Semiparametric Econometrics

European University Institute, Spring 2006.

Preliminary syllabus.

These lectures will be organized around the following (very big) 'picture.' Data is generated from a joint distribution function $F(z)$ where z is finite dimensional; the estimation and interpretation of $F(z)$ is the goal. I add 'interpretation' because if we know $F(z)$, then all statistical functionals of $F(z)$ (such as conditional expectation functions) can be calculated, but this doesn't usually completely satisfy our curiosity: we also typically want to know the correspondence between models and distributions, i.e. the restrictions on $F(z)$ implied by certain interpretations, and the restriction on the possible interpretations implied by certain $F(z)$'s. This problem of interpretation is studied under the rubric of 'identification.'

Given this picture, I interpret the field of semiparametric econometrics as consisting largely of a grab-bag of special techniques for special situations (which nonetheless illustrate deep and fundamental methods and principles) plus attempts to get at the issues of identification and estimation at some pretty good level of generality.

The lectures will cover three topics in ten lectures. The first topic is identification and quantile methods. The motivation for grouping these together is that many contemporary developments in 'local' identification theory are couched in terms of quantiles and the models these developments motivate are most easily implemented using various types of quantile regressions. In addition, quantile methods are interesting in their own right as being the most general way of describing entire distributions (as opposed to moments, including conditional moments, etc.) once parametric specification is abandoned.

The second topic consists of those elements of the 'grab-bag of techniques' that fall outside of the other two topics, but which are essential to understanding the field. We concentrate on the estimation side, where most methods are interpretable as combinations of kernel smoothing and series or sieve estimation. For one lecture we consider the bootstrap, which in most versions is attempting to conduct inference under minimal parametric assumptions.

The third topic is empirical likelihood and its generalizations. These techniques employ moment conditions to construct a 'nonparametric' likelihood; their promise is that they achieve efficiency in estimation when the models implied by the relevant theory can be expressed in terms of moments.

Assessment will be based on a project that will require some computer work. Sometime around Lecture 5, for each of the three topics, I will provide data and suggest a problem that can be treated with the methods that the course covers. The project assignment will be to pick *one* of the three problems, *or* design a similar one of your own choosing, and carry out an appropriate analysis and write up a short report of 8-10 pages (including illustrations and tables.)

There are two textbooks for the course:

(1) Pagan and Ullah: *Nonparametric Econometrics,* Cambridge University Press, 1999, available in paperback.
(2) Hastie, Tibshirani, and Friedman: *The Elements of Statistical Learning: Data Mining, Inference and Prediction,* Springer, 2001. (Handsomely produced and reasonably priced for a lavish Springer book, maybe 80 Euros.) ('HTF').

Schedule of Lectures.

Lecture 1. Quantile estimation methods**.** (This will be review for students who took the compulsory block 2 econometric course in November.)

- Koenker, R.(2003) cemmap master class: 'Short course on quantile regression'; "Exercises in quantile regression":
  (It's not intended that you do the exercises, just read about them.)

Lecture 2. Identification and interpretation of 'structural' effects.

- Chesher, A., (2001) "Exogenous impact and conditional quantile functions," CEMMAP Discussion Paper 01/01, 2001.
- Chesher, A., (2003) "Identification in nonseparable models," *Econometrica,* 71, 1405–42.
- Chesher, A., (2005) "Nonparametric identification under discrete variation," *Econometrica,* 73, 1525-1550.

The basic idea of Chesher's work is exposited within a 2 equation framework in his first paper; it is highly readable.

Lecture 3. Quantile estimation methods for structural effects: nonparametric and semiparametric estimators in the presence of endogeneity.

- Chernozhukov, V. and C. Hansen, (2005) "An IV model of quantile treatment effects," *Econometrica,* 73, 245-61.

- Chernozhukov, V. and C. Hansen, (2004) "Inference on the instrumental quantile regression process for structural and treatment effect models," MIT working paper, forthcoming *Journal of Econometrics.*
- Ma, L. and R. Koenker, (2004) "Quantile regression methods for recursive structural models," http://cemmap.ifs.org.uk/docs/cwp0104.pdf; forthcoming *Journal of Econometrics.*

Ma and Koenker build explicitly on Chesher's framework and develop some new methods for implementation: they show one (very elegant) way to do it.

Lecture 4. Density estimation: Chapter 2, Pagan and Ullah. Conditional moment estimation ('nonparametric regression'): Chapter 3, Pagan and Ullah; Chapter 6, HTF.

Lecture 5. Spline methods and 'regularization' (fitting a function subject to smoothness constraints): Chapter 5, HTF. Model assessment: Chapter 7, HTF.

Lecture 6. Discrete choice models; censored regression models; sample selection models: Chapters 7 and 9, Pagan and Ullah.

Lecture 7. The bootstrap, readings to be determined.

Lectures 8–10. Contemporary GMM theory: empirical likelihood and its generalizations. (This list to be updated and reorganized:)

- Newey and Smith, (2004) "Higher order properties of GMM and generalized empirical likelihood estimators," *Econometrica,* 72, 219-256.
- Schennach, S. "Exponentially tilted empirical likelihood," University of Chicago, mimeo, 2/2003, probably to appear in *Annals of Statistics.*
- Kitamura, Y., Tripathi, G., and Ahn, H. (2004) "Empirical likelihood–based inference in conditional moment restriction models," *Econometrica,* forthcoming.
- Imbens, G., Spady, R. and Johnson, P., (1998) "Information theoretic approaches to inference in moment condition models," *Econometrica,* 66, 333-358.

  The technical details of Newey and Smith will not be covered, but the result and setup are important, as is the basic idea in Schennach, which will be covered in some detail. Imbens et al. is a local favorite, but is much less important, serving as background.

R. H. Spady 16 February 2006