COSC 6323 - Statistical Methods in Research
Project Phase - 2

Members: Team-8

1. Farah Naz Chowdhury, ID:1798957, fchowdhury4@uh.edu
2. Md Rafiqul Islam Rabin, ID:1797648, mrabin@central.uh.edu
3. S M Salah Uddin Kadir , ID:1800503, ssalahuddinkadir@uh.edu

April 05, 2019.

Contributions: Contribution paragraph goes here.

Fig. 4:

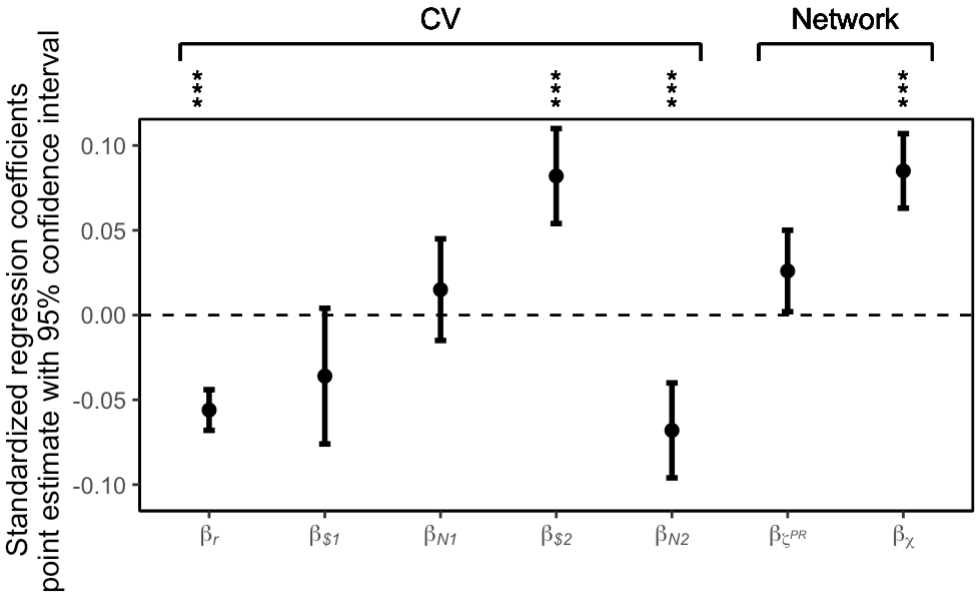

Fig. 4. Career cross-sectional regression model.

Description of figure content:

The figure shows the standardized coefficient values of (CV + Network) model. Each coefficient values has corresponding error ranges. All of the values have been taken after applying the logarithms. We have specified the CV coefficients and Network coefficients using the annotation. We have used the three star (***) annotation to specify the coefficients which probability of t-value is less than 0.001.

Observations, conclusions, and hypotheses:

From the figure, we can see that the total amount of funding of NIH is greater than the total amount of NSF funding. The p-value of NIH is less than 0.001 which indicates the NIH coefficients are significant. Although, the correlation with the number of NIH grants (β_{N2}) is negative, so the cost is related to the management of several smaller grants versus fewer bigger grants. On the other hand, the p-values of NSF coefficients are greater than 0.001, so we can say that the relation of NSF variables are not significant in the OLS model. So, there are different levels of dependency on the funding between the biology and computer science faculty.

Table S2:

	CV		CV + Network		CV + Network [Standardized]	
CV parameters						
Departmental rank, β_r	-0.052***	(0.006)	-0.047***	(0.006)	-0.056***	(0.007)
Productivity (h-index), β_h	1.857***	(0.016)	1.866***	(0.018)	1.179***	(0.012)
Total NSF funding, $\beta_{\$1}$	-0.005*	(0.002)	-0.005*	(0.002)	-0.035*	(0.015)
# of NSF grants, β_{N1}	0.024	(0.014)	0.013	(0.015)	0.015	(0.016)
Total NIH funding, $\beta_{\$2}$	0.016***	(0.004)	0.014***	(0.003)	0.083***	(0.021)
# of NIH grants, β_{N2}	-0.067***	(0.019)	-0.061**	(0.019)	-0.069**	(0.022)
Network parameters						
PageRank Centrality, $\beta_{\mathscr{C}^{PR}}$			0.041**	(0.014)	0.026**	(0.009)
Cross-disciplinarity, β_x			0.571***	(0.061)	0.086***	(0.009)
Discipline ($\mathscr{O}$) dummy	Y		Y		Y	
5-year cohort ($y^0_{i,5}$) dummy	Y		Y		Y	
Constant	1.398***	(0.234)	1.706***	(0.271)	7.744***	(0.216)
n	4190		3900		3900	
adj. R^2	0.883		0.882		0.882	

Standard errors in parentheses.
* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

Table S2. Career data set: Pooled cross-sectional model.

Description of figure content:

Descriptions goes here.

Observations, conclusions, and hypotheses:

observations goes here.

Table S3:

	(a) $\mathscr{C}^{PR}$	(b) $\mathscr{C}^{B}$	(c) $\mathscr{C}^{D}$	(d) ~~β_{N1}, β_{N2}~~	(e) ~~β_r~~
CV parameters					
Departmental rank, β_r	-0.047***	-0.042***	-0.044***	-0.046***	
	(0.006)	(0.006)	(0.006)	(0.006)	
Productivity (h-index), β_h	1.866***	1.901***	1.848***	1.862***	1.892***
	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)
Total NSF funding, $\beta_{\$1}$	-0.005*	-0.004	-0.005*	-0.003**	-0.005*
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
# of NSF grants, β_{N1}	0.013	0.009	0.007		0.006
	(0.015)	(0.015)	(0.015)		(0.015)
Total NIH funding, $\beta_{\$2}$	0.014***	0.014***	0.014***	0.003*	0.013***
	(0.003)	(0.004)	(0.003)	(0.001)	(0.004)
# of NIH grants, β_{N2}	-0.061**	-0.065**	-0.062**		-0.059**
	(0.019)	(0.020)	(0.019)		(0.019)
Network parameters					
PageRank centrality, $\beta_{\mathscr{C}^{PR}}$	0.041**			0.042**	0.057***
	(0.014)			(0.014)	(0.014)
Betweenness centrality, $\beta_{\mathscr{C}^{B}}$		-0.000			
		(0.005)			
Degree centrality, $\beta_{\mathscr{C}^{D}}$			0.052***		
			(0.010)		
Cross-disciplinarity, β_{χ}	0.571***	0.562***	0.530***	0.579***	0.555***
	(0.061)	(0.062)	(0.061)	(0.061)	(0.061)
Discipline ($\mathscr{O}$) dummy	Y	Y	Y	Y	Y
5-year cohort ($y^0_{i,5}$) dummy	Y	Y	Y	Y	Y
Constant	1.706***	1.200***	1.344***	1.711***	1.615***
	(0.271)	(0.225)	(0.226)	(0.271)	(0.273)
n	3900	3387	3900	3900	3900
adj. R^2	0.882	0.873	0.883	0.882	0.881

Standard errors in parentheses, listed below coefficient estimate.
* $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

Table S3. Career data set: Pooled cross-sectional model — robustness check.

Description of figure content:

The parameter estimates for the variants of the 'CV + Network' pooled cross-sectional models reported in Table S3: (a) Model with PageRank centrality, (b) Model with betweenness centrality, (c) Model with degree centrality, (d) Model without the number of grants variables, and (e) Model without the departmental rank variable. The model (a) of the Table S3 is same as the (CV+Network) model of the Table S2. In next two models (b) and (c), the PageRank Centrality ($\mathscr{C}^{PR}$) has been replaced respectively by the Betweenness Centrality ($\mathscr{C}^{B}$) and the Degree Centrality ($\mathscr{C}^{D}$), those Betweenness and Degree Centrality are two alternative centrality along with the PageRank Centrality. In the model (d), the variables (N^{NSF} & N^{NIH}) related to the number of grants have been removed. In the model (e), the department rank variable (β_r) has been removed.

Observations, conclusions, and hypotheses:

In this Table S3, the robustness of the cross-sectional model has been tested by exploring several variations, through model (a) to (e).

- Model (a,b,c): The Betweenness and Degree Centrality are two alternative centrality along with the PageRank Centrality. Replacing the PageRank Centrality with those two centralities respectively doesn't indicate any significant differences.
- Model (d): This model suspects the correlation effects, and so consider only the effect of total funding with removing the effect of number of grants. The result also isn't significantly different.
- Model (e): This model assumes that the most recent university affiliation could inaccurately represent the career of faculties. The estimates after removing the department rank variable doesn't indicate any significant differences too.

Therefore, in all cases, the results of the modified regression estimates are not significantly different, that indicates the robustness of the cross-sectional model.