The Harris School at the University of Chicago

Survey Research Methodology and Analysis

PP41600

Winter Quarter 2018

Meeting Days: Thursday

Meeting Time: 3:00-5:50 PM CT

Meeting Place: **Harris School 140C**

Instructors: Michael Davern, Executive Vice President, NORC at the University of Chicago

davern@uchicago.edu (phone) 312-357-3770

Office Hours: By Appointment

Felicia LeClere, Senior Fellow, NORC at the University of Chicago

LeClere-Felicia@norc.org (phone) 312-759-4029

Office Hours: By Appointment

Teaching Assistant: Yuvraj Pathak

Office Hours Location: Student Lounge (Basement of Harris School),

Office Phone: NA

Email: ypathak@uchicago.edu

Office Hours: Wednesdays from 12 – 1 pm or by appointment

Commented [MD1]: Yuvraj please update

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 60% lecture, 25% hands on lab work, and about 15% 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 instructors, TA and from one another how to solve their analysis problems.

V. Software, Readings, Course Text

Software

INSTRUCTIONS FOR ACCESSING the UCHICAGO VLAB COMPUTERS

This document provides instructions for accessing vlab.uchicago.edu, which provides Stata, R, SPSS and ArcGIS for student use. To access the server, you will need the Chicago VPN software (if working off campus) and a Remote Desktop Connection to the lab.

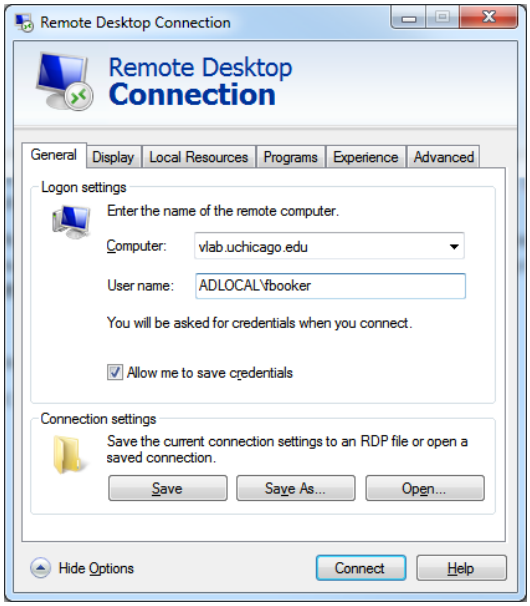
Chicago VPN

cVPN is the University's Virtual Private Network (VPN) concentrator that provides secure access to University network resources as if you were on campus, no matter where you are in the world. cVPN is necessary to access restricted resources, such as the Virtual Lab, when you are working off campus.

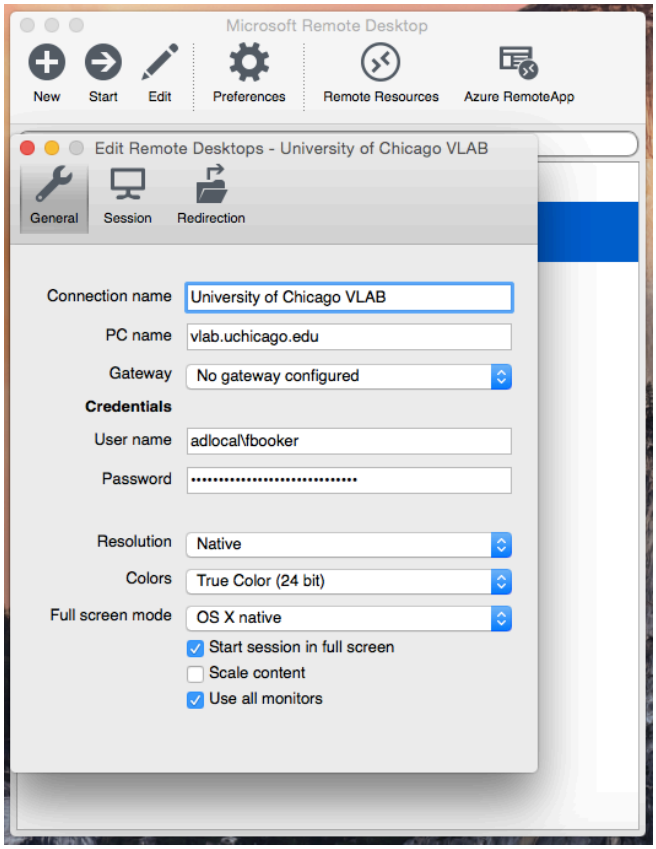
The Cisco AnyConnect VPN software can be downloaded from https://cvpn.uchicago.edu and installed on Windows, Apple, or Linux systems. Once installed, launch Cisco AnyConnect Secure Mobility Client. You will connect to **cvpn.uchicago.edu** using your CnetID and password to authenticate.

Remote Desktop Connection (RDC)

On Windows systems, RDC is part of the standard install. You should be able to search for the application in the search window. A sample configuration for connecting to the vlab is shown below, note that ADLOCAL\username is ADLOCAL\your-cnetID.



On Apple systems, you will need to download Microsoft Remote Desktop from the Apple Store. After it is installed, launch the Application and create a “New” remote desktop. The standard configuration is shown below.



Once you have configured your session, click start and your session should begin. Subsequently, when you start Microsoft Remote Desktop, you will see your connection displayed on the opening screen and you can simply click on the session name.

If you are using a **linux operating system**, your system should have the rdesktop application installed already. You need to open a terminal and run the following command:

rdesktop -g 1024x768 vlab.uchicago.edu

When you have connected, you will be prompted for your CnetID and password. For example:

Login: ADLOCAL\fbooker
Password: **********

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**

Commented [MD2]: Yuvraj… Can you check on the best way to do this? Also is the best way to distribute the data set to everyone through CHALK?

***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://fcsm.sites.usa.gov/files/2014/04/spwp31.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

Commented [MD3]: Add a measurement error reading(s) and perhaps the National academies admin data/organic data to survey linkage report.

Boudreuax, 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.

Short, Kathleen. 2014. The Supplemental Poverty Measure: 2013. U.S. Census Bureau: Washington DC. http://www.census.gov/content/dam/Census/library/publications/2014/demo/p60-251.pdf

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.

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

Urban Institute. TRIM Poverty Model Version 4.1. Urban Institute: Washington DC. http://trim3.urban.org/documentation/Poverty.html

Week	Date	Class Topics	Due at the Start of Class	Readings and Course Materials
1	1/4/2018	(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 Links to Survey Data handout
2	1/11/2018	Introduction to surveys, survey quality and survey error Complex survey sample designs, sampling error, variance estimation and survey data analysis	**Group List.** One person per group submit the list; If not in a group one will be assigned by Instructor	Federal Committee Statistical Methodology (FCSM) (2001): Chapter 1 and FCSM Chapter 3
2	Lab 1	STATA basics, accessing data, creating usable data files, data management; Show IPUMS site and structure of the files		
3	1/18/2018	Unit level non-response, coverage error and response rates	**Lab 1; 1st group paper due (1a) on**	FCSM (2001) Chapters 4 and 5 and Groves (2006) paper
3	Lab 2	Additional data management and basic processing		
4	1/25/2018	Complex Sample design and determining a necessary sample size for a survey	**Lab 2**	Davern et al 2006, Boudreaux et al (2012)
4	Lab 3	Survey estimates basics in STATA		
5	2/1/2018	Measurement Error in Surveys	**Lab 3; Out of Class Assignment Due (Sample Size)**	FCSM (2001) Chapter 6 and Moore et al (2000) paper
5	Lab 4	Tabling STATA output		
6	2/8/2018	Measurement Error in Surveys (cont'd)	**Lab 4; 2nd group paper due (1b)**	ADD SOME NEW READINGS? Felicia anything you want to inlcude?
6	Lab 5	Survey means and t-tests in STATA		
7	2/15/2018	Processing Error and Total Survey Error	**Lab 5**	FCSM (2001) Chapters 6 and 7
7	Lab 6	**STATA** survey regressions procedures		

8	2/22/2018	Using Survey Estimates to Make Policy: The Supplemental Poverty Measure and other examples	**Lab 6**	Short 2014
8	Lab 7	STATA Replicate weight variance estimation		
9	3/1/2018	The future of surveys/Surveys in the context of other data	**Lab 7; 3rd group paper due (1c)**	Couper 2013
9	Lab 8	STATA Hotdeck Imputation		
10	3/8/2018	FIRST HALF OD CLASS OPEN TOPIC: PERHAPS SOMEHTING FROM THE RECENT NATIONAL ACADEMIEIS REPORTS?	**Lab 8; One copy of Group presentaiton which counts as Lab 9;**	TBD
10	Lab 9	Group Presentations		
11	3/15/2018	NO FINAL EXAM	**Final (OPTIONAL) paper re-write due at 1 pm 3/15/18**	

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 of 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) -- 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.

Submitting assignments: Please submit assignments in hard-copy in class or on-line using the course Chalk site. Instructions on submitting assignments via Chalk are available here: ***http://answers.uchicago.edu/page.php?id=20069***.

Commented [MD4]: Felicia… Should we sdd in MCBS public Use File?

Commented [MD5]: Yuvraj…. Do you want to use CHALK or something else this year? I think U of C has a new system, right?

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 (12%).
   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 (12%).
   c. Revised Assignments (a), and (b), a description of the analytic research framework employed, and a table with key analytic statistics and a description of what the findings tell us about the research question (30%).
   d. A revised "final assignment" paper can be resubmitted by March 15th at 1:00 pm if the group wishes to improve upon the initial assessment. Instructors can award up to half of the points deducted on the first submission (for example on submission one if the group received 80/100 points; by rewriting the assignment and turning it back in the group could get up to 10 points more or 90/100 total for the final paper).
2. There will be **9 lab assignments** and the highest eight scores will count toward your final grade (the lowest score will be dropped so only **8 labs count in your final grade**) (5% each) and **1 out-of-class assignment** (6%). These assignments will make up a total of 46% of the final grade. Under some circumstances, in-class assignments can be made up if instructors are **notified in advance about an absence**.

Commented [MD6]: In the past we have had two out of class assignments. Will talk with Felicia about whether to have a second again this year.

Late work will be docked <u>10% of the total</u> for each day it is late. Assignments are <u>due by the beginning of class on the date due</u>. 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.