VRIJE UNIVERSITEIT AMSTERDAM

Statistical and Econometric Analysis of Network Data

SYLLABUS

VU Amsterdam Summer School

July 2024

VRIJE
UNIVERSITEIT
AMSTERDAM

Any general questions for the Summer School support team? Contact amsterdamsummerschool@vu.nl.

Statistical and Econometric Analysis of Network Data

Contact Info
Prof. Michael D. König
Department of Spatial Economics, School of Business and Economics, Vrije Universiteit Amsterdam
Email: m.d.konig@vu.nl
URL: https://research.vu.nl/en/persons/michael-konig

Course Description
Networks play an increasingly dominant role in many social, business, and economic environments. Moreover, network data becomes increasingly important and available due to the rise of online social media and digitization. This course offers a concise introduction into the most recent econometric methods developed for processing, visualizing and learning from network data. The course will combine lectures with hands-on empirical and programming exercises.

Learning Objectives
Upon successful completion of the course, students will:

- become acquainted with different statistical methodologies for analyzing networks while learning how to see these different methodologies complementing each other.
- learn to model network problem situations mathematically, and adapt the methods learned to new situations at hand.
- be able to recognize, understand, and analyze societal and business problems in which networks are central.
- learn how networks affect demand and supply in markets, how this leads to market failures, and how government policies can address these.

Reading List
All relevant material will be covered in the lecture slides. The slides will be made available to the students on the course website before the start of the course. The following literature is complementary to the course slides and covers some additional relevant material for further reading:

- Graham, Bryan, and De Paula, Aureo. *Econometric Analysis of Network Data*. Elsevier, 2020.
- Kolaczyk, Eric, *Statistical Analysis of Network Data: Methods and Models*, Springer, 2009.
- LeSage, James, and Robert Kelley Pace. *Introduction to Spatial Econometrics*. Chapman and Hall/CRC, 2009.
- Bramoulle, Yann, Andrea Galeotti, and Brian Rogers. *The Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016.
- Jackson, Matthew. O. *Social and Economic Networks*. Princeton University Press, 2010.
- Carrington, Peter, John Scott, and Stanley Wasserman. *Models and Methods in Social Network Analysis*, Cambridge University Press, 2005.

Tutorials and Exercises
The course consists of five days in which you will practice the material of the lectures in the first three days with tutorials in the last two days using exercises that will be implemented in Matlab. During the tutorials, you are required to complete an exercise which will build on the skills you have obtained in the lectures. You can complete the exercises in groups of two or three students (in work groups or remotely) but are required to submit the solutions individually by the 18th of August 2024. Exercise solutions and assignments should be submitted via email to: m.d.konig@vu.nl. If you successfully completed the exercises you obtain a certificate for this course.

Grading
Tutorial and programming exercises: 100%

Daily Course Schedule

	Monday	Tuesday	Wednesday	Thursday	Friday
10:00-11:00	**Lecture 1, Part** 1: Networks – Introduction and Examples	**Lecture 3**: Econometrics of Network Formation	**Lecture 6:** Network Panel Estimation	**Tutorial 1:** Matlab Intro, Vis. Netw. & Fitting Deg. Distr.	**Tutorial 3:** DMH Algorithm
11:00-11:30	Break				
11:30-12:30	**Lecture 1, Part 2**: Basic Definitions and Characterizations	**Lecture 4**: Coevolution of Networks and Behavior	**Lecture 7:** Big Data Meets Networks	**Tutorial 2:** SAR Model and Logistic Regression	**Student Presentations**
12:30-14:00	Lunch Break				
14:00-15:00	**Lecture 2:** Econometrics of Interactions in Networks (finish at 14:45)	**Lecture 5:** Network Formation with Multiple Activities	Social Programme	**Work groups**	**Work groups** (finish at 14:45)

Detailed Schedule

Day 1

Lecture 1: Networks: Basic Definitions and Characterizations

1. Examples of Networks and Data

2. Network Statistics, Visualization and Graphs

- Elements of Graph Theory
- Graphs and Matrices
- Bipartite Graphs
- Core-periphery Networks and Nested Split Graphs
- Network Statistics: Average path length, clustering and assortativity
- Centrality in Networks: Degree, eigenvector, Katz-Bonacich centrality and Google's Page Rank
- Network Visualization: Force-directed, circular and layered layout

Lecture 2: Econometrics of Interactions in Networks

3. Econometrics of Interactions in Networks

- Spatial Autoregressive (SAR) Model
- Linear Quadratic Utility
- Endogeneity of the Spatial Lag
- Two-Stage Least Squares (2SLS)
- Maximum Likelihood Estimation (MLE)
- Identification Issues

– Correlated Effects, Sorting and Selection
– Endogenous Link Formation
• Multiple Spatial Weight Matrices
• Spatial Panel Data

Day 2

Lecture 3: Econometrics of Network Formation

4. Econometrics of Network Formation

• Exponential Random Graph Model (ERGM)
• Conditional Edge-Independence
– Erdös-Rényi Random Graph
– Logistic Regression
– Unobservable Characteristics (beta-model)
– Tetrad Logit Estimator
• Random Utility Model
• Maximum Likelihood Estimation (MLE)
• Markov Chain Monte Carlo
– Gibbs Sampling
– Metropolis Hastings Algorithm
• Stochastic Block Model (SBM)
• Temporal ERGM

Lecture 4: Coevolution of Networks and Behavior

5. Joint Estimation of Outcomes and Network Formation

5.1. Coevolution of Networks and Behavior: An application to R&D collaboration networks

• Structural Model: Utility and the potential game
• Estimation
– Computational Problem and the Exchange Algorithm
– Double Metropolis-Hastings (DMH) Algorithm
– Unobserved Heterogeneity
• Empirical Illustration: R&D collaborations

Lecture 5: Network Formation with Multiple Activities

5.2*. Network Formation with Multiple Activities: An application to team production and co-authorship networks*

• Bipartite Network, Production Function, and Utility
• Equilibrium Characterization and Line Graphs
• Estimation with Endogenous Matching
• Empirical Illustration: Co-authorship networks

Day 3

Lecture 6: Network Panel Estimation

6. Spatial Modeling Approach for Dynamic Network Formation and Interactions

- Spatial Dynamic Panel Data (SDPD) Model
- A General Dynamic Network Formation Model
- Combining SDPD with the Network Formation Model: Joint likelihood function
- An Empirical Application to Peer Effects in Academic Performance

Lecture 7: Big Data meets Networks

7. Big Data meets Networks

- The Digital Layer: How innovative firms relate on the Web
- Automated Robot for Generic Universal Scraping (ARGUS)
- Input, Interface and Output of ARGUS
- Sectoral Hyperlink Network
- Hyperlink Types

8. Riot networks and the Tullock Paradox: An application to the Egyptian Arab Spring (if time permits)

- Benchmark Model and the Law of Motion
- Equilibrium Characterization (Stochastically Stable States)
- The Belief-based Model
- Comparing Equilibria and Transition Times
- Empirical Application: The fall of Morsi on Twitter
- Text-as-Data Approach
- Structural Estimation
- Counterfactuals

Day 4

Tutorial 1:
- Introduction to Matlab
- Visualizing Networks
- Fitting Degree Distributions

Tutorial 2:
- Spatial Autoregressive (SAR) Model
- Logistic Regression

Work groups

Day 5

Tutorial 3:
• Double Metropolis-Hastings (DMH) Algorithm
• Application to R&D Networks

Student presentations
• 5 minutes presentations of a topic on networks

Work groups