

HCMG 900: ProSeminar in Health Services Research

Spring 2011

JMHH 304

Scott Harrington
harring@wharton.upenn.edu

206 Colonial Penn Center
<http://www.scottharringtonphd.com/>

Overview

This course explores econometric methods widely used in health care research with a focus on applications in health economics. Concepts, applications, and practice are emphasized, as opposed to technical derivations of estimators and their properties. Methods covered include models of treatment effects; models with qualitative, limited, and transformed dependent variables; two-part, GLM, and sample selection models; count data and duration models; stochastic frontier models; regression discontinuity models; and matching / propensity score methods. The last part of the course considers testing for adverse selection and special topics.

Prerequisites: Microeconomics, econometrics

Readings

The reading list consists of published articles and working papers, which should be read prior to class (approximately two papers per week). The papers are available on-line through PennText or posted on WebCafé. The course also will employ Andrew Jones's primer, *Applied Econometrics for Health Economists: A Practical Guide* (OHE Research, 2nd ed., 2007). A pre-print version is available online and will be posted to WebCafé room, or you can buy a paperback copy. Jones's chapter, "Health Econometrics," in Volume 1A of the *Handbook of Health Economics* (Cutler and Newhouse, eds.; available in Lippincott reference) might be a useful resource. Useful econometric texts include W. Greene, *Econometric Analysis*; J. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, and A. Cameron and P. Trivedi, *Microeconometrics*. J. Angrist and J-S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion* is interesting. The *STATA* reference manuals also provide useful background on particular methods.

Other Requirements and Grading

In addition to reading the assigned papers prior to class, you are required to:

- Attend selected HCMG and LDI visiting speaker seminars as requested and possible.
- Complete several hands-on data analysis assignments using data from the RAND Health Insurance Experiment and other sources (to be provided), preferably using *STATA*.
- Conduct a research project and present the results to the class on an assigned subject. The focus should be on the appropriate use of data and estimation.
- Present a brief evaluative summary in class of two papers from those starred on the reading list.
- Take a final exam.

Grading: Project – 40%; participation, assignments, and presentation – 40%; final – 20%

Course Outline and Readings

I. Course introduction and background / review on research methodology and estimation (Jan. 13)

- A. Theory / hypothesis / test paradigm
- B. Classical linear regression with stochastic regressors
- C. Violation of assumptions

II. Introduction to estimating treatment effects / program evaluation (Jan. 20 and 27)

- A. Panel data estimation
- B. Difference-in-difference (D-D) analysis
- C. D-D with panel data
- D. Instrumental variables estimation
- E. Potential outcomes framework
- F. Standard errors (WLS, grouped data)

Sharon Long, Alshadye Yemane, and Karen Stockley, Disentangling the Effects of Health Reform in Massachusetts: How Important are the Special Provisions for Young Adults, *American Economic Review* 100 (2010): 297-302.

Phillip Levine, Parental Involvement Laws and Fertility Behavior, *Journal of Health Economics* 22 (2003): 861-878.

*Jason Fletcher, Social Interactions and Smoking: Evidence Using Multiple Student Cohorts, Instrumental Variables, and School Fixed Effects,” *Health Economics* 19 (2010): 466-484.

*John Cawley, John Moran, and Kosali Simon, “The Impact of Income on the Weight of Elderly Americans,” *Health Economics* 19 (2010): 979-993.

Robert LaLonde, Evaluating the Econometric Evaluations of Training Programs with Experimental Data, *American Economic Review* (1986): 604-620.

Joshua Angrist and J-S Pischke, The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con Out of Econometrics, *Journal of Economic Perspectives* 24 (2010): 3-30.

III. Qualitative dependent variables (Feb. 3, 10)

- A. Binary response models
- B. Ordered multiple response models
- C. Unordered multinomial response models

Andrew Jones, *Applied Econometrics for Health Economists* (“Jones”), Chapters 2-5.

*Didem Bernard, Jessica Banthin, and William Encinosa, Wealth, Assets, and the Affordability of Health Insurance, Economic Research Initiative on the Uninsured, University of Michigan, June 2007. (Also see the shorter article in *Health Affairs* 28 (2009): 887–896.)

*Amy Davidoff, Linda Blumberg, and Len Nichols, State Health Insurance Market Reforms and Access to Insurance for High-Risk Employees, *Journal of Health Economics* 24 (2005): 725-750.

Bryan Dowd, et al., Health Plan Enrollment and Mortality in the Medicare Program, *Health*

Economics, June 21, 2010.

Justin Wang, et al., Do Bad Report Cards Have Consequences? Impacts of Publicly Reported Provider Quality Information on the CABG Market in Pennsylvania, *Journal of Health Economics*. December 10, 2010.

IV. Limited and transformed dependent variables (Feb. 17, 24)

- A. Modeling health expenditures with two-part and GLM models
- B. Choosing a model
- C. Sample selection models
- D. Nonlinear models with endogenous variables
- E. Quantile regression

Jones, Chapters 6-8, 11.

(Optional further background: Borislava Mihaylova, et al., Review of Statistical Methods for Analyzing Healthcare Resources and Costs, *Health Economics*, August 21, 2010.; and Steven Hill and G. Edward Miller, Health Expenditure Estimation and Function Form: Applications of the Generalized Gamma and Extended Estimating Equations Models, *Health Economics*, 2009.)

Melinda Buntin and Alan Zaslavsky, Too Much Ado about Two-Part Models and Transformation? Comparing Methods of Modeling Medicare Expenditures, *Journal of Health Economics* 23 (2004): 525-542.

Siu Fai Leung and Shihti Yu, On the Choice between Sample Selection and Two-Part Models, *Journal of Econometrics* 72 (1996): 197-229.

Joseph Terza, Anirban Basu, and Paul Rathouz, Two-State Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling, *Journal of Health Economics* 27 (2008): 531-543.

Jay Bhattacharya, Dana Goldman, and Daniel McCaffrey, Estimating Probit Models with Self-Selected Treatments, *Statistics in Medicine* 25 (2006): 389-413.

*David Madden, Sample Selection Versus Two-Part Models Revisited: The Case of Female Smoking and Drinking, *Journal of Health Economics* 27 (2008): 300-307.

*Christoph Schmidt and Harald Tauchmann, Heterogeneity in the Intergenerational Transmission of Alcohol Consumption: A Quantile Regression Approach, *Journal of Health Economics*. October 8, 2010.

V. Count data models (March 3)

- A. Poisson and negative binomial models
- B. Zero inflated models

Jones, Chapter 9.

Darius Lakdawalla and Neeraj Sood, HIV Breakthroughs and Risky Sexual Behavior, *Quarterly Journal of Economics* 121 (2006): 1063.

VI. Survival/duration analysis (March 17)

- A. Survival, hazard, and cumulative hazard functions
- B. Nonparametric methods
- C. Proportional hazard models
- D. Parametric models

Jones, Chapter 10.

*Andrew Wilper, et al., Health Insurance and Mortality in US Adults, *American Journal of Public Health* 99 (2009).

*Paul Frijters, et al., Childhood Conditions and Length of Life: Evidence from the UK Boyd Orr Cohort: 1937-2005, *Journal of Health Economics* 29 (2010): 39-47.

VII. Other methods (March 24, March 31, April 7 or TBA) and more applications

- A. Regression discontinuity models
- B. Stochastic frontier models
- C. Diagnostics for influential observations and collinearity
- D. Matching and propensity score methods (Jeff Silber, date tba)

*Christopher David Card and Carlos Dobkin, The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age, *American Economic Journal: Applied Economics* 1 (2009): 164-182.

Michael Rosko and Ryan Mutter, Stochastic Frontier Analysis of Hospital Inefficiency: A Review of Empirical Issues, *Medical Care Research and Review* 65 (2008): 131-166.

Scott Harrington, Rachel Werner, and Emi Terasawa, Stochastic Frontier Analysis of Hospital Mortality, draft, June 2010.

*Amy Finkelstein, The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare, *Quarterly Journal of Economics* 122 (2007).

*Leemore Dafny, Mark Duggan, Subramaniam Ramanarayanan, Paying a Premium on Your Premium? Consolidation in the U.S. Health Insurance Industry NBER Working Paper w15434, October 2009.

Daniel McCaffrey, et al., Marijuana Use and High School Dropout: The Influence of Unobservables, *Health Economics* 19 (2010): 1281-1299.

VIII. Testing for adverse selection (April 14)

Alma Cohen and Peter Siegelman, Testing for Adverse Selection in Insurance Markets, NBER working paper 15586, December 2009.

*Hanming Fang, Michael Keane, and Dan Silverman, Sources of Advantageous Selection: Evidence from the Medigap Insurance Market, *Journal of Political Economy* 116 (2008): 303-350.

April 21: Research presentations

Homework Assignments (subject to change)

1. IV estimation and testing (tentative, tbd, due Feb. 3)
2. Qualitative dependent variables (Rand HIE data, due Feb. 10)
3. Conditional logit (Kaiser/HRET health plan survey, due Feb. 17)
4. Health expenditure modeling (Rand HIE data, due Feb. 24)
5. Two-part vs. sample selection models (Rand HIE data, due March 3)
6. Count data models (Rand HIE data, due March 17)
7. Survival models (Denmark cancer survival data, due Mar 24)

Research Project

Topic areas

1. Health and behavior
2. Regional variation in health expenditures
3. Uninsurance on health
4. Limited liability, safety net, and health insurance demand
5. Defensive medicine
6. Changes in FDA behavior (approvals, durations)
7. Replication / extension of published study
8. Monte Carlo analysis of estimator properties
9. Other if approved

Schedule

Task	Completion date
Select topic	January 27
Submit literature review	February 24
Hypotheses, data	March 17
Present results	April 21

Meetings

You should meet with me individually every few weeks for consultation and advice.