ECO 7377 Microeconometrics

Instructor Information

Instructor:	Zhan Gao	**Office:**	ULEE 301U
Email:	zhangao@smu.edu	**Office Hours:**	MW 2:30-3:30pm

Course Details

- **Meeting Days/Times/Location:**

Start Date	End Date	Location	Meeting Days	Time
Jan. 20, 2026	May 5 2026	ULEE301Q	Mon., Wed.	1 - 2:20pm

Grading

Grades will be based on the following components:

- **Presentations (60%):** Each student should pick three papers from the reference list provided at the beginning of the semester and deliver a 20-minute presentation for each paper in class. Slides are required.
- **Referee Report (10%):** Each student should pick one paper from the reference list (different from those selected for presentation) and write a two-page report that critically reviews the paper as an academic journal reviewer would.
- **Research Proposal (30%):** Each student will submit a 10–15 page project proposal related to topics covered in the course. The project can be empirical or theoretical and should be suitable for development as a second-year or third-year paper.

If you wish to present papers outside the reference list, please consult with the instructor in advance.

General Advice

- Python and R are must-have, not optional.
- Build your workflow. Keep iterating. An example.

  - Git(hub)
  - VS Code, Cursor, Anti Gravity...
  - LaTeX: Overleaf, TexSlide, VS code extensions...
  - other useful tools
- Embrace AI: AI-augmented $\succ$ AI-substitute human capital
- Presentation and writing are IMPORTANT
- Subscribe working paper series; pay attention to job market candidates

Course Outline and References

Causal Inference Basics

- Potential Outcome Framework
- Randomized Trials
- Two Chracterizations of ATE
- Doubly Robust and Efficient Estimation

Required Reading

- Wager (2025, Chapter 1–3)
- Chernozhukov et al. (2025, Chapter 2, 5)
- Gaillac and L'Hour (2025, Chapter 3)

Further Reading

- Bai et al. (2025)

Supervised Learning

- Regularization: Ridge, Lasso, Dantzig Selector, nonconvex penalties
- Nerest neighbors
- Boosting

- Tree-based methods
- Support Vector Machines (SVM)
- Basics of (Deep) neural networks
- Numerical Optimization

Required Reading

- Athey and Imbens (2019); Mullainathan and Spiess (2017)
- Chernozhukov et al. (2025, Chapter 3, 8)
- Dell (2025)
- Refer to relevant chapters of James et al. (2021) for details

Debiasd Inference

- Lasso Inference
- Inference with High-dimensional IV
- Debiased/Double Machine Learning (DML) with Partially Linear Models
- DML in General Frameworks

Required Reading

- Chernozhukov et al. (2025, Chapter 4, 9)
- Gaillac and L'Hour (2025, Chapter 4, 5, 6, 7)
- Ahrens et al. (2025); Shi et al. (2025); Chernozhukov et al. (2018)

Heterogeneous Treatment Effects

- Outcome Regression and Regularization Bias
- Causal Forest
- DML for Heterogeneous Effects
- Policy Learning / Optimal Treatment Rule

Required Reading

- Chernozhukov et al. (2025, Chapter 14)
- Gaillac and L'Hour (2025, Chapter 8, 9)
- Wager (2025, Chapter 4, 5)

Unsupervised Learning

- Clustering
- Principle Component Analysis (PCA)
- Autoencoder and beyond

Required Reading

- James et al. (2021, Chapter 12)
- Chernozhukov et al. (2025, Chapter 10)

Unstructured Data

- Tokenization
- Topic models
- Word Embeddings
- From RNN, seq2seq, LSTM to Transformer, BERT, GPT
- Image Embeddings and CNN
- Causal Inference with Unstructured Data

Required Reading

- Gaillac and L'Hour (2025, Chapter 12, 13, 14)
- Chernozhukov et al. (2025, Chapter 10)
- Ash and Hansen (2023)
- Gentzkow et al. (2019)

Large Language Models (LLM)

- Pre- and Post Training
- LLM for Economics Research

Required Reading

- Korinek (2023, 2024, 2025)

If time permits, we may cover the following topics:

- Recent development in panel data models
- Beyond two way fixed effect and difference-in-difference
- Quantile Treatment Effects
- Application reinforcement learning in economics
- Weak instruments

References

Ahrens, A., V. Chernozhukov, C. Hansen, D. Kozbur, M. Schaffer, and T. Wiemann (2025). An introduction to double/debiased machine learning. *arXiv preprint arXiv:2504.08324*.

Ash, E. and S. Hansen (2023). Text algorithms in economics. *Annual Review of Economics 15*, 659–688.

Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economics 11*(1), 685–725.

Bai, Y., A. M. Shaikh, and M. Tabord-Meehan (2025). A primer on the analysis of randomized experiments and a survey of some recent advances. *arXiv preprint arXiv:2405.03910*.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018, 01). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal 21*(1), C1–C68.

Chernozhukov, V., C. Hansen, N. Kallus, M. Spindler, and V. Syrgkanis (2025). Applied causal inference powered by ML and AI. *arXiv preprint arXiv:2403.02467*.

Dell, M. (2025). Deep learning for economists. *Journal of Economic Literature 63*(1), 5–58.

Gaillac, C. and J. L'Hour (2025). *Machine Learning for Econometrics*. Oxford University Press.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–574.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2021). *An introduction to statistical learning: with applications in R*. Springer.

Korinek, A. (2023). Generative ai for economic research: Use cases and implications for economists. *Journal of Economic Literature 61*(4), 1281–1317.

Korinek, A. (2024). Llms learn to collaborate and reason: December 2024 update to "generative ai for economic research: Use cases and implications for economists," published in the journal of economic. *Journal of Economic Literature 61*(4).

Korinek, A. (2025). Ai agents for economic research: August 2025 update to 'generative ai for economic research: Use cases and implications for economists'. *Journal of Economic Literature 61*(4).

Mullainathan, S. and J. Spiess (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

Shi, B., X. Mao, M. Yang, and B. Li (2025). What, why, and how: An empiricist's guide to double/debiased machine learning. *Information Systems Research*.

Wager, S. (2025). *Causal inference: A statistical learning approach*.

Suggested Further Reading

Causal Inference

Bai, Y. (2022). Optimality of matched-pair designs in randomized controlled trials. *American Economic Review 112*(12), 3911–3940.

Breunig, C., R. Liu, and Z. Yu (2025). Double robust bayesian inference on average treatment effects. *Econometrica 93*(2), 539–568.

Bruns-Smith, D. (2025). Two-stage machine learning for nonparametric instrumental variable regression. *Available at SSRN: https://ssrn.com/abstract=5788782*.

Bruns-Smith, D., O. Dukes, A. Feller, and E. L. Ogburn (2025). Augmented balancing weights as linear regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, qkaf019.

Chen, X., P. H. Sant'Anna, and H. Xie (2025). Efficient difference-in-differences and event study estimators. *arXiv preprint arXiv:2506.17729*.

Cheng, Y., J. Wang, X. Cao, Z.-J. M. Shen, and Y. Zhang (2026). A deep-did method to estimate heterogeneous treatment effects: Application to content creator selection. *Marketing Science*.

Demirkaya, E., Y. Fan, L. Gao, J. Lv, P. Vossler, and J. Wang (2024). Optimal nonparametric inference with two-scale distributional nearest neighbors. *Journal of the American Statistical Association 119*(545), 297–307.

Epanomeritakis, A. and D. Viviano (2025). Choosing what to learn: Experimental design when combining experimental with observational evidence. *arXiv preprint arXiv:2510.23434*.

Farrell, M. H. (2015). Robust inference on average treatment effects with possibly more covariates than observations. *Journal of Econometrics 189*(1), 1–23.

Heckman, J. and R. Pinto (2024). Econometric causality: The central role of thought experiments. *Journal of Econometrics 243*(1-2), 105719.

Hirshberg, D. A. and S. Wager (2021). Augmented minimax linear estimation. *The Annals of Statistics 49*(6), 3206–3227.

Ida, T., T. Ishihara, K. Ito, D. Kido, T. Kitagawa, S. Sakaguchi, and S. Sasaki (2025). Choosing who chooses: Selection-driven targeting in energy rebate programs.

Kancharla, M. and H. Kang (2021). A robust, differentially private randomized experiment for evaluating online educational programs with sensitive student data. *arXiv preprint arXiv:2112.02452*.

Leung, M. P. (2022). Causal inference under approximate neighborhood interference. *Econometrica 90*(1), 267–293.

Leung, M. P. and P. Loupos (2025). Graph neural networks for causal inference under network confounding. *arXiv preprint arXiv:2211.07823*.

Munro, E., X. Kuang, and S. Wager (2025). Treatment effects in market equilibrium. *American Economic Review 115*(10), 3273–3321.

Obradović, F. (2024). Identification of long-term treatment effects via temporal links, observational, and experimental data. *arXiv preprint arXiv:2411.04380*.

Shi, C., X. Wang, S. Luo, H. Zhu, J. Ye, and R. Song (2023). Dynamic causal effects evaluation in a/b testing with a reinforcement learning framework. *Journal of the American Statistical Association 118*(543), 2059–2071.

Wang, J. and S. Yang (2025). Estimation of heterogeneous treatment effects in network-based quasi-experiments. *Available at SSRN 5220294*.

Xiong, R., A. Chin, and S. J. Taylor (2024). Data-driven switchback experiments: Theoretical tradeoffs and empirical bayes designs. *arXiv preprint arXiv:2406.06768*.

Yuan, Q., X. Li, L. Du, M. Chen, M. Sun, Y. Gao, S. He, J. Chen, and Z. Zhang (2025). Private: Differentially private average treatment effect estimation for observational data. *arXiv preprint arXiv:2512.14557*.

Debiased/Double Machine Learning

Argañaraz, F. and J. C. Escanciano (2025). Debiased machine learning for unobserved heterogeneity: High-dimensional panels and measurement error models. *arXiv preprint arXiv:2507.13788*.

Bjelac, J., V. Chernozhukov, P.-A. Klotz, J. Kueck, and T. Schmitz (2026). Automatic debiased machine learning and sensitivity analysis for sample selection models. *arXiv preprint arXiv:2601.08643*.

Bonhomme, S., K. Jochmans, and M. Weidner (2025). A neyman-orthogonalization approach to the incidental parameter problem. *arXiv preprint arXiv:2412.10304*.

Cao, J. and M. P. Leung (2025). Neighborhood stability in double/debiased machine learning with dependent data. *arXiv preprint arXiv:2511.10995*.

Chernozhukov, V., J. C. Escanciano, H. Ichimura, W. K. Newey, and J. M. Robins (2022). Locally robust semiparametric estimation. *Econometrica 90*(4), 1501–1535.

Chernozhukov, V., W. K. Newey, and R. Singh (2022). Automatic debiased machine learning of causal and structural effects. *Econometrica 90*(3), 967–1027.

Chiang, H. D., K. Kato, Y. Ma, and Y. Sasaki (2022). Multiway cluster robust double/debiased machine learning. *Journal of Business & Economic Statistics 40*(3), 1046–1056.

DiTraglia, F. J. and L. Liu (2025). Bayesian double machine learning for causal inference. *arXiv preprint arXiv:2508.12688*.

Farrell, M. H., T. Liang, and S. Misra (2020). Deep learning for individual heterogeneity: An automatic inference framework. *arXiv preprint arXiv:2010.14694*.

Farrell, M. H., T. Liang, and S. Misra (2021). Deep neural networks for estimation and inference. *Econometrica 89*(1), 181–213.

Ma, Y., P. H. Sant'Anna, Y. Sasaki, and T. Ura (2023). Doubly robust estimators with weak overlap. *arXiv preprint arXiv:2304.08974*.

Velez, A. (2024). On the asymptotic properties of debiased machine learning estimators. *arXiv preprint arXiv:2411.01864*.

Econometrics of ML-generated Data / LLM

Battaglia, L., T. Christensen, S. Hansen, and S. Sacher (2024). Inference for regression with variables generated by ai or machine learning. *arXiv preprint arXiv:2402.15585*.

Carlson, J. and M. Dell (2025). A unifying framework for robust and efficient inference with unstructured data. *arXiv preprint arXiv:2505.00282*.

Chen, H., A. Didisheim, and L. Somoza (2025). Out of the black box: Uncertainty quantification for llms via conditional probabilities. *Available at SSRN: https://ssrn.com/abstract=5012852*.

Chen, X., H. Hong, and E. Tamer (2005). Measurement error models with auxiliary data. *The Review of Economic Studies 72*(2), 343–366.

Chen, X., H. Hong, and A. Tarozzi (2008). Semiparametric efficiency in GMM models with auxiliary data. *The Annals of Statistics 36*(2), 808 – 843.

Compiani, G., I. Morozov, and S. Seiler (2025). Demand estimation with text and image data. *arXiv preprint arXiv:2503.20711*.

Ding, Y. and R. Shi (2024). Deep learning for individual heterogeneity with generated regressors by adversarial training.

Hahn, J. and G. Ridder (2013). Asymptotic variance of semiparametric estimators with generated regressors. *Econometrica 81*(1), 315–340.

He, S., L. Lv, A. Manela, and J. Wu (2025). Chronologically consistent large language models. *arXiv preprint arXiv:2502.21206*.

Jean, N., M. Burke, M. Xie, W. M. Alampay Davis, D. B. Lobell, and S. Ermon (2016). Combining satellite imagery and machine learning to predict poverty. *Science 353*(6301), 790–794.

Ji, W., W. Yuan, E. Getzen, K. Cho, M. I. Jordan, S. Mei, J. E. Weston, W. J. Su, J. Xu, and L. Zhang (2025). An overview of large language models for statisticians. *arXiv preprint arXiv:2502.17814*.

Kluger, D. M., K. Lu, T. Zrnic, S. Wang, and S. Bates (2025). Prediction-powered inference with imputed covariates and nonuniform sampling. *arXiv preprint arXiv:2501.18577*.

Ludwig, J., S. Mullainathan, and A. Rambachan (2025). Large language models: An applied econometric framework. Technical report, National Bureau of Economic Research.

Modarressi, I., J. Spiess, and A. Venugopal (2025). Causal inference on outcomes learned from text. *arXiv preprint arXiv:2503.00725*.

Rambachan, A., R. Singh, and D. Viviano (2024). Program evaluation with remotely sensed outcomes. *arXiv preprint arXiv:2411.10959*.

Empirical Research Unsing Unstructured Data

Adukia, A., A. Eble, E. Harrison, H. B. Runesha, and T. Szasz (2023, 08). What we teach about race and gender: Representation in images and text of children's books*. *The Quarterly Journal of Economics 138*(4), 2225–2285.

Ash, E., D. L. Chen, and A. Ornaghi (2024). Gender attitudes in the judiciary: Evidence from us circuit courts. *American Economic Journal: Applied Economics 16*(1), 314–350.

Autor, D., C. Chin, A. Salomons, and B. Seegmiller (2024). New frontiers: The origins and content of new work, 1940–2018. *The Quarterly Journal of Economics 139*(3), 1399–1465.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Blumenstock, J., G. Cadamuro, and R. On (2015). Predicting poverty and wealth from mobile phone metadata. *Science 350*(6264), 1073–1076.

Garg, P. and T. Fetzer (2025). Causal claims in economics. *arXiv preprint arXiv:2501.06873*.

Gebru, T., J. Krause, Y. Wang, D. Chen, J. Deng, E. L. Aiden, and L. Fei-Fei (2017). Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences 114*(50), 13108–13113.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). Measuring group differences in high-dimensional choices. *Econometrica 87*(4), 1307–1340.

Hansen, S., M. McMahon, and A. Prat (2018). Transparency and deliberation within the fomc. *Quarterly Journal of Economics 133*(2), 801–870.

Juhász, R., N. J. Lane, E. Oehlsen, and V. C. Perez (2025). Measuring industrial policy: A text-based approach. Technical report, National Bureau of Economic Research.

Khachiyan, A., A. Thomas, H. Zhou, G. Hanson, A. Cloninger, T. Rosing, and A. K. Khandelwal (2022). Using neural networks to predict microspatial economic growth. *American Economic Review: Insights 4*(4), 491–506.

Lee, K. (2025). Generative brand choice.

Liu, J. (2025). Managing by feedback.

Liu, J. Z. and Y. Tang (2024). Live broadcasting the courtroom: A field experiment in real trials. *Journal of Legal Studies*.

Michalopoulos, S. and M. M. Xue (2021). Folklore. *Quarterly Journal of Economics 136*(4), 1993–2046.

Vafa, K., S. Athey, and D. M. Blei (2025). Estimating wage disparities using foundation models. *Proceedings of the National Academy of Sciences 122*(22), e2427298122.

Generative AI Policy

Generative AI may be used in this course. We will cover how Generative AI is used within this discipline, including how to navigate its potential uses and abuses, how and when to attribute sources, and other developing topics. When using Generative AI, follow these parameters:

- Take responsibility for the content (e.g., written and digital/interactive media assignments and projects). AI can produce content that contains inaccurate information, offensive language/images, and biased or unethical representations. What you submit is your responsibility across these dimensions.
- Do not enter private, sensitive, or copyrighted data from others into AI tools without their consent.
- Provide clear attribution of your sources. Any assignments that utilize Generative AI without attribution per the guidelines shared in this course can be seen as potential academic dishonesty and treated at the undergraduate level within the SMU Student Honor Code and at the graduate and professional level within the honor codes found in their respective school policies.

Title IX and Disability Accommodations

Disability Accommodations

Students who need academic accommodations for a disability must first register with **Disability Accommodations & Success Strategies (DASS)**. Students can call 214-768-1470 or visit smu.edu/DASS to begin the process. Once they are registered and approved, students then submit a DASS Accommodation Letter through the electronic portal, DASS Link, and then communicate directly with each of their instructors to make appropriate arrangements. Please note that accommodations are not retroactive, but rather require advance notice in order to implement.

Sexual Harassment (Title IX)

All forms of sexual harassment, including sexual assault, dating violence, domestic violence and stalking, are violations of SMU's **Title IX Sexual Harassment Policy** and may also violate Texas law. Students who wish to file a complaint or to receive more

information about the grievance process may contact **Samantha Thomas**, SMU's Title IX Coordinator, at accessequity@smu.edu or 214-768-3601. Please note that faculty and staff are mandatory reporters. If students notify faculty or staff of sexual harassment, they must report it to the Title IX Coordinator. For more information about sexual harassment, including resources available to assist students, please visit smu.edu/sexualharassment.

Pregnant Students

Students who are pregnant may request academic adjustments or work accommodations by contacting the **Office of Student Advocacy and Support** by calling 214-768-4564. Students seeking assistance must schedule an appointment with their professors as early as possible, present a letter from the Office of the Dean of Students, and make appropriate arrangements. Please note that academic adjustments are not retroactive and, when feasible, require advance notice to implement.

Academic Policies

Religious Observance

Religiously observant students wishing to be absent on holidays that require missing class should notify their professors in writing within the first two weeks of the semester and should discuss with them, in advance, acceptable ways of making up any work missed because of the absence. Click here for a list of holidays.

Medical-Related Absences

To ensure academic continuity and avoid any course penalties, students should follow procedures described by their instructors in order to be provided with appropriate modifications to assignments, deadlines, and exams.

Excused Absences for University Extracurricular Activities

Students participating in an officially sanctioned, scheduled university extracurricular activity should be given the opportunity to make up class assignments or other graded assignments that were missed as a result of their participation. It is the responsibility of the student to make arrangements for make-up work with the instructor prior to any

missed scheduled examinations or other missed assignments. (See current Catalog under heading of "Academic Records/Excused Absences.")

Final Exams

Final course examinations shall be given in all courses where appropriate, and some form of final assessment is essential. Final exams and assessments must be administered as specified in the official examination schedule and cannot be administered or due during the last week of classes or during the Reading Period. Syllabi must state clearly the form of the final exam or assessment, and the due date and time must match the official SMU exam schedule. SMU policy states that all exceptions to the examination schedule may be made only upon written recommendation of the chair of the department sponsoring the course and with the concurrence of the dean of that school, who will allow exceptions only in accordance with guidelines from the Office of the Provost.

Academic Dishonesty

Students are expected to embrace and uphold the SMU Honor Code. Violations of the Honor Code will be acted upon in accordance with the policies and procedures outlined in the Mustang Student Handbook.

Generative AI Guidance

Generative AI may be used in this course. We will cover how Generative AI is used within this discipline, including how to navigate its potential uses and abuses, how and when to attribute sources, and other developing topics. When using Generative AI, follow these parameters:

- Take responsibility for the content you submit (e.g., written and digital/interactive media assignments and projects). AI can produce content that contains inaccurate information, offensive language/images, and biased or unethical representations. You are responsible for what you submit across these dimensions.
- Do not enter private, sensitive, or copyrighted data from others into AI tools without their consent.
- Provide clear attribution of your sources. Any assignments that utilize Generative AI without attribution, per the guidelines shared in this course, may be considered

academic dishonesty and treated at the undergraduate level within the SMU Student Honor Code and at the graduate and professional level within the honor codes of their respective schools.

Zoom Use Guidelines

SACSCOC and university policy require that all in-person classes be taught in person and not moved to Zoom or other technology-mediated modalities, except for extenuating circumstances as laid out here.

Student Support Services

Student Academic Success Programs

Students needing assistance with writing assignments for SMU courses may schedule an appointment with the **Writing Center** through Writing Center Online. Students who would like support for subject-specific tutoring should access the tutor schedule. Students can schedule individual academic counseling sessions and attend learning strategies workshops on a variety of skills. The A-LEC is located on the second floor of the Loyd All Sports Center, Suite 202; 214-768-3648; smu.edu/a-lec and smu.edu/asd.

Caring Community Connections Program

CCC is a resource for anyone in the SMU community to refer students of concern to the Office of the Dean of Students. The online referral form can be found at smu.edu/deanofstudentsccc. After a referral form is submitted, students will be contacted to discuss the concern, strategize options, and be connected to appropriate resources. Anyone who is unclear about what steps to take if they have concerns about students should contact the Office of the Dean of Students at 214-768-4564.

Mental Health Resources: Counseling Services and Teletherapy

Throughout the academic year, students may encounter different stressors or go through life experiences which impact their mental health and academic performance. Students who are in distress or have concerns about their mental health can schedule a same-day or next-day appointment to speak with a counselor by calling **Counseling**

Services at 214-768-2277. Counselors are available at any time, day or night, for students in crisis at this number (select option 2 to be connected with a counselor immediately). Students seeking ongoing counseling should call the same number (select option 1) during normal business hours to schedule an initial appointment. **SMU Teletherapy** provides another free option for on-demand counseling and video appointments with a medical professional.

Campus Carry Law

In accordance with Texas Senate Bill 11, also known as the "campus carry" law, and following consultation with the entire University community, SMU chooses to remain a weapons-free campus. Specifically, SMU prohibits possession of weapons (either openly or in a concealed manner) on campus. For more information, please see: smu.edu/campuscarrylaw.