RTN-MICRODATA COURSE ON DURATION ANALYSIS

Gerard J. van den Berg
(Amsterdam, IFAU, IZA, IFS)

Copenhagen, Dec 2005

Contact information:

see http://staff.feweb.vu.nl/gberg/

Schedule and topics:

1. *Thursday Dec 1, 10:00-12:00 and 13:30-14:30:*

   Basic concepts of duration analysis.
   Models for single duration variables.
   Single spell and multiple spell data.
   Models for multiple duration variables and competing risks models.

2. *Thursday Dec 1, 15:00-17:00:*

   Stock sampling.

3. *Friday Dec 2, 10:00-13:00:*

   Effects of treatments on duration outcomes.

Summary:

This course provides an overview of duration models with microdata, with a focus on (1) the specification and identification of models for unemployment durations, and (2) the analysis of treatment effects on duration variables, notably policy effects on the exit rate from unemployment to work. We start by examining the specification and identification of the (Mixed) Proportional Hazard (MPH) model as a semi-parametric model for single-spell durations. We provide intuition on what drives identification, and we infer to what extent biases may occur because of misspecifications. Subsequently, we examine extensions of single-spell duration models, notably competing risks models and models for multiple unemployment spells. We also examine to what extent stock samples (i.e., samples of durations among those currently in the state of interest) can be used for inference.

After this, we pay special attention to the analysis of treatment effects on duration variables, like the effect of participation in a training program on the unemployment duration. We first focus on cases in which a treatment can be summarized by the point in time at which it is initiated and we are interested in the effect of this treatment on some outcome duration, like the effect of participation in a training course on the unemployment duration. We define "no anticipation of treatment" and argue that it is an essential condition for identification in a continuous-time model framework. We prove identification of semi-parametric models based on the MPH model framework. These allow for dependent unobserved heterogeneity, i.e. for selectivity. Exclusion restrictions on covariates are not required. We provide results for both single-spell and multiple-spell data. We show that the timing of events conveys useful information on the treatment effect. We illustrate things with empirical examples. Provided enough time is left, we also examine the empirical analysis of treatment effects on duration outcomes if there is an instrumental variable. Here we distinguish between cases where the treatment starts at the moment of randomization and cases where it starts at a later point in time. It turns out that randomization and instrumental variation by themselves are often insufficient for inference on interesting effects, and they need to be augmented by a semi-parametric structure.

Literature etc.:

Main reference for the course:

Van den Berg , G.J. (2001), "Duration models: Specification, identification, and multiple durations", in J.J. Heckman and E. Leamer, editors, Handbook of Econometrics, Volume V, North Holland, Amsterdam....see http://staff.feweb.vu.nl/gberg/handbook.zip

Other references for the topics listed under 1 and 2:

Cameron, A.C. and P.K. Trivedi (2005), Microeconometrics, Princeton University Press, Princeton. (This brand new book has some chapters on duration analysis.)

Lancaster, T. (1990), The Econometric Analysis of Transition Data , Cambridge University Press, Cambridge. (The only book fully dedicated to econometric duration analysis, with lots on estimation issues.)

Kalbfleisch, J.D. and R.L. Prentice (1980), The Statistical Analysis of Failure Time Data, Wiley, New York. (Biostatistically oriented book.)

Kiefer, N.M. (1988), "Economic duration data and hazard functions", Journal of Economic Literature, 26, 646--679. (A lucid basic introduction on the basics of econometric duration analysis.)

Klein, J.P. and M.L. Moeschberger (1999), Survival Analysis, Springer, New York. (Nonparametric, biostatistically oriented book.)

Ridder, G. (1984), "The distribution of single-spell duration data", in G.R. Neumann and N. Westergaard-Nielsen, editors, Studies in Labor Market Dynamics, Springer-Verlag, Heidelberg. (best piece ever written on stock samples).

Main reference for the topics listed under 3:

Abbring, J.H. and G.J. van den Berg (2003), "The non-parametric identification of treatment effects in duration models", Econometrica 71, 1491-1517.

Other references for the topics listed under 3:

Abbring, J.H. and G.J. van den Berg (2004), "Analyzing the effect of dynamically assigned treatments using duration models, binary treatment models, and panel data models", Empirical Economics 29, 5-20.

Abbring, J.H. and G.J. van den Berg (2004), "Social experiments and instrumental variables with duration outcomes", Working paper.

Ham, J.C. and R.J. LaLonde (1996), “The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training”, Econometrica, 64, 175-205.

On software:

For econometric analysis of rather standard “reduced-form” duration models, STATA is undoubtedly the best package. Estimation of non-standard models may need to be done in packages like GAUSS or Ox, where you have to program the likelihood function yourself.

Pelz, C.J. and J.P. Klein (1996), "Analysis of survival data: a comparison of three major statistical packages ({SAS, SPSS, BMDP})", Working paper, Medical College of Wisconsin, Milwaukee.

Stephen Jenkins' course on "Survival analysis with Stata", see
http://www.iser.essex.ac.uk/teaching/stephenj/ec968/