**Sociology 7200: Techniques of Demographic Analysis II**

TR 9:30-10:45 in Eppler S 306 (lab in Williams 215)

Spring 2014

Professor Kara Joyner

237 Williams Hall, 419-372-7166
e-mail: kjoyner@bgsu.edu
Office Hours: MF 2:00-3:00 (or by appointment)

*Required Reading for the Course:* Students are advised to buy one of the two books below (depending on whether they plan to use SAS or Stata) and required to read articles listed under scheduled topics (distributed in class or available electronically). Note that additional articles may be added during the semester.


*Course Description:* This course builds on the material in Demographic Analysis I, which covers the basic methods of formal demography; however, this course does not require students to have taken the first course in the sequence. Students in this course will learn about two major techniques for analyzing longitudinal and multi-level data: survival analysis and multi-level models. These methods are useful for examining the timing of events and the clustering of lower-level units within higher-level units (e.g., individuals within schools). In addition to covering recent applications of these methods, this course reviews relevant material covered in the first course. The assignments and class examples rely on public-use data from the NSFG.

*Format:* I will lecture and generate discussion during some parts of the class. During other parts, students will share their answers to assignments and manipulate data in a computer lab using Excel and SAS/Stata. Students will lead the discussion of the readings that apply these methods (indicated below by a *). In the event that students are not participating (or participating but obviously not prepared) exams will be added to the schedule.

*Grading:* Grades will be based on two assignments (40%), class participation (10%), and an individual project (50%). The guidelines for the assignments are discussed below. The assignments require students to properly code variables and create appropriate file structures on their own. They should attempt to complete the assignments prior to the labs. (An inappropriate file structure or error in the coding of variables may lead to a model that is essentially producing garbage.) Students are expected to read the assigned material before it is covered in the lectures. The individual project requires students to incorporate feedback they receive from others.

*Absences:* Students are responsible for finding out about announcements made and material covered during their absence. Regular class attendance, including the lab component, is very critical. If you experience extenuating circumstances during the semester, please contact me.
For Assignments 1 and 2, you are required to focus on men or women from the 2006 to 2010 Cycle of the National Survey of Family Growth. You can choose between two events: the transition to first birth, or alternatively, the transition to first co-residential union. For the project, you may choose your data and event. More detail will be given later about the assignments, but here are the key issues they must address. You may work with other students on Assignment 1 and 2, but your coding, writing, and focal variable(s) must be distinctive.

Assignment 1 (due 2-20): Assignment 1 requires you to create the key variables needed to conduct competing risk life estimates of your outcome: a timing variable that indicates the age when respondent first experiences the event or reaches the interview (if they fail to experience the event) and a censoring variable that indicates the type of event experienced (e.g., no birth, first non-marital birth, and first marital birth). In addition, you need to create a variable (or set of variables) that address a substantive question that interests you. Ideally, this variable(s) will correspond to a time prior to the risk period. If not, then your interpretations of the results should address this fact. Your discussion of the analyses must include the following: the hypothesis and rationale; the source of data; a description of how you coded your key variables; the number of cases dropped and the reasons they were dropped; descriptive statistics for all variables used in your analyses (one set for the analytical sample and another set stratified by the censoring variable); a hazard graph of the competing risk events; a cumulative failure graph of the competing risk events; and additional analyses addressing how life table estimates differ across categories of your focal variable(s). The assigned journal articles offer examples on how to present and discuss the analyses. All tables and graphs (generated in Excel) must be copied into a single word document.

Assignment 2 (due 4-1): Assignment 2 builds on your work in Assignment 1. First, you must correct any problems found in Assignment 2. Next, you are required to estimate both continuous-time and discrete-time competing risk models of your focal event. The discrete-time models require the creation of a person-years file and variables that capture how the likelihood of experiencing the events changes over time. The latter will require you to compare model fit using various specifications of time (e.g., age indicator variables). The final models will highlight the effects of your focal variable(s) on the events of interest. Beware not to run your Cox models and life table estimates using the person-years file. In addition to presenting results based on your models, you will need to present a new set of descriptive statistics of all variables used to generate discrete-time models for your analytic sample (one set for the overall sample of person-years and another set selecting respondents in their first year of risk).

Individuals Project Proposals (due date TBA): Your individual project proposal will focus on your outcome and data set of choice. It is essentially the middle parts of a journal article addressing: the hypotheses and rational for the key focal variable(s); description of your data and sample; presentation of descriptive statistics and life table estimates, and details on the models you plan on running. (The project does not require estimation of models.) For the project, you need not utilize an outcome with competing risks. Nor are you required to estimate both continuous-time and discrete-time models. However, you must propose to extend the survival models in some way (e.g., addressing left-truncation, including time-varying covariates, or
estimating a fixed-effects model). A one-page proposal of the individual project is due 2-13 for circulation to the entire class. At this time, it is fine if you have not yet determined the model.

Discussion of Journal Articles. Students will be responsible for guiding the discussion of a journal article that applies the techniques of this course. The discussion should include a one-page summary that addresses the motivation for using the technique, the statistical program used, any important details about the file structure, results for key variables, any issues that are not clearly addressed, and any questions about the models.

Schedule:

Week 1 (1/14 & 1/16): Introduction to Survival Analysis


Week 2 (1/21 & 1/23): Introduction to Multilevel Models


Week 3 (1/28 & 1/30): Review of the Life Table


Chapter 3 from Allison (2010); Chapter 8 from Cleves et al. (2010)
Week 4 (2/4 & 2/6): Competing Risk Life Table Estimates

Handout on computing competing risk life table estimates in SAS (to be circulated)


Week 5 (2/11 & 2/13): In-class work on Assignment 1

Week 6 (2/18 & 2/20): Student presentations of topic for individual project

Week 7 (2/25 & 2/27): Discrete-Time Models


Chapter 7 (specifically the material on the discrete-time model) from Allison (2010); see the CFDR workshop on event history analysis for the Stata codes needed for these models.

Week 8 (3/4 & 3/6): Continuous-Time Models


Chapter 5 from Allison (2010); Chapter 9 from Cleves et al. (2010)

Week 9 (3/18 & 3/20): Time-Varying Covariates and Competing-Risk Models


See relevant sections from Allison (2010) and Cleves et al. (2010)

**Week 10 (3/25 & 3/27):** In-class work on Assignment 2

**Week 11 (4/1 & 4/3):** Extensions of Survival Models


**Week 12 (4/8 & 4/10):** Multilevel Models


**Week 13 (4/15 & 4/17):** Multilevel Models


**Week 14 (4/22 & 4/24):** Integration of Survival and Multilevel Models / In-class work


**Week 15 (4/29):** In-class work on projects

*Notes:* No class May 1st due to academic travel.

Useful links:

German Rodriguez: [http://data.princeton.edu/](http://data.princeton.edu/)

Programming for Applied Longitudinal Data Analysis:  
[http://www.ats.ucla.edu/stat/examples/aldal](http://www.ats.ucla.edu/stat/examples/aldal)

Centre for Multilevel Modeling: [http://www.bristol.ac.uk/cmm/](http://www.bristol.ac.uk/cmm/)

Programming for Multilevel Models:  
[http://www.ats.ucla.edu/stat/examples/ma_snijders/default.htm](http://www.ats.ucla.edu/stat/examples/ma_snijders/default.htm)