ECONOMETRIC ANALYSIS OF NETWORK DATA

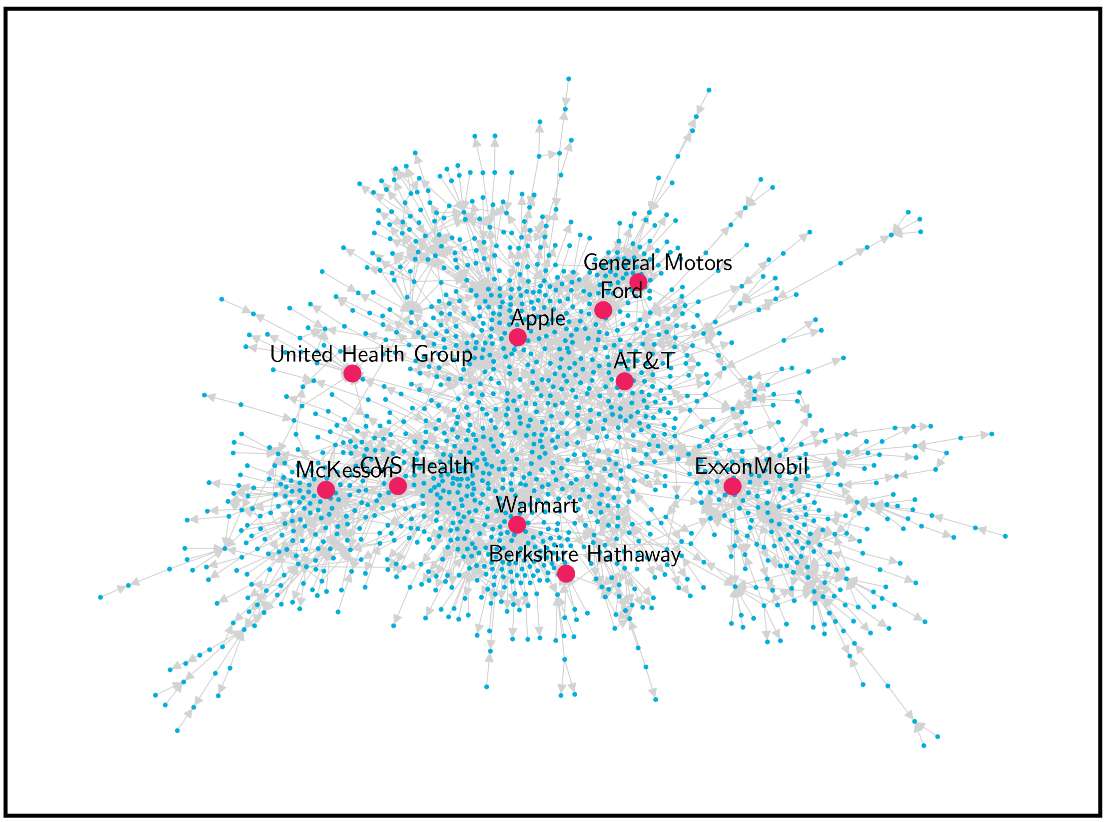
Econometrics Summer Masterclass and Workshop

University of Warwick

May, 2018

Course Description

This course will provide a short overview of econometric methods appropriate for the analysis of social and economic networks. Many social and economic activities are embedded in networks. Furthermore, datasets with natural graph theoretic (i.e., network) structure are increasingly available to researchers. We will review (i) how to describe, summarize and visually present network data and (ii) formal econometric models of network formation that admit heterogeneity, strategic behavior, and/or dynamics. The focus will be on the formal development of methods, but selected empirical examples will also be covered, as will methods of practical computation.



Course Logistics

Instructor: Bryan Graham, Professor of Economics, University of California – Berkeley

Email: `bgraham@econ.berkeley.edu`

Time: May 31st & June 1st, Four sessions of two hours each.

Prerequisites: The equivalent of a first year Ph.D. level sequence in econometrics. Specifically an understanding of probability and statistical inference at the level of Casella and Berger (1990, *Statistical Inference*), linear regression analysis at the level of Goldberger (1991, *A Course in Econometrics*) and some exposure to non-linear models (e.g., maximum likelihood, M-estimation). I will also assume a basic knowledge of applied matrix algebra. Familiarity with the theory of U-Statistics is helpful.

Textbook: Readings preceded by a [r] in the course outline are "required" (i.e., should ideally be read prior to class), while those preceded by a [b] are for "background" (i.e., may be useful for students interested in additional material). Two introductory books, written from different disciplinary perspectives, are Jackson (2006) and Newman (2010). The survey by Goldenberg et al. (2009) covers much of the technical literature in statistics and machine learning, but is now somewhat dated.

Computation: The bulk of class will be devoted to the formal development of the material, albeit with empirical illustrations as well as ample discussions of the various practicalities of implementation. However some class time may also be spent actual computation. Computational examples will be done using Python. Python is a widely used general purpose programming language with good functionality for scientific computing. I recommend the Anaconda distribution, which is available for download at http://continuum.io/downloads. Some basic tutorials on installing and using Python, with a focus on economic applications, can be found online at `http://quant-econ.net`. Good books for learning Python, with some coverage of statistical applications, are Guttag (2013), McKinney (2017) and VanderPlas (2017). The first is an excellent introduction to computer science as well as Python, the last two are cookbook type references with lots of examples. The code I will provide will execute properly in Python 3.6. This is the latest Python release. Graphviz is a free graph visualization program that is also useful (`http://www.graphviz.org/`).

Course Outline

Date	Topic	Readings
Lec 1 - 5/31	**Describing**	[r] de Paula (2017); Graham (2015); Jackson (2008, Ch. 2)
(2 hours)	**Networks**	[b] Goldenberg et al. (2009)
	Examples of networks	[b] Atalay et al. (2011); Mizuno et al. (2014)
		[b] Apicella et al. (2012); Acemoglu et al. (2012)
	Small worlds	[b] Milgram (1967)
	Degree distributions	[b] Mitzenmacher (2004)
	Homophily	[b] McPherson et al. (2001)
	Triads	[b] Granovetter (1973); Jackson et al. (2012)
Lec 2 - 5/31	**Nonparametrics:**	[r] Holland & Leinhardt (1976)
(2 hours)	**Network Moments**	
	Theory	[r] Diaconis & Janson (2008); Bickel et al. (2011)
	Computation	[r] Bhattacharya & Bickel (2015)
		[r] Blitzstein & Diaconis (2011)
Lec 3 - 6/1	**Dyadic Regression**	[r] Fafchamps & Gubert (2007); Aronow et al. (2017)
(2 hours)		[r] Chatterjee et al. (2011); Graham (2017)
Lec 4 - 6/1	**Dynamic &:**	[r] Graham (2016)
	Strategic Models	[r] Miyauchi (2016); Leung (2015)
(2 hours)		[b] Christakis et al. (2010); Mele (2017)

References

Acemoglu, D., Carvalho, V., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5), 1977 – 2016.

Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, 481(7382), 497 – 501.

Aronow, P. M., Samii, C., & Assenova, V. A. (2017). Cluster-robust variance estimation for dyadic data. *Political Analysis*, 23(4), 564 – 577.

Atalay, E., Hortaçsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences*, 108(13), 5199 – 5202.

Bhattacharya, S. & Bickel, P. J. (2015). Subsampling bootstrap of count features of networks. *Annals of Statistics*, 43(6), 2384 – 2411.

Bickel, P. J., Chen, A., & Levina, E. (2011). The method of moments and degree distributions for network models. *Annals of Statistics*, 39(5), 2280 – 2301.

Blitzstein, J. & Diaconis, P. (2011). A sequential importance sampling algorithm for generating random graphs with prescribed degrees. *Internet Mathematics*, 6(4), 489 – 522.

Chatterjee, S., Diaconis, P., & Sly, A. (2011). Random graphs with a given degree sequence. *Annals of Applied Probability*, 21(4), 1400 – 1435.

Christakis, N. A., Fowler, J. H., Imbens, G. W., & Kalyanaraman, K. (2010). *An empirical model for strategic network formation*. NBER Working Paper 16039, National Bureau of Economic Research.

de Paula, Á. (2017). *Advances in Economics and Econometrics, Eleventh World Congress*, volume 1, chapter Econometrics of network models, (pp. 268 – 323). Cambridge University Press: Cambridge.

Diaconis, P. & Janson, S. (2008). Graph limits and exchangeable random graphs. *Rendiconti di Matematica*, 28(1), 33 – 61.

Fafchamps, M. & Gubert, F. (2007). The formation of risk sharing networks. *Journal of Development Economics*, 83(2), 326 – 350.

Goldenberg, A., Zheng, A., Fienberg, S. E., & Airoldi, E. M. (2009). A survey of statistical network models. *Foundations and Trends in Machine Learning*, 2(2), 129–333.

Graham, B. S. (2015). Methods of identification in social networks. *Annual Review of Economics*, 7(1), 465–485.

Graham, B. S. (2016). *Homophily and transitivity in dynamic network formation.* NBER Working Paper 22186, National Bureau of Economic Research.

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica*, 85(4), 1033 – 1063.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360 – 1380.

Guttag, J. V. (2013). *Introduction to Computation and Programming Using Python.* Cambridge, MA: The MIT Press.

Holland, P. W. & Leinhardt, S. (1976). Local structure in social networks. *Sociological Methodology*, 7, 1 – 45.

Jackson, M. (2006). The economics of social networks. In R. Blundell, W. Newey, & T. Persson (Eds.), *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society.* Cambridge: Cambridge University Press.

Jackson, M. O. (2008). *Social and Economic Networks.* Princeton: Princeton University Press.

Jackson, M. O., Rodriguez-Barraquer, T., & Tan, X. (2012). Social capital and social quilts: network patterns of favor exchange. *American Economic Review*, 102(5), 1857–1897.

Leung, M. (2015). Two-step estimation of network-formation models with incomplete information. *Journal of Econometrics*, 188(1), 182 – 195.

McKinney, W. (2017). *Python for Data Analysis.* Cambridge: O'Reilly.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: homophily in social networks. *Annual Review of Sociology*, 27(1), 415 – 444.

Mele, A. (2017). A structural model of dense network formation. *Econometrica*, 85(3), 825 – 850. John Hopkins University.

Milgram, S. (1967). The small-world problem. *Psychology Today*, 1(1), 61 – 67.

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1(2), 226 – 251.

Miyauchi, Y. (2016). Structural estimation of a pairwise stable network with nonnegative externality. *Journal of Econometrics*, 195(2), 224 – 235.

Mizuno, T., Souma, W., & Watanabe, T. (2014). The structure and evolution of buyer-supplier networks. *Plos One*, 9(7), e100712.

Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford: Oxford University Press.

VanderPlas, J. (2017). *Python Data Science Handbook*. Boston: O'Reilly.