Social Capital Inequalities among Postgraduate Students and Social Selection Processes

Neylson J. B. F. Crepalde[1]

[1]Sociology Postgraduation Program – Minas Gerais Federal University (UFMG)
Belo Horizonte, MG – Brazil

neylsoncrepalde@gmail.com

Abstract. *This paper aims to discuss social capital inequalities between post-graduate students enrolled in a social sciences program in a Brazilian university. 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.*
Keywords*: Social Capital; Social Network Analysis; Social Selection Model.*

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 social selection processes, i.e., the processes by which people make ties with each other forming groups, are not so acknowledged yet although they are central to understand how the access to social resources is achieved. This has only recently been investigated.

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, I will focus our investigation in two main aspects: academic productivity and social selection/tie formation. I used data collected online from 47 social sciences postgraduate students (masters and doctoral). To perform the analysis, I used linear models, social network analysis, stochastic blockmodelling and the highly modern *social selection model* (SSM).

[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 findings: first, 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 I aim to investigate.

2. 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[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, Lusher et al. 2013, Lazega and Higgins 2014]. 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 et al. 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 and Daraganova 2013].

3. Results

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

---

[1]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?

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

The results are presented in Table 1. A Likelihood test showed that model 2 has a better adjust to the data. 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 I 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.

3.1. Collaboration and association among postgraduate students

3.1.1. 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 and Higgins 2014, De et al. 2011, Wasserman and 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 et al. 2008].

The algorithm got a best result with 4 blocks. Blocks 1 (two students) and 2 (seven students) are constituted by individuals who are affiliated to two very close research groups. These groups deal with quantitative sociology and social network analysis. Block 3 (five students) shows people who belong to other research group that focus on sociology of crime. The remaining students were allocated to Block 4. 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, it is an important feature of social reality. Now, I will focus on the Collaboration network seeking an statistical explanation for its emergence.

Table 1. 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

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

3.1.2. Social Selection Model

In Table 2 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 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 et al. 2013, Lazega and Higgins 2014].

The adjustment measures suggest that model 3 is the one best suited for the network I 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 Indi-

Table 2. 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)

cation 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.

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.

4. Discussion

[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 I found the same pattern with the difference that it was not the research specialty itself that connects students but, essencialy, the research groups. The SSM showed that methodological habilities lead to collaboration more than theoretical ones. This is in consonance with [Moody 2004]'s findings about quan-

titative work. In fact, the research groups that appear in the blockmodel are essentially quantitative researchers.

References

Daudin, J.-J., Picard, F., and Robin, S. (2008). A mixture model for random graphs. *Statistics and computing*, 18(2):173–183.

De, Nooy, W., Mrvar, A., and Batagelj, V. (2011). *Exploratory social network analysis with Pajek*. Cambridge University Press, Cambridge.

Erickson, B. H. (2001). Good networks and good jobs: The value of social capital to employers and employees. In Lin, N., Cook, K., and Burt, R., editors, *Social capital: Theory and research*, pages 127–158. Aldine De Gruyter, New York.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, pages 1360–1380.

Granovetter, M. S. (1995). *Getting a job*. University of Chicago Press, Chicago.

Higgins, S. S. (2005). *Fundamentos teóricos do capital social*. Argos, Chapecó.

Lazega, E. and Higgins, S. S. (2014). *Redes sociais e estruturas relacionais*. Fino Traço, Belo Horizonte.

Lin, N. (1999). Building a network theory of social capital. *Connections*, 22(1):28–51.

Lusher, D., Koskinen, J., and Robins, G., editors (2013). *Exponential random graph models for social networks*. Cambridge University Press, Cambridge.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American sociological review*, 69(2):213–238.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, pages 1–24.

Portes, A. (2000). The two meanings of social capital. In *Sociological forum*, volume 15, pages 1–12. Springer.

Robins, G. and Daraganova, G. (2013). Social selection, dyadic covariates, and geospatial effects. In Lusher, D., Koskinen, J., and Robins, G., editors, *Exponential random graph models for social networks*. Cambridge University Press, Cambridge.

Robins, G., Elliott, P., and Pattison, P. (2001). Network models for social selection processes. *Social networks*, 23(1):1–30.

Robins, G., Pattison, P., Kalish, Y., and Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social networks*, 29(2):173–191.

Wang, P., Robins, G., Pattison, P., and Lazega, E. (2016). Social selection models for multilevel networks. *Social Networks*, 44:346–362.

Wasserman, S. and Faust, K. (1994). *Social network analysis*, volume 8. Cambridge university press, Cambridge.