UNIVERSITY OF MARYLAND
Department of Economics

Part A:	**Guido Kuersteiner**	**Econ 722**
Part B:	**Ingmar R. Prucha**	Spring 2024

Lecture: Part A, Tu 5:00-7:30pm LEF 1221 (if the course needs to be moved online Zoom links will be provided).
Part B, Tu 5:00-7:30pm LEF 1221 (if the course needs to be moved online Zoom links will be provided).

If needed, some lectures may be shifted to Th 5:00-7:30pm

Kuersteiner Office: Tydings Hall 3145
gkuerste@umd.edu
http://econweb.umd.edu/~kuersteiner/
Office Hours: Tu 11:30am-12:30pm (and by appointment)

Prucha Office: Tydings Hall 3147A
prucha@umd.edu
http://econweb.umd.edu/~prucha/
Office Hours: Tu 2-4pm (and by appointment)

Course Website
Part A: All course related materials are posted on ELMS. Grades for the midterm exam are posted on ELMS. The graded midterm exam can be viewed upon sending an email request to Prof. Kuersteiner.
Part B: Most course related material will be made available on Prof. Prucha's teaching web site http://econweb.umd.edu/~prucha/econometricsIV_Part2.html.
Please send an email to Prof. Prucha for login/password information.

Course Communication
Part A: You should check the course web-site regularly for new material. You will receive information about problem sets and materials posted on the web-site from a Google group email list.
Part B: Email will be used as needed for course communication.

Emergency Protocol
In case the university closes for weather related or other emergencies, lectures will be held virtually by Zoom if feasible. Exams will also be conducted as remote Zoom session if physical locations on campus are unavailable during the announced exam times.

ECONOMETRICS IV
Microeconometrics

COURSE OVERVIEW

Part A of the course will cover the following topics
- Binary Response Models
- Multinomial Response Models
- Series Estimation
- Machine Learning
- Semiparametric Inference using Series and ML

Part B of the course will cover the following topics
- Quantile Regression
- Non-parametric and Semi-parametric Estimation Methods
- Spatial and Social Interaction Network Models
- Dynamic Panel Data Models
- Weak Instruments
- Cluster and Stratified Sampling (if not covered elsewhere)
- Boot strap and Jack Knife methods (if time permits)

COURSE AIMS

The course is oriented to provide students with a rigorous and broad knowledge of econometric methods especially important for conducting empirical research in micro-economics. The course is not geared towards training econometric theorists, although this course would be necessary training for such a specialization. In particular, the aim of the course is to provide students with the necessary tools to (i) read intelligently all empirical research (with a proper understanding of the underlying methodology of inference), and (ii) to conduct empirical research suitable for publication in **any** economics or econometrics journal. The course builds on Econometrics I and II, and complements Empirical Microeconomics.

ASSUMED REQUIREMENTS

Students are assumed to have knowledge of the material covered in Econometrics I and II.

PRINCIPAL TEXTS

Cameron, A.C., and P.K. Trivedi, Microeconometrics, Methods and Applications, Cambridge, 2005.
Wooldridge, J.M., Econometric Analysis of Cross Section and Panel Data, MIT Press, 2010.

SUPPLEMENTARY TEXTS

Arellano, M. Panel Data Econometrics, Oxford University Press, 2003.
Baltagi, B.H., Econometric Analysis of Panel Data, Wiley, Fourth Edition, 2013.
Li, Q., and J.S. Racine, Nonparametric Econometrics, Theory and Practice, Princeton University Press, 2007.
Pagan, A., and A. Ullah, Nonparametric Econometrics, Cambridge University Press, 1999.

PROBLEM SETS

Part A of the course includes a set of optional problem sets that offer training in the computational implementation of the methods discussed in the course. 'Optional' means that problem sets are not graded, do not have to be handed in or done at all and that solutions may not always be provided.

For some problem sets there are solutions in the form of Python code. Some resources for Python are listed below. Python is available as an open source, freely downloadable software development environment and programming language. There are over 10,000 add on packages available for the language, including packages with machine learning subroutines such as scikit-learn, PyTorch, TensorFlow (Google) and Apache Spark. Spark contains an SQL implementation for handling and analyzing large distributed data sets. In addition, Stata 17 contains API functions to run Stata commands and access Stata data directly from within Python code. Anaconda is an integrator software that integrates a large number of the Python packages and handles installation and updating. The latest currently available language version for Python is 3.12. However, my preferred IDE Spyder currently only supports Python 3.11. If you are using Spyder, you need to make sure to read the corresponding tutorials and reference texts for the correct Python version. The problem sets also contain step by step installation instructions for a Windows Subsystem Linux installation and some basic exposure to the Linux/Unix operating system environment.

General:

https://www.python.org/ for software downloads, tutorials and language reference. Python 3.10 and Python 3.11 distributions are available from this web-site but using them may require more effort installing a workable software distribution.

Integrated Development Environments (IDE):

Note that the Anaconda distribution contains the Spyder IDE as well as scikit-learn. The package PyTorch can be installed through the Anaconda.Navigator interface which comes with the Anaconda distribution.

https://www.anaconda.com/
https://www.spyder-ide.org/

Machine Learning Packages:

https://scikit-learn.org/stable/
https://pytorch.org/
https://spark.apache.org/

Language Tutorials:

https://docs.python.org/3.9/tutorial/index.html

Books:
Beazly, D. M. , Python. Essential Reference. Fourth Edition. Addison Wesley, 2009.
Raschka, S, Liu, Y. and Mirjalili, V., Machine Learning with PyTorch and Scikit-Learn, Packt, 2022
Ramalho, L., Fluent Python, 2nd Edition. O'Reilly, 2022

Part B of the course will likely include optional problem sets intended for a deeper understanding of the course material. 'Optional' means that problem sets are not graded, do not have be handed in or done at all and that solutions may not always be provided.

GRADING POLICY

The final grade in Econ 722 will be based on the performance in Part A and B of the course, each component getting equal weights:

Part A:	Exam	50.0%*
Part B:	Exam	50.0%*

* No makeup exams will be given except in cases of illness (confirmed by a doctor's certificate), religious observance, participation in University activities at the request of the University authorities, or compelling circumstances beyond the student's control. If at all possible, the student must inform me (or the Economics Department) of her/his situation before the exam.

In case the University is closed during (part of) the official scheduled time period for the final exam, the exam will be rescheduled according to the instructions that will be given by the University in that eventuality.

MIDTERM EXAM: Thursday, March 28, 2024, take home exam.
FINAL EXAM: Thursday, May 16, 2024, 4-6pm, or alternatively take-home exam. More details will be provided at the beginning of part B of the course.

UNIVERSITY AND GRADUATE SCHOOL RULES AND REGULATIONS:

University policies can be found here: https://policies.umd.edu/
In particular:

- https://policies.umd.edu/general-administration/university-of-maryland-disability-accessibility-policy-and-procedures
- https://policies.umd.edu/general-administration/university-of-maryland-policy-and-procedures-on-sexual-harassment-and-other-sexual-misconduct
- https://policies.umd.edu/student-affairs/university-of-maryland-policy-on-excused-absence
- https://policies.umd.edu/research/university-of-maryland-intellectual-property-policy

All graduate school policies can be found here: https://gradschool.umd.edu/course-related-policies

In particular note the following items:

Academic Integrity

The student-administered University Honor Code and Honor Pledge (shc.umd.edu/code.html) prohibits students from cheating on exams, plagiarizing papers, submitting the same paper for credit in two courses without authorization, buying papers, submitting fraudulent documents and forging signatures.

Compliance with the code is administered by the Student Honor Council, which strives to promote a community of trust on the College Park campus.

University policy of the Code of Academic Integrity, including procedures that handle violations can be found here: https://president.umd.edu/administration/policies/section-iii-academic-affairs/iii-100a

COPYRIGHT PROTECTION FOR CLASS MATERIALS

The lecture class and all other course materials that exist in a tangible medium, such as written or recorded lectures, Power Point presentations, handouts and tests, problem sets and solutions, are copyright protected. Students may not copy and distribute such materials except for personal use and with the instructor's permission. Obtaining and using such materials from courses taught in previous years without the instructor's explicit permission constitutes a copy right breach. In addition, unauthorized use of video or audio recordings may be in violation of state and federal law.

ATTENDANCE

By signing up for this class you agree to exam formats, course requirements and timing of exams and due dates of work to be handed in. Attendance in all lectures is expected except when excused for health or other reasons permitted by university policies. Absences need to be reported by email to the instructor at least one hour before class.

HEALTH AND MASK MANDATES

Covid related policies: https://umd.edu/4Maryland

STUDENTS WITH DISABILITIES

UMD guarantees appropriate accommodations for students with disabilities. If you require accommodations, please contact me as soon as possible. If you need further clarification, the link to ADS is: https://www.counseling.umd.edu/ads/

COURSE EVALUATIONS

Students are encouraged to submit course evaluations through CourseEvalUM (www.courseevalum.umd.edu).

READING LIST FOR PART A

Binary Choice Models
*Wooldridge Chapter 15
Cameron and Trivedi Chapter 14

Rivers, D. and Q.H. Voung (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models," Journal of Econometrics, 39, 347-366.

Multinomial Response Models
*Wooldridge Chapter 16
Cameron and Trivedi Chapter 15

Dempster, A.P., N.M. Laird and D.B. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," JOuranl of the Royal Statistical Society, Series B, 39, 1-38

McFadden, D.L. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior", in Frontiers of Econometrics, ed. P Arembka. New York: Academic Press, 105-142.

McFadden, D.L. (1978), "Modeling the Choice of Residential Location," in Spatial Interaction Theory and Residential Location, ed. A Karlqvist, Amseterdam: North Holland, 75-96.

Hajivassiliou, V.A. and D. L. McFadden (1998): “The Method of Simulated Scores for the Estimation of LDV Models,” Econometrica, 66, 863-896.

Hajivassiliou, V.A and P. Ruud (1994), “Classical Estimation methods for LDV Models using Simulation,” in Handbook of Econometrics, Vol IV, ed. By R.F. Engle and D.L. McFadden, Elsevier Science, Ch. 40, 2383-2441.

Lerman, S. and C. Manski (1981), “"On the Use of Simulated Frequencies to Approximate Choice Probabilities," in Structural Analysis of Discrete Data with Econometric Applications, ed. by C. Manski and D. McFadden. Cambridge, MA: MIT Press, 305-319.

Tierney, L. (1994), “Markov Chains for Exploring Posterior Distributions,” The Annals of Statistics, 22, 1701-1762.

Geman, S. and D. Geman (1984), “Stochastic Relaxation, Gibbs distributions and the Bayesian Restoration of Images. IEEE Transactions of Pattern Analysis and Machine Intelligence, 6, 721-741.

McFadden, D.L. and K. Train (2000), “Mixed MNL Model for Discrete Response,” Journal of Applied Econometrics, 15, 447-470.

Train, K.E. (2009), “Discrete Choice Methods with Simulation,” New York: Cambridge University Press. Chapter 14.

Wu, J. (1983), “On the Convergence Properties of the EM Algorithm,” The Annals of Statistics, 11, 95-103.

Series and Sieve Estimation, Semiparametric Inference, Nonlinear Instrumental Variables

Ackerberg, D. X. Chen and J. Hahn (2012) “A practical asymptotic variance estimator for two step semiparametric estimators,” *The Review of Economics and Statistics,* 481-498.

Andrews, D.W.K. (1994) “Asymptotics for Semiparametric Econometric Models via Stochastic Equicontinity”, *Econometrica,* 43-72.

Newey, W.K. (1994): “The Asymptotic Variance of Semiparametric Estimators,” *Econometrica,* Vol. 62, No. 6 (Nov., 1994), pp. 1349-1382.

Newey, W.K. (1997): “Convergence rates and asymptotic normality for series estimators,” Journal of Econometrics 79 (1997) 147-168.

Powell, M.J.D. (1981): Approximation theory and methods, Cambridge University Press, New York.

Machine Learning Algorithms

Athey, S. and G. Imbens (2019), "Machine Learning Methods Economists Should Know About," manuscript.

Buhlmann, P. and van de Geer, S. Statistics for High-Dimensional Data, Springer 2011

Efron, B. and Hastie, T., Computer Age Statistical Inference, Cambridge, 2016 Computer Age Statistical Inference, Cambridge, 2016

Goodfellow, I., Bengio, Y. and Courville, A. Deep Learning, MIT Press, 2016

Hansen, B. E (2007): Least Squares Model Averaging, *Econometrica,* Vol 75, 1175-1189.

Hastie, T., R. Tibshirani and J. Friedman (2008): "The elements of Statistical Learning, 2nd Edition", Springer Verlag

Hastie, T., R. Tibshirani and J. Wainwright (2015): "Statistical Learning with Sparsity. The Lasso and Generalizations", Chaphman & Hall/CRC Press.

Poetscher, B.M (2009): Confidence Sets Based on Sparse Estimators Are Necessarily Large, Sankhyā: The Indian Journal of Statistics, Series A Vol. 71, No. 1, pp. 1-18

Poetscher, B.M and U. Schneider (2009): On the distribution of the adaptive LASSO estimator, Journal of Statistical Planning and Inference, 2775-2790.

Poetscher, B.M and U. Schneider (2009): Confidence sets based on penalized maximum likelihood estimators in Gaussian regression, Electronic Journal of Statistics, 334–360.

Amann, N. and U. Schneider (2021): Uniform Asymptotics and Confidence Regions Based on the Adaptive Lasso With Partially Consistent Tuning,

Data Science and Applications

Blei, D.M. and P. Smyth (2017): Science and Data Science, Proceedings of the National Academy of Sciences, vol 114, 8689-8692.

Murdoch, W.J., C. Singh, K. Kumbier, R. Abbasi-Asl and B. Yu (2019): Definitions, methods, and applications in interpretable machine learning.

Yu, B. and K Kumbier (2020): Veridical data science, Proceedings of the National Academy of Sciences, vol 117, 3920-3929.

Socher, R. A. Perelygin, J.Y.Wu, J Chuang, C. Manning, A.Y. Ng, and C. Potts (2013): Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank, Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, 1631–1642.

Pennington, J. R. Socher, C. D. Manning (2014): GloVe: Global Vectors forWord Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP),1532–1543.

Applications of Series Estimators and Machine Learning

Athey, S., J. Tibshirani and S. Wager (2019): Generalized Random Forests, The Annals of Statistics, Vol. 47(2), 1149-1178.

Ai, C. and X. Chen (2003) "Efficient Estimation of Models with Conditional Moment restrictions Containing Unknown Functions," *Econometrica,* Vol. 71, No. 6 (Nov., 2003), pp. 1795-1843.

Antonelli, J., M. Cefalu, N. Palmer, and D. Agniel (2018): "Doubly Robust Matching Estimators for High Dimensional Confounding Adjustment, *Biometrics*, Vol 74, 1171-1179.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012), "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain", *Econometrica,* Vol. 80, 2369–2429.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, and W. Newey (2017), "Double/Debiased/Neyman Machine Learning of Treatment Effects", *American Economic Review: Papers & Proceedings* 2017, 107(5): 261–265

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, J. Robins (2017), "Double/Debiased Machine Learning for Treatment and Causal Parameters", manuscript.

Chernozhukov, V., J.C. Escanciano, H. Ichimura, W.K. Newey and J. Robins (2021): "Locally Robust Semiparametric Estimation," Econometrica, Vol.90(4),1501-1535.

Chernozhukov, Newey, Singh (2022): "Automatic Debiased Machine Learning for Causal and Structural Effects, Econometrica, Vol 90(3), 967-1027.

Geman, S. and C.R. Hwang (1982): "Nonparametric maximum Likelihood Estimation by the Method of Sieves," Annals of Statistics, Vol. 10, pp. 401-414.

Hahn, J. (1998): “On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects,” *Econometrica,* Vol. 66, pp. 315-331.

Hansen, B.B. (2008): “The prognostic analogue of the propensity score,”, *Biometrika*, Vol 95, pp. 481-488.

Hansen, B. (2007): “Least Squares Model Averaging,” *Econometrica,* Vol. 75, pp. 1175-1189.

Heckman, J.J. and E. J. Vytlacil (1999) “Local instrumental variables and latent variable models for identifying and bounding treatment effects”, Proc. Natl. Acad. Sci. USA, Vol. 96, pp. 4730–4734.

Hirano, K., G.W. Imbens and G. Ridder (2003): “Efficient Estimation of Average Treatment Effects,” *Econometrica,* Vol. 71, pp. 1161-1189.

Horvitz, D., and Thompson, D. (1952), “A Generalization of Sampling Without Replacement From a Finite Population,” *Journal of the American Statistical Association*, 47, 663–685

Ichimura, H. and W. Newey (2022), “The Influence Function of Semiparametric Estimators”, *Quantitative Economics,* 13, 29-61.

Kuersteiner, G. and R. Okui (2010): “Constructing Optimal Instruments by First Stage Prediction Averaging,” *Econometrica,* Vol. 78, pp. 697-718.

Newey, W.K. (2013): “Nonparametric Instrumental Variables,” *American Economic Review*, May 2013 Vol. 103, pp. 550-556.

Newey, W.K. and J.L. Powell (2003): “Instrumental Variable Estimation of Nonparametric Models,” *Econometrica,* Vol. 71, pp. 1565-1578.

Wager, S and S. Athey (2018): “Estimation and Inference of Heterogenous Treatment Effects using Random Forests”, Journal of the American Statistical Association, Vol 113, 2018, 1228-1242.

READING LIST FOR PART B

Nonparametric and Semiparametric Estimation

Prucha, I.R., Handout on Nonparametric and Semiparametric Estimation

Below is a list of some texts and review articles. References to research articles are given in the handout.

Cameron, A.C., and P.K. Trivedi, 2005, Microeconometrics, Methods and Applications, Cambridge University Press, Cambridge, Ch. 9.

Fan, J., and I. Gijbels, 1996, Local Polynomial Modeling and Its Applications, Chapman & Hall, New York.

Haerdle, W., 1990, Applied Nonparametric Regression, Cambridge University Press, Cambridge.

Haerdle, W., and O. Linton, 1994, Applied Nonparametric Methods, in E.F. Engle and D.L. McFadden, Handbook of Econometrics, Vol. IV, Elsevier, New York, pp. 2297-2339.

Horowitz, J.L., 1998, Semiparametric Methods in Econometrics, Spinger, New York.

Ichimura, H., and P. E. Todd, 2007, Implementing Nonparametric and Semiparametric Estimators, in J. Heckman and E. Leamer, eds., Handbook of Econometrics, Vol. VI B, Elsevier, New York, pp. 5360-5468.

Li, Q., and J.S. Racine, 2007, Nonparametric Econometrics, Theory and Practice, Princeton University Press, Princeton.

Pagan, A., and A. Ullah, 1999, Nonparametric Econometrics, Cambridge University Press, Cambridge.

Powell, J.L., 1994, Estimation of Semiparametric Models, in E.F. Engle and D.L. McFadden, Handbook of Econometrics, Vol. IV, Elsevier, New York, pp. 2444-2521.

Prakasa Rao, B.L.S., 1983, Nonparametric Functional Estimation, Academic Press, New York.

Silverman, B.W., 1986, Density Estimation for Statistics and Data Analysis, Chapman and Hall, New York.

Wassermann, L., 2006, All of Nonparametric Statistics, Springer, New York.

Yatchew, A., 2003, Semiparametric Regression for the Applied Econometrician, Cambridge University Press, Cambridge.

Spatial/Cross Sectional Interaction Models

Prucha, I.R., Handout on Estimation of Spatial Models

Below is a list of some texts and review articles. References to research articles are given in the handout.

Anselin, L., 1988, Spatial Econometrics: Methods and Models (Kluwer Academic Publishers, Boston).

Anselin L. 2001. Spatial econometrics. In A Companion in Theoretical Econometrics. Baltagi B. (Ed.). Basil Blackwell: New York, NY.

Arbia, G., 2006, Spatial Econometrics, Statistical Foundations and Applications to Regional Convergence (Springer, New York)

Cressie, N., 1993, Statistics of Spatial Data (Wiley, New York).

Cliff, A. and J. Ord., 1973, Spatial Autocorrelation (Pion, London).

Cliff, A. and J. Ord., 1981, Spatial Processes, Models and Applications (Pion, London).

Haining, R., 2003, Spatial Data Analysis, Theory and Practice (Cambridge University Press: Cambridge).

Below are some recent articles that explicitly connect spatial and social interaction models

Kuersteiner, G.M. and I.R. Prucha, 2020, Dynamic spatial panel models: Networks, common shocks, and sequential exogeneity. Econometrica, 88, 2109-2146.

Lee, L.-F., 2007, Identification and Estimation of Econometric Models with Group Interactions, Contectual Factors and Fixed Effects, Journal of Econometrics, 140, 333-374.

Lee, L.-F., X. Liu and X. Lin, 2010, Specification and Estimation of Social Interaction Models with Network Structure, Contextual Factors, Correlation and Fixed Effects, The Econometrics Journal 13, 145-17.

Liu, X., and L.-F. Lee, 2010, GMM Estimation of Social Interaction Models with Centrality", Journal of Econometrics 159, 99-115.

Quantile Regression

Prucha, I.R., Handout on LAD and Quantile Regression

Below is a list of some texts and review articles. References to research articles are given in the handout.

Koenker, R. and K. Hallock, 2001, Quantile Regression, Journal of Economic Perspectives, 15, 143-156.

Cade, B. and B. Noon, 2003, A Gentle Introduction to Quantile Regression for Ecologists, Frontiers in Ecology and the Environment, 1, 412-420.

A more extended treatment of the subject is now also available:

Koenker, R., 2005, Quantile Regression, Econometric Society Monograph Series, Cambridge

University Press. Errata list (http://www.econ.uiuc.edu/~roger/research/rq/errata.pdf)

Weak Instruments

Prucha, I.R., Handout on Weak Instruments

Below is a list of some review articles. References to research articles are given in the handout.

Andrews, DWK., and J.H. Stock , "Inference with Weak Instruments," with, in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. III, ed. by R. Blundell, W.K. Newey and T. Persson. Cambridge, UK: Cambridge University Press, 2007. (http://cowles.econ.yale.edu/P/cd/d15a/d1530.pdf)

Hansen, C., J. Hausman, and W. Newey, Estimation with Many Instrumental Variables, Working Paper, 2006 (http://faculty.chicagobooth.edu/christian.hansen/research/manyiv3jun14.pdf)

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," Journal of Business and Economic Statistics, 20, 518-529.

Dynamic Panel Data Models

Prucha, I.R., Handout on Panel Data Models

Below is a list of some texts and reviews. References to research articles are given in the handout.

Arellano, M., 2003, Panel Data Econometrics, Oxford University Press, Part III.

Arrelano, M., and B. Honore, Panel Data Models: Some recent Development, in in J. Heckman and E. Leamer, eds., Handbook of Econometrics, Vol. V, Elsevier, New York, pp. 3229-3296.

Baltagi, B.H., 2013, Econometric Analysis of Panel Data, Wiley, Ch. 8.

Hsiao, C., 2014, Analysis of Panel Data, Cambridge University Press, Ch.4.

Wooldridge, J.M., 2010, Econometric Analysis of Cross Section and Panel Data, MIT Press, Ch. 10, 11.