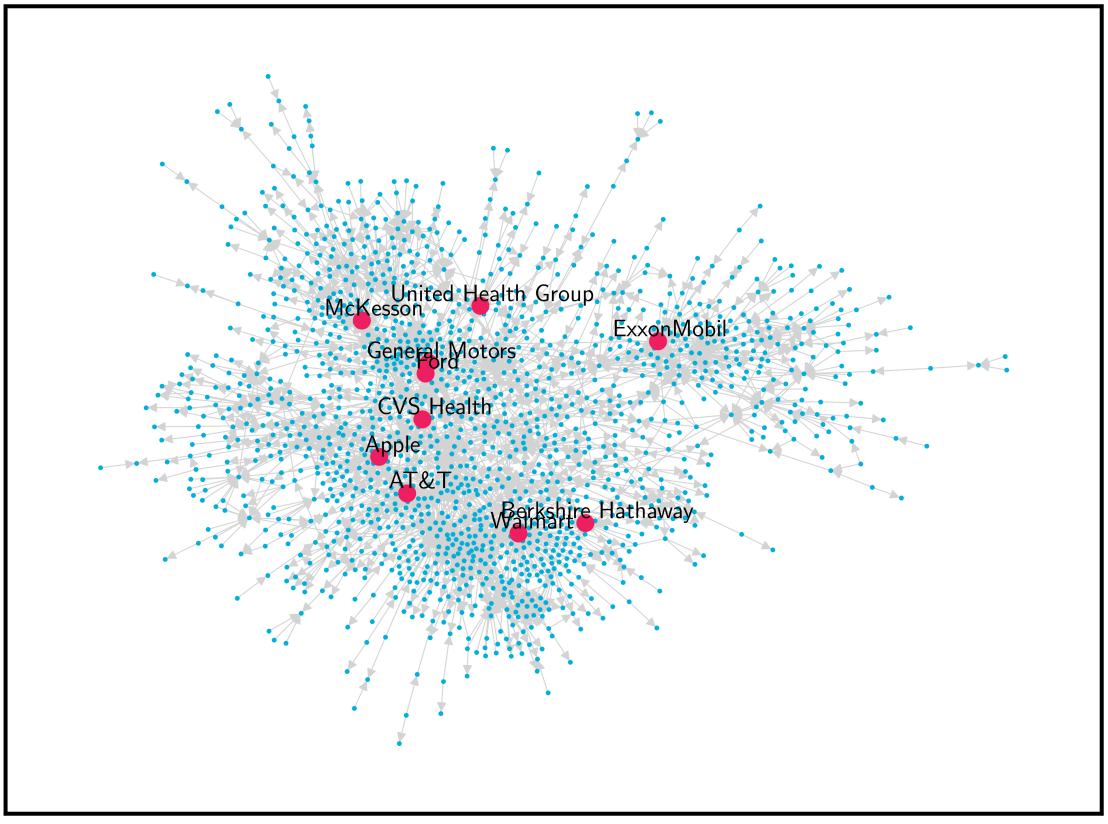
Econometric Methods for Networks

Econometrics Camp at CEMP/IESR Jinan University December 9-12

Course Description

This course will provide an 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, including ones that admit heterogeneity and/or strategic behavior. Special emphasis will be placed on parametric and non-parametric methods appropriate for dyadic analysis – as arises in, for example, the analysis of international trade or migration flows.



Course Logistics

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

Email: bgraham@econ.berkeley.edu

Time: December 11th & 12th, 8:30 to 10AM & 10:15 to 11:45AM.

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, GMM). I will also assume a basic knowledge of applied linear/matrix algebra.

Textbook: My draft *Handbook of Econometrics* chapter contains material on many of the topics we will cover (Graham, TBD). 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 or empirical applications). Students who anticipate doing research involving networks may consider purchasing the textbooks by Jackson (2006) and Newman (2010), but doing so is not necessary. The survey by Goldenberg et al. (2009) covers much of the technical literature in statistics and machine learning, but is now somewhat dated. All papers authored by me may be found online at http://bryangraham.github.io/econometrics/research/.

GitHub: Supplemental course materials, including slides, lecture nodes and computer programs, will be made available on GitHub at https://github.com/bryangraham/short_courses.

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 I do intend to reserve some class time for actual practice with computation. Computational examples will be done using Python. Python is a widely used general purpose programming language with good functionality for scientific computing. For those wishing to manage a Python environment on their personal computer, the Anaconda distribution, which is available for download at https://www.anaconda.com/distribution/, is a convenient way to get started. 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), VanderPlas (2017), and McKinney (2017). The code I will provide will execute properly in Python 3.6, which is (close to) the latest Python release. Graphviz is a free graph visualization program that is also useful (http://www.graphviz.org/).

Course Outline

Date	Topic	Readings
Topic 1	**Describing Networks**	[r] Jackson et al. (2017)
	Examples of networks	[b] Atalay et al. (2011); Mizuno et al. (2014)
		[b] Apicella et al. (2012); Glitz (2017)
	Small worlds	[b] Milgram (1967)
	Degree distributions	[b] Mitzenmacher (2004)
	Homophily	[b] McPherson et al. (2001)
	Triads	[b] Granovetter (1973); Jackson et al. (2012)
		[b] Holland & Leinhardt (1976)
Topic 2	**Dyadic Analysis & Regression**	[r] Fafchamps & Gubert (2007); Aronow et al. (2017)
		[r] Menzel (2017)
		[r] Graham et al. (2019)
		[b]Santos Silva & Tenreyro (2006); König et al. (2019)
		[b]Anderson (2011)
Topic 3	**Heterogeneity**	[r] Chatterjee et al. (2011)
		[r] Graham (2017)
		[b] Dzemski (2018); Jochmans (2018)
Topic 4	**Strategic Interaction: Testing & Estimation**	[r] Pelican & Graham (2019); Graham & Pelican (2020)
		[r] Blitzstein & Diaconis (2011)
		[r] McDonald et al. (2007)
		[r] Miyauchi (2016)

References

Anderson, J. E. (2011). The gravity model. *Annual Review in Economics*, 3(1), 133 – 160.

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.

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.

Dzemski, A. (2018). An empirical model of dyadic link formation in a network with unobserved heterogeneity. *Review of Economics and Statistics*. University of Mannheim.

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

Glitz, A. (2017). Coworker networks in the labour marke. *Labour Economics*, 44(1), 218 – 230.

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

Graham, B. S. (TBD). *Handbook of Econometrics*, volume 7, chapter Network data. North-Holland: Amsterdam.

Graham, B. S., Niu, F., & Powell, J. L. (2019). *Kernel density estimation for undirected dyadic data*. Technical report, University of California - Berkeley.

Graham, B. S. & Pelican, A. (2020). *The Econometrics of Social and Economic Networks*, chapter Testing for externalities in network formation using simulation",. Elsevier: Amsterdam.

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

Jackson, M. O., Rogers, B. W., & Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, 55(1), 49 – 95.

Jochmans, K. (2018). Semiparametric analysis of network formation. *Journal of Business and Economic Statistics*, 36(4), 705 – 713.

König, M. D., Liu, X., & Zenou, Y. (2019). R&d networks: theory, empirics and policy implications. *Review of Economics and Statistics*, 101(3), 476 – 491.

McDonald, J. W., Smith, P. W. F., & Forster, J. J. (2007). Markov chain monte carlo exact inference for social networks. *Social Networks*, 29(1), 127 – 136.

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.

Menzel, K. (2017). *Bootstrap with clustering in two or more dimensions*. Technical Report 1703.03043v2, arXiv.

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.

Pelican, A. & Graham, B. S. (2019). *Testing for strategic interaction in social and economic network formation*. Technical report, University of California - Berkeley.

Santos Silva, J. & Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88(4), 641 – 658.

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