Teaching from *Statistical Models* D.A. Freedman
Statistics Department U.C. Berkeley

In fall 2006, I taught a graduate class from *Statistical Models*. There were about 50 students, including many statisticians and biostatisticians. However, the majority were from other departments—engineering, political science, economics, ..., plant biology. The course met twice a week for 80-minute lectures, and there was a weekly two-hour discussion section focused on the computer labs. The schedule was about like this.

Chapter 1. Two lectures.

Chapter 2. Two lectures.

Chapter 3. Assigned for reading (this would not be a good idea for an undergraduate class).

Chapter 4. Six lectures.

1. Theorems 4.1 and 4.2, with a little matrix algebra.
2. Theorem 4.3 and the beginning of Theorem 4.4, with a little more matrix algebra.
3. Finished Theorem 4.4. Explained variance. Regression in the case $p = 1$, first with variable x then with $x \equiv 1$.
4. OLS is BLUE (sketch proof only). GLS, including brief review of positive definite matrices. Example: telephone lines reduce anomie and increase life expectancy—or do they? Reference:
   www.stat.berkeley.edu/users/census/poliscih.pdf
5. Normal theory and the t-test (sketch proof only).
6. F-test (sketch proof only). Data snooping.

Chapter 5. Four lectures

1. Blau & Duncan on social status. When to standardize. (The answer is, seldom.)
2. Gibson on McCarthy. Interaction terms.
3. Types of variables, including dummy variables. Response schedules in the univariate case.
4. Response schedules in the multivariate case; response schedules for path diagrams.

Chapter 6. Five lectures

1. Binomial, Poisson. Skipped the normal distribution (this would not be a good idea for an undergraduate class). Probit/logit with one explanatory variable.
2. Probit model for reading, with multiple explanatory variables. The latent variable formulation. Identification.
3. Tables 1, 2, 3 in Evans & Schwab on the effects of Catholic schools. Postponed marginal effects and average causal effects. Discussed the model and the choice problem.

4. The choice problem in the context of rational choice theory (aka expected utility theory). Critique. The response schedule formulation for the probit model; marginal effects.
5. Average treatment effects; the biprobit model to handle endogeneity of attending Catholic school.

Chapter 7. Three lectures.

Chapter 8. Four lectures.

1. Supply and demand; the butter model.
2. IVLS and 2SLS; instrumenting the butter model.
3. Rindfuss et al on education and age at first birth.
4. Schneider et al on school choice and social capital.

I omitted the proof in Section 8.4, and the technical material in Section 8.8.

I tend to get a lot of student discussion in class. If discussion flagged, or lectures ended early for some other reason, I would work a problem or two from the book. I assigned projects, although that was a lot of work: there were ten target papers to read, and ten student papers to grade. Each group come in for a half-hour discussion with me, rather than presenting in class. I gave supplemental lectures on survival analysis. Other options might include ecological regression, or census adjustment.