Econometric Inference with Many Regressors or Many Instruments

Mikkel Sølvsten | PhD course | November–December 2026

This six-hour PhD course examines estimation and inference when the number of regressors or instruments is large relative to the sample. Using judge-design IV and fixed-effects variance decompositions, we study why conventional methods can fail, how leave-out corrections work, and what valid inference requires. Lectures combine worked examples, numerical illustrations, and short discussion exercises.

Prerequisites: First-year PhD econometrics, including matrix OLS, instrumental variables, projection matrices, and basic asymptotic inference. Prior familiarity with AKM models is not required.

By the end of the course, students should be able to:

- Explain the sources of bias and inferential failure in the two applications.
- Interpret conventional and corrected estimates, standard errors, and confidence sets.
- Identify the assumptions and design features relevant to choosing an appropriate method.

Schedule

Lecture	Date in 2026	Time
1	Monday 23 November	11:00–13:00
2	Wednesday 25 November	14:00–16:00
3	Wednesday 2 December	14:00–16:00

Course content

Lecture 1 Many instruments and leave-out estimation. The incidental-parameters problem of Neyman and Scott (1948) motivates the course: growing nuisance dimension can invalidate familiar large-sample arguments. We then turn to Bekker's (1994) many-instrument framework and examine why estimating a large first stage can bias 2SLS. Angrist, Imbens, and Krueger (1999) provide the central solution: JIVE constructs instruments using leave-one-out fitted values. Judge-design examples and numerical comparisons show how instrument count, collective strength, and own-observation influence affect estimation.

Lecture 2 Many regressors and variance components. Huber (1973) provides an early perspective on regression with growing parameter dimension. The worker–firm model of Abowd, Kramarz, and Margolis (1999) makes the problem concrete: estimation noise in fixed effects contaminates variance decompositions. Following Kline, Saggio, and Sølvsten (2020), we study a leave-out correction and distinguish bias in quadratic functions of coefficients from uncertainty about individual coefficients. Kline (2024) places these methods in the broader literature on firm wage effects and discusses the role of mobility and connectedness.

Lecture 3 Valid inference and practical choices. Mikusheva and Sun (2024) explain how valid tests and confidence sets can be constructed with many weak instruments, including by inverting jackknife Anderson–Rubin tests. Yap (2026) demonstrates why treatment-effect heterogeneity requires further care. We compare the assumptions behind alternative procedures and discuss practical choices in the two applications. A brief extension to Mikusheva and Sølvsten (2025) shows how feedback in time-series regressions can produce OLS coefficient bias with many controls, a different mechanism from variance-component bias.

Preparation and further reading

Read the following excerpts before each lecture, focusing on intuition and the proposed methods. Proofs, detailed variance calculations, and simulation studies are optional. Page numbers refer to the published articles.

- **Before Lecture 1:** Angrist, Imbens, and Krueger (1999), Sections 1–3, pp. 57–61.
- **Before Lecture 2:** Kline, Saggio, and Sølvsten (2020), from the Introduction through Section 1.1, pp. 1859–1863, including the model setup at the start of Section 1.
- **Before Lecture 3:** Mikusheva and Sun (2024), the opening of Section 4 and Section 4.1, pp. C10–C13.

For optional background, Anatolyev (2019) surveys both literatures. Kline (2024), Section 3 and Sections 4.1–4.2, expands the AKM discussion. Sølvsten (2020) provides a complementary perspective on robustness to outliers and thick-tailed errors.

Selected references

- Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). "High Wage Workers and High Wage Firms." *Econometrica*, 67(2), 251–333.
- Angrist, J. D., G. W. Imbens, and A. B. Krueger (1999). "Jackknife Instrumental Variables Estimation." *Journal of Applied Econometrics*, 14(1), 57–67.
- Anatolyev, S. (2019). "Many Instruments and/or Regressors: A Friendly Guide." *Journal of Economic Surveys*, 33(2), 689–726.
- Bekker, P. A. (1994). "Alternative Approximations to the Distributions of Instrumental Variable Estimators." *Econometrica*, 62(3), 657–681.
- Huber, P. J. (1973). "Robust Regression: Asymptotics, Conjectures and Monte Carlo." *The Annals of Statistics*, 1(5), 799–821.
- Kline, P. (2024). "Firm Wage Effects." *Handbook of Labor Economics*, 5, 115–181.
- Kline, P., R. Saggio, and M. Sølvsten (2020). "Leave-out Estimation of Variance Components." *Econometrica*, 88(5), 1859–1898.
- Mikusheva, A., and M. Sølvsten (2025). "Linear Regression with Weak Exogeneity." *Quantitative Economics*, 16(2), 367–403.
- Mikusheva, A., and L. Sun (2024). "Weak Identification with Many Instruments." *The Econometrics Journal*, 27(2), C1–C28.
- Neyman, J., and E. L. Scott (1948). "Consistent Estimates Based on Partially Consistent Observations." *Econometrica*, 16(1), 1–32.
- Sølvsten, M. (2020). "Robust Estimation with Many Instruments." *Journal of Econometrics*, 214(2), 495–512.
- Yap, L. (2026). "Inference with Many Weak Instruments and Heterogeneity." Working paper, May 13 version.