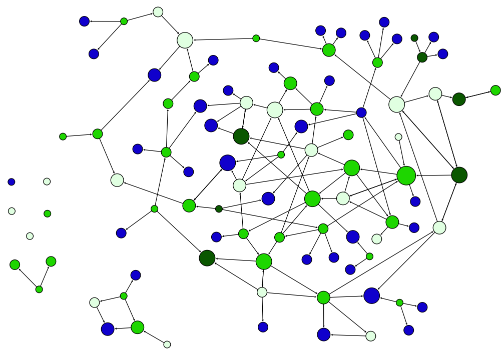
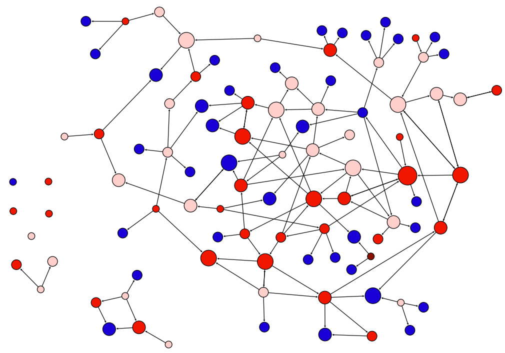


Figure 14: Diagram: school 4 - according to outcomes in Mathematics and Language



(a) Outcomes in Language



(b) Outcomes in Mathematics

**School 5**

Figure 15: Diagram: school 5 - according to gender

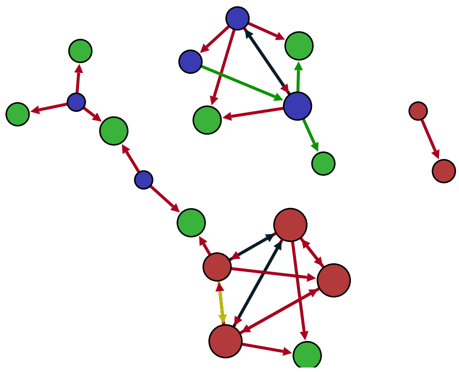


Figure 16: Diagram: school 5 - according to self declared race

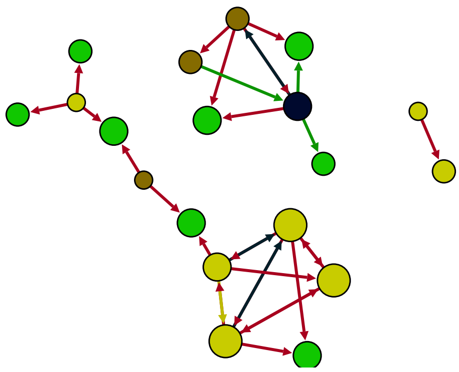
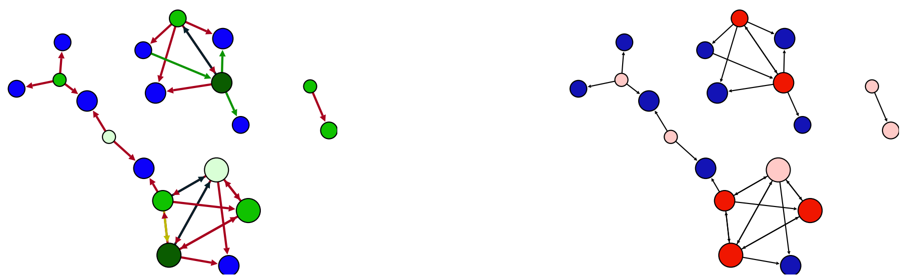


Figure 17: Diagram: school 5 - according to outcomes in Mathematics and Language



(a) Outcomes in Language

(b) Outcomes in Mathematics

**School 6**

Figure 18: Diagram: school 6 - according to gender

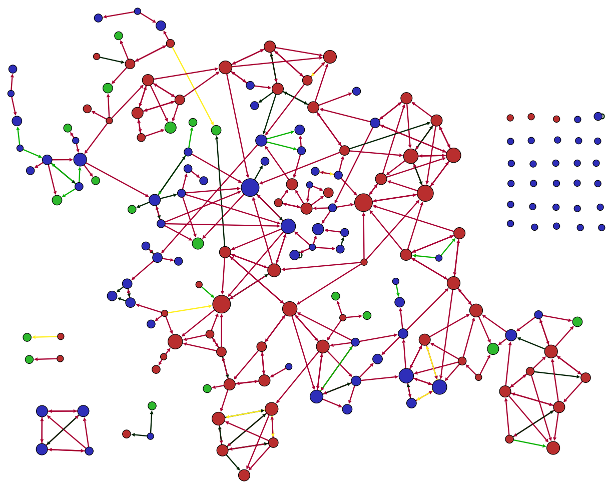


Figure 19: Diagram: school 6 - according to self declared race

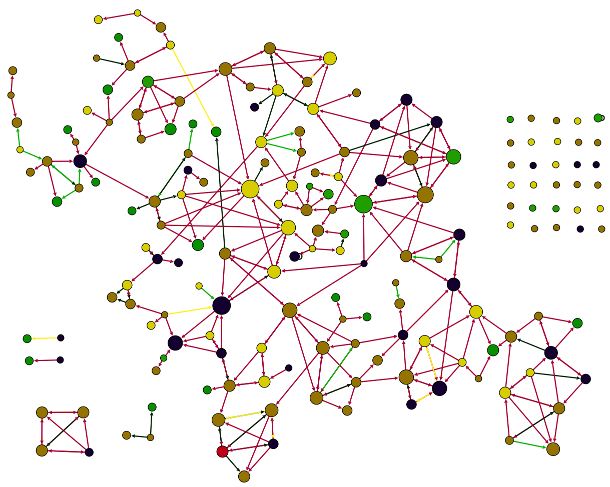
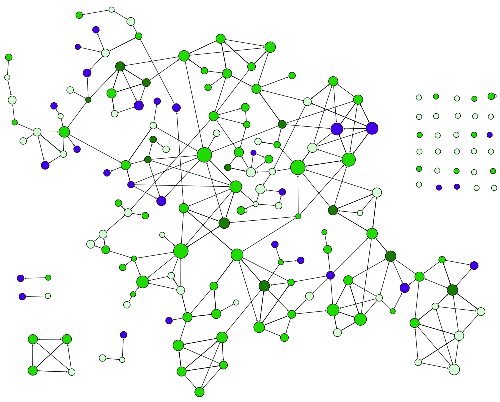
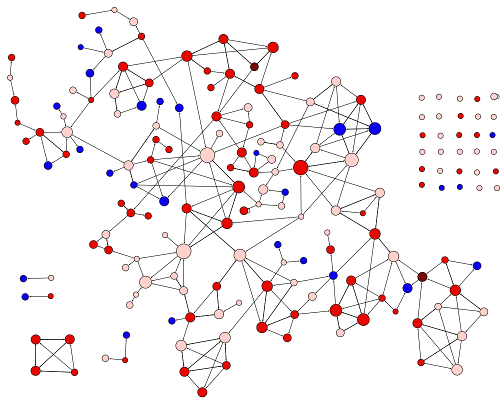


Figure 20: Diagram: school 6 - according to outcomes in Mathematics and Language



(a) Outcomes in Language



(b) Outcomes in Mathematics

**School 8**

Figure 21: Diagram: school 7 - according to gender

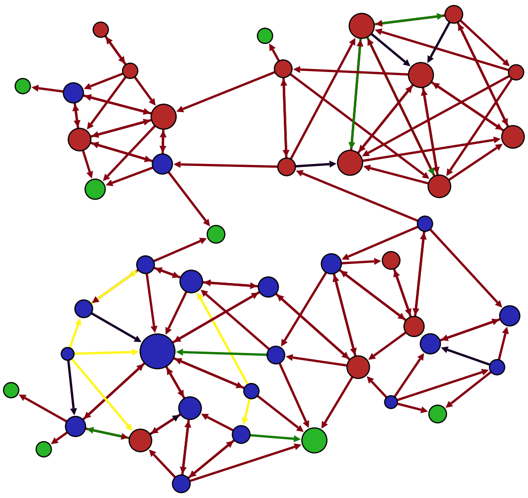


Figure 22: Diagram: school 7 - according to self declared race

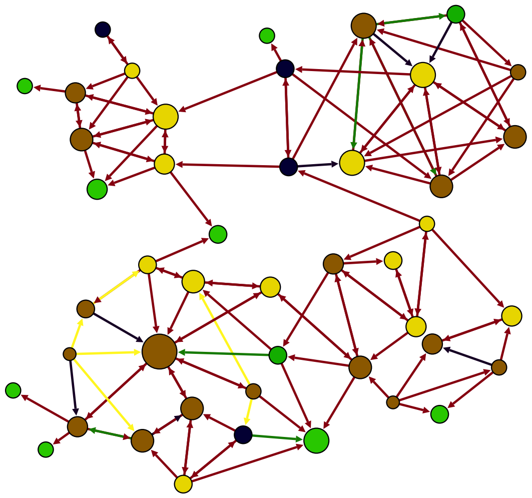
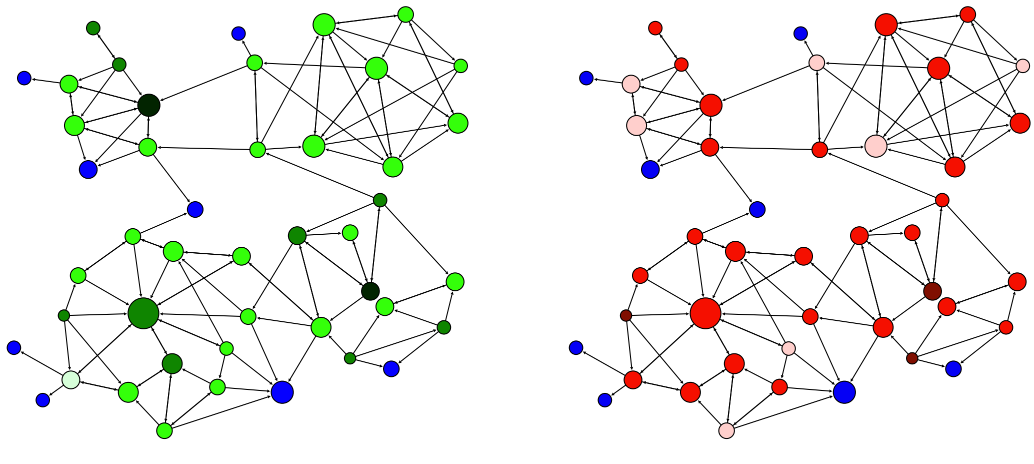


Figure 23: Diagram: school 7 - according to outcomes in Mathematics and Language



(a) Outcomes in Language

(b) Outcomes in Mathematics

# 2 NETWORKS' STATISTICS

Figure 24: Distribution of the number of citations - number of named friends

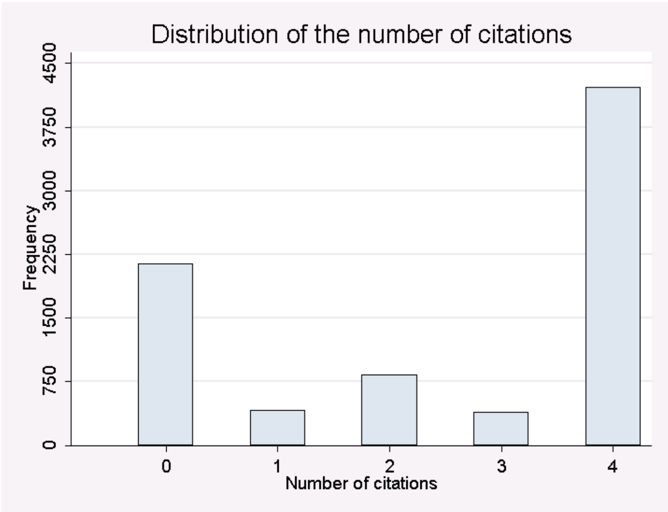



Figure 25: *Outdegree* distribution - number of found friends

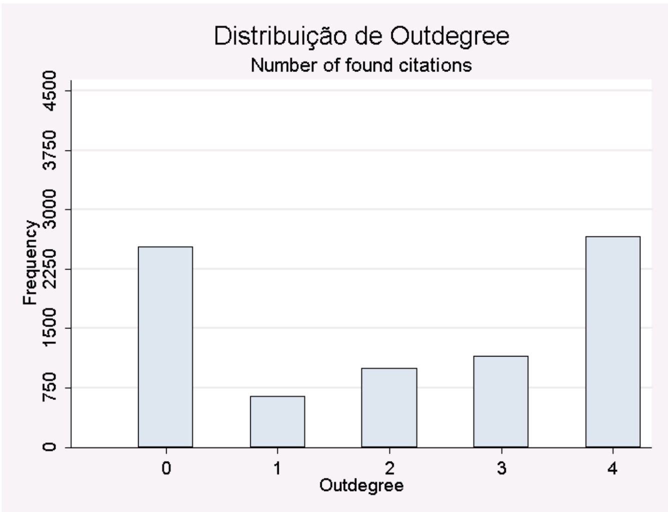


Figure 26: *Outdegree* distribution - number friends who were not found

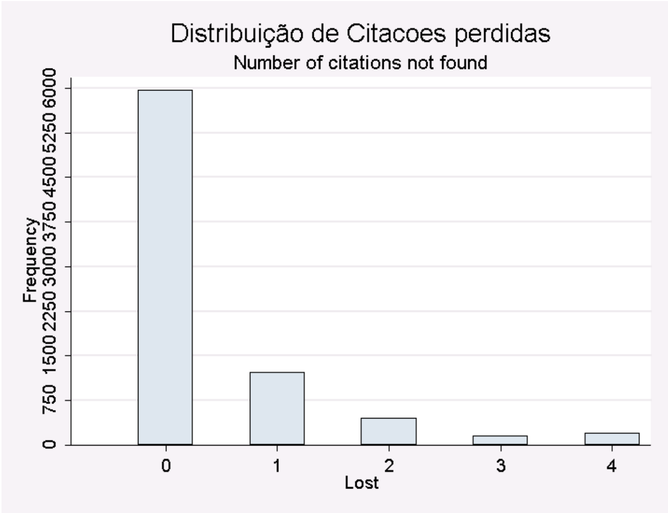



Figure 27: *Indegree* distribution - number of times one student was mentioned by the others among those who answered the questionnaire

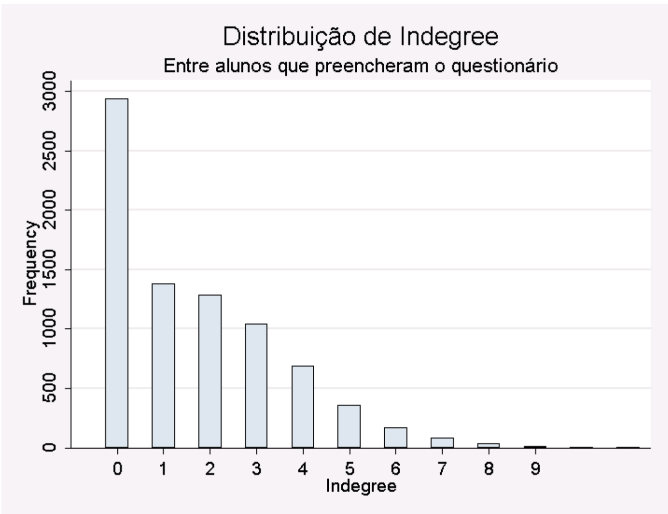



Figure 28: *Indegree* distribution - number of times one student was mentioned by the others among those who did not answer the questionnaire

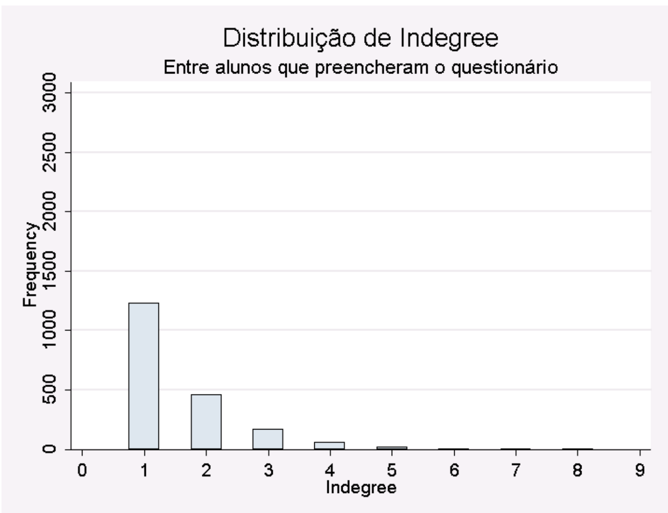


Table 6: Networks' statistics according to individual characteristics

| | N | Average number of citations made | Average number of found friends | Average number of citations received | Degree undirected | Clustering | Clustering directed |
|---|---|---|---|---|---|---|---|
| Gender | | | | | | | |
| F | 3285 | 3.13 | 2.76 | 2.27 | 3.64 | 0.39 | 0.43 |
| M | 3347 | 2.74 | 2.29 | 1.85 | 3.16 | 0.31 | 0.32 |
| Total | 6632 | 2.93 | 2.52 | 2.06 | 3.39 | 0.35 | 0.38 |
| Race | | | | | | | |
| White | 2268 | 2.83 | 2.46 | 2.05 | 3.34 | 0.36 | 0.39 |
| Pardo | 3257 | 3.07 | 2.65 | 2.1 | 3.52 | 0.35 | 0.38 |
| Black | 842 | 2.92 | 2.44 | 2.03 | 3.29 | 0.33 | 0.35 |
| Asian | 24 | 2.46 | 1.83 | 2 | 2.38 | 0.47 | 0.38 |
| Indian | 98 | 3.04 | 2.52 | 2.15 | 3.64 | 0.36 | 0.41 |
| NA | 143 | 1.65 | 1.27 | 1.29 | 1.89 | 0.39 | 0.35 |
| Total | 6632 | 2.93 | 2.52 | 2.06 | 3.39 | 0.35 | 0.38 |
| Age | | | | | | | |
| 12 | 6 | 4 | 1.67 | 0.17 | 1.5 | 0 | 0 |
| 13 | 19 | 3 | 2.74 | 2.05 | 3.95 | 0.35 | 0.33 |
| 14 | 4100 | 3.04 | 2.66 | 2.19 | 3.56 | 0.37 | 0.4 |
| 15 | 1742 | 2.85 | 2.41 | 1.91 | 3.22 | 0.33 | 0.36 |
| 16 | 560 | 2.58 | 2.08 | 1.77 | 2.94 | 0.31 | 0.31 |
| 17 | 169 | 2.46 | 1.91 | 1.53 | 2.78 | 0.29 | 0.3 |
| 18 and above | 36 | 2.42 | 2 | 1.58 | 2.63 | 0.21 | 0.21 |
| Total | 6632 | 2.93 | 2.52 | 2.06 | 3.39 | 0.35 | 0.38 |
| Outcome in Language | | | | | | | |
| Below basic | 2319 | 2.81 | 2.28 | 1.75 | 3.03 | 0.34 | 0.33 |
| Basic | 3389 | 3.03 | 2.67 | 2.2 | 3.6 | 0.36 | 0.4 |
| Adequate | 601 | 3.15 | 2.93 | 2.65 | 4.02 | 0.4 | 0.45 |
| Advanced | 44 | 3.45 | 3.14 | 3.16 | 4.52 | 0.35 | 0.42 |
| Total | 6353 | 2.96 | 2.56 | 2.09 | 3.44 | 0.36 | 0.38 |
| Outcome in Mathematics | | | | | | | |
| Below basic | 2826 | 2.84 | 2.37 | 1.87 | 3.17 | 0.35 | 0.36 |
| Basic | 3231 | 3.06 | 2.69 | 2.24 | 3.62 | 0.36 | 0.4 |
| Adequate | 283 | 3.04 | 2.86 | 2.6 | 3.99 | 0.38 | 0.43 |
| Advanced | 13 | 3.54 | 3.23 | 2.23 | 3.85 | 0.38 | 0.41 |
| Total | 6353 | 2.96 | 2.56 | 2.09 | 3.44 | 0.36 | 0.38 |

Table 7: Networks' statistics according to individual characteristics - II

| | N | Average shortest path | Centrality | Decay centrality 0.25 | Decay centrality 0.50 | Decay centrality 0.75 |
|---|---|---|---|---|---|---|
| Gender | | | | | | |
| F | 3285 | 4.89 | 24.34 | 1.55 | 6.79 | 27.03 |
| M | 3347 | 4.74 | 26.61 | 1.41 | 6.3 | 25.09 |
| Total | 6632 | 4.82 | 25.45 | 1.48 | 6.54 | 26.05 |
| | | | | | | |
| Race | | | | | | |
| White | 2268 | 4.76 | 25.02 | 1.46 | 6.41 | 25.2 |
| Pardo | 3257 | 4.85 | 25.64 | 1.53 | 6.76 | 27.11 |
| Black | 842 | 4.81 | 26.05 | 1.45 | 6.46 | 25.87 |
| Asian | 24 | 4.94 | 25.64 | 1.13 | 5.56 | 22.68 |
| Indian | 98 | 4.8 | 25.4 | 1.58 | 6.94 | 27.37 |
| NA | 143 | 5.1 | 22.83 | 0.86 | 3.94 | 16.31 |
| Total | 6632 | 4.82 | 25.45 | 1.48 | 6.54 | 26.05 |
| | | | | | | |
| Age | | | | | | |
| 12 | 6 | 3.51 | 66.31 | 0.53 | 2.19 | 9.48 |
| 13 | 19 | 5.83 | 25.15 | 1.76 | 7.65 | 30.88 |
| 14 | 4100 | 4.83 | 25.41 | 1.55 | 6.78 | 26.81 |
| 15 | 1742 | 4.81 | 25.23 | 1.41 | 6.3 | 25.51 |
| 16 | 560 | 4.81 | 25.41 | 1.32 | 5.96 | 23.85 |
| 17 | 169 | 4.55 | 27.15 | 1.2 | 5.27 | 21.06 |
| 18 and above | 36 | 4.57 | 26.99 | 1.21 | 5.7 | 24.38 |
| Total | 6632 | 4.82 | 25.45 | 1.48 | 6.54 | 26.05 |
| | | | | | | |
| Outcome in Language | | | | | | |
| Below basic | 2319 | 4.85 | 26.4 | 1.35 | 6.09 | 24.82 |
| Basic | 3389 | 4.83 | 24.71 | 1.56 | 6.85 | 27.16 |
| Adequate | 601 | 4.7 | 25.76 | 1.71 | 7.24 | 27.37 |
| Advanced | 44 | 4.82 | 23.74 | 1.92 | 8.13 | 30.63 |
| Total | 6353 | 4.82 | 25.4 | 1.5 | 6.62 | 26.35 |
| | | | | | | |
| Outcome in Mathematics | | | | | | |
| Below basic | 2826 | 4.82 | 26.01 | 1.39 | 6.26 | 25.3 |
| Basic | 3231 | 4.84 | 25.03 | 1.57 | 6.88 | 27.22 |
| Adequate | 283 | 4.57 | 24.19 | 1.7 | 7.18 | 26.8 |
| Advanced | 13 | 5.52 | 18.31 | 1.66 | 6.96 | 26.71 |
| Total | 6353 | 4.82 | 25.4 | 1.5 | 6.62 | 26.35 |

Table 8: Average indegree according to race and other characteristics

| | White | Brown | Black | Asian | Indigenous | NA | Total |
|---|---|---|---|---|---|---|---|
| Gender | | | | | | | |
| F | 1.88 | 1.99 | 1.97 | 2 | 2.13 | 1.52 | 1.94 |
| M | 1.56 | 1.66 | 1.62 | 1.44 | 1.53 | 0.84 | 1.6 |
| Total | 1.66 | 1.77 | 1.68 | 1.66 | 1.74 | 1.13 | 1.77 |
| Age | | | | | | | |
| 12 | 0.5 | 0 | 0 | | | | 0.17 |
| 13 | 1.13 | 2.31 | 0 | | | 0 | 1.7 |
| 14 | 1.85 | 1.95 | 1.9 | 1.5 | 2.06 | 1.43 | 1.9 |
| 15 | 1.55 | 1.71 | 1.55 | 1.88 | 1.36 | 1.07 | 1.62 |
| 16 | 1.47 | 1.61 | 1.6 | 1.33 | 1.1 | 1.17 | 1.55 |
| 17 | 1.24 | 1.03 | 1.86 | 0 | 2.25 | 0.25 | 1.23 |
| Above 18 | 0.8 | 1.4 | 2.75 | 5 | | 0 | 1.36 |
| Total | 1.66 | 1.77 | 1.68 | 1.66 | 1.74 | 1.13 | 1.77 |
| Outcome in Language | | | | | | | |
| Below Basic | 1.34 | 1.57 | 1.55 | 1.18 | 1.5 | 0.98 | 1.48 |
| Basic | 1.77 | 1.9 | 1.87 | 1.83 | 2.02 | 1.4 | 1.84 |
| Adequate | 2.19 | 2.33 | 1.9 | 2.67 | 3.25 | 1.6 | 2.23 |
| Advanced | 2.07 | 2.73 | 2.5 | | | | 2.4 |
| Total | 1.66 | 1.77 | 1.68 | 1.66 | 1.74 | 1.13 | 1.75 |
| Outcome in Mathematics | | | | | | | |
| Below Basic | 1.53 | 1.65 | 1.66 | 1.93 | 1.37 | 1.07 | 1.6 |
| Basic | 1.78 | 1.93 | 1.78 | 1.44 | 2.12 | 1.42 | 1.85 |
| Adequate | 2.01 | 2.3 | 2.15 | 1.5 | 5.75 | 1.5 | 2.19 |
| Advanced | 1.89 | 1.11 | 2 | 0 | | | 1.45 |
| Total | 1.66 | 1.77 | 1.68 | 1.66 | 1.74 | 1.13 | 1.75 |

Figure 29: Distribuição do *indegree* de acordo com o desempenho em Português e com a raça autodeclarada

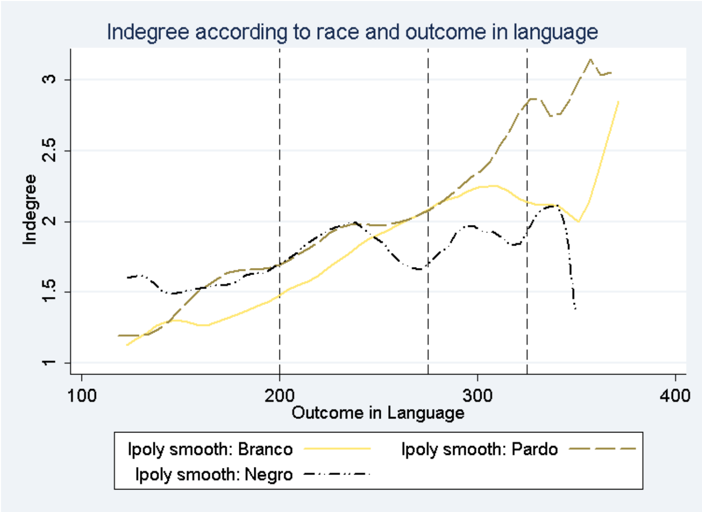



Figure 30: Distribuição do *indegree* de acordo com o desempenho em Português e com a raça autodeclarada - incluindo intervalos de confiança

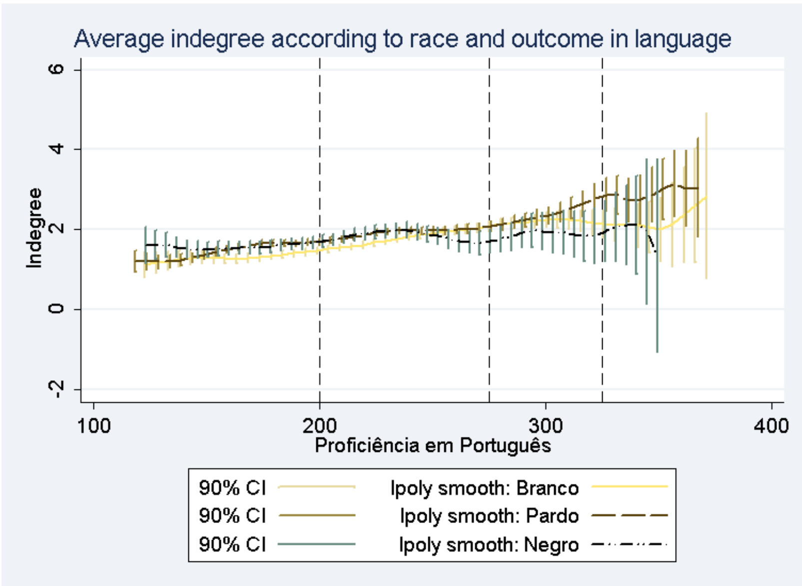


Figure 31: Distribuição do *indegree* de acordo com o desempenho em Matemática e com a raça autodeclarada

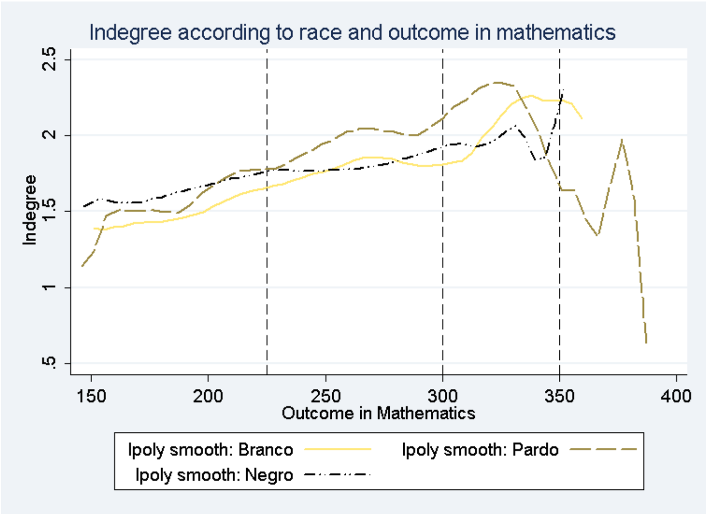



Figure 32: Distribuição do *indegree* de acordo com o desempenho em Matemática e com a raça autodeclarada - incluindo intervalos de confiança

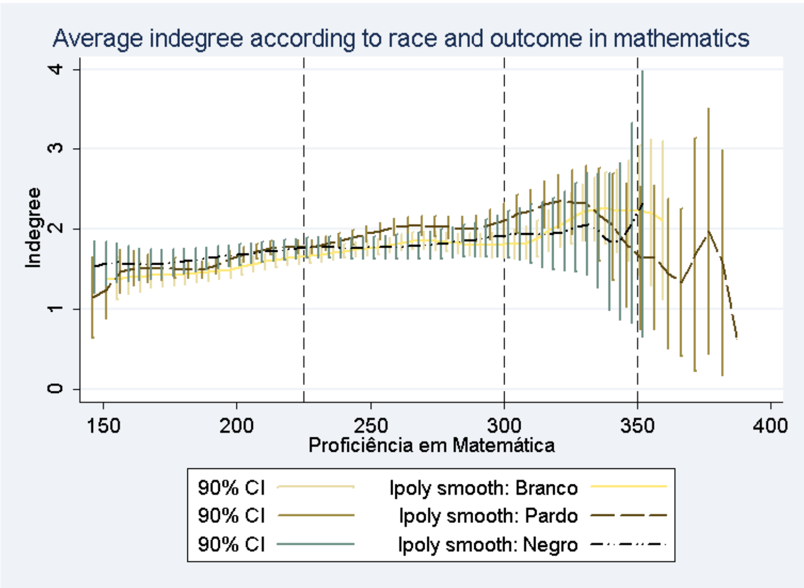


Table 9: Evidences of homophily

| | N | Male friends | Male students | White friends | White students | Brown friends | Brown students | Black friends | Black students |
|---|---|---|---|---|---|---|---|---|---|
| Gender | | | | | | | | | |
| F | 3835 | 0.17 | 0.49 | 0.28 | 0.35 | 0.41 | 0.49 | 0.09 | 0.13 |
| M | 3865 | 0.81 | 0.51 | 0.27 | 0.35 | 0.4 | 0.49 | 0.11 | 0.13 |
| Total | 7700 | 0.47 | 0.5 | 0.28 | 0.35 | 0.41 | 0.49 | 0.1 | 0.13 |
| | | | | | | | | | |
| Race | | | | | | | | | |
| White | 2792 | 0.47 | 0.48 | 0.34 | 0.38 | 0.38 | 0.46 | 0.09 | 0.12 |
| Brown | 3872 | 0.47 | 0.48 | 0.25 | 0.33 | 0.43 | 0.5 | 0.1 | 0.13 |
| Black | 1018 | 0.5 | 0.48 | 0.22 | 0.33 | 0.4 | 0.48 | 0.17 | 0.15 |
| Asian | 29 | 0.67 | 0.52 | 0.28 | 0.42 | 0.48 | 0.41 | 0.12 | 0.11 |
| Indigenous | 121 | 0.48 | 0.48 | 0.23 | 0.34 | 0.38 | 0.48 | 0.1 | 0.13 |
| NA | 161 | 0.4 | 0.49 | 0.29 | 0.33 | 0.45 | 0.46 | 0.05 | 0.14 |
| Total | 7993 | 0.47 | 0.48 | 0.28 | 0.35 | 0.41 | 0.48 | 0.1 | 0.13 |
| | | | | | | | | | |
| Age | | | | | | | | | |
| 12 | 6 | 0.5 | 0.49 | 0.08 | 0.32 | 0.08 | 0.54 | 0 | 0.11 |
| 13 | 23 | 0.22 | 0.5 | 0.17 | 0.33 | 0.43 | 0.49 | 0.14 | 0.13 |
| 14 | 4726 | 0.45 | 0.5 | 0.29 | 0.35 | 0.41 | 0.49 | 0.1 | 0.12 |
| 15 | 2054 | 0.51 | 0.5 | 0.26 | 0.35 | 0.4 | 0.48 | 0.1 | 0.13 |
| 16 | 640 | 0.55 | 0.51 | 0.26 | 0.34 | 0.4 | 0.49 | 0.12 | 0.13 |
| 17 | 210 | 0.58 | 0.5 | 0.19 | 0.33 | 0.4 | 0.49 | 0.14 | 0.14 |
| Above 18 | 41 | 0.58 | 0.52 | 0.15 | 0.33 | 0.24 | 0.48 | 0.13 | 0.14 |
| Total | 7700 | 0.47 | 0.5 | 0.28 | 0.35 | 0.41 | 0.49 | 0.1 | 0.13 |
| | | | | | | | | | |
| Outcome in Language | | | | | | | | | |
| Below Basic | 2741 | 0.55 | 0.5 | 0.25 | 0.34 | 0.4 | 0.49 | 0.11 | 0.13 |
| Basic | 4052 | 0.44 | 0.49 | 0.28 | 0.35 | 0.41 | 0.48 | 0.1 | 0.13 |
| Adequate | 714 | 0.39 | 0.48 | 0.34 | 0.36 | 0.4 | 0.48 | 0.07 | 0.12 |
| Advanced | 58 | 0.45 | 0.48 | 0.41 | 0.39 | 0.44 | 0.47 | 0.04 | 0.11 |
| Total | 7565 | 0.47 | 0.49 | 0.28 | 0.35 | 0.41 | 0.49 | 0.1 | 0.13 |
| | | | | | | | | | |
| Outcome in Mathematics | | | | | | | | | |
| Below Basic | 3302 | 0.46 | 0.49 | 0.25 | 0.34 | 0.41 | 0.49 | 0.11 | 0.13 |
| Basic | 3907 | 0.47 | 0.49 | 0.29 | 0.35 | 0.41 | 0.49 | 0.1 | 0.12 |
| Adequate | 336 | 0.52 | 0.48 | 0.35 | 0.38 | 0.39 | 0.47 | 0.08 | 0.11 |
| Advanced | 20 | 0.65 | 0.47 | 0.28 | 0.4 | 0.37 | 0.48 | 0.1 | 0.09 |
| Total | 7565 | 0.47 | 0.49 | 0.28 | 0.35 | 0.41 | 0.49 | 0.1 | 0.13 |

Figure 33: Average shortest path according to school size

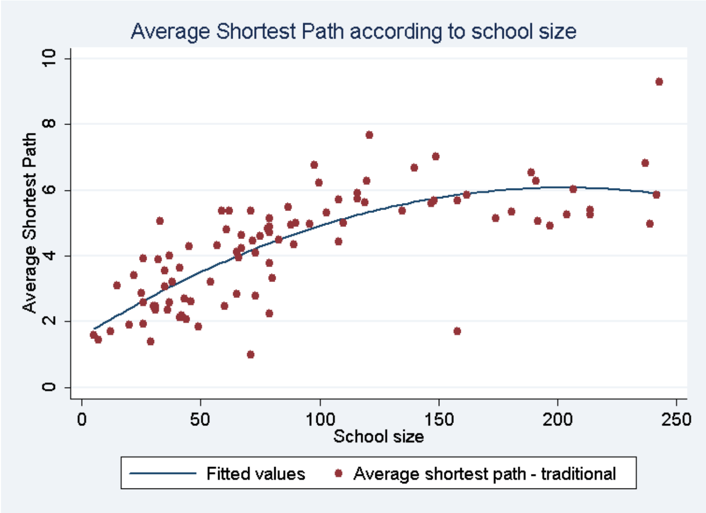



Figure 34: Diameter of the largest component according to school size

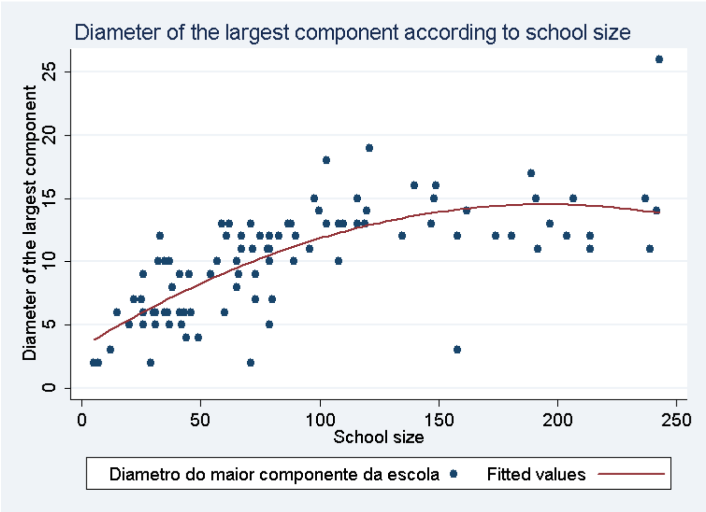


Figure 35: Number of components according to school size

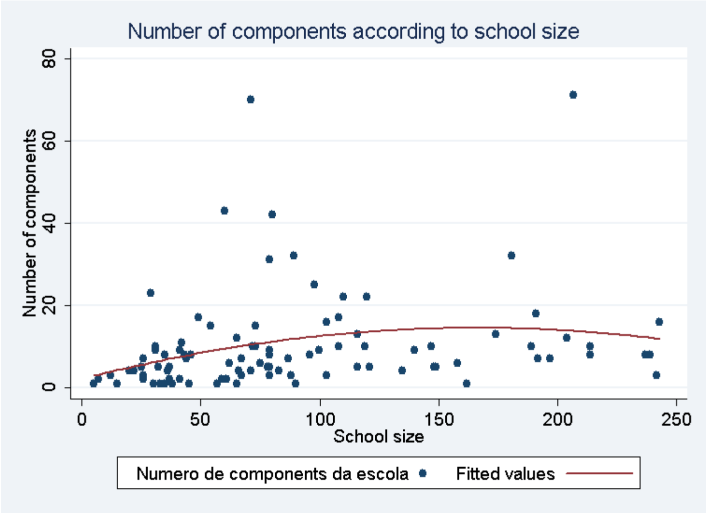



Figure 36: Size of the largest component according to the school size

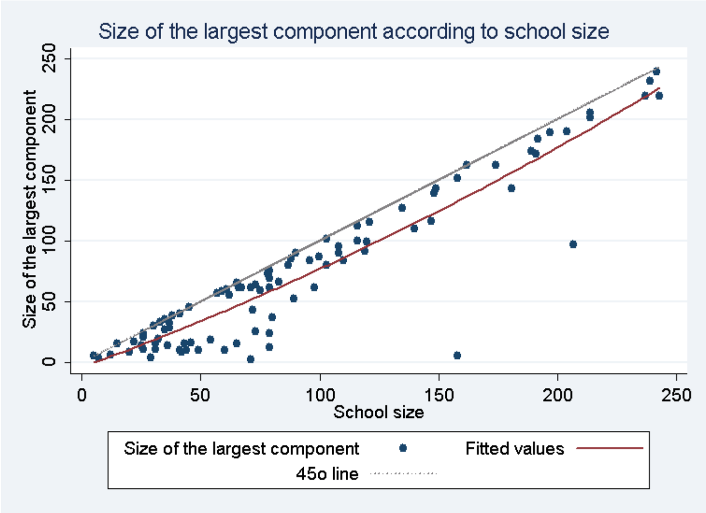



Figure 37: Percentage of the students on the largest component according to the school size

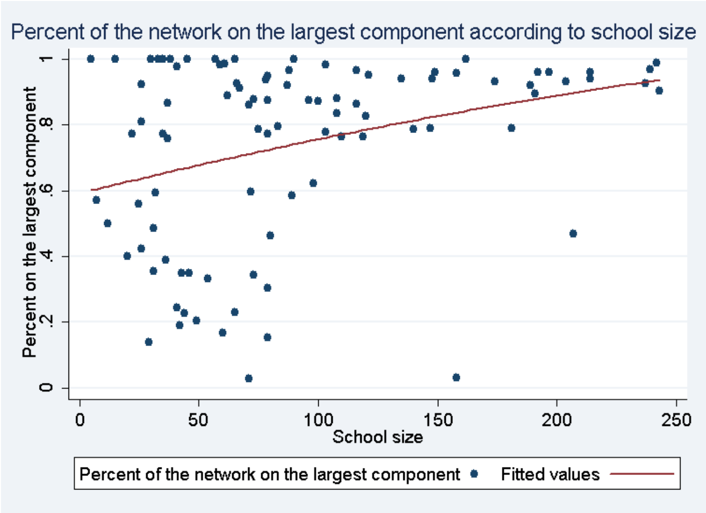


Figure 38: Overall clustering according to school size

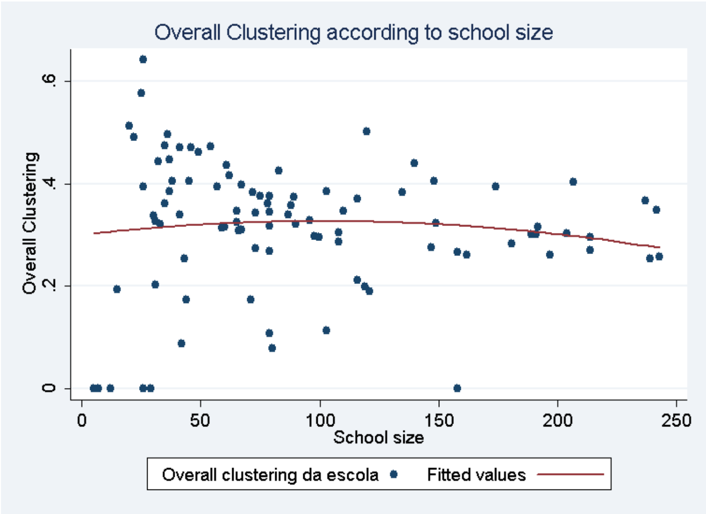



Figure 39: Individual clustering according to school size

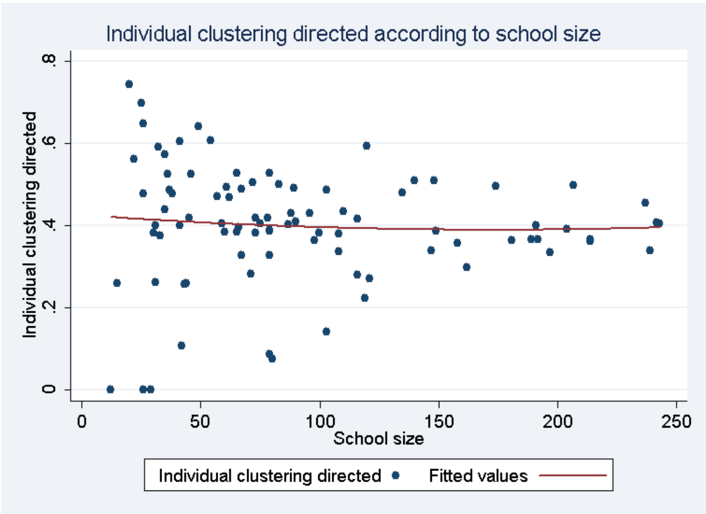


Figure 40: Positive Assortativity According to School Size

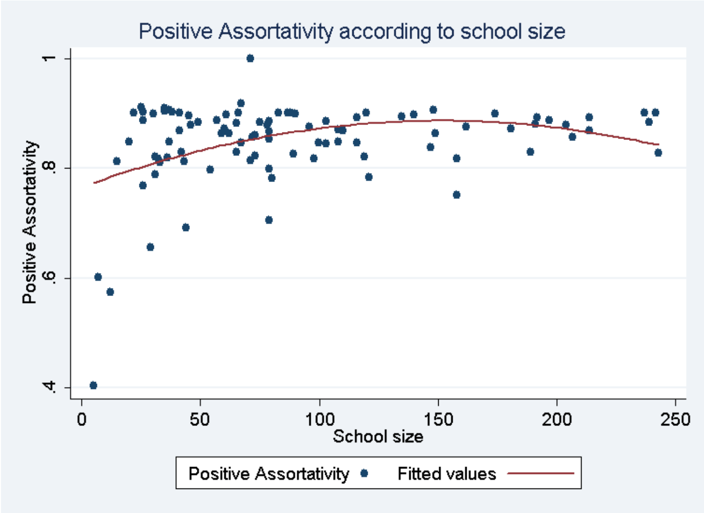



Figure 41: Positive Assortativity between White and Non-White According to School Size

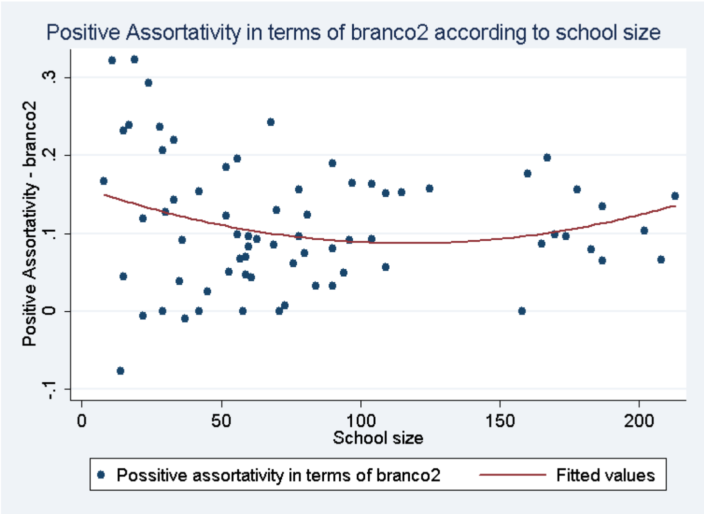


Figure 42: Positive Assortativity between Brown and Non-Brown According to School Size

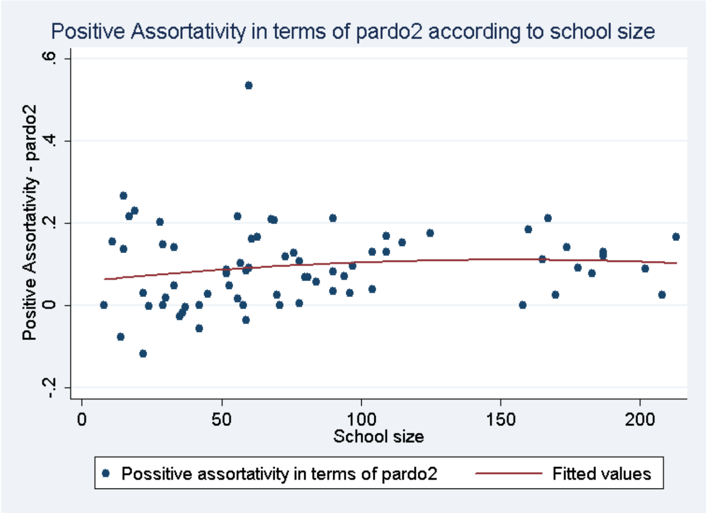



Figure 43: Positive Assortativity between Black and Non-Black According to School Size

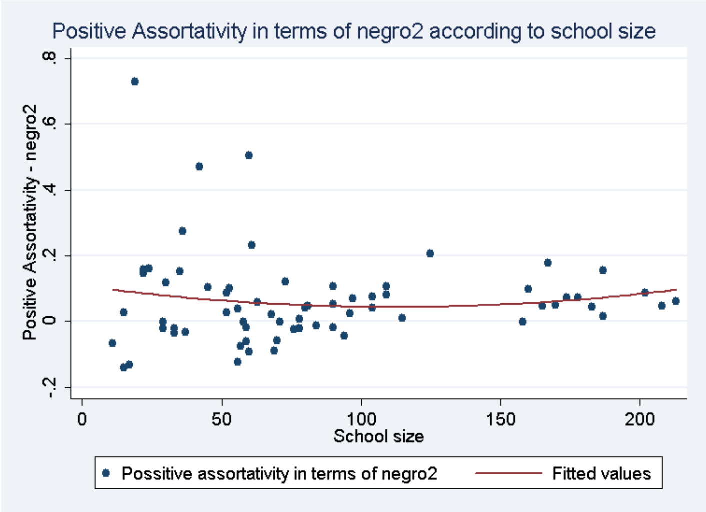



Figure 44: Positive Assortativity between Genders According to School Size

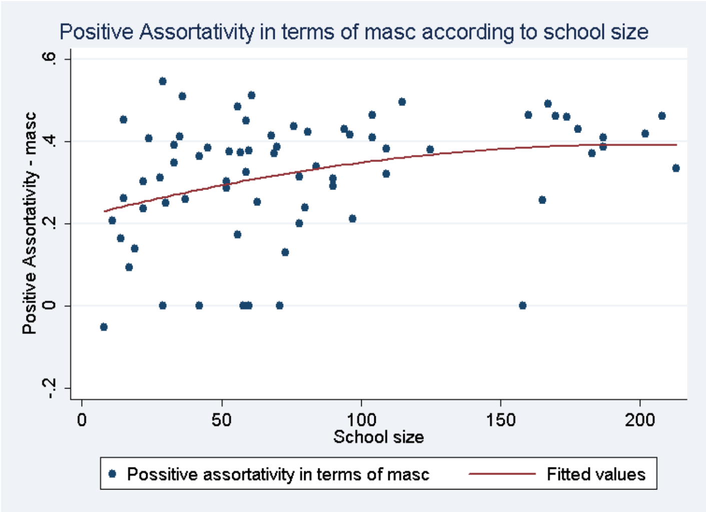
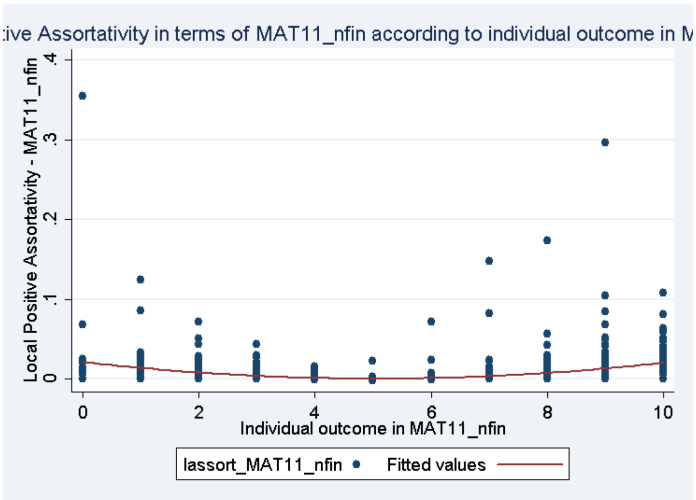


Figure 45: Positive Assortativity in terms of Outcome According to Individual Outcome



# 3 REGRESSIONS

Labels:
t2_el Interaction between dummies for treatment and elegibility
t24 Interaction between dummies for treatment and indicating that the student?s previous outcome in math was below 4
t25 - Interaction between dummies for treatment and indicating that the student?s previous outcome in math was between 4 and 5
t27 - Interaction between dummies for treatment and indicating that the student?s previous outcome in math was between 6 and 7
t210 - Interaction between dummies for treatment and indicating that the student?s previous outcome in math was above 7

Controls: Contols 1 - Final score in mathematics and language on the previous year (except for regressions with dummies for outcome ranges interacted with treatment, where they were replaced by dummies for each outcome range with less than 4 as a baseline) and dummy for gender
Controls 2- Dummies for race and number of students at the school
Controls 3 - Dummies indicating whether parents work and if the child?s responsible was a single mother (indicators of social status)

## 3.1 All students

Table 10: All Students: Probit - probability of being an isolated node

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | 0.424*** | 0.430*** | 0.413*** | 0.380*** | 0.391** | 0.327* | 0.308 | 0.179 | | | | |
| | (0.0767) | (0.0798) | (0.0801) | (0.0984) | (0.171) | (0.196) | (0.197) | (0.231) | | | | |
| elegivel | 0.290*** | -0.0767 | -0.0515 | -0.193 | 0.266* | -0.147 | -0.123 | -0.326 | | | | |
| | (0.0907) | (0.141) | (0.142) | (0.176) | (0.142) | (0.185) | (0.186) | (0.220) | | | | |
| t2_el | | | | | 0.0397 | 0.118 | 0.120 | 0.239 | | | | |
| | | | | | (0.183) | (0.206) | (0.206) | (0.247) | | | | |
| t24 | | | | | | | | | 0.753*** | 0.764*** | 0.762*** | 1.010*** |
| | | | | | | | | | (0.172) | (0.173) | (0.174) | (0.219) |
| t25 | | | | | | | | | 0.377*** | 0.374*** | 0.352*** | 0.173 |
| | | | | | | | | | (0.115) | (0.116) | (0.117) | (0.149) |
| t27 | | | | | | | | | 0.375*** | 0.359** | 0.341** | 0.394** |
| | | | | | | | | | (0.141) | (0.141) | (0.142) | (0.186) |
| t210 | | | | | | | | | 0.392** | 0.290 | 0.274 | 0.164 |
| | | | | | | | | | (0.172) | (0.196) | (0.197) | (0.232) |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |
| Constant | -1.171*** | 0.158 | 0.259 | 0.908** | -1.152*** | 0.220 | 0.322 | 1.024** | -1.127*** | -0.393 | -0.237 | 0.437 |
| | (0.175) | (0.323) | (0.328) | (0.436) | (0.195) | (0.340) | (0.344) | (0.451) | (0.196) | (0.357) | (0.365) | (0.486) |
| Observations | 3,167 | 3,075 | 3,075 | 2,254 | 3,167 | 3,075 | 3,075 | 2,254 | 3,167 | 3,075 | 3,075 | 2,254 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: All Students: Probit - probability of having positive outdegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.203*** | -0.181** | -0.155** | -0.0971 | -0.143 | 0.0598 | 0.0965 | 0.0764 | | | | |
| | (0.0717) | (0.0743) | (0.0748) | (0.0913) | (0.156) | (0.176) | (0.178) | (0.218) | | | | |
| elegivel | -0.319*** | -0.0410 | -0.0831 | -0.130 | -0.279** | 0.104 | 0.0677 | -0.0241 | | | | |
| | (0.0835) | (0.129) | (0.131) | (0.163) | (0.124) | (0.159) | (0.160) | (0.201) | | | | |
| t2_el | | | | | -0.0720 | -0.277 | -0.289 | -0.202 | | | | |
| | | | | | (0.166) | (0.184) | (0.185) | (0.231) | | | | |
| t24 | | | | | | | | | -0.254 | -0.283* | -0.269 | -0.388* |
| | | | | | | | | | (0.163) | (0.165) | (0.166) | (0.206) |
| t25 | | | | | | | | | -0.266** | -0.259** | -0.226** | -0.0817 |
| | | | | | | | | | (0.106) | (0.108) | (0.108) | (0.136) |
| t27 | | | | | | | | | -0.177 | -0.143 | -0.123 | -0.0824 |
| | | | | | | | | | (0.125) | (0.126) | (0.126) | (0.158) |
| t210 | | | | | | | | | -0.142 | 0.0919 | 0.123 | 0.0896 |
| | | | | | | | | | (0.156) | (0.176) | (0.178) | (0.219) |
| Constant | 1.144*** | -0.0471 | -0.218 | -0.223 | 1.113*** | -0.170 | -0.347 | -0.316 | 1.103*** | 0.0502 | -0.222 | -0.372 |
| | (0.167) | (0.303) | (0.307) | (0.408) | (0.182) | (0.312) | (0.317) | (0.421) | (0.182) | (0.326) | (0.334) | (0.448) |
| Observations | 3,148 | 3,048 | 3,048 | 2,263 | 3,148 | 3,048 | 3,048 | 2,263 | 3,148 | 3,048 | 3,048 | 2,263 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: All Students: Probit - probability of having positive indegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.337*** | -0.349*** | -0.346*** | -0.304*** | -0.277** | -0.282* | -0.279* | -0.173 | | | | |
| | (0.0603) | (0.0622) | (0.0623) | (0.0733) | (0.129) | (0.148) | (0.148) | (0.169) | | | | |
| Eligible | -0.270*** | -0.0202 | -0.0266 | 0.0381 | -0.231** | 0.0210 | 0.0148 | 0.117 | | | | |
| | (0.0697) | (0.109) | (0.110) | (0.130) | (0.102) | (0.137) | (0.137) | (0.158) | | | | |
| Treated x Elegible | | | | | -0.0734 | -0.0771 | -0.0774 | -0.156 | | | | |
| | | | | | (0.139) | (0.156) | (0.156) | (0.180) | | | | |
| Treated x Math below 4 | | | | | | | | | -0.490*** | -0.507*** | -0.508*** | -0.536*** |
| | | | | | | | | | (0.145) | (0.146) | (0.146) | (0.176) |
| Treated x Math between 4 and 5 | | | | | | | | | -0.374*** | -0.373*** | -0.367*** | -0.259** |
| | | | | | | | | | (0.0896) | (0.0903) | (0.0904) | (0.110) |
| Treated x Math between 6 and 7 | | | | | | | | | -0.282*** | -0.281*** | -0.282*** | -0.317*** |
| | | | | | | | | | (0.102) | (0.103) | (0.103) | (0.122) |
| Treated x Math above 7 | | | | | | | | | -0.278** | -0.257* | -0.256* | -0.163 |
| | | | | | | | | | (0.129) | (0.148) | (0.148) | (0.169) |
| Constant | 0.605*** | -0.515* | -0.565** | -0.959*** | 0.573*** | -0.549** | -0.599** | -1.024*** | 0.559*** | -0.0638 | -0.131 | -0.543 |
| | (0.151) | (0.265) | (0.269) | (0.341) | (0.162) | (0.274) | (0.277) | (0.349) | (0.163) | (0.281) | (0.288) | (0.364) |
| Observations | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 13: All Students : Degree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.501*** | -0.523*** | -0.514*** | -0.469*** | -0.285* | -0.315* | -0.302* | -0.277 | | | | |
| | (0.0787) | (0.0792) | (0.0793) | (0.0898) | (0.156) | (0.168) | (0.168) | (0.183) | | | | |
| elegivel | -0.617*** | -0.113 | -0.124 | -0.0961 | -0.485*** | 0.00269 | -0.00512 | 0.00935 | | | | |
| | (0.0844) | (0.132) | (0.132) | (0.148) | (0.118) | (0.155) | (0.155) | (0.172) | | | | |
| t2_el | | | | | -0.270 | -0.250 | -0.255 | -0.235 | | | | |
| | | | | | (0.168) | (0.178) | (0.178) | (0.196) | | | | |
| t24 | | | | | | | | | -1.038*** | -1.135*** | -1.140*** | -1.294*** |
| | | | | | | | | | (0.203) | (0.203) | (0.203) | (0.238) |
| t25 | | | | | | | | | -0.560*** | -0.551*** | -0.538*** | -0.390*** |
| | | | | | | | | | (0.117) | (0.117) | (0.117) | (0.135) |
| t27 | | | | | | | | | -0.411*** | -0.404*** | -0.395*** | -0.395*** |
| | | | | | | | | | (0.125) | (0.124) | (0.124) | (0.141) |
| t210 | | | | | | | | | -0.282* | -0.266 | -0.253 | -0.240 |
| | | | | | | | | | (0.155) | (0.167) | (0.167) | (0.183) |
| Constant | 2.604*** | 0.153 | 0.0584 | -0.144 | 2.501*** | 0.0667 | -0.0308 | -0.219 | | | | |
| | (0.212) | (0.349) | (0.353) | (0.428) | (0.221) | (0.355) | (0.358) | (0.433) | | | | |
| Observations | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 |
| R-squared | 0.165 | 0.195 | 0.196 | 0.184 | 0.166 | 0.195 | 0.196 | 0.185 | 0.792 | 0.798 | 0.798 | 0.819 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14: All Students : Indegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.394*** | -0.424*** | -0.422*** | -0.452*** | -0.217 | -0.296* | -0.293* | -0.251 | | | | |
| | (0.0726) | (0.0733) | (0.0734) | (0.0855) | (0.143) | (0.155) | (0.155) | (0.175) | | | | |
| elegivel | -0.486*** | -0.0671 | -0.0664 | 0.00745 | -0.378*** | 0.00427 | 0.00573 | 0.118 | | | | |
| | (0.0779) | (0.122) | (0.122) | (0.141) | (0.109) | (0.144) | (0.144) | (0.164) | | | | |
| t2_el | | | | | -0.221 | -0.154 | -0.155 | -0.247 | | | | |
| | | | | | (0.155) | (0.164) | (0.165) | (0.187) | | | | |
| t24 | | | | | | | | | -0.723*** | -0.808*** | -0.810*** | -0.912*** |
| | | | | | | | | | (0.188) | (0.188) | (0.188) | (0.227) |
| t25 | | | | | | | | | -0.430*** | -0.420*** | -0.420*** | -0.414*** |
| | | | | | | | | | (0.108) | (0.108) | (0.108) | (0.129) |
| t27 | | | | | | | | | -0.376*** | -0.379*** | -0.376*** | -0.452*** |
| | | | | | | | | | (0.116) | (0.115) | (0.115) | (0.135) |
| t210 | | | | | | | | | -0.219 | -0.256* | -0.252 | -0.219 |
| | | | | | | | | | (0.143) | (0.155) | (0.155) | (0.174) |
| Constant | 1.643*** | -0.479 | -0.478 | -0.974** | 1.559*** | -0.533 | -0.532 | -1.053** | | | | |
| | (0.196) | (0.323) | (0.327) | (0.408) | (0.204) | (0.328) | (0.332) | (0.412) | | | | |
| Observations | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 |
| R-squared | 0.133 | 0.162 | 0.162 | 0.165 | 0.134 | 0.162 | 0.162 | 0.166 | 0.635 | 0.645 | 0.645 | 0.665 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15: All Students : Decay centrality 1

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.162*** | -0.174*** | -0.171*** | -0.164*** | -0.147** | -0.190*** | -0.185** | -0.175** | | | | |
| | (0.0340) | (0.0343) | (0.0344) | (0.0391) | (0.0672) | (0.0727) | (0.0727) | (0.0799) | | | | |
| elegivel | -0.268*** | -0.0737 | -0.0775 | -0.0804 | -0.259*** | -0.0825 | -0.0855 | -0.0862 | | | | |
| | (0.0365) | (0.0570) | (0.0571) | (0.0645) | (0.0509) | (0.0674) | (0.0674) | (0.0751) | | | | |
| t2_el | | | | | -0.0180 | 0.0191 | 0.0172 | 0.0129 | | | | |
| | | | | | (0.0724) | (0.0770) | (0.0770) | (0.0857) | | | | |
| t24 | | | | | | | | | -0.409*** | -0.448*** | -0.449*** | -0.522*** |
| | | | | | | | | | (0.0879) | (0.0879) | (0.0879) | (0.104) |
| t25 | | | | | | | | | -0.155*** | -0.151*** | -0.146*** | -0.0848 |
| | | | | | | | | | (0.0508) | (0.0506) | (0.0507) | (0.0590) |
| t27 | | | | | | | | | -0.101* | -0.102* | -0.0992* | -0.125** |
| | | | | | | | | | (0.0542) | (0.0539) | (0.0539) | (0.0614) |
| t210 | | | | | | | | | -0.146** | -0.169** | -0.165** | -0.159** |
| | | | | | | | | | (0.0669) | (0.0725) | (0.0725) | (0.0797) |
| Constant | 0.871*** | -0.0589 | -0.0939 | -0.175 | 0.864*** | -0.0523 | -0.0879 | -0.171 | | | | |
| | (0.0916) | (0.151) | (0.153) | (0.187) | (0.0957) | (0.154) | (0.155) | (0.189) | | | | |
| Observations | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 |
| R-squared | 0.252 | 0.273 | 0.274 | 0.266 | 0.252 | 0.273 | 0.274 | 0.266 | 0.800 | 0.804 | 0.805 | 0.823 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 16: All Students: Decay centrality 2

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.424*** | -0.483*** | -0.472*** | -0.469*** | -0.580** | -0.834*** | -0.819*** | -0.767** | | | | |
| | (0.144) | (0.146) | (0.146) | (0.167) | (0.284) | (0.309) | (0.309) | (0.341) | | | | |
| elegivel | -1.003*** | -0.276 | -0.289 | -0.329 | -1.099*** | -0.473* | -0.483* | -0.493 | | | | |
| | (0.154) | (0.242) | (0.243) | (0.275) | (0.215) | (0.286) | (0.286) | (0.320) | | | | |
| t2_el | | | | | 0.196 | 0.423 | 0.417 | 0.367 | | | | |
| | | | | | (0.306) | (0.327) | (0.327) | (0.365) | | | | |
| t24 | | | | | | | | | -1.529*** | -1.669*** | -1.676*** | -1.961*** |
| | | | | | | | | | (0.372) | (0.374) | (0.374) | (0.441) |
| t25 | | | | | | | | | -0.251 | -0.241 | -0.225 | 0.0103 |
| | | | | | | | | | (0.215) | (0.215) | (0.215) | (0.251) |
| t27 | | | | | | | | | -0.164 | -0.184 | -0.173 | -0.320 |
| | | | | | | | | | (0.230) | (0.229) | (0.229) | (0.262) |
| t210 | | | | | | | | | -0.573** | -0.756** | -0.741** | -0.707** |
| | | | | | | | | | (0.283) | (0.308) | (0.308) | (0.339) |
| Constant | 2.198*** | -1.191* | -1.312** | -1.710** | 2.273*** | -1.044 | -1.166* | -1.593** | | | | |
| | (0.387) | (0.643) | (0.650) | (0.795) | (0.405) | (0.653) | (0.660) | (0.804) | | | | |
| Observations | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 |
| R-squared | 0.435 | 0.446 | 0.446 | 0.450 | 0.435 | 0.446 | 0.446 | 0.450 | 0.823 | 0.826 | 0.826 | 0.842 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17: All Students: Decay centrality 3

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -1.173** | -1.370*** | -1.333*** | -1.207** | -1.883** | -2.714*** | -2.658*** | -2.393** | | | | |
| | (0.466) | (0.474) | (0.475) | (0.535) | (0.921) | (1.005) | (1.005) | (1.092) | | | | |
| elegivel | -2.453*** | -0.0416 | -0.0852 | -0.0445 | -2.888*** | -0.792 | -0.825 | -0.698 | | | | |
| | (0.500) | (0.789) | (0.789) | (0.881) | (0.698) | (0.931) | (0.931) | (1.025) | | | | |
| t2_el | | | | | 0.888 | 1.614 | 1.591 | 1.457 | | | | |
| | | | | | (0.993) | (1.064) | (1.064) | (1.170) | | | | |
| t24 | | | | | | | | | -5.079*** | -5.503*** | -5.526*** | -6.327*** |
| | | | | | | | | | (1.206) | (1.215) | (1.215) | (1.413) |
| t25 | | | | | | | | | -0.244 | -0.235 | -0.180 | 0.697 |
| | | | | | | | | | (0.696) | (0.699) | (0.700) | (0.804) |
| t27 | | | | | | | | | -0.503 | -0.594 | -0.555 | -0.879 |
| | | | | | | | | | (0.744) | (0.745) | (0.745) | (0.838) |
| t210 | | | | | | | | | -1.862** | -2.484** | -2.428** | -2.206** |
| | | | | | | | | | (0.917) | (1.002) | (1.002) | (1.087) |
| Constant | 2.413* | -8.295*** | -8.700*** | -10.61*** | 2.753** | -7.735*** | -8.143*** | -10.15*** | | | | |
| | (1.256) | (2.093) | (2.115) | (2.549) | (1.312) | (2.125) | (2.147) | (2.576) | | | | |
| Observations | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 | 3,300 | 3,199 | 3,199 | 2,370 |
| R-squared | 0.657 | 0.661 | 0.661 | 0.675 | 0.657 | 0.661 | 0.661 | 0.676 | 0.883 | 0.884 | 0.884 | 0.897 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18: All Students : Overall Clustering

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.0891*** | -0.0865*** | -0.0855*** | -0.0972*** | -0.0363 | -0.0218 | -0.0201 | -0.0283 | | | | |
| | (0.0144) | (0.0147) | (0.0147) | (0.0165) | (0.0269) | (0.0290) | (0.0291) | (0.0316) | | | | |
| elegivel | -0.0105 | 0.000973 | -0.000176 | 0.00951 | 0.0208 | 0.0361 | 0.0352 | 0.0469 | | | | |
| | (0.0146) | (0.0235) | (0.0235) | (0.0263) | (0.0198) | (0.0271) | (0.0271) | (0.0300) | | | | |
| t2_el | | | | | -0.0678** | -0.0802*** | -0.0810*** | -0.0871** | | | | |
| | | | | | (0.0291) | (0.0310) | (0.0310) | (0.0340) | | | | |
| t24 | | | | | | | | | -0.0985** | -0.0944** | -0.0957** | -0.144*** |
| | | | | | | | | | (0.0398) | (0.0401) | (0.0401) | (0.0465) |
| t25 | | | | | | | | | -0.118*** | -0.116*** | -0.114*** | -0.122*** |
| | | | | | | | | | (0.0217) | (0.0217) | (0.0217) | (0.0247) |
| t27 | | | | | | | | | -0.0943*** | -0.0927*** | -0.0919*** | -0.102*** |
| | | | | | | | | | (0.0224) | (0.0225) | (0.0225) | (0.0252) |
| t210 | | | | | | | | | -0.0356 | -0.0209 | -0.0194 | -0.0266 |
| | | | | | | | | | (0.0269) | (0.0290) | (0.0290) | (0.0315) |
| Constant | 0.340*** | 0.301*** | 0.284*** | 0.204** | 0.315*** | 0.276*** | 0.258*** | 0.179** | | | | |
| | (0.0434) | (0.0670) | (0.0678) | (0.0812) | (0.0446) | (0.0677) | (0.0684) | (0.0817) | | | | |
| Observations | 2,661 | 2,585 | 2,585 | 1,994 | 2,661 | 2,585 | 2,585 | 1,994 | 2,661 | 2,585 | 2,585 | 1,994 |
| R-squared | 0.058 | 0.059 | 0.060 | 0.076 | 0.060 | 0.061 | 0.062 | 0.079 | 0.612 | 0.612 | 0.612 | 0.627 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19: All Students : Positive assortiviness in terms of mathematics

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.0147 | -0.0143 | -0.0154 | -0.0161 | -0.00630 | -0.00337 | -0.00534 | -0.0164 | | | | |
| | (0.0141) | (0.0145) | (0.0145) | (0.0151) | (0.0254) | (0.0276) | (0.0277) | (0.0281) | | | | |
| elegivel | 0.0457*** | 0.0560** | 0.0579*** | 0.0406* | 0.0507*** | 0.0618** | 0.0632** | 0.0405 | | | | |
| | (0.0136) | (0.0221) | (0.0221) | (0.0229) | (0.0184) | (0.0254) | (0.0254) | (0.0260) | | | | |
| t2_el | | | | | -0.0109 | -0.0135 | -0.0124 | 0.000393 | | | | |
| | | | | | (0.0271) | (0.0291) | (0.0291) | (0.0297) | | | | |
| t24 | | | | | | | | | 0.00943 | 0.00229 | 0.00387 | -0.00223 |
| | | | | | | | | | (0.0394) | (0.0399) | (0.0399) | (0.0435) |
| t25 | | | | | | | | | -0.0268 | -0.0248 | -0.0278 | -0.0247 |
| | | | | | | | | | (0.0208) | (0.0209) | (0.0209) | (0.0222) |
| t27 | | | | | | | | | -0.00978 | -0.00958 | -0.00980 | -0.00904 |
| | | | | | | | | | (0.0212) | (0.0213) | (0.0213) | (0.0220) |
| t210 | | | | | | | | | -0.00701 | -0.00431 | -0.00614 | -0.0190 |
| | | | | | | | | | (0.0254) | (0.0276) | (0.0276) | (0.0281) |
| Constant | 0.730*** | 0.666*** | 0.689*** | 0.747*** | 0.726*** | 0.661*** | 0.684*** | 0.747*** | | | | |
| | (0.0459) | (0.0684) | (0.0692) | (0.0738) | (0.0470) | (0.0691) | (0.0699) | (0.0744) | | | | |
| Observations | 2,305 | 2,232 | 2,232 | 1,740 | 2,305 | 2,232 | 2,232 | 1,740 | 2,305 | 2,232 | 2,232 | 1,740 |
| R-squared | 0.105 | 0.111 | 0.113 | 0.143 | 0.105 | 0.111 | 0.113 | 0.143 | 0.918 | 0.919 | 0.919 | 0.931 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 20: All Students : Average friends' outcome in mathematics

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | | | | |
| t24 | -0.0418 | -0.108 | -0.114 | -0.374* |
| | (0.177) | (0.176) | (0.176) | (0.202) |
| t25 | -0.103 | -0.0697 | -0.0658 | 0.00736 |
| | (0.0977) | (0.0966) | (0.0967) | (0.109) |
| t27 | -0.241** | -0.233** | -0.232** | -0.296*** |
| | (0.100) | (0.0987) | (0.0988) | (0.109) |
| t210 | -0.174 | -0.135 | -0.131 | -0.202 |
| | (0.118) | (0.126) | (0.126) | (0.136) |
| mat4 | | | | |
| mat5 | 1.244 | -0.0673 | -0.0710 | -0.355** |
| | (0.813) | (0.148) | (0.148) | (0.170) |
| mat7 | 1.655** | 0.207 | 0.200 | -0.0252 |
| | (0.816) | (0.157) | (0.158) | (0.181) |
| mat10 | 1.970** | -0.134 | -0.169 | 0.0733 |
| | (0.818) | (0.867) | (0.869) | (0.973) |
| Observations | 2,369 | 2,293 | 2,293 | 1,787 |
| R-squared | 0.963 | 0.965 | 0.965 | 0.967 |
| Controls 1 | n | y | y | y |
| Controls 2 | n | n | y | y |
| Controls 3 | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## 3.2 Students that enrolled for the extra classes

Table 21: Students that enroll : Probit - probability of being an isolated node

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | 0.875*** | 0.872*** | 0.862*** | 0.972** | 0.929* | 0.568 | 0.386 | 0.300 | | | | |
| | (0.300) | (0.308) | (0.311) | (0.404) | (0.517) | (0.550) | (0.557) | (0.660) | | | | |
| elegivel | -0.0109 | 0.0221 | 0.0663 | 0.637 | 0.0390 | -0.267 | -0.390 | -0.181 | | | | |
| | (0.263) | (0.456) | (0.463) | (0.634) | (0.468) | (0.626) | (0.638) | (0.904) | | | | |
| t2_el | | | | | -0.0730 | 0.402 | 0.634 | 1.083 | | | | |
| | | | | | (0.564) | (0.609) | (0.627) | (0.856) | | | | |
| t24 | | | | | | | | | 0.705 | 0.953 | 0.847 | 5.386 |
| | | | | | | | | | (0.856) | (0.896) | (0.905) | (419.6) |
| t25 | | | | | | | | | 1.066*** | 1.170*** | 1.265*** | 1.459** |
| | | | | | | | | | (0.411) | (0.423) | (0.440) | (0.670) |
| t27 | | | | | | | | | 0.658 | 0.764 | 0.855 | 5.308 |
| | | | | | | | | | (0.492) | (0.507) | (0.524) | (298.2) |
| t210 | | | | | | | | | 0.916* | 0.575 | 0.390 | 0.204 |
| | | | | | | | | | (0.514) | (0.551) | (0.559) | (0.662) |
| Constant | -2.563*** | -2.746** | -2.757** | -2.720 | -2.598*** | -2.470* | -2.360* | -2.178 | -2.556*** | -3.031** | -2.954** | -1.677 |
| | (0.525) | (1.272) | (1.311) | (2.028) | (0.592) | (1.333) | (1.363) | (2.090) | (0.583) | (1.342) | (1.395) | (2.200) |
| Observations | 298 | 288 | 288 | 180 | 298 | 288 | 288 | 180 | 298 | 288 | 288 | 180 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 22: Students that enroll : Probit - probability of having positive outdegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.219 | -0.0973 | -0.0935 | 0.00729 | -0.902* | -0.300 | -0.198 | -0.0878 | | | | |
| | (0.262) | (0.274) | (0.276) | (0.351) | (0.523) | (0.586) | (0.591) | (0.716) | | | | |
| elegivel | -0.192 | -0.343 | -0.377 | -0.987 | -0.780 | -0.500 | -0.458 | -1.065 | | | | |
| | (0.261) | (0.451) | (0.452) | (0.623) | (0.483) | (0.609) | (0.609) | (0.808) | | | | |
| t2_el | | | | | 0.892 | 0.251 | 0.130 | 0.126 | | | | |
| | | | | | (0.576) | (0.640) | (0.647) | (0.829) | | | | |
| t25 | | | | | | | | | -0.0506 | -0.112 | -0.147 | -0.103 |
| | | | | | | | | | (0.384) | (0.390) | (0.396) | (0.583) |
| t27 | | | | | | | | | -0.0601 | -0.0644 | -0.0955 | 0.107 |
| | | | | | | | | | (0.419) | (0.426) | (0.430) | (0.634) |
| t210 | | | | | | | | | -0.872* | -0.256 | -0.147 | -0.00397 |
| | | | | | | | | | (0.525) | (0.591) | (0.596) | (0.744) |
| Constant | 2.200*** | 1.967* | 1.881 | 2.821 | 2.639*** | 2.109* | 1.954 | 2.891 | 2.604*** | 2.421** | 2.147* | 2.370 |
| | (0.459) | (1.186) | (1.207) | (1.795) | (0.577) | (1.246) | (1.263) | (1.854) | (0.585) | (1.218) | (1.261) | (1.899) |
| Observations | 347 | 336 | 336 | 194 | 347 | 336 | 336 | 194 | 334 | 323 | 323 | 177 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 23: Students that enroll : Probit - probability of having positive indegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.387** | -0.364** | -0.366** | -0.238 | -0.859** | -0.638* | -0.696* | -0.662 | | | | |
| | (0.174) | (0.178) | (0.179) | (0.217) | (0.355) | (0.372) | (0.377) | (0.440) | | | | |
| elegivel | -0.0838 | 0.127 | 0.116 | -0.0853 | -0.458 | -0.0837 | -0.138 | -0.440 | | | | |
| | (0.187) | (0.311) | (0.313) | (0.378) | (0.314) | (0.403) | (0.408) | (0.500) | | | | |
| t2_el | | | | | 0.609 | 0.348 | 0.417 | 0.563 | | | | |
| | | | | | (0.395) | (0.413) | (0.418) | (0.504) | | | | |
| t24 | | | | | | | | | 0.542 | 0.471 | 0.460 | 0.685 |
| | | | | | | | | | (0.491) | (0.496) | (0.497) | (0.565) |
| t25 | | | | | | | | | -0.415 | -0.464* | -0.439 | -0.241 |
| | | | | | | | | | (0.270) | (0.273) | (0.275) | (0.347) |
| t27 | | | | | | | | | -0.367 | -0.371 | -0.366 | -0.306 |
| | | | | | | | | | (0.280) | (0.284) | (0.284) | (0.361) |
| t210 | | | | | | | | | -0.862** | -0.648* | -0.706* | -0.686 |
| | | | | | | | | | (0.355) | (0.372) | (0.377) | (0.436) |
| Constant | 0.822*** | -0.448 | -0.589 | -0.584 | 1.100*** | -0.263 | -0.364 | -0.329 | 1.138*** | -0.147 | -0.235 | -0.898 |
| | (0.305) | (0.830) | (0.850) | (1.105) | (0.362) | (0.862) | (0.883) | (1.136) | (0.365) | (0.824) | (0.862) | (1.133) |
| Observations | 447 | 435 | 435 | 326 | 447 | 435 | 435 | 326 | 447 | 435 | 435 | 326 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 24: Students that enrolled : Degree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.663*** | -0.622*** | -0.627*** | -0.567** | -0.737** | -0.566 | -0.596 | -0.615 | | | | |
| | (0.195) | (0.200) | (0.200) | (0.227) | (0.364) | (0.394) | (0.396) | (0.435) | | | | |
| elegivel | -0.304 | -0.0166 | -0.0127 | -0.293 | -0.358 | 0.0207 | 0.00828 | -0.326 | | | | |
| | (0.198) | (0.328) | (0.328) | (0.366) | (0.298) | (0.399) | (0.400) | (0.448) | | | | |
| t2_el | | | | | 0.0977 | -0.0711 | -0.0401 | 0.0625 | | | | |
| | | | | | (0.404) | (0.432) | (0.434) | (0.487) | | | | |
| t2_mat | | | | | | | | | | | | |
| t24 | | | | | | | | | -0.648 | -0.681 | -0.693 | -0.707 |
| | | | | | | | | | (0.597) | (0.602) | (0.603) | (0.642) |
| t25 | | | | | | | | | -0.847*** | -0.864*** | -0.861*** | -0.732** |
| | | | | | | | | | (0.295) | (0.297) | (0.298) | (0.345) |
| t27 | | | | | | | | | -0.536* | -0.510* | -0.508* | -0.341 |
| | | | | | | | | | (0.300) | (0.302) | (0.302) | (0.356) |
| t210 | | | | | | | | | -0.734** | -0.543 | -0.569 | -0.586 |
| | | | | | | | | | (0.362) | (0.392) | (0.395) | (0.435) |
| Constant | 2.609*** | 1.123 | 1.211 | 0.918 | 2.646*** | 1.095 | 1.195 | 0.935 | | | | |
| | (0.363) | (0.920) | (0.943) | (1.118) | (0.395) | (0.937) | (0.960) | (1.128) | | | | |
| Observations | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 |
| R-squared | 0.317 | 0.322 | 0.323 | 0.355 | 0.317 | 0.322 | 0.323 | 0.355 | 0.847 | 0.848 | 0.849 | 0.871 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 25: Students that enrolled : Indegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.387** | -0.339* | -0.346* | -0.338 | -0.752** | -0.659* | -0.724* | -0.759* | | | | |
| | (0.194) | (0.199) | (0.198) | (0.231) | (0.362) | (0.392) | (0.393) | (0.443) | | | | |
| elegivel | -0.286 | 0.193 | 0.196 | 0.184 | -0.551* | -0.0214 | -0.0562 | -0.109 | | | | |
| | (0.197) | (0.326) | (0.326) | (0.372) | (0.297) | (0.396) | (0.396) | (0.455) | | | | |
| t2_el | | | | | 0.481 | 0.408 | 0.480 | 0.553 | | | | |
| | | | | | (0.402) | (0.429) | (0.430) | (0.495) | | | | |
| t2_mat | | | | | | | | | | | | |
| t24 | | | | | | | | | -0.317 | -0.330 | -0.374 | -0.243 |
| | | | | | | | | | (0.597) | (0.603) | (0.603) | (0.658) |
| t25 | | | | | | | | | -0.316 | -0.309 | -0.282 | -0.177 |
| | | | | | | | | | (0.295) | (0.298) | (0.298) | (0.353) |
| t27 | | | | | | | | | -0.271 | -0.243 | -0.241 | -0.267 |
| | | | | | | | | | (0.299) | (0.303) | (0.302) | (0.364) |
| t210 | | | | | | | | | -0.746** | -0.640 | -0.705* | -0.761* |
| | | | | | | | | | (0.362) | (0.393) | (0.394) | (0.446) |
| Constant | 1.522*** | -0.617 | -0.608 | -0.693 | 1.705*** | -0.457 | -0.416 | -0.537 | | | | |
| | (0.362) | (0.916) | (0.936) | (1.139) | (0.393) | (0.932) | (0.951) | (1.147) | | | | |
| Observations | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 |
| R-squared | 0.231 | 0.247 | 0.253 | 0.298 | 0.234 | 0.249 | 0.255 | 0.301 | 0.693 | 0.695 | 0.698 | 0.728 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 26: Students that enrolled : Decay centrality 2

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.800** | -0.818** | -0.811** | -0.783** | -1.506** | -1.764*** | -1.740** | -1.809** | | | | |
| | (0.336) | (0.342) | (0.343) | (0.386) | (0.626) | (0.673) | (0.678) | (0.738) | | | | |
| elegivel | -0.416 | -0.0787 | -0.0847 | -0.878 | -0.928* | -0.709 | -0.703 | -1.591** | | | | |
| | (0.341) | (0.562) | (0.563) | (0.622) | (0.513) | (0.681) | (0.683) | (0.759) | | | | |
| t2_el | | | | | 0.929 | 1.202 | 1.180 | 1.346 | | | | |
| | | | | | (0.695) | (0.738) | (0.742) | (0.826) | | | | |
| t24 | | | | | | | | | -0.549 | -0.490 | -0.481 | -0.814 |
| | | | | | | | | | (1.031) | (1.031) | (1.034) | (1.091) |
| t25 | | | | | | | | | -1.022** | -1.014** | -1.012** | -0.396 |
| | | | | | | | | | (0.509) | (0.509) | (0.511) | (0.586) |
| t27 | | | | | | | | | -0.265 | -0.275 | -0.277 | -0.434 |
| | | | | | | | | | (0.517) | (0.517) | (0.518) | (0.604) |
| t210 | | | | | | | | | -1.486** | -1.715** | -1.687** | -1.820** |
| | | | | | | | | | (0.626) | (0.673) | (0.677) | (0.739) |
| Constant | 1.997*** | 0.709 | 0.553 | 0.235 | 2.352*** | 1.180 | 1.025 | 0.613 | | | | |
| | (0.627) | (1.578) | (1.617) | (1.902) | (0.680) | (1.601) | (1.641) | (1.911) | | | | |
| Observations | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 |
| R-squared | 0.502 | 0.513 | 0.513 | 0.540 | 0.504 | 0.516 | 0.516 | 0.544 | 0.850 | 0.854 | 0.854 | 0.872 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 27: Students that enrolled : Decay centrality 3

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -2.910*** | -2.962*** | -2.918*** | -3.019** | -5.664*** | -6.510*** | -6.325*** | -6.306*** | | | | |
| | (1.091) | (1.113) | (1.114) | (1.258) | (2.031) | (2.188) | (2.200) | (2.406) | | | | |
| elegivel | -0.347 | 0.601 | 0.571 | -1.206 | -2.344 | -1.765 | -1.696 | -3.489 | | | | |
| | (1.109) | (1.827) | (1.829) | (2.029) | (1.665) | (2.214) | (2.219) | (2.476) | | | | |
| t2_el | | | | | 3.622 | 4.510* | 4.325* | 4.312 | | | | |
| | | | | | (2.255) | (2.398) | (2.411) | (2.693) | | | | |
| t24 | | | | | | | | | -1.930 | -1.821 | -1.736 | -2.482 |
| | | | | | | | | | (3.351) | (3.356) | (3.362) | (3.564) |
| t25 | | | | | | | | | -3.237* | -3.206* | -3.221* | -1.621 |
| | | | | | | | | | (1.654) | (1.656) | (1.662) | (1.914) |
| t27 | | | | | | | | | -1.251 | -1.259 | -1.271 | -2.281 |
| | | | | | | | | | (1.681) | (1.683) | (1.684) | (1.974) |
| t210 | | | | | | | | | -5.618*** | -6.394*** | -6.197*** | -6.391*** |
| | | | | | | | | | (2.034) | (2.188) | (2.199) | (2.415) |
| Constant | 2.524 | -1.501 | -2.228 | -3.086 | 3.908* | 0.269 | -0.495 | -1.875 | | | | |
| | (2.037) | (5.133) | (5.254) | (6.203) | (2.208) | (5.204) | (5.329) | (6.234) | | | | |
| Observations | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 | 478 | 466 | 466 | 353 |
| R-squared | 0.691 | 0.697 | 0.698 | 0.718 | 0.693 | 0.700 | 0.701 | 0.720 | 0.889 | 0.892 | 0.892 | 0.903 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 28: Students that enroled : Overall Clustering

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.0174 | -0.00700 | -0.00739 | -0.0480 | -0.0302 | -0.00818 | -0.00591 | -0.0983 | | | | |
| | (0.0393) | (0.0406) | (0.0406) | (0.0464) | (0.0705) | (0.0768) | (0.0769) | (0.0873) | | | | |
| elegivel | -0.0313 | -0.0960 | -0.0970 | -0.0561 | -0.0398 | -0.0967 | -0.0961 | -0.0885 | | | | |
| | (0.0375) | (0.0605) | (0.0605) | (0.0692) | (0.0542) | (0.0726) | (0.0727) | (0.0841) | | | | |
| t2_el | | | | | 0.0167 | 0.00148 | -0.00185 | 0.0646 | | | | |
| | | | | | (0.0763) | (0.0815) | (0.0818) | (0.0950) | | | | |
| t24 | | | | | | | | | 0.0392 | 0.0445 | 0.0390 | 0.0110 |
| | | | | | | | | | (0.111) | (0.112) | (0.113) | (0.123) |
| t25 | | | | | | | | | 0.0496 | 0.0561 | 0.0585 | -0.0232 |
| | | | | | | | | | (0.0589) | (0.0596) | (0.0597) | (0.0692) |
| t27 | | | | | | | | | -0.0766 | -0.0685 | -0.0721 | -0.0464 |
| | | | | | | | | | (0.0591) | (0.0599) | (0.0601) | (0.0720) |
| t210 | | | | | | | | | -0.0354 | -0.0199 | -0.0188 | -0.0996 |
| | | | | | | | | | (0.0707) | (0.0772) | (0.0773) | (0.0881) |
| Constant | 0.151* | 0.277 | 0.225 | 0.246 | 0.158* | 0.277 | 0.224 | 0.264 | | | | |
| | (0.0806) | (0.177) | (0.182) | (0.220) | (0.0861) | (0.181) | (0.186) | (0.221) | | | | |
| Observations | 397 | 387 | 387 | 302 | 397 | 387 | 387 | 302 | 397 | 387 | 387 | 302 |
| R-squared | 0.153 | 0.155 | 0.159 | 0.236 | 0.153 | 0.155 | 0.159 | 0.237 | 0.665 | 0.662 | 0.664 | 0.705 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 29: Students that enrolled : Individual Clustering

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | 0.00968 | 0.0188 | 0.0191 | 0.00418 | -0.000172 | 0.0175 | 0.0146 | -0.0343 | | | | |
| | (0.0424) | (0.0433) | (0.0434) | (0.0495) | (0.0747) | (0.0799) | (0.0802) | (0.0900) | | | | |
| elegivel | -0.0336 | -0.0339 | -0.0342 | 0.0387 | -0.0401 | -0.0347 | -0.0370 | 0.0140 | | | | |
| | (0.0387) | (0.0629) | (0.0630) | (0.0717) | (0.0558) | (0.0751) | (0.0754) | (0.0866) | | | | |
| t2_el | | | | | 0.0126 | 0.00158 | 0.00566 | 0.0495 | | | | |
| | | | | | (0.0787) | (0.0835) | (0.0838) | (0.0966) | | | | |
| t24 | | | | | | | | | 0.0827 | 0.0845 | 0.0819 | 0.0580 |
| | | | | | | | | | (0.121) | (0.122) | (0.123) | (0.135) |
| t25 | | | | | | | | | 0.0725 | 0.0772 | 0.0807 | 0.0206 |
| | | | | | | | | | (0.0621) | (0.0626) | (0.0629) | (0.0724) |
| t27 | | | | | | | | | -0.0588 | -0.0506 | -0.0497 | 0.00224 |
| | | | | | | | | | (0.0609) | (0.0615) | (0.0616) | (0.0734) |
| t210 | | | | | | | | | -0.00625 | 0.00602 | 0.00305 | -0.0324 |
| | | | | | | | | | (0.0750) | (0.0802) | (0.0804) | (0.0909) |
| Constant | 0.150* | -0.0103 | -0.00722 | 0.0106 | 0.155* | -0.00961 | -0.00467 | 0.0227 | | | | |
| | (0.0787) | (0.185) | (0.191) | (0.232) | (0.0850) | (0.189) | (0.195) | (0.234) | | | | |
| Observations | 382 | 374 | 374 | 292 | 382 | 374 | 374 | 292 | 382 | 374 | 374 | 292 |
| R-squared | 0.222 | 0.230 | 0.232 | 0.276 | 0.222 | 0.230 | 0.232 | 0.276 | 0.691 | 0.693 | 0.693 | 0.723 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 30: Students that enrolled : Positive assortiviness in terms of mathematics

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.0484 | -0.0341 | -0.0349 | -0.0563 | -0.116* | -0.0800 | -0.0800 | -0.158** | | | | |
| | (0.0352) | (0.0361) | (0.0362) | (0.0406) | (0.0599) | (0.0649) | (0.0652) | (0.0737) | | | | |
| elegivel | 0.0231 | -0.00678 | -0.00619 | 0.0200 | -0.0238 | -0.0345 | -0.0335 | -0.0434 | | | | |
| | (0.0313) | (0.0516) | (0.0517) | (0.0596) | (0.0461) | (0.0610) | (0.0612) | (0.0707) | | | | |
| t2_el | | | | | 0.0876 | 0.0573 | 0.0563 | 0.130* | | | | |
| | | | | | (0.0631) | (0.0673) | (0.0675) | (0.0786) | | | | |
| t24 | | | | | | | | | 0.0662 | 0.0753 | 0.0800 | 0.148 |
| | | | | | | | | | (0.0961) | (0.0966) | (0.0970) | (0.107) |
| t25 | | | | | | | | | -0.0338 | -0.0275 | -0.0318 | -0.0572 |
| | | | | | | | | | (0.0502) | (0.0505) | (0.0509) | (0.0581) |
| t27 | | | | | | | | | -0.0683 | -0.0580 | -0.0581 | -0.0712 |
| | | | | | | | | | (0.0512) | (0.0516) | (0.0517) | (0.0588) |
| t210 | | | | | | | | | -0.125** | -0.0918 | -0.0917 | -0.169** |
| | | | | | | | | | (0.0602) | (0.0651) | (0.0653) | (0.0738) |
| Constant | 0.752*** | 0.862*** | 0.872*** | 0.823*** | 0.787*** | 0.878*** | 0.889*** | 0.839*** | | | | |
| | (0.0616) | (0.145) | (0.150) | (0.185) | (0.0664) | (0.146) | (0.152) | (0.185) | | | | |
| Observations | 355 | 346 | 346 | 271 | 355 | 346 | 346 | 271 | 355 | 346 | 346 | 271 |
| R-squared | 0.352 | 0.355 | 0.356 | 0.356 | 0.355 | 0.356 | 0.357 | 0.364 | 0.944 | 0.945 | 0.945 | 0.949 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 31: Students that enrolled : Average friends' outcome in mathematics

| VARIABLES | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| t24 | -0.799* | -0.702 | -0.770 | -1.049** |
| | (0.482) | (0.477) | (0.474) | (0.521) |
| t25 | -0.162 | -0.137 | -0.0581 | 0.208 |
| | (0.258) | (0.255) | (0.255) | (0.304) |
| t27 | -0.292 | -0.314 | -0.311 | -0.0457 |
| | (0.261) | (0.259) | (0.258) | (0.301) |
| t210 | -0.206 | -0.440 | -0.432 | -0.438 |
| | (0.299) | (0.317) | (0.315) | (0.353) |
| Observations | 363 | 354 | 354 | 277 |
| R-squared | 0.972 | 0.973 | 0.974 | 0.977 |
| Controls 1 | n | y | y | y |
| Controls 2 | n | n | y | y |
| Controls 3 | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## 3.3 Students that did not enrol for the extra classes

Table 32: Students that did not enrol : Probit - probability of being an isolated node

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | 0.493*** | 0.497*** | 0.482*** | 0.458*** | 0.444** | 0.424* | 0.411* | 0.356 | | | | |
| | (0.0847) | (0.0883) | (0.0887) | (0.109) | (0.190) | (0.221) | (0.222) | (0.265) | | | | |
| elegivel | 0.328*** | -0.0481 | -0.0267 | -0.183 | 0.294* | -0.0962 | -0.0740 | -0.248 | | | | |
| | (0.0996) | (0.153) | (0.154) | (0.191) | (0.155) | (0.203) | (0.204) | (0.244) | | | | |
| t2_el | | | | | 0.0569 | 0.0813 | 0.0797 | 0.117 | | | | |
| | | | | | (0.200) | (0.228) | (0.229) | (0.277) | | | | |
| t24 | | | | | | | | | 0.796*** | 0.820*** | 0.823*** | 1.033*** |
| | | | | | | | | | (0.178) | (0.180) | (0.181) | (0.227) |
| t25 | | | | | | | | | 0.410*** | 0.390*** | 0.369*** | 0.165 |
| | | | | | | | | | (0.125) | (0.127) | (0.128) | (0.165) |
| t27 | | | | | | | | | 0.479*** | 0.454*** | 0.436*** | 0.498** |
| | | | | | | | | | (0.153) | (0.155) | (0.155) | (0.198) |
| t210 | | | | | | | | | 0.446** | 0.379* | 0.369* | 0.325 |
| | | | | | | | | | (0.191) | (0.221) | (0.222) | (0.265) |
| Constant | -0.651*** | 0.701** | 0.779** | 1.515*** | -0.624*** | 0.741** | 0.819** | 1.567*** | -0.643*** | 0.179 | 0.304 | 0.955* |
| | (0.210) | (0.350) | (0.353) | (0.473) | (0.229) | (0.367) | (0.370) | (0.488) | (0.230) | (0.388) | (0.394) | (0.528) |
| Observations | 2,723 | 2,642 | 2,642 | 1,937 | 2,723 | 2,642 | 2,642 | 1,937 | 2,723 | 2,642 | 2,642 | 1,937 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 33: Students that did not enrol : Probit - probability of having positive outdegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.275*** | -0.255*** | -0.226*** | -0.201** | -0.151 | -0.00378 | 0.0351 | -0.0434 | | | | |
| | (0.0790) | (0.0819) | (0.0824) | (0.101) | (0.171) | (0.192) | (0.194) | (0.244) | | | | |
| elegivel | -0.347*** | -0.0308 | -0.0707 | -0.131 | -0.266** | 0.119 | 0.0839 | -0.0361 | | | | |
| | (0.0908) | (0.138) | (0.140) | (0.176) | (0.133) | (0.170) | (0.172) | (0.219) | | | | |
| t2_el | | | | | -0.147 | -0.287 | -0.298 | -0.180 | | | | |
| | | | | | (0.180) | (0.199) | (0.200) | (0.254) | | | | |
| t24 | | | | | | | | | -0.312* | -0.351** | -0.343** | -0.419** |
| | | | | | | | | | (0.169) | (0.171) | (0.171) | (0.213) |
| t25 | | | | | | | | | -0.344*** | -0.322*** | -0.283** | -0.153 |
| | | | | | | | | | (0.116) | (0.118) | (0.119) | (0.150) |
| t27 | | | | | | | | | -0.268** | -0.220 | -0.195 | -0.226 |
| | | | | | | | | | (0.136) | (0.137) | (0.137) | (0.170) |
| t210 | | | | | | | | | -0.154 | 0.0303 | 0.0642 | -0.0290 |
| | | | | | | | | | (0.171) | (0.192) | (0.194) | (0.244) |
| Constant | 0.664*** | -0.557* | -0.700** | -0.843* | 0.602*** | -0.676** | -0.826** | -0.920** | 0.620*** | -0.478 | -0.716** | -0.940* |
| | (0.202) | (0.327) | (0.331) | (0.443) | (0.215) | (0.337) | (0.340) | (0.455) | (0.215) | (0.353) | (0.360) | (0.487) |
| Observations | 2,703 | 2,615 | 2,615 | 1,934 | 2,703 | 2,615 | 2,615 | 1,934 | 2,703 | 2,615 | 2,615 | 1,934 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 34: Students that did not enrol : Probit - probability of having positive indegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.369*** | -0.383*** | -0.379*** | -0.344*** | -0.216 | -0.245 | -0.239 | -0.132 | | | | |
| | (0.0672) | (0.0695) | (0.0697) | (0.0824) | (0.143) | (0.166) | (0.166) | (0.193) | | | | |
| elegivel | -0.293*** | -0.0506 | -0.0560 | 0.0213 | -0.195* | 0.0330 | 0.0286 | 0.145 | | | | |
| | (0.0763) | (0.119) | (0.119) | (0.142) | (0.110) | (0.149) | (0.149) | (0.173) | | | | |
| t2_el | | | | | -0.183 | -0.157 | -0.159 | -0.245 | | | | |
| | | | | | (0.151) | (0.172) | (0.172) | (0.201) | | | | |
| t24 | | | | | | | | | -0.599*** | -0.626*** | -0.628*** | -0.706*** |
| | | | | | | | | | (0.153) | (0.155) | (0.155) | (0.190) |
| t25 | | | | | | | | | -0.406*** | -0.395*** | -0.389*** | -0.289** |
| | | | | | | | | | (0.0980) | (0.0992) | (0.0995) | (0.121) |
| t27 | | | | | | | | | -0.310*** | -0.308*** | -0.304*** | -0.329** |
| | | | | | | | | | (0.113) | (0.113) | (0.113) | (0.134) |
| t210 | | | | | | | | | -0.215 | -0.212 | -0.207 | -0.108 |
| | | | | | | | | | (0.143) | (0.166) | (0.166) | (0.193) |
| Constant | 0.412** | -0.634** | -0.665** | -1.104*** | 0.336* | -0.698** | -0.731** | -1.195*** | 0.352* | -0.218 | -0.267 | -0.645 |
| | (0.191) | (0.296) | (0.298) | (0.386) | (0.200) | (0.304) | (0.306) | (0.393) | (0.200) | (0.314) | (0.319) | (0.412) |
| Observations | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 35: Students that did not enrol : Degree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| Treated | -0.458*** | -0.483*** | -0.471*** | -0.416*** | -0.166 | -0.231 | -0.215 | -0.172 | | | | |
| | (0.0886) | (0.0891) | (0.0893) | (0.102) | (0.174) | (0.187) | (0.187) | (0.207) | | | | |
| elegivel | -0.643*** | -0.120 | -0.133 | -0.0982 | -0.468*** | 0.0176 | 0.00747 | 0.0330 | | | | |
| | (0.0935) | (0.145) | (0.145) | (0.163) | (0.129) | (0.171) | (0.170) | (0.190) | | | | |
| t2_el | | | | | -0.361* | -0.299 | -0.304 | -0.296 | | | | |
| | | | | | (0.185) | (0.196) | (0.196) | (0.218) | | | | |
| t24 | | | | | | | | | -1.027*** | -1.147*** | -1.155*** | -1.327*** |
| | | | | | | | | | (0.218) | (0.217) | (0.217) | (0.258) |
| t25 | | | | | | | | | -0.484*** | -0.461*** | -0.443*** | -0.268* |
| | | | | | | | | | (0.130) | (0.130) | (0.130) | (0.152) |
| t27 | | | | | | | | | -0.398*** | -0.390*** | -0.376*** | -0.382** |
| | | | | | | | | | (0.140) | (0.139) | (0.139) | (0.156) |
| t210 | | | | | | | | | -0.170 | -0.183 | -0.168 | -0.131 |
| | | | | | | | | | (0.172) | (0.187) | (0.187) | (0.206) |
| Constant | 2.648*** | 0.223 | 0.123 | 0.0448 | 2.514*** | 0.125 | 0.0237 | -0.0399 | | | | |
| | (0.273) | (0.398) | (0.400) | (0.499) | (0.281) | (0.403) | (0.405) | (0.503) | | | | |
| | | | | | | | | | | | | |
| Observations | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 |
| R-squared | 0.153 | 0.186 | 0.188 | 0.173 | 0.154 | 0.187 | 0.189 | 0.173 | 0.786 | 0.793 | 0.793 | 0.814 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 36: Students that did not enrol : Indegree

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.388*** | -0.426*** | -0.425*** | -0.458*** | -0.108 | -0.220 | -0.219 | -0.155 | | | | |
| | (0.0810) | (0.0819) | (0.0820) | (0.0959) | (0.159) | (0.172) | (0.172) | (0.195) | | | | |
| elegivel | -0.501*** | -0.118 | -0.116 | -0.0769 | -0.332*** | -0.00473 | -0.00321 | 0.0859 | | | | |
| | (0.0855) | (0.133) | (0.133) | (0.154) | (0.118) | (0.157) | (0.157) | (0.179) | | | | |
| t2_el | | | | | -0.347** | -0.246 | -0.245 | -0.367* | | | | |
| | | | | | (0.169) | (0.180) | (0.180) | (0.206) | | | | |
| t2_mat | | | | | | | | | | | | |
| t24 | | | | | | | | | -0.720*** | -0.826*** | -0.834*** | -0.989*** |
| | | | | | | | | | (0.199) | (0.200) | (0.200) | (0.244) |
| t25 | | | | | | | | | -0.418*** | -0.402*** | -0.402*** | -0.397*** |
| | | | | | | | | | (0.119) | (0.119) | (0.120) | (0.143) |
| t27 | | | | | | | | | -0.412*** | -0.418*** | -0.413*** | -0.488*** |
| | | | | | | | | | (0.128) | (0.127) | (0.128) | (0.148) |
| t210 | | | | | | | | | -0.113 | -0.182 | -0.181 | -0.121 |
| | | | | | | | | | (0.158) | (0.172) | (0.172) | (0.195) |
| Constant | 1.662*** | -0.307 | -0.318 | -0.775 | 1.533*** | -0.387 | -0.398 | -0.880* | | | | |
| | (0.249) | (0.366) | (0.368) | (0.471) | (0.257) | (0.370) | (0.372) | (0.475) | | | | |
| Observations | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 |
| R-squared | 0.126 | 0.156 | 0.157 | 0.157 | 0.127 | 0.157 | 0.158 | 0.158 | 0.629 | 0.640 | 0.641 | 0.661 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 37: Students that did not enrol : Decay centrality 1

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | r1 | r2 | r3 | r5 | r6 | r7 | r8 | r10 | r16 | r17 | r18 | r20 |
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.140*** | -0.152*** | -0.148*** | -0.141*** | -0.0915 | -0.138* | -0.133 | -0.115 | | | | |
| | (0.0385) | (0.0389) | (0.0389) | (0.0446) | (0.0753) | (0.0818) | (0.0818) | (0.0908) | | | | |
| elegivel | -0.279*** | -0.0769 | -0.0811 | -0.0670 | -0.250*** | -0.0695 | -0.0729 | -0.0531 | | | | |
| | (0.0406) | (0.0630) | (0.0631) | (0.0717) | (0.0562) | (0.0744) | (0.0744) | (0.0834) | | | | |
| t2_el | | | | | -0.0599 | -0.0162 | -0.0177 | -0.0314 | | | | |
| | | | | | (0.0802) | (0.0856) | (0.0856) | (0.0959) | | | | |
| t24 | | | | | | | | | -0.409*** | -0.459*** | -0.461*** | -0.540*** |
| | | | | | | | | | (0.0947) | (0.0948) | (0.0947) | (0.113) |
| t25 | | | | | | | | | -0.115** | -0.107* | -0.101* | -0.0403 |
| | | | | | | | | | (0.0567) | (0.0566) | (0.0567) | (0.0665) |
| t27 | | | | | | | | | -0.0970 | -0.0970 | -0.0924 | -0.118* |
| | | | | | | | | | (0.0608) | (0.0604) | (0.0605) | (0.0686) |
| t210 | | | | | | | | | -0.0919 | -0.118 | -0.113 | -0.0968 |
| | | | | | | | | | (0.0750) | (0.0816) | (0.0816) | (0.0905) |
| Constant | 0.916*** | -0.0165 | -0.0490 | -0.0942 | 0.894*** | -0.0218 | -0.0548 | -0.103 | | | | |
| | (0.118) | (0.174) | (0.175) | (0.219) | (0.122) | (0.176) | (0.177) | (0.221) | | | | |
| Observations | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 |
| R-squared | 0.245 | 0.268 | 0.269 | 0.257 | 0.245 | 0.268 | 0.269 | 0.257 | 0.795 | 0.801 | 0.801 | 0.819 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 38: Students that did not enrol : Decay centrality 2

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| Treated | -0.339** | -0.396** | -0.384** | -0.381** | -0.342 | -0.584* | -0.568 | -0.482 | | | | |
| | (0.163) | (0.166) | (0.166) | (0.190) | (0.319) | (0.348) | (0.349) | (0.388) | | | | |
| elegivel | -1.052*** | -0.296 | -0.309 | -0.258 | -1.053*** | -0.399 | -0.410 | -0.313 | | | | |
| | (0.172) | (0.268) | (0.269) | (0.306) | (0.238) | (0.317) | (0.317) | (0.356) | | | | |
| t2_el | | | | | 0.00314 | 0.224 | 0.219 | 0.123 | | | | |
| | | | | | (0.340) | (0.365) | (0.365) | (0.410) | | | | |
| t24 | | | | | | | | | -1.565*** | -1.743*** | -1.750*** | -2.046*** |
| | | | | | | | | | (0.401) | (0.403) | (0.404) | (0.483) |
| t25 | | | | | | | | | -0.101 | -0.0791 | -0.0602 | 0.141 |
| | | | | | | | | | (0.240) | (0.241) | (0.242) | (0.284) |
| t27 | | | | | | | | | -0.150 | -0.165 | -0.151 | -0.278 |
| | | | | | | | | | (0.258) | (0.257) | (0.258) | (0.293) |
| t210 | | | | | | | | | -0.340 | -0.503 | -0.487 | -0.408 |
| | | | | | | | | | (0.318) | (0.347) | (0.347) | (0.387) |
| Constant | 2.429*** | -1.008 | -1.106 | -1.415 | 2.430*** | -0.935 | -1.034 | -1.380 | | | | |
| | (0.501) | (0.739) | (0.744) | (0.936) | (0.517) | (0.749) | (0.753) | (0.944) | | | | |
| | | | | | | | | | | | | |
| Observations | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 |
| R-squared | 0.433 | 0.443 | 0.444 | 0.445 | 0.433 | 0.443 | 0.444 | 0.445 | 0.823 | 0.826 | 0.826 | 0.841 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 39: Students that did not enrol : Decay centrality 3

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.881* | -1.071** | -1.032* | -0.848 | -1.006 | -1.759 | -1.706 | -1.309 | | | | |
| | (0.528) | (0.539) | (0.540) | (0.610) | (1.035) | (1.134) | (1.135) | (1.243) | | | | |
| elegivel | -2.635*** | -0.107 | -0.150 | 0.152 | -2.710*** | -0.485 | -0.520 | -0.0955 | | | | |
| | (0.558) | (0.874) | (0.875) | (0.982) | (0.773) | (1.032) | (1.032) | (1.142) | | | | |
| t2_el | | | | | 0.155 | 0.818 | 0.802 | 0.558 | | | | |
| | | | | | (1.102) | (1.187) | (1.188) | (1.312) | | | | |
| t2_mat | | | | | | | | | | | | |
| t24 | | | | | | | | | -5.234*** | -5.775*** | -5.797*** | -6.593*** |
| | | | | | | | | | (1.301) | (1.313) | (1.314) | (1.548) |
| t25 | | | | | | | | | 0.209 | 0.256 | 0.318 | 1.112 |
| | | | | | | | | | (0.779) | (0.785) | (0.786) | (0.910) |
| t27 | | | | | | | | | -0.415 | -0.497 | -0.453 | -0.615 |
| | | | | | | | | | (0.836) | (0.838) | (0.838) | (0.939) |
| t210 | | | | | | | | | -0.998 | -1.514 | -1.462 | -1.067 |
| | | | | | | | | | (1.030) | (1.130) | (1.131) | (1.237) |
| Constant | 2.856* | -7.995*** | -8.313*** | -10.11*** | 2.914* | -7.729*** | -8.051*** | -9.954*** | | | | |
| | (1.626) | (2.407) | (2.423) | (3.000) | (1.678) | (2.438) | (2.454) | (3.024) | | | | |
| Observations | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 | 2,822 | 2,733 | 2,733 | 2,017 |
| R-squared | 0.654 | 0.656 | 0.657 | 0.670 | 0.654 | 0.657 | 0.657 | 0.670 | 0.884 | 0.885 | 0.885 | 0.898 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## Table 40: Students that did not enrol : Overall Clustering

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1c(16) | (17) | (18) | (20) | | | | | | | | | |
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.0906*** | -0.0873*** | -0.0858*** | -0.0949*** | -0.0221 | -0.00325 | -0.00122 | -0.00127 | | | | |
| | (0.0160) | (0.0163) | (0.0163) | (0.0182) | (0.0296) | (0.0320) | (0.0320) | (0.0349) | | | | |
| elegivel | -0.00546 | 0.0181 | 0.0169 | 0.0237 | 0.0347 | 0.0632** | 0.0623** | 0.0736** | | | | |
| | (0.0161) | (0.0256) | (0.0256) | (0.0286) | (0.0217) | (0.0295) | (0.0295) | (0.0327) | | | | |
| t2_el | | | | | -0.0873*** | -0.104*** | -0.104*** | -0.117*** | | | | |
| | | | | | (0.0318) | (0.0339) | (0.0339) | (0.0372) | | | | |
| t24 | | | | | | | | | -0.110** | -0.107** | -0.108** | -0.159*** |
| | | | | | | | | | (0.0430) | (0.0433) | (0.0433) | (0.0507) |
| t25 | | | | | | | | | -0.138*** | -0.135*** | -0.133*** | -0.134*** |
| | | | | | | | | | (0.0238) | (0.0238) | (0.0239) | (0.0271) |
| t27 | | | | | | | | | -0.0839*** | -0.0823*** | -0.0810*** | -0.0963*** |
| | | | | | | | | | (0.0246) | (0.0246) | (0.0246) | (0.0273) |
| t210 | | | | | | | | | -0.0206 | 0.000236 | 0.00194 | 0.00284 |
| Constant | 0.420*** | 0.368*** | 0.354*** | 0.271*** | 0.390*** | 0.338*** | 0.324*** | 0.243*** | | | | |
| | (0.0548) | (0.0766) | (0.0771) | (0.0942) | (0.0559) | (0.0771) | (0.0776) | (0.0944) | | | | |
| Observations | 2,264 | 2,198 | 2,198 | 1,692 | 2,264 | 2,198 | 2,198 | 1,692 | 2,264 | 2,198 | 2,198 | 1,692 |
| R-squared | 0.058 | 0.061 | 0.062 | 0.072 | 0.061 | 0.065 | 0.066 | 0.078 | 0.611 | 0.612 | 0.612 | 0.624 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 41: Students that did not enrol : Individual Clustering

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.101*** | -0.107*** | -0.106*** | -0.111*** | -0.0432 | -0.0564* | -0.0550 | -0.0441 | | | | |
| | (0.0173) | (0.0175) | (0.0176) | (0.0198) | (0.0311) | (0.0335) | (0.0335) | (0.0369) | | | | |
| elegivel | -0.0290* | 0.0271 | 0.0267 | 0.0151 | 0.00610 | 0.0549* | 0.0548* | 0.0515 | | | | |
| | (0.0169) | (0.0269) | (0.0269) | (0.0304) | (0.0229) | (0.0312) | (0.0312) | (0.0348) | | | | |
| t2_el | | | | | -0.0755** | -0.0628* | -0.0636* | -0.0842** | | | | |
| | | | | | (0.0334) | (0.0355) | (0.0356) | (0.0394) | | | | |
| t24 | | | | | | | | | -0.0983** | -0.107** | -0.109** | -0.208*** |
| | | | | | | | | | (0.0477) | (0.0480) | (0.0480) | (0.0567) |
| t25 | | | | | | | | | -0.130*** | -0.127*** | -0.126*** | -0.121*** |
| | | | | | | | | | (0.0255) | (0.0255) | (0.0255) | (0.0294) |
| t27 | | | | | | | | | -0.121*** | -0.124*** | -0.123*** | -0.124*** |
| | | | | | | | | | (0.0262) | (0.0261) | (0.0262) | (0.0294) |
| t210 | | | | | | | | | -0.0428 | -0.0525 | -0.0513 | -0.0398 |
| | | | | | | | | | (0.0310) | (0.0335) | (0.0335) | (0.0368) |
| Constant | 0.524*** | 0.284*** | 0.275*** | 0.282*** | 0.496*** | 0.265*** | 0.255*** | 0.259** | | | | |
| | (0.0590) | (0.0821) | (0.0827) | (0.101) | (0.0602) | (0.0828) | (0.0834) | (0.102) | | | | |
| Observations | 2,080 | 2,016 | 2,016 | 1,560 | 2,080 | 2,016 | 2,016 | 1,560 | 2,080 | 2,016 | 2,016 | 1,560 |
| R-squared | 0.075 | 0.091 | 0.091 | 0.096 | 0.078 | 0.092 | 0.093 | 0.099 | 0.640 | 0.645 | 0.645 | 0.653 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |
| Standard errors in parentheses | | | | | | | | | | | | |
| *** p<0.01, ** p<0.05, * p<0.1 | | | | | | | | | | | | |

Table 42: Students that did not enrol : Positive assortiviness in terms of mathematics

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | | | | | | | | | | | | |
| Treated | -0.00495 | -0.00758 | -0.00877 | -0.00589 | 0.0238 | 0.0200 | 0.0177 | 0.0156 | | | | |
| | (0.0152) | (0.0156) | (0.0156) | (0.0162) | (0.0277) | (0.0302) | (0.0302) | (0.0307) | | | | |
| elegivel | 0.0455*** | 0.0576** | 0.0598** | 0.0420* | 0.0622*** | 0.0721*** | 0.0736*** | 0.0528* | | | | |
| | (0.0149) | (0.0242) | (0.0242) | (0.0250) | (0.0201) | (0.0277) | (0.0277) | (0.0282) | | | | |
| t2_el | | | | | -0.0369 | -0.0340 | -0.0326 | -0.0268 | | | | |
| | | | | | (0.0297) | (0.0319) | (0.0319) | (0.0325) | | | | |
| t24 | | | | | | | | | -0.0126 | -0.0253 | -0.0238 | -0.0474 |
| | | | | | | | | | (0.0427) | (0.0432) | (0.0432) | (0.0476) |
| t25 | | | | | | | | | -0.0181 | -0.0154 | -0.0183 | -0.0103 |
| | | | | | | | | | (0.0227) | (0.0228) | (0.0228) | (0.0242) |
| t27 | | | | | | | | | -0.00259 | -0.00468 | -0.00508 | -0.00274 |
| | | | | | | | | | (0.0230) | (0.0230) | (0.0230) | (0.0237) |
| t210 | | | | | | | | | 0.0226 | 0.0186 | 0.0165 | 0.0128 |
| | | | | | | | | | (0.0277) | (0.0302) | (0.0302) | (0.0307) |
| Constant | 0.645** | 0.569** | 0.593** | 0.664*** | 0.636** | 0.562** | 0.586** | 0.659*** | | | | |
| | (0.255) | (0.262) | (0.262) | (0.244) | (0.255) | (0.262) | (0.262) | (0.244) | | | | |
| Observations | 1,950 | 1,886 | 1,886 | 1,469 | 1,950 | 1,886 | 1,886 | 1,469 | 1,950 | 1,886 | 1,886 | 1,469 |
| R-squared | 0.110 | 0.119 | 0.121 | 0.146 | 0.110 | 0.119 | 0.122 | 0.146 | 0.918 | 0.918 | 0.919 | 0.932 |
| Controls 1 | n | y | y | y | n | y | y | y | n | y | y | y |
| Controls 2 | n | n | y | y | n | n | y | y | n | n | y | y |
| Controls 3 | n | n | n | y | n | n | n | y | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 43: Students that did not enrol : Average friends' outcome in mathematics

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | r1 | r2 | r3 | r5 |
| VARIABLES | avg_mat_am | avg_mat_am | avg_mat_am | avg_mat_am |
| t24 | 0.107 | 0.00803 | 0.00165 | -0.271 |
| | (0.192) | (0.190) | (0.191) | (0.222) |
| t25 | -0.0903 | -0.0545 | -0.0526 | -0.0260 |
| | (0.107) | (0.105) | (0.106) | (0.120) |
| t27 | -0.244** | -0.239** | -0.237** | -0.352*** |
| | (0.109) | (0.107) | (0.107) | (0.119) |
| t210 | -0.123 | -0.0590 | -0.0561 | -0.153 |
| | (0.129) | (0.139) | (0.139) | (0.151) |
| Observations | 2,006 | 1,939 | 1,939 | 1,510 |
| R-squared | 0.962 | 0.964 | 0.964 | 0.966 |
| Controls 1 | n | y | y | y |
| Controls 2 | n | n | y | y |
| Controls 3 | n | n | n | y |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1