Social Capital Inequalities among Postgraduate Students and Social Selection Processes

Neylson J. B. F. Crepalde
⟨neylson.crepalde@izabelahendrix.edu.br⟩

2016, July

Abstract

This paper aims to discuss social capital inequalities between postgraduate students enrolled in a social sciences program in a Brazilian university. I argue that "social capital" concept has been widely and carelessly mobilized which may lead to confusion and spu rious conclusions. A solution would be to operationalize social capital with social network analysis. I analyze data from 47 postgraduate students using linear models, stochastic blockmodelling and the Social Selection Model (SSM). The analysis shows that social formations occur mainly from participation in research groups and from methodological perceived habilities. The variable that has the biggest effect on academic productivity is also participating in a research group.

Keywords: Social Capital; Social Network Analysis; Social Selection Model; Educational Inequalities.

1 Introduction

The idea of a fully egalitarian world is an impossible abstraction. As Lin (1999) shows, people are born in an already stratified social space where individuals in different positions have different access, and therefore, different mobilizing capacities of their social capital. This is due to individual positioning in social structures, family socioeconomic background, and other person related attributes. This is no news to sociologists. However, the process of social groups formation is not so acknowledged yet although it is central to understand how the access to social resources is achieved. This has only recently been investigated.

Here, I intend to show that the concept of social capital has been widely mobilized in a confusing and careless way, specially within studies that deal with educational inequalities. This, according to Portes (2000), can lead to spurious conclusions. I will present a methodological framework that operationalizes the concept very well obtaining many gains in objectivity and reliability, social network analysis.

Therefore, this paper aims to discuss the inequalities that are held within the academic system, specifically within a social sciences postgraduate program in a Brazilian university. To do this, we will focus our investigation in two main aspects: academic productivity and social ties formation. We used data collected online from 47 social sciences psotgraduate students (masters and doctoral). To perform the analysis, we used linear models, social network analysis, stochastic blockmodelling and the highly modern *social selection model* (SSM).

This research has the merit of presenting new data and new evidence on the subject and it also operationalizes the concept of social capital in a different and more straightforward way that we see in the literature, i.e., with formal relational data. Let us begin reviewing social capital concept and how this concept is mobilized within the educational inequalities field.

2 Social Capital and Inequalities

Lin (1999, p. 35) defines social capital as "resources embedded in a social structure which are accessed and/or mobilized in purposive actions". This standpoint leads to three other fidings: first, that there are resources embedded in a social structure. Second, individuals have different accessibility to these resources and, third, they can mobilize this resources in purposive actions. For the purposes of this paper, this definition is sufficient as it accounts for the inequalities that we aim to investigate.

Portes (1998) presents an excellent review on the origin and applicability of social capital in sociology. So does Higgins (2005). The later considers three conceptualizations of the term, namely one by Putnam, another by Fukuyama and another by Portes. Putnam investigates the democratic institutions and, for him, social capital is important because "it is regarded as the source from which spring cooperative interactions that express themselves in different forms of association of the civic community" (HIGGINS, 2005, p. 63). Social capital in this context is related to some specific characteristics of social organization that make the joint action possible, such as norms, trust and symbolic systems.

> Putnam's analysis on social capital as an explaining factor of civic community, which constitutes the context for good institutional performance, concludes with the idea that the stocks of trust, norms and systems of participation tend to be cumulative and reinforce themselves mutually. Virtuous circles are created that result in social balances with high levels of cooperation, reciprocity, civism and collective well-being, characteristics that define civic community. (HIGGINS, 2005, p. 67)

On the other hand, to Fukuyama, "social capital is an acting and informal norm that promotes cooperation among two or more individuals" (HIGGINS, 2005, p. 67). This can be simply a reciprocity or a huge complex system like Christianity. Every other aspect related to social capital such as norms, trust, etc. are resulting epiphenomena, not constitutive parts of it. According to this author, social capital assumes an economic and a political function reducing transaction costs on formal negotiations and acting as a counterweight to excessive individualism. In Fukuyama's point of view, it emerges from repeated games of the prisoner's dilemma.

Portes (2000) anchors his discussion of social capital on two authors, namely, Bourdieu and Coleman. To Bourdieu, in very general lines, people build their relations taking into account the benefits they could obtain later. Coleman, on the other hand, focus on social capital as social control source. The author's main focus to show how this concept has been poorly managed giving space to a great amount of confusion and spurious conclusions. To do this, after reviewing in detail both Bourdieu's and Coleman's perspectives on the subject, he analyzes the educational attainment of immigrant children. Social capital indexes are operationalized in these terms:

> Closure of parental networks is measured by the number of parents of child's friends known to her or his parents; parental school involvement is a composite index of parent's participation in school activities and frequency of meetings with school staff about their child's academic progress. The content of these items correspond closely to Coleman's conceptual description of social capital. (PORTES, 2000, p. 7)

Up to this point, it is easy to perceive that the concept is operationalized in a variety of ways in the literature. In fact, Portes (2000, p. 2) states that "much of the controversy surrounding social capital has to do with its application to different types of problems and its use in theories involving different units of analysis". However, since the 1970's, social capital has been conceived within a theoretical framework that accounts both for the structured and the stochastic parts of social life. In this sense, recent researches developed under the relational paradigm in the social sciences pinpoint to the relevance of the mobilization of tangible and intangible resources from imediate social structure of each individual to the possibility of social mobility. In his most famous article "The Strength of Weak Ties", Granovetter (1973) showed the importance of "distant" contacts to have access to new information. According to him, "from the individual's point of view, then, weak ties are an importante resource in making possible mobility opportunity" (GRANOVETTER, 1973, p. 1373). In another work, Granovetter (1995) operationalized the social capital concept magisterally to explain the success within the labor market. He shows that most of the first jobs were obtained through a weak tie. For Lin (1999, p. 470), "the strength of weak ties might lie in their accessing social positions vertically higher in the social hierarchy, which had the advantage in facilitating the instrumental action".

Erickson (2001) develops her paper in close dialogue with Granovetter. She investigates the effect of social capital on the hiring process focusing both on employee and employer. Social capital here is regarded as the diversity and extension of personal network of the individuals. The author states that social capital can be desirable, irrelevant or undesirable. "From an employer's point of view, employee social capital can be an asset if used for the firm but a threat if used by the employee to set up another rival firm, or used by the employee after defecting to another firm" (ERICKSON, 2001, p. 131-2). According to the author, "looking at the social and human capital that workers have, and the level and income of the jobs they have, will show whether good networks make their own contribution to getting a better job" (ERICKSON, 2001, p. 133). The main point of the article is that regardless of the contact point by which the job was taken, having a good social capital will lead to better jobs and better pay.

Now we will take a closer look on how the concept of "social capital" is mobilized within educational inequalities studies.

2.1 Social Capital and Education

Most studies reviewed within the field of education inequalities do not deal with social capital (RIBEIRO, 2009b; RIBEIRO, 2009a; PRATES; COLLARES, 2014; VILELA; COLLARES, 2009; FERNANDES, 2005).

Silva (2003, p. 110) studies educational stratification within the brazilian educational expansion context. This author argues that "a more convenient form to analyze the determination of scholarity is to concept it in terms of a sequence of transitions between educational levels". Thus, Silva (2003) divides his analysis in three moments, namely, the completion of the 1st grade, the completion of the 4th grade and the completion of the 8th grade[1]. He uses an approach sugested by Mare (1980) that employs logistic models to estimate success probabilities for each one of the mentioned transitions. According to Silva (2003, p. 109), to use a linear model is problematic because "we discard the possibility that its coefficients vary in a systematic and significative way through the diverse scholarity levels"[2]. The author uses these variables as predictors: *area*, *region*, *gender*, *skin colour*, *age* (and age^2), *education of the householder*, *income*, *female householder* and *number of children*. It is curious, though, that the mobilization of the variable *female householder* as a proxy of social capital does not seem a good choice because it is based on a highly questionable assumption (that families with a female householder have poorer social capital) and it does not tell us anything apparently relevant about tie mobilization and its resources. Silva (2003) observes the same educational stratification pattern contained in Brazilian specialized literature with the exception of the variable *per capita family income* "which grows as one goes further inside school system; this pattern is explicable by the growing incidence of private and paid teaching in higher instruction levels of Brazilian educational system" (SILVA, 2003, p. 132).

Mont'Alvão (2011) applied multinomial logistic regression models to understand the completion of secondary school and entering university. This author found that, in Brazil from 2001 to 2007, the impact of social origins on the access of graduate education persists or increase when compared to secondary school. He states that "although the private university system is more unequal than the public system, the later has shown a considerable increase ineducational stratification. Inequalities based on geographic region, race, gender, and social capital are also present" (MONT'ALVÃO, 2011, p. 430). Unfortunately, the author used the same variable (*female householder*) as a proxy of social capital without any discussion on the subject. Indeed, the only place he mentions social capital is on the articles abstract.

[1] Grades of Brazilian school system.
[2] All translations within this paper were made by the author.

Silva & Hasenbalg (2000) also use *female householder* along with the number of *14 or less years old children* in the families studied, the percentage of *mothers that work* and whether *children from 10 to 14 years old work* or not. According to them

> Two changes in familiar structure observed in the last decades can represent a decrease of social capital at children's disposal and less favorable contexts to their schooling. The first, resultant from the growing instability of marriages and unions, is the increase of the proportion of families that have a female householder (. . .). The second change is due simply to the growing participation of women in the labor force, which results in the increase of the proportion of mothers that work. (SILVA; HASENBALG, 2000)

Although the authors are worried about explaining the social context of the chosen variables, why are they good indexes for social capital or not is never discussed.

The papers reviewed here are sufficient to show that the research field of educational inequalities in Brazil does not seem to give much attention to how "social capital" is operationalized. In short, the main argument I want to delineate in this section is that this concept has been widely and inadvertently used. I believe that the most useful tool to handle it is social network analysis by which social capital can be objectly measured using a repertoire of indexes like Burt's constraint, centrality measures (both indexes for popularity and activity within a given group), structural equivalence and modelled using up-to-date statistical techniques. This methodological framework will be presented in next section.

3 Data and Methods

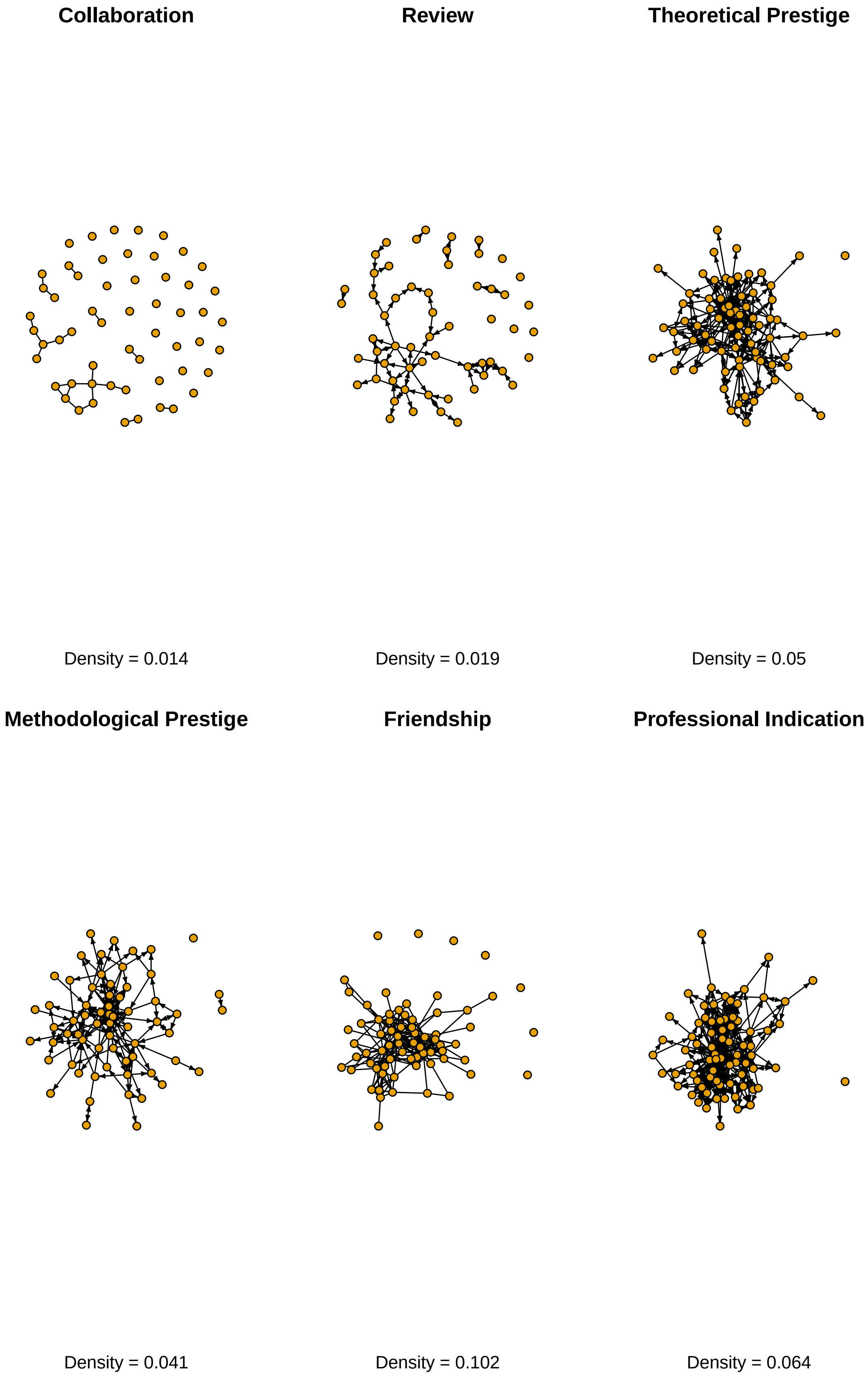
In conducting this research, data from 47 postgraduate students was collected through an online survey in May, 2016. This sample corresponds to 60% of the total number of students enrolled in a postgraduation program at a Brazilian university and it is representative of the total. Therefore, for the purpose of the analysis conducted here, we will assume this group as a "complete network".

The developed questionnaire had four parts.The first was an invitation and an agreement term with which the subjects must have agreed to proceed to the questions. In the second part, the subjects were questioned about personal and socialeconomic information. In the third part, they were questioned about their academic life, productivity and impressions. In the last part, the sociometric part, they were present to the complete list of grad colleagues and they were asked to indicate persons regarding 6 issues: scientific collaboration, paper reviewing, theoretical and methodological advisement, friendship and professional indication[3]. The sociometric questions generated the 6 networks in Figure 1.

To perform the analysis, social network analysis metrics were used as well as a statistic model from the *exponential random graph models* family, or p* models (ROBINS et al., 2007;

[3] The questions used were: (1) With which of these colleagues did you write or publish a scientific work?; (2) To which of these colleagues did you ask to revise a paper?; (3) If you had a theoretical doubt, whom would you ask for help?; (4) If you had a methodological doubt, whom would you ask for help?; (5) Which of these colleagues do you get together on social occasions?; (6) If you knew about a job vacancy in your acting field, which of these colleagues would you indicate?

Figure 1 – Six Networks



Note: Developed by the author with **R** (R CORE TEAM, 2016) and **igraph** (CSARDI; NEPUSZ, 2006).

LUSHER; KOSKINEN; ROBINS, 2013; LAZEGA; HIGGINS, 2014). This models were developed specially to deal with the observations independence problem. Classic statistical models assume that observations are independent one from another. This is not the case when one is working with network data. “The fact that I choose Peter as a friend is not necessarily independent of the fact that I also choose Paul because they can be friends with each other” (LAZEGA; HIGGINS, 2014, p. 76). Lazega & Higgins (2014) state that a tie presupposes a triad which leads to dependency among observations.

The p* model can be defined by

$$Pr(Y = y) \quad = \quad \left(\frac{1}{k}\right) exp\left\{\sum_{A} \eta_A g_A(\mathbf{y})\right\} \tag{1}$$

where Y is the theoretical estimated graph, y is the observed graph, $\sum_A$ is the summation of all configurations A, η_A is the estimated parameter corresponding to the configuration A, $g_A(\mathbf{y})$ is the network statistic corresponding to the configuration A of the graph $\mathbf{y}$ and k is a constant which ensures the proper probability distribution (ROBINS et al., 2007).

Here, I use an extension of the p* models known as *Social Selection Model* (SSM). The SSM was proposed by Robins, Elliott & Pattison (2001) with the goal of accounting for heterogeneity within the social structures using nodal attributes as exogenous covariates. So, in addition to modelling endogenous variables, i.e., network configurations that explain self-organizing processes, the SSM accounts for exogenous variables that also have an effect of structure emergence (WANG et al., 2016). Beyond that, I also analyze the effect of dyadic covariates, i.e., the effect of the existence of an i–j tie in another relation network on the existence of an i–j tie in the modelled network (ROBINS; DARAGANOVA, 2013).

In the next section, I will analyze how the attributes (exogenous variables) relate to each other, I will describe and present the network measures that are important for this study and I also present the SSM results.

4 Results

4.1 Individual related variables

If we observe some descriptive statistics of the data in Tables 1 and 2, we can notice that the respondent students of this postgraduate program have a mean age of 30,32 years and a mean income of R$2.636,00 (Brazilian reais). Most of them are women, they are predominantly white, 63,8% do not have a job, 51% are doctoral students and 70,2% receive academic scholarship.

When asked “which is your main occupation?”, 61,7% of the respondents declare themselves as “students”. The second most cited occupation was “professor” (10,6%) followed by “sociologist” (8,5%) as we can see in Table 3.

The subjects were also asked to answer about their perception on their on grades and productivity (Figure 2) regarding paper publication. Most students have a good self-evaluation

Table 1 – Descriptive Statistics - Quantitative

Age	Income (R$)
Min. : 23	Min. : 0
1st Qu.: 27	1st Qu.: 1500
Median : 29	Median : 2200
Mean : 30.32	Mean : 2636
3rd Qu.: 33	3rd Qu.: 3000
Max. : 43	Max. : 8200

Note: Elaborated by the author

Table 2 – Descriptive Statistics - Qualitative (%)

Gender	Race	Formal Work	Academic situation	Scholarship
Female: 63,83	White: 55.32	No: 63.83	Doctor: 06.38	No: 29.79
Male: 36,17	Mestizo: 02.13	Yes: 36.17	Doctorate Std.: 51.06	Yes: 70.21
	Dark Skinned: 02.13		Master Std.: 40.43	
	Black: 04.26		Master: 02.13	
	Brown: 36.17			

Note: Elaborated by the author

Table 3 – Occupations

	Frequency	Percentage
Research Assistant	1	2.13
Social Worker	1	2.13
Housewife	1	2.13
Student	29	61.70
Researcher	3	6.38
Professor	5	10.64
Psychologist	1	2.13
Government Employee	2	4.26
Sociologist	4	8.51
Total	47	100.00

Note: Elaborated by the author

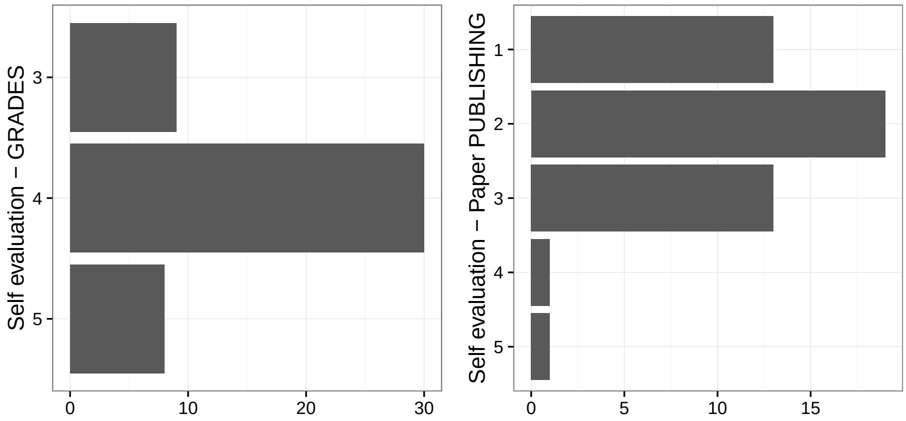
about their grades but a bad one about publishing[4]. In Figure 3 we can check respondents productivity measured by the number of published papers, the number of papers presented in congresses and the number of classes given (all since 2014). As a preliminar analysis of the relations between students perceptions (*Grades* and *Pubs*) and their actual productivity (*Papers, Congress* and *Classes*), I correlated this variables. Age and Income were also added to the correlation matrix (Table 4) along with a *productivity score* (*Prod.*) made by summing the productivity variables with weight 2 assigned to published papers.

In Table 4 we can notice that the variable pairs that have the highest correlation are *Self Evaluation Publications – Papers*, and *Self Evaluation Publications – Productivity Score*. Age and Income had very low correlation coefficients with the other variables.

To investigate if age, income, scholarship and occupation have an effect on academic productivity within the studied universe, I estimated two linear models, one by OLS and a Generalized

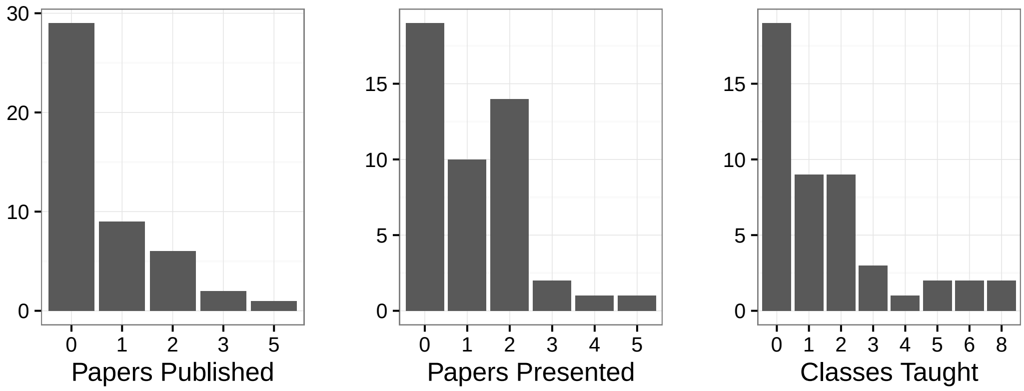
[4] For "grades", the perception was measured in a scale from 1 (I have horrible grades) through 5 (I have excellent grades). For "productivity", the perception was also measured in a scale from 1 (I do not publish anything ever) through 5 (I am a publishing machine).

Figure 2 – Self-evaluation



Note: Elaborated by the author using **ggplot2** (WICKHAM, 2009).

Figure 3 – Productivity



Note: Elaborated by the author

Table 4 – Correlations

	Age	Income	Grades	Pubs	Papers	Congress	Classes	Prod.
Age	1.00							
Income	0.39	1.00						
Grades	-0.19	-0.03	1.00					
Pubs	0.09	0.16	0.32	1.00				
Papers	-0.01	0.12	0.09	0.54	1.00			
Congress	0.16	0.19	0.24	0.49	0.57	1.00		
Classes	0.21	0.45	-0.16	0.15	0.18	0.07	1.00	
Prod.	0.16	0.37	0.04	0.53	0.82	0.65	0.66	1.00

Note: Elaborated by the author using **xtable** (DAHL, 2016).

Linear Model with gamma distribution and identity link, with the following specification:

$$\widehat{Productivity} = \beta_0 + \beta_1 ResearchGroup + \beta_2 Gender + \beta_3 White + \beta_4 SelfEval(Pub) + \beta_5 Age(centralized) + \beta_6 Work + \beta_7 Income + \beta_8 Scholarship + \epsilon \quad (2)$$

The results are presented in Table 5 and a Likelihood test is presented in Table 6. The

Table 5 – Regression Models

	Model 1 (OLS)	Model 2 (GLM – Gamma)
Intercept	−1.14 (2.17)	0.19 (1.88)
Research Group (Yes)	0.86 (1.06)	1.52 (0.91)
Gender (Male)	0.43 (1.17)	−0.38 (0.71)
White (Yes)	0.70 (1.06)	0.79 (0.73)
Self-eval (Pub)	1.93 (0.60)**	1.80 (0.54)**
Age (centralized)	0.00 (0.12)	0.13 (0.10)
Work (Yes)	−1.69 (1.37)	−1.43 (0.89)
Income	0.00 (0.00)*	0.00 (0.00)
Scholarship (Yes)	−0.52 (1.34)	−0.55 (0.99)
R^2	0.41	
Adj. R^2	0.29	
Num. obs.	47	47
RMSE	3.33	
AIC		238.57
BIC		257.07
Log Likelihood	-118.29	-109.29
Deviance	422.34	19.16

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Note: Elaborated by the author.

Likelihood test shows that model 2 has a better adjust to the data. We are not trying to build a prediction model or making unscrupulous inferences. Therefore, p-values here are irrelevant. We will now analyze the estimated effect of the variables over productivity inside this particular context. Further inferences must respect the self-organizing charasteristics of any empirical social structure.

What first comes to attention is the big effect that participating in a research group has over productivity. This is the second biggest effect found on the model and, therefore, a central variable to understand academic productivity. At first, men seem to be more productive than women. A *t test* showed this relation. However, in the GLM, when controlling for the other available variables we find that womem are more productive, *ceteris paribus*. This is very interesting although not possible to deal with here. I intend to deepen into the academic production mechanisms in another work. Also white people seem to be more productive than non-white in this network. Self-evaluation regarding publishing has a big positive effect on productivity. This shows us that students tend to be honest about their own academic performance, nothing more. We can be tented to take hasty conclusions about expectations and productivity but it is very difficult to talk about causality in this case since we have no further information on the mechanisms that involve these variables. All we can affirm is that students expectations are highly correlated with their academic performance. To be working on a formal job seems to reduce productivity which is somewhat obvious. Age and Income have very low effects and, curiously, scholarship has a negative effect. This is also very interesting because it is expected that scholarship students have more time to dedicate to their research projects and, therefore, a higher productivity performance which is not the case. I will come back to this point in the last section.

Table 6 – Likelihood Ratio Test

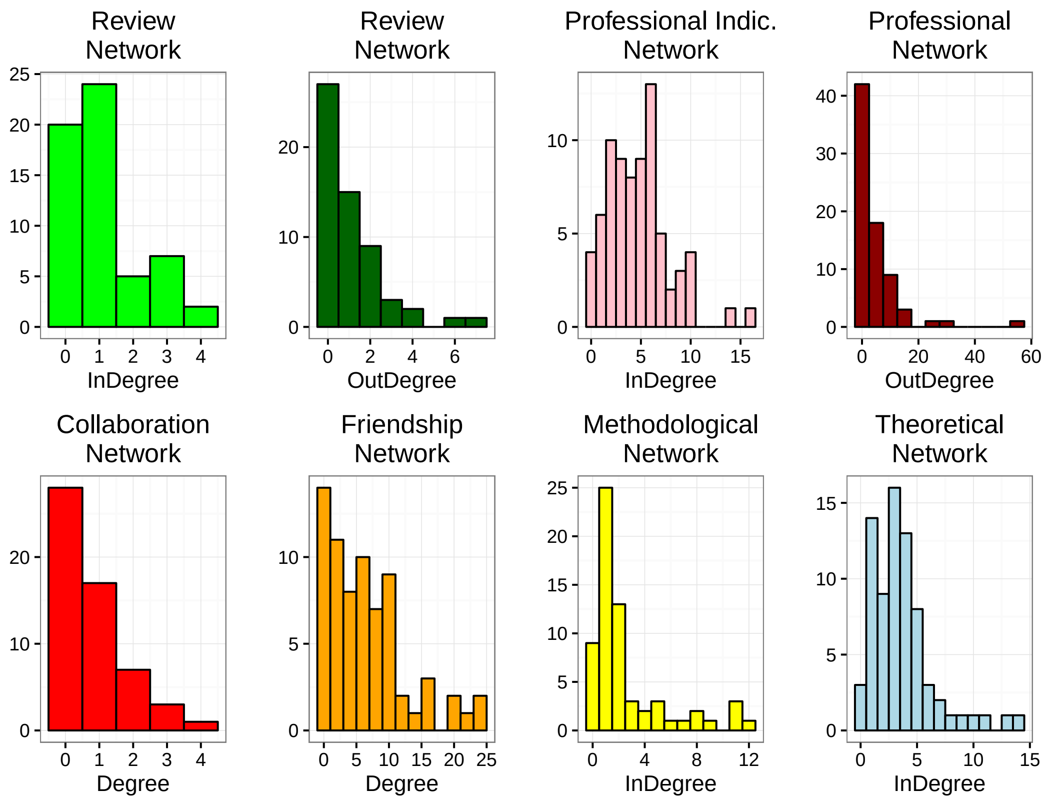
	#Df	LogLik	Df	Chisq	Pr(>Chisq)
OLS	10	-118.29			
GLM (Gamma)	10	-109.29	0	18.01	0.0000

4.2 Relational Data

Here, we begin to operationalize the social capital concept as resources embbeded in the social structure. The main purposive action of the postgraduate student in Brazilian academic system is to publish. In the mobilization of one's immediate social structure, this can be facilitated by collaborating, asking for a pair review on a ready article or simply by taking advisement from people that have more theoretical or methodological habilities. Also, popularity and a good social positioning, for example, are powerful social resources. Let us begin the analysis.

In Figure 4, I present the degree distributions for the six networks.

Figure 4 – Degree Distributions



Note: Elaborated by the author

The network measures that interest us here are, predominantly, indegree centrality and outdegree centrality which are indicators respectively for popularity and activity. We can notice that Review and Collaboration networks are not so connected; they have a high amout of nodes with degrees 0 and 1. The Professional Indication network is relatively decentralized with popularity and it has some nodes who are extremely active (with more than 20 indications). The Friendship network is also highly decentralized. Regarding the Methodological and the

Theoretical advisement, popularity is what is most important for this investigation rather than activity, i.e., who is seen by the pairs as the most competent student regarding methodological and theoretical issues and in what extent this popularity relates with productivity. It is important to observe that the Methodological network is more concentrated than the Theoretical one. The later has more nodes with indegrees 3, 4 and 5 than the former.

I estimated again the generalized linear model presented in equation 2 with "log" link and the gamma distribution and including centrality measures for the collaboration, review (in) and methodological (in) networks, one at a time; although no one of them had a significant effect, again, the p-value here is not relevant. The exponential coefficients for these measures were $Cent(Collaboration) = 1.17$, $Cent(Review) = 1.10$ and $Cent(Methodological) = 1.00$ showing that people who collaborate tend to be more productive. The same is applied to people who are regarded as methodologically competent.

4.3 Collaboration and association among postgraduate students

4.3.1 Stochastic Blockmodel

One of the most important concepts of social network analysis is *structural equivalence*. Two individuals are considered structurally equivalent when they present the same relational profile, i.e., the same tie patterns (LAZEGA; HIGGINS, 2014; DE NOOY; MRVAR; BATAGELJ, 2011; WASSERMAN; FAUST, 1994).

I looked after blocks of structurally equivalent nodes within the Review network. I used the Erdös-Rényi mixture model, a special case of binary stochastic blockmodels. It was fit with the algorithm presented by Daudin, Picard & Robin (2008). The results for the model are shown in Figure 5 and the Review network coloured by block is in Figure 6.

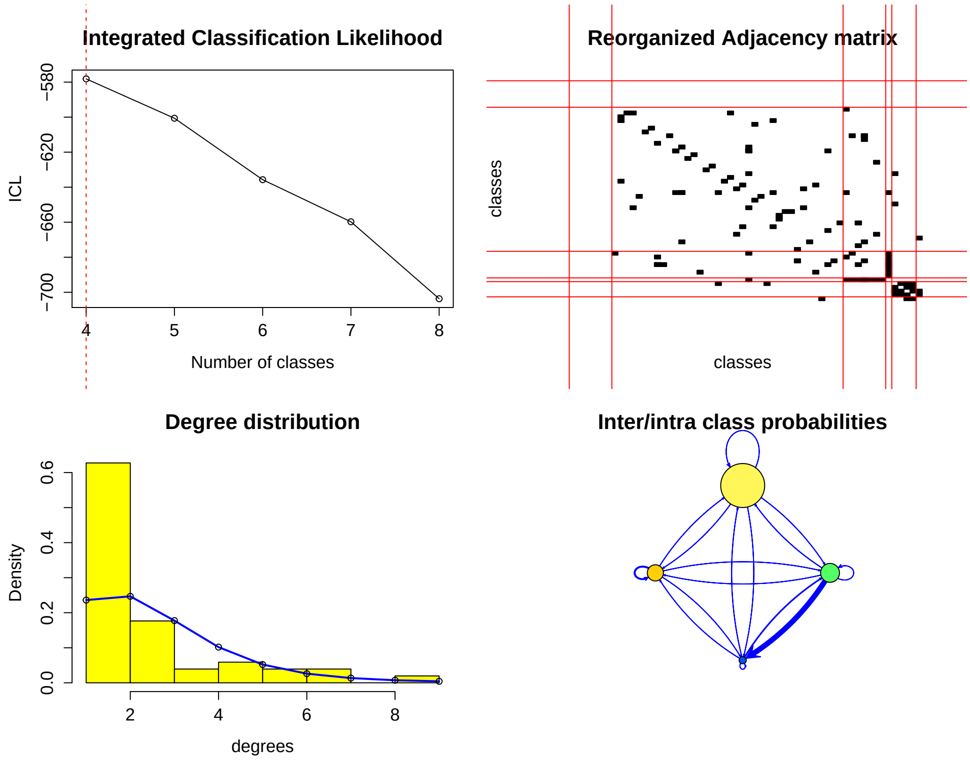
Blocks 1 and 2 are constituted by students who are affiliated to two very close research groups. These groups deal with quantitative sociology and social network analysis. Block 3 shows people who belong to other research group that focus on sociology of crime. The model shows that people within strong research groups tend to be structurally equivalent regarding requests for review.

The blockmodel clustered nodes according to some strong research groups within the program. This shows us that being in a group shapes your relational pattern and, therefore, are an important feature of social reality. Now, we will focus on the Collaboration network seeking an statistical explanation for its emergence.

4.3.2 Social Selection Model

In Table 7 are presented the results for the p* models estimated to analyze the Collaboration Network. First I estimated a model only with edges and controlling for the covariate network effects. Then, I inserted three more structure configurations, namely, isolates, triangulation and connectivity to estimate model 2. Model 3 was fit inserting five attribute related statistics, namely, being in the same research group, being of the same sex, being white, the productivity

Figure 5 – Blockmodel results



Note: Elaborated by the author with **mixer**

score and income. Statistical significance is measured here by Wald test, i.e., the coefficient will be statistic significant if it is bigger than two times the standard error (LUSHER; KOSKINEN; ROBINS, 2013; LAZEGA; HIGGINS, 2014).

All adjustment measures suggest that model 3 is the one best suited for the network we want to explain. We will focus on it. The edges parameter is often compared to the intercep of a regular regression model. It indicates that this network has a lot less edges than it would be expected in a "random world". The isolates and the connectivity coefficients were not significant which indicates us that these are not important configurations for the emergence of this network. Triangulation coefficient indicates that this network has a tendency for group formation. The estimates for the covariate network effects show that all other measured relations have influence on how postgraduate students collaborate. The effect of the Review net is the biggest one which shows us that students who ask for paper review have a lot more probability to actually write or publish in partnership. The second biggest effect is found in Methodological net. This result tell us that students tend to collaborate more with people they see as methodological competent than with they see as theoretically competent. The positive and significative coefficient for Professional Indication net explicits a different and more generic kind of prestige. The Friendship net had a negative and significant estimate showing that these postgraduate students build different relations with regard to friendship and academic work. They tend to write and publish with some people and develop friendship relations with different people, i.e., these two social features do not coincide.

Figure 6 – Review Network with Structural Equivalence

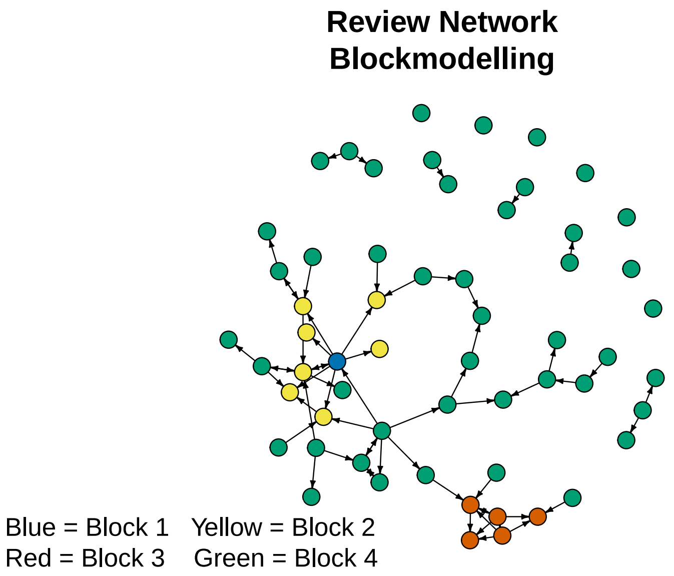


Table 7 – ERGM's – Dependent: **Collaboration Network**

	Model 1	Model 2 (ERGM)	Model 3 (SSM)
Purely structural effects			
Edges	-4.56 (0.28)*	-4.05 (1.10)*	-4.22 (0.13)*
Isolates		0.52 (0.74)	0.37 (0.22)
Triangulation (gwesp)		1.04 (0.36)*	1.07 (0.29)*
Conectivity(twopath)		-0.16 (0.29)	-0.14 (0.15)
Covariate network effects			
Review net	1.81 (0.70)*	1.87 (0.68)*	2.31 (0.14)*
Theoretical net	0.40 (0.66)	0.40 (0.65)	0.54 (0.14)*
Methodological net	1.26 (0.68)	1.27 (0.66)	1.24 (0.13)*
Professional Indic. net	0.81 (0.52)	0.73 (0.50)	0.65 (0.23)*
Friendship net	-0.46 (0.67)	-0.50 (0.66)	-0.54 (0.09)*
Actor-relation effects			
Same Research Group			1.94 (0.29)*
Homophily (Gender)			-0.65 (0.12)*
Homophily (White)			-0.73 (0.31)*
Absolute Difference (Productivity)			-0.00 (0.05)
Absolute Difference (Income)			0.00 (0.00)
AIC	223.33	220.81	220.26
BIC	255.15	268.54	294.51
Log Likelihood	-105.66	-101.41	-96.13

*$significant$ (Wald test)

Note: Elaborated by the author using **statnet** (HANDCOCK et al., 2008) and **texreg** (LEIFELD, 2013).

The SSM shows that, again, participating in a research group is a very important variable to understand academic collaboration. The research group big effect tells us that the students have greater probabilities of writing and publishing together within groups. I also tested gender and race homophily; both were refused by model results. Students in this program tend to collaborate with people of different race and gender. I did not find any significant estimates for productivity score and income which shows these are not important variables to explain collaboration ties formation.

5 Discussion

James Moody (2004, p. 213) stated that the scientific collaboration network in social sciences is moved by research specialty and that “quantitative work is more likely to be coauthored than non-quantitative work”. In this research we found the same pattern with the difference that it was not the research specialty itself that connects students but, essencialy, the research groups. This is very clear from both the blockmodel and the SSM results. The main cement that glues this postgraduate students together in scientific collaboration is the research groups. Furthermore, the second main aggregator was methodological advisement. The SSM showed that methodological habilities lead to collaboration more than theoretical ones. This is in consonance with Moody’s findings about quantitative work. In fact, the research groups that appear in the blockmodel are essentially quantitative researchers.

The linear models did not present big productivity differences by gender, race or income. In this sense, a postgraduate program seems to provide equal opportunities for all its students, once they are in. The models pointed for the centrality of research groups to productivity.

A quite amazing finding is that scholarship students are less productive than non-scholarship ones. This deserves deeper investigation for it contrasts all university funding policies. One hypothesis would be that most non-scholarship students work as professors in other universities and, therefore, are more productive considering the amount of classes to give and final works to orient. That was not the case since most non-scholarship students are not working. Also, that hypothesis is not in agreement with the linear models results which show that, on average, students who work are less productive than those who do not. This subject is out of the scope of this paper and I will limit myself to point the relevance of further investigation.

Finally, I hope to have shown the strength of social network analysis on the operationalization of social capital. Although I have addressed a different area from the educational inequalities articles reviewed, I could deal with the same concept, i.e., “resources embedded in a social structure which are accessed and/or mobilized in purposive actions” (LIN, 1999, p. 35) in a more straightfoward and objective way. We can actually visualize the social structure and mathematically measure popularity, social positioning and social roles. The educational inequalities field would gain a lot with the use of this methodological framework and a more reflexive way of thinking on social capital.

Bibliography

CSARDI, Gabor; NEPUSZ, Tamas. The igraph software package for complex network research. *InterJournal, Complex Systems*, v. 1695, n. 5, p. 1–9, 2006.

DAHL, David B. *xtable: Export Tables to LaTeX or HTML*. [S.l.], 2016. R package version 1.8-2. Disponível em: ⟨https://CRAN.R-project.org/package=xtable⟩.

DAUDIN, J-J; PICARD, Franck; ROBIN, Stéphane. A mixture model for random graphs. *Statistics and computing*, Springer, v. 18, n. 2, p. 173–183, 2008.

DE NOOY, Wouter; MRVAR, Andrej; BATAGELJ, Vladimir. *Exploratory social network analysis with Pajek*. Cambridge: Cambridge University Press, 2011.

ERICKSON, Bonnie H. Good networks and good jobs: The value of social capital to employers and employees. In: LIN, Nan; COOK, Karen; BURT, Ronald (Ed.). *Social capital: Theory and research*. New York: Aldine De Gruyter, 2001. p. 127–158.

FERNANDES, Danielle C. Estratificação educacional, origem socioeconômica e raça no brasil: as barreiras da cor. *Prêmio Ipea*, v. 40, p. 21–72, 2005.

GRANOVETTER, Mark S. The strength of weak ties. *American journal of sociology*, JSTOR, p. 1360–1380, 1973.

GRANOVETTER, Mark S. *Getting a job*: A study of contacts and careers. Chicago: University of Chicago Press, 1995.

HANDCOCK, Mark S. et al. statnet: Software tools for the representation, visualization, analysis and simulation of network data. *Journal of Statistical Software*, v. 24, n. 1, p. 1–11, 2008. Disponível em: ⟨http://www.jstatsoft.org/v24/i01⟩.

HIGGINS, Silvio Salej. *Fundamentos teóricos do capital social*. Chapecó: Argos, 2005.

LAZEGA, Emmanuel; HIGGINS, Silvio Salej. *Redes sociais e estruturas relacionais*. Belo Horizonte: Fino Traço, 2014.

LEIFELD, Philip. texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software*, v. 55, n. 8, p. 1–24, 2013. Disponível em: ⟨http://www.jstatsoft.org/v55/i08/⟩.

LIN, Nan. Building a network theory of social capital. *Connections*, v. 22, n. 1, p. 28–51, 1999.

LUSHER, Dean; KOSKINEN, Johan; ROBINS, Garry (Ed.). *Exponential random graph models for social networks*: Theory, methods, and applications. Cambridge: Cambridge University Press, 2013.

MARE, Robert D. Social background and school continuation decisions. *Journal of the American Statistical Association*, Taylor & Francis, v. 75, n. 370, p. 295–305, 1980.

MONT'ALVÃO, Arnaldo. Estratificação educacional no brasil do século XXI. *Dados*, SciELO Brasil, v. 52, n. 2, p. 389–430, 2011.

MOODY, James. The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American sociological review*, Sage Publications, v. 69, n. 2, p. 213–238, 2004.

PORTES, Alejandro. Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, p. 1–24, 1998.

PORTES, Alejandro. The two meanings of social capital. In: SPRINGER. *Sociological forum*. [S.l.], 2000. v. 15, n. 1, p. 1–12.

PRATES, Antonio Augusto Pereira; COLLARES, Ana Cristina Murta. *Desigualdade e expansão do ensino superior na sociedade contemporânea*: O caso brasileiro no final do século XX ao princípio do século XXI. Belo Horizonte, MG: Fino Traço, 2014.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2016. Disponível em: ⟨https://www.R-project.org/⟩.

RIBEIRO, Carlos Antonio Costa. Classe, raça e mobilidade social no brasil. In: _____. *Desigualdade de oportunidades no Brasil*. Belo Horizonte, MG: Argvmentvm, 2009a. p. 151–186.

RIBEIRO, Carlos Antonio Costa. Desigualdade de oportunidades educacionais no brasil: Raça, classe e gênero. In: _____. *Desigualdade de oportunidades no Brasil*. Belo Horizonte, MG: Argvmentvm, 2009b. p. 21–74.

ROBINS, Garry; DARAGANOVA, Galina. Social selection, dyadic covariates, and geospatial effects. In: LUSHER, Dean; KOSKINEN, Johan; ROBINS, Garry (Ed.). *Exponential random graph models for social networks*: Theory, methods, and applications. Cambridge: Cambridge University Press, 2013.

ROBINS, Garry; ELLIOTT, Peter; PATTISON, Philippa. Network models for social selection processes. *Social networks*, Elsevier, v. 23, n. 1, p. 1–30, 2001.

ROBINS, Garry et al. An introduction to exponential random graph (p*) models for social networks. *Social networks*, Elsevier, v. 29, n. 2, p. 173–191, 2007.

SILVA, Nelson do Valle; HASENBALG, Carlos. Tendências da desigualdade educacional no Brasil. *Dados*, scielo, v. 43, p. 423 – 445, 2000. Disponível em: ⟨http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582000000300001&nrm=iso⟩.

SILVA, Nelson do Valle. Expansão escolar e estratificação educacional no brasil. In: HASENBALG, Carlos Alfredo; SILVA, Nelson do Valle (Org.). *Origens e destinos*: desigualdades sociais ao longo da vida. Rio de Janeiro: IUPERJ/UCAM, 2003.

VILELA, Elaine M.; COLLARES, Ana Cristina. Origens e destinos sociais: pode a escola quebrar essa ligação? *Teoria & Sociedade (UFMG)*, v. 17, n. 62-91, 2009.

WANG, Peng et al. Social selection models for multilevel networks. *Social Networks*, Elsevier, v. 44, p. 346–362, 2016.

WASSERMAN, Stanley; FAUST, Katherine. *Social network analysis*: Methods and applications. Cambridge: Cambridge university press, 1994. v. 8.

WICKHAM, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2009. ISBN 978-0-387-98140-6. Disponível em: ⟨http://ggplot2.org⟩.