

# **The Harris School at the University of Chicago**

## **Survey Research Methodology and Analysis**

**PP41600**

**Winter Quarter 2014**

---

**Meeting Days:** Monday

**Meeting Time:** 1:30-4:30

**Meeting Place:** The Harris School 1155 East 60<sup>th</sup> Street PBPL 142

**Instructor:** Michael Davern, Ph.D. Senior Vice President and Director of Public Health Research

NORC at the University of Chicago

**Office Address:** 55E Monroe Suite 3000 (office number 3048) Chicago IL 60603

**Office Phone:** 312-357-3770

**E-mail:** [davern@uchicago.edu](mailto:davern@uchicago.edu)

**Office Hours:** By Appointment

**Teaching Assistant:** Pablo Celhay B.

**Office Address:** Harris School of Public Policy, Cubicle 164

**Office Phone:** NA

**E-mail:** [pcelhay@uchicago.edu](mailto:pcelhay@uchicago.edu)

**Office Hours:** Wednesdays from 9:00 am to 10:00 am in room 140 B

---

### **I. Course Description**

The goal of this course is to learn about the methods used to collect publicly available survey data that can be used for policy research so that students can appropriately use these data to answer policy relevant questions. Students will learn about the methods used to collect survey data, how to develop researchable policy questions that can be answered with the survey data, and about the limitations of the survey data for answering policy research questions. In order to analyze policy questions using available survey data, students will also learn about actual survey instruments, survey sample designs, survey data processing, and survey data systems that the major public policy relevant surveys use. The course will also examine specific measurement and analysis issues that are of interest to policy research (e.g., measuring public program enrollment and public program eligibility simulation). By the end of the course each student will understand the methods used to collect survey data, have developed a researchable policy question, carried out the appropriate analysis to answer the question, produced high quality analytical tables, and written up descriptions of

the methods used to produce the numbers in the tables in a style that is consistent with professional policy research.

**Who should take this course:**

This course is targeted to students interested in obtaining skills for developing policy-relevant research questions that can be answered by using existing survey data for empirical evidence. We will use a variety of publicly available survey datasets, including the decennial census, the Current Population Survey, the American Community Survey, the General Social Survey, the Behavioral Risk Factor Surveillance System, and the National Health Interview Survey to illustrate important issues in demographic data collection and analysis. Students may use this course to develop analytic projects that will assist them in meeting Master's or Ph.D. research requirements. The course expands on required statistics and research methods courses within various disciplines (sociology, political Science, public policy, and public health among others). The tools learned in basic statistics and research methods are essential but not sufficient to work with the complex realities of using existing survey data resources to conduct policy relevant analyses.

## II. Course Prerequisites

Graduate level research methods course, basic graduate level statistics course, or permission of instructor.

## III. Course Goals and Objectives

### *Specific Skills Students Will Acquire*

1. Develop a better understanding of survey methodology.
2. Learn to ask research questions that can be answered with available survey data.
3. Understand the strengths and limitations of using survey data to answer policy research questions.
4. Perform analysis of survey data using complex survey data using *STATA* statistical software.
5. Produce publication quality survey methodology descriptions and tables.

## IV. Methods of Instruction and Work Expectations

The class will be run with a mixture of lecture, discussion and computer lab work (about 40% lecture, 40% hands on lab work, and about 20% group discussion). All three are essential for students to gain the skills and knowledge necessary to work with publicly available demographic data. Students will learn directly from the instructor and TA and from one another how to solve their analysis problems.

## V. Software, Readings, Course Text

### Software

#### **Note to Non-Harris Students:**

Non-Harris students should contact ([hsit@lists.uchicago.edu](mailto:hsit@lists.uchicago.edu)) to obtain access to the computer lab and/or remote server. Non-Harris students will need to pay for printing if they want to print. (Harris accepts Mastercard and Visa and can setup a printing account with \$10.)

The Harris School Student STATA Server is now available for registered users. Please follow the directions below for your computer's operating system. Problems report problems and suggestions to HSIT ([hsit@lists.uchicago.edu](mailto:hsit@lists.uchicago.edu)).

#### **Instructions for Accessing the Student STATA Server (Windows XP, Vista, 7)**

- Open **Remote Desktop Connection** that is located in **All Programs** □ **Accessories**
- Enter the following IP address for the Computer: **128.135.46.245** and then select **Connect**.

**Note:** Continue to connect to the server even if you are prompted about an invalid security certificate.

- When you have connected to the server, you will be prompted for your **Harris username** and **password**. Your username must be preceded by: " HARRISSCHOOL\ ", for example, HARRISSCHOOL\fboker

**Note:** *if you want your computer to remember your login, check the box to remember your credentials.*

You should now be logged onto the student STATA server. Your U: drive should be available just like in the computer lab. You have access to STATA, Microsoft Word and Excel, Adobe Reader, Mozilla Firefox and Internet Explorer. You are not currently enabled to upload files from your computer to the STATA server, however, you can download files from chalk and other locations on the internet.

**Note about adofiles:**

*Currently, students cannot install ado files in the STATA folder(s) or run updates; however, they can install files in the PLUS or PERSONAL folder if they redirect those locations to their U: drive, e.g. sysdir set PLUS "u:\ado"*

*or*

*sysdir set PERSONAL "u:\ado"*

*You need to change the relevant sysdir each time you want to use the ado.*

- When finished, don't forget to logoff. Click the **Start** button and select **Log off**

**Instructions for Accessing the Student STATA Server (MAC OS 10.5 and Higher)**

- Open **Remote Desktop Connection** application located in your **Applications** folder.

**Note:**

*If you don't have this application, it can be freely downloaded from*

*<http://www.microsoft.com/download/en/details.aspx?id=18140>.*

- Enter the following for the Computer name: **128.135.46.245** and then select **Connect**.
- You will be connected to a Windows server. When prompted, enter your **Harris username** and **password**. The username must be preceded by "HARRISSCHOOL", for example, HARRISSCHOOL\fboker.

You should now be logged onto the student STATA server. Your U: drive is available just like in the computer lab. You have access to Stata, Microsoft Word and Excel, Adobe Reader, Mozilla Firefox and Internet Explorer. You are not currently enabled to upload files from your computer to the STATA server, however, you can download files from chalk and other locations on the internet.

**Note about adofiles:**

*Currently, students cannot install ado files in the Stata folder(s) or run updates; however, they can install files in the PLUS or PERSONAL folder if they redirect those locations to their U: drive, e.g.*

*sysdir set PLUS "u:\ado"*

*or*

*sysdir set PERSONAL "u:\ado"*

*You need to change the relevant sysdir each time you want to use the ado.*

- When finished, don't forget to logoff. Click the **Start** button and select **Log off**

**Course Readings:**

Most of the required readings are available on the Internet or will be handed out.

### **Course Reference Materials Available On-Line (alphabetical by author)**

Federal Committee on Statistical Methodology (FCSM). 2001. "Measuring and Reporting Sources of Error in Surveys." Washington DC: Statistical Policy Office, Office of the Management and Budget. [http://www.fcsm.gov/01papers/SPWP31\\_final.pdf](http://www.fcsm.gov/01papers/SPWP31_final.pdf)

Moore, Jeffrey, Linda L. Stinson, and Edward J. Welniak Jr. 2000. "Income Measurement Error in Surveys." *Journal of Official Statistics*. 16(4): 331-361.

<http://www.jos.nu/Articles/abstract.asp?article=164331>

### **Course Reference Materials Available From Instructor**

Boudreaux, Michel, Peter Graven and Michael Davern. "Alternative Variance Estimators for the American Community Survey and Current Population Survey."

Couper, Mick. 2013. "Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys." *Survey Research Methods*. 7(3):145-156.

Davern, Michael, Arthur Jones Jr., James Lepkowski, Gestur Davidson, and Lynn A. Blewett. 2006. "Unstable inferences? An Examination of Complex Survey Sample Design Adjustments Using the Current Population Survey for Health Services Research." *Inquiry*. 43(3): 283-97.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(4), 646-675.

### **Supplemental Readings**

Davern, Michael, Arthur Jones Jr., James Lepkowski, Gestur Davidson, and Lynn A. Blewett. 2007. "Estimating Standard Errors for Regression Coefficients Using the Current Population Survey's Public Use File." *Inquiry*. 44(2):211-224.

Davern, Michael, Jacob Alex Klerman, David Baugh, Kathleen Call, and George Greenberg. 2009. "An Examination of the Medicaid Undercount in the Current Population Survey (CPS): Preliminary Results from Record Linking." *Health Services Research*. 44(23) 965-87.

Kalton, Graham. (2002). Models in the Practice of Survey Sampling (Revisited). *Journal of Official Statistics*. 18(2):129-54. <http://www.jos.nu/Articles/abstract.asp?article=182129>

Kish, Leslie. (1995). Methods for Design Effects. *Journal of Official Statistics*. 11(1):55-77.

<http://www.jos.nu/Articles/abstract.asp?article=11155>

National Center for Educational Statistics. 2002. *National Center for Educational Statistics Statistical Standards*. Washington DC: National Center for Educational Statistics.

[http://nces.ed.gov/statprog/stat\\_standards.asp](http://nces.ed.gov/statprog/stat_standards.asp)

Short, Kathleen. 2001. Experimental Poverty Measures: 1999. US Census Bureau: Washington DC.

<http://www.census.gov/prod/2001pubs/p60-216.pdf>

Urban Institute. TRIM Poverty Model Version 4.1. Urban Institute: Washington DC.

<http://trim3.urban.org/documentation/Poverty.html>

Wk	Date	Class Topic	Due Dates	Readings
1	13-Jan-14	(1) Syllabus, (2) Review list of public data sources (3) Fitting research questions to existing data (4) basic background on US federal survey data		Review web sites on handout distributed on first day of class
	13-Jan-14	<b>LAB In-class Assignment 1:</b> STATA basics, accessing data, creating usable data files, data management		Go over lab in the classroom; Show IPUMS site and structure of the files; Move to lab
2	1/17	<b>Send Pablo and Mike a list of people in your group for the course assignment</b>	<b>Have one person in your group send Pablo and myself a note</b>	
	22 and 21-Jan-14	<b>LAB In-class Assignment 2:</b> Additional data management and basic processing		Make up labs will be held on 1/21 and 1/22 to make sure all students can attend at least one session. No lecture. Pablo will coordinate
3	20-Jan-14	<b>NO CLASS Martin Luther King Day Holiday</b>		
	20-Jan-14	<b>NO CLASS Martin Luther King Day Holiday</b>		
4	27-Jan-14	Introduction to surveys, survey quality and survey error Complex survey sample designs, sampling error, variance estimation and survey data analysis	<b>1st group paper due</b>	Federal Committee Statistical Methodology (FCSM) (2001): Chapter 1 and FCSM Chapter 3
	27-Jan-14	<b>LAB In-class Assignment 3:</b> Survey estimates		
5	3-Feb-14	Complex Sample design and determining a necessary sample size for a survey	Hand back first paper	Davern et al 2006, Boudreaux et al (2012)
	3-Feb-14	<b>LAB in-class Assignment 4:</b> Tabling Stata output		
6	10-Feb-14	Unit level non-response, coverage error and response rates		FCSM (2001) Chapters 4 and 5 and Groves (2006) paper
	10-Feb-14	<b>LAB In-class Assignment 5:</b> survey means and t-tests		
7	17-Feb-14	Measurement Error in Surveys	<b>2nd group paper due</b>	FCSM (2001) Chapter 6 and Moore et al (2000) paper
	17-Feb-14	<b>LAB In-class Assignment 6:</b> survey regressions procedures		
8	24-Feb-14	Measurement Error in Surveys	<b>Take Home Assignment</b>	No new readings... Same as last week. (FCSM (2001) Chapter 6 and Moore et al

			<b>due</b> Hand back 2 <sup>nd</sup> paper	(2000) paper
	24-Feb-14	<b>LAB In-class Assignment 7:</b> Replicate weight variance estimation		

Wk	Date	Class Topic	Due Dates	Readings
9	3-March-14	Processing Error and Total Survey Error		FCSM (2001) Chapters 6 and 7
	3-March-14	<b>LAB 8 In class Assignment 8:</b> Imputation		
10	10-March-14	The future of surveys	<b>3<sup>rd</sup> group paper due</b>	Couper 2013
	10-March-14	Short student presentations		
11	14-March-14	<b>NO CLASS</b>	Hand back 3 <sup>rd</sup> paper (will be distributed on Campus by 1 pm CT)	
	20-March-14	NO CLASS: And Final (optional) rewrite due by 1 pm	<b>Final rewrite of paper due</b>	<b>Send Via Email to Pablo and I through email</b>

## VII. Evaluation and Grading

### *Course Projects*

Course projects are to be completed in a groups of 3 (one or two groups may have only 2 or as many as 4 members). Students will complete a series of short papers and data tables that present empirical data informing a specific policy-relevant research question using data from IPUMS family of data sets ([www.ipums.org](http://www.ipums.org)) -- which include the Current Population Survey, Census Data (both US and many other countries), the National Health Interview Survey -- the Behavioral Risk Factor Surveillance System (BRFSS) or the General Social Survey. Other datasets can potentially be used but **will require instructor permission** before use. These papers will build off one another to form the core of an empirical research paper or report. Each paper will be organized and formatted in accordance with the criteria specified in the relevant assignment. Revised versions of all preceding sections are to be turned in with each new section. Upon instructor review of the final assignment, the groups may wish to revise the entire paper and turn it back in for further consideration. The goal of this project is to have a fully assembled empirical research paper or report that uses actual analysis of survey data.

**Students will be evaluated as follows:**

1. **3 short papers** representing an empirical research study's components. Each paper should have 1-inch margins, use 11-point Arial or 12-point Times font, and be double-spaced.
  - a. Background Paper (3-5 pages) define policy issue, research question, and data set to be used (10%).
  - b. Revised assignment (a) and a description of the data set, description of the key variables to be used, and a descriptive statistics table of key variables (10%).
  - c. Revised Assignments (a), and (b), a description of the analytic research framework to be employed, and a table with key analytic statistics and a description of what the findings tell us about the research question (40%).
2. There will be **8 lab assignments** and the highest seven scores will count toward your final grade (the lowest score will be dropped so only **7 labs count in your final grade**) (5% each) and **1 out-of-class assignment** (5%). These assignments will make up a total of 40% of the final grade. Under some circumstances, in-class assignments can be made up if instructors are **notified in advance about an absence**.
3. A revised "final assignment" paper can be resubmitted by March 20<sup>th</sup> at 1:00 pm if the group wishes to improve upon the initial assessment.

**Late work will be docked 10% of the total for each day it is late.** Assignments are due by the beginning of class on the date due. Exceptions may be granted by the instructors on a case-by-case basis. The student must contact the instructors *in advance* of the deadline, and the instructors *must agree* that the student's circumstances warrant a new deadline for the student. If you think an excuse may be tenuous, it likely is. Make every effort to turn assignments in on time.

**Grading**

**A/F letter grade will be determined by total effort as follows:**

A	95-100%	(4.0) Represents achievement that is outstanding relative to the level necessary to meet course requirements
A-	90-94%	
B+	87-89%	(3.0) Represents achievement that is significantly above the level necessary to meet course requirements
B	83-86%	
B-	80-82%	
C+	77-79%	(2.0) Represents achievement that meets the minimum course requirements
C	73-76%	
C-	70-72%	
D+	67-69%	(1.0) Achievement below minimum course expectations but sufficient to be awarded credit
D	63-66%	
D-	60-62%	
F	below 60%	Represents failure (no credit) and signifies that the work was either (1) completed at a level of achievement that is not worthy of credit or (2) was not completed and there was no agreement between the instructor and the student that the student would be awarded an I