UNC
GILLINGS SCHOOL
OF GLOBAL
PUBLIC HEALTH

Syllabus
HPM 881: Linear Regression Models
Spring 2019
3 Credits | Residential

Table of Contents

Course Overview

Course Description

This course provides an introduction to one of the primary tools of statistical analysis, linear regression models. We will study regression models in which the dependent variable is continuous, although extensions to binary dependent variables will also be examined. After introducing the classical linear regression model ('ordinary least squares' or OLS) and its major assumptions, the remainder of the course explores how to test each of the underlying assumptions and how to correct the model when the assumptions fail. Topics include least squares regression, multicollinearity, heteroscedasticity, autocorrelation, and hypothesis testing. We will use Stata software in the course for examples and assignments.

Prerequisites

There are three prerequisites to HPM 881:
1. Bios 600 or graduate-level equivalent
2. Math background in topics such as matrix algebra and calculus. Students should understand terms such as determinants, rank, logarithm, exponential function, and be able to multiply matrices and take derivatives of polynomials, logs, and exponential functions.
3. Working knowledge of Stata. Regression based commands will be covered in 881, but students entering the class should be able to read in raw data, define variables, merge data sets and conduct basic descriptive statistics.

HPM 880 satisfies all three prerequisites.

Instructor

Justin Trogdon, PhD
Associate Professor
Department of Health Policy and Management
McGavran-Greenberg 1101-B
Phone: 919-962-2433
Email: trogdonj@email.unc.edu

Teaching Assistant

Yunwei Chen
[Office location TBA]
Email: ywchen@live.unc.edu

Course Website

https://sakai.unc.edu/welcome/. Use your ONYEN and password. This will be the primary resource for course-related communication. Course materials including this syllabus, presentations, assignments, optional readings, and announcements will be posted periodically. Students should check the website regularly.

Class Days, Times, Location

Mondays and Wednesdays, 2:30p – 4:25p,
Rosenau 228

Office Hours

Dr. Trogdon will hold office hours Wednesdays 1-2p. The TA will hold office hours Tuesdays 9:30-10:30a. Students may also request an appointment by email.

Course Texts   There is no required textbook.

I highly recommend "Wooldridge, Jeffrey M. 2016. Introductory Econometrics: A Modern Approach. Cenage Learning." For those of you in a PhD program or planning on taking 882 and 883, you should probably consider Wooldridge required. For those of you who will not take another statistics course, you can probably get by with another regression textbook, although one using an "econometric" perspective will likely prove more useful. For example, "Kennedy, Peter. 2008. A Guide to Econometrics. Wiley-Blackwell" has an intuitive treatment of the subjects covered in this course.

The instructor's Lecture Notes will be made available to students through the class web page on Sakai.

Additional readings from recent journal articles may be required for the course. I will provide you the citation for these readings and whenever possible will use readings available electronically through the UNC library.

Manuals for Stata (StataCorp. Stata Statistical Software: Release 15.0 College Station, TX: Stata Corporation. (www.stata.com) are available in various computer labs and on-line. Students who plan to use Stata from home or on a regular basis are encouraged to purchase their own copy. Students can get this at a substantial discount by purchasing through UNC's Software Acquisition. Again, if this is the last course, you can probably get by with a one-year purchase, but those in PhD or 882/883 will want to buy the perpetual license. Explanations of many of the procedures and tests are surprisingly good. Virtual Lab is a particularly attractive option based on its price, but reliable access can be an issue.

Web Sources: Among the many useful websites is that of the Medical Expenditure Panel Survey (MEPS), which will be used for class examples and homework assignments. These data are available for free from the Agency for Healthcare Research and Quality (AHRQ) through their web site (https://meps.ahrq.gov/mepsweb/). The MEPS is a very comprehensive (and complicated) dataset that may be useful to students as they write papers even beyond this class.

Course Format   This is primarily a lecture-based course. The lecture will be supplemented with small group in-class exercises, case studies, and examples from the public health literature. Small group exercises may be both critical, in terms of analyzing manuscripts, and constructive, in terms of developing estimates. Small group exercises are intended to give students further practice applying the course materials and experience working in a research team.

Course Policies and Resources

Recognizing, Valuing and Encouraging Inclusion and Diversity in the Classroom

We share the School`s commitment to diversity. We are committed to ensuring that the School is a diverse, inclusive, civil and welcoming community. Diversity and inclusion are central to our mission — to improve public health, promote individual well-being and eliminate health inequities across North Carolina and around the world. Diversity and inclusion are assets that contribute to our strength, excellence and individual and institutional success. We welcome, value and learn from individual differences and perspectives. These include but are not limited to: cultural and racial/ethnic background; country of origin; gender; age; socioeconomic status; physical and learning abilities; physical appearance; religion; political perspective; sexual identity and veteran status. Diversity, inclusiveness and civility are core values we hold, as well as characteristics of the School that we intend to strengthen.

We are committed to expanding diversity and inclusiveness across the School—among faculty, staff, students, on advisory groups, and in our curricula, leadership, policies and practices. We measure diversity and inclusion not only in numbers, but also by the extent to which students, alumni, faculty and staff members perceive the School's environment as welcoming, valuing all individuals and supporting their development."

In this class, we practice these commitments in the following ways:

- Develop classroom participation approaches that acknowledge the diversity of ways of contributing in the classroom and foster participation and engagement of *all* students.
- Structure assessment approaches that acknowledge different methods for acquiring knowledge and demonstrating proficiency.
- Encourage and solicit feedback from students to continually improve inclusive practices.

As a student in the class, you are also expected to understand and uphold the following UNC policies:

- **Diversity and Inclusion at the Gillings School of Global Public Health**: http://sph.unc.edu/resource-pages/diversity/
- **UNC Non-Discrimination Policies**: http://policy.sites.unc.edu/files/2013/04/nondiscrim.pdf
- **Prohibited Discrimination, Harassment, and Related Misconduct at UNC**: https://deanofstudents.unc.edu/incident-reporting/prohibited-harassmentsexual-misconduct

Accessibility	UNC-CH supports all reasonable accommodations, including resources and services, for students with disabilities, chronic medical conditions, a temporary disability, or a pregnancy complication resulting in difficulties with accessing learning opportunities. All accommodations are coordinated through the UNC Office of Accessibility Resources & Services (ARS), https://ars.unc.edu/; phone 919-962-8300; email ars@unc.edu. Students must document/register their need for accommodations with ARS before accommodations can be implemented.
Counseling and Psychological Services	CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Go to their website: https://caps.unc.edu or visit their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more.
UNC Honor Code	As a student at UNC-Chapel Hill, you are bound by the university's Honor Code, through which UNC maintains standards of academic excellence and community values. It is your responsibility to learn about and abide by the code. All written assignments or presentations (including team projects) should be completed in a manner that demonstrates academic integrity and excellence. Work should be completed in your own words, but your ideas should be supported with well-cited evidence and theory. To ensure effective functioning of the Honor System at UNC, students are expected to:

    a. Conduct all academic work within the letter and spirit of the Honor Code, which prohibits the giving or receiving of unauthorized aid in all academic processes.

    b. Learn the recognized techniques of proper attribution of sources used in written work; and to identify allowable resource materials or aids to be used during completion of any graded work.

    c. **Sign a pledge on all graded academic work certifying that no unauthorized assistance has been received or given in the completion of the work.**

    d. Report any instance in which reasonable grounds exist to believe that a fellow student has violated the Honor Code.

Instructors are required to report suspected violations of the Honor Code, including inappropriate collaborative work or problematic use of secondary materials, to the Honor Court. Honor Court sanctions can include receiving a zero for the assignment, failing the course and/or suspension from the university. If you have any questions about your rights and responsibilities, please consult the Office of Student Conduct at https://studentconduct.unc.edu/, or consult these other resources:

- Honor system module.
- UNC library's plagiarism tutorial.
- UNC Writing Center handout on plagiarism.

Instructor Expectations

Overall

The instructor's primary role is to help create an effective learning environment, act as an expert resource, help you find other resources, and facilitate learning.

Email

The instructor will typically respond to email within 24 hours or less if sent Monday through Friday. The instructor may respond to weekend emails, but it is not required of them. If you receive an out of office reply when emailing, it may take longer to receive a reply. The instructor will provide advance notice, if possible, when they will be out of the office.

Feedback

Exams and paper drafts will receive written feedback that coincides with the assessment rubric. Feedback is meant to be constructive and help the student continue to build upon their skills. The types of feedback you may receive are descriptive feedback, evaluative feedback, and motivational feedback. Feedback is a tool that you as a learner can use to understand the areas that you are succeeding in and what you can do to improve in other areas.

Grading

Assignments will be graded no more than two weeks after the due date. Early submissions will not be graded before the final due date.

Syllabus Changes

The instructor reserves to right to make changes to the syllabus, including project due dates and test dates. These changes will be announced as early as possible.

Telephone Messages

The instructor will respond to telephone messages within 24 hours Monday through Thursday. Calls left on a Friday will be returned that day if possible; if not, they will be returned on Monday. Email is the preferred means of communication.

Student Expectations

Appropriate Use of Course Resources:

The materials used in this class, including, but not limited to, syllabus, exams, quizzes, and assignments are copyright protected works. Any unauthorized copying of the class materials is a violation of federal law and may result in disciplinary actions being taken against the student. Additionally, the sharing of class materials without the specific, express approval of the instructor may be a violation of the University's Student Honor Code and an act of academic dishonesty, which could result in further disciplinary action. This includes, among other things, uploading class materials to websites for the purpose of sharing those materials with other current or future students.

Assignments

Submit all assignments through Sakai or in person during class. Emailing assignments is not acceptable unless prior arrangements have been made. If you are having issues submitting assignments, try a different web browser first. If switching browsers does not work, email or call the instructor for guidance.

Attendance/ Participation

Your attendance and active participation are an integral part of your learning experience in this course. If you are unavoidably absent, please notify the course instructor (and Teaching Assistant).

Communication

You are expected to follow common courtesy in all communication to include email, discussion boards, and face-to-face. All electronic communications sent

should follow proper English grammar rules to include complete sentences. This is a professional course, and you are expected to communicate as a professional.

Be sure to check that the email address Sakai has for you is correct as this will be the primary way I will communicate with you outside of class. I assume you will check your email at least once per day for messages. Note that you may need to change your email address in both Sakai and the On-Line Campus Directory.

Contributions   You are expected to offer individual contributions in class and on individual assignments, and collaborate with fellow students on assignments for which students may work together, such as group assignments.

Email   All email correspondence between student/instructor and peer/peer will be conducted in a professional manner following email etiquette.

- View the following link for more information on email etiquette: http://metropolitanorganizing.com/etiquette-professional-organizing-services/essential-email-etiquette-tips/

Late Work   Late work will not be accepted.

Readings   Readings for a particular class should be completed before the class session and before completing associated activities

Technical support   The UNC Information Technology Services (ITS) department provides technical support 24-hours per day, seven days per week. If you need computer help, please contact the ITS Help Desk by phone at +1-919-962-HELP (919-962-4357), or by email at help@unc.edu, or by visiting their website at http://help.unc.edu, or by UNC Live Chat at http://its.unc.edu/itrc/chat.

Competencies, Learning Objectives, and Assessment

Map
Competencies taught in this course, learning objectives mapped to these competencies, and assignments that assess attainment of these competencies.

Competency	Learning Objectives	Assessment Assignment
Understand and appropriately apply analytical strategies used in health services/health policy research	1. To understand the regression framework and the basic assumptions of ordinary least squares	Midterm 1 and Final Exam
	2. To estimate and interpret multiple regression models	Midterm 2, Final Exam and Empirical Paper
	3. To test model specification	Final Exam and Empirical Paper
	4. To conduct original research using OLS, including describing a data set, development and testing of an appropriate OLS model, and the coherent, concise description of the results	Empirical Paper

Course Assignments and Assessments

This course will include graded assignments and/or exams.

Assignments	Points/Percentages
1. Problem sets	10
2. Midterm 1	20
3. Midterm 2	20
4. Final Exam	30
5. Empirical paper	20
TOTAL	100

Grading Scale

Final course grades will be determined using the following UNC Graduate School grading scale. The relative weight of each course component is shown in the table above.

H	Greater than or equal to 90	High Pass: Clearly excellent graduate work
P	Greater than or equal to 70	Pass: Entirely satisfactory graduate work
L	Greater than or equal to 60	Low Pass: Inadequate graduate work
F	Less than 60	Fail

Assignment Descriptions

Descriptions of the assignments follow.

1. Problem Sets (10%)

Problem sets will be required throughout the course. Answers to problems will be reviewed as warranted. Unless otherwise indicated, assignments for the class are due at the beginning of class.

We reserve the right to select only a subset of questions from each problem set for grading. Questions on problem sets will be graded on a two-point scale:

   2 = perfect answer
   0 = incorrect or mostly incorrect
   1 = somewhere in between

2. Midterm 1 (20%)

Midterm exams will be taken individually (not in teams). Midterm 1 will be scheduled during class time. Please make sure you know the dates of this exam; alternate dates will only be scheduled in cases of serious illness. Midterm 1 will assess Learning Objective #1. Students are expected to be able to recognize and explain the basic assumptions of OLS in a multivariate regression context. Explanations should clearly state each assumption, using math where appropriate, describe the importance of the assumption in plain English, and provide examples for violations of the assumption. Midterm 1 can include true/false, short answer and brief essay questions to assess Learning Objective #1 and other topics covered in lectures prior to the exam.

3. Midterm 2 (20%)

Midterm exams will be taken individually (not in teams). Midterm 2 will be scheduled during class time. Please make sure you know the dates of this exam; alternate dates will only be scheduled in cases of

serious illness. Midterm 2 will assess Learning Objective #2. Students are expected to be able to derive estimators for regression coefficients and standard errors and to interpret coefficients and standard errors from multivariate regressions. Derivations and interpretations should use math where appropriate and explain the meaning of the estimates in plain English. Midterm 2 can include true/false, short answer and brief essay questions to assess Learning Objective #2 and other topics covered in lectures after Midterm 1 and prior to Midterm 2.

4. Final Exam (30%)

The final exam will be cumulative. The final will occur during the scheduled time on the University's final exam schedule. Also, identify the final exams for all your cases and assess for conflict by the second week of the semester. The final exam will assess Learning Objectives #1-3. In addition to the expectations for Learning Objectives #1-2 described above for the midterm exams, students are expected to identify and apply the appropriate model specification tests for basic assumptions of OLS, where possible. Students should be able to explain the purpose of the test in plain English, provide the mathematical formula for the test, and interpret the results by selecting the appropriate model specification among several options. The final exam can include true/false, short answer and brief essay questions to assess Learning Objectives #1-3 and other topics covered in lectures after Midterm 2.

5. Empirical Paper (20%)

Each student will write an empirical paper that will use an OLS regression model and specification tests learned during the semester. The empirical paper will assess Learning Objectives #2-4. In addition to the expectations for Learning Objectives #2-3 described above for the final exam, students are expected to conduct original research using OLS, including describing a data set, development and testing of an appropriate OLS model, and the coherent, concise description of the results. Please refer to the grading rubric for specific expectations to demonstrate the competency. Two intermediary deliverables will be due on the dates of the midterm exams: a one-page proposal and Table 1 of descriptive statistics of key variables. Both intermediary deliverables will be scored using the same 2-point scale as the problem set questions. Each intermediary deliverable will count as a problem set for the final grade. Students should start looking now for an interesting empirical question to answer and a data set that can be used to answer that question (the MEPS is strongly encouraged for this purpose, but other data sources are allowed).

Course-at-a-Glance

The instructor reserves to right to make changes to the syllabus, including project due dates and test dates. These changes will be announced as early as possible.

Date/Session/Week/Unit	Topic and Competency	Assignment Due
Wed, 1/9; Unit 1	Course Introduction – Structure, expectations and assignments; MEPS overview	
Mon, 1/14; Unit 2	Introduction to Regression Analysis	
Wed, 1/16; Unit 3	Multiple Regression	
Mon, 1/21	NO CLASS – MLK Day	
Wed, 1/23; Unit 4	Properties and Assumptions of OLS	PS 1 Due
Mon, 1/28; Unit 4	Properties and Assumptions of OLS	
Wed, 1/30; Unit 5	Interpretation of OLS	PS 2 Due
Mon, 2/4; Unit 5	Interpretation of OLS	
Wed, 2/6; Unit 6	Specification Tests	PS 3 Due
Mon, 2/11; Unit 6	Specification Tests	
Wed, 2/13	MIDTERM 1	PROPOSAL DUE
Mon, 2/18; Unit 7	Multicollinearity	
Wed, 2/20; Unit 7	Multicollinearity	
Mon, 2/25; Unit 8	Heteroskedasticity	
Wed, 2/27; Unit 8	Heteroskedasticity	PS 4 Due
Mon, 3/4; Unit 9	Linear Probability Model	
Wed, 3/6; Unit 9	Linear Probability Model	PS 5 Due
Mon, 3/11	NO CLASS – SPRING BREAK	
Wed, 3/13	NO CLASS – SPRING BREAK	

Mon, 3/18	Catch up / recovery	
Wed, 3/20	Catch up / recovery	
Mon, 3/25; Unit 10	Transformations	
Wed, 3/27; Unit 10	Transformations	PS 6 Due
Mon, 4/1; Unit 11	Autocorrelation	
Wed, 4/3; Unit 11	Autocorrelation	PS 7 Due
Mon, 4/8	MIDTERM 2	PAPER TABLE 1 DUE
Wed, 4/10; Unit 12	Measurement Error	
Mon, 4/15; Unit 12	Measurement Error	
Wed, 4/17; Unit 13	Endogeneity	PS 8 Due
Mon, 4/22; Unit 13	Endogeneity	
Wed, 4/24	Wrap up / loose ends	PAPER DUE
Tue, 5/7; 8a-11a	**FINAL EXAM**	**Location TBA**

Course Schedule

The instructor reserves to right to make changes to the syllabus, including project due dates and test dates. These changes will be announced as early as possible.

Session 1	
Topic	Unit 1: Course Introduction
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares
Required Readings	Wooldridge (5th Ed) - Appendix A: basic math tools – e.g. summation notation, functions, logs, derivatives, etc. - Appendix B: probability – e.g. random variables, distributions, etc. - Appendix C: statistics – e.g. populations, parameters, sampling, variance, bias, etc. - Appendix D: basic matrix algebra – e.g. definitions, operations, etc.
Assignments/Deadlines	None

Session 2	
Topic	Unit 2: Introduction to Regression Analysis
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares
Required Readings	Wooldridge (5th Ed) - Chapter 2: Section 2.1, 2.2 - Appendix C: C.4 (high level overview of method of moments, maximum likelihood, least squares)
Assignments/Deadlines	None

Session 3	
Topic	Unit 3: Multiple Regression
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research

Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares
Required Readings	Wooldridge (5th Ed) - Chapter 3: 3.1, 3.2 (through p. 74) - Appendix E: E.1 (linear regression in matrix form)
Assignments/Deadlines	None

Session 4	
Topic	Unit 4: Properties and Assumptions of OLS
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares
Required Readings	Wooldridge (5th Ed) - Chapter 2: 2.5 (simple regression) - Chapter 3: 3.3, 3.4, 3.5 (multiple regression)
Assignments/Deadlines	Problem Set 1 due

Session 5	
Topic	Unit 4: Properties and Assumptions of OLS
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares
Required Readings	Wooldridge (5th Ed) - Chapter 2: 2.5 (simple regression) - Chapter 3: 3.3, 3.4, 3.5 (multiple regression)
Assignments/Deadlines	None

Session 6	
Topic	Unit 5: Interpretation of OLS
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research

Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models
Required Readings	Wooldridge (5th Ed) - Chapter 2: 2.3, 2.4 - Chapter 3: 3.2 - Chapter 6: 6.3 (more on goodness-of-fit) - Chapter 7: 7.1-7.4 (through p.245)
Assignments/Deadlines	Problem Set 2 due

Session 7	
Topic	Unit 5: Interpretation of OLS
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models
Required Readings	Wooldridge (5th Ed) - Chapter 2: 2.3, 2.4 - Chapter 3: 3.2 - Chapter 6: 6.3 (more on goodness-of-fit) - Chapter 7: 7.1-7.4 (through p.245)
Assignments/Deadlines	None

Session 8	
Topic	Unit 6: Specification Tests
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	3. To test model specification
Required Readings	Wooldridge (5th Ed) - Chapter 4 - Chapter 7: 7.4 (p.245-248)
Assignments/Deadlines	Problem Set 3 due

Session 9	
Topic	Unit 6: Specification Tests
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	3. To test model specification
Required Readings	Wooldridge (5th Ed) - Chapter 4 - Chapter 7: 7.4 (p.245-248)
Assignments/Deadlines	None

Session 10	
Topic	Midterm 1
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares
Required Readings	None
Assignments/Deadlines	None

Session 11	
Topic	Unit 7: Multicollinearity
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models
Required Readings	None
Assignments/Deadlines	None

Session 12	
Topic	Unit 7: Multicollinearity

Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models
Required Readings	None
Assignments/Deadlines	None

Session 13	
Topic	Unit 8: Heteroskedasticity
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Wooldridge (5th Ed) - Chapter 8: 8.1-8.4
Assignments/Deadlines	Problem Set 4 due

Session 14	
Topic	Unit 8: Heteroskedasticity
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Wooldridge (5th Ed) - Chapter 8: 8.1-8.4
Assignments/Deadlines	None

Session 15	

Topic	Unit 9: Linear Probability Model
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Woodridge (5th Ed) - Chapter 7: 7.5
Assignments/Deadlines	Problem Set 5 due

Session 16	
Topic	Unit 9: Linear Probability Model
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Woodridge (5th Ed) - Chapter 7: 7.5
Assignments/Deadlines	None

Session 17	
Topic	Catch up / recovery
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	None
Assignments/Deadlines	Problem Set 6 due

Session 18	
Topic	Catch up / recovery

Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	None
Assignments/Deadlines	None

Session 19	
Topic	Unit 10: Transformations
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Woodridge (5th Ed) - Chapter 6: 6.2 (through p. 194)
Assignments/Deadlines	Problem Set 7 due

Session 20	
Topic	Unit 10: Transformations
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Woodridge (5th Ed) - Chapter 6: 6.2 (through p. 194)
Assignments/Deadlines	None

Session 21	
Topic	Unit 11: Autocorrelation
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research

Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Woodridge (5th Ed) - Chapter 12: 12.1, 12.2 (note that some of the discussion related to time-series models may be beyond the scope of this class)
Assignments/Deadlines	Problem Set 8 due

Session 22	
Topic	Unit 11: Autocorrelation
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	Woodridge (5th Ed) - Chapter 12: 12.1, 12.2 (note that some of the discussion related to time-series models may be beyond the scope of this class)
Assignments/Deadlines	None

Session 23	
Topic	Midterm 2
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	2. To estimate and interpret multiple regression models
Required Readings	None
Assignments/Deadlines	None

Session 24	
Topic	Unit 12: Measurement Error

Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	None
Assignments/Deadlines	None

Session 25	
Topic	Unit 12: Measurement Error
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification
Required Readings	None
Assignments/Deadlines	None

Session 26	
Topic	Unit 13: Endogeneity
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 3. To test model specification
Required Readings	None
Assignments/Deadlines	Problem Set 9 due

Session 27	
Topic	Unit 13: Endogeneity

Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 3. To test model specification
Required Readings	None
Assignments/Deadlines	None

Session 28	
Topic	Wrap up / loose ends
Competency Addressed	Understand and appropriately apply analytical strategies used in health services/health policy research
Learning Objective(s)	1. To understand the regression framework and the basic assumptions of ordinary least squares 2. To estimate and interpret multiple regression models 3. To test model specification 4. To conduct original research using OLS, including describing a data set, development and testing of an appropriate OLS model, and the coherent, concise description of the results
Required Readings	None
Assignments/Deadlines	Empirical Paper due

Appendix: Rubrics

Assignment Rubric: Midterm 1, Midterm 2, and Final Exam (20%, 20%, 30%)

Each exam can consist of true/false, short answer, and brief essay questions. Point values for each question will be provided on the exam. The table below describes the rubric for each type of question.

Criteria	Fully Met	Partially Met	Not met
True/False and Explain	(90-100%) • Correct answer for T/F • Complete and clear explanation for T/F answer, including math where needed	(70-80%) • Correct answer for T/F • Mostly complete explanation	(0-60%) • Incorrect answer for T/F • Incomplete explanation
Short Answer	(90-100%) • Correctly define terms • Clearly and concisely interpret regression output (e.g., marginal effects or statistical tests) • Identify potential limitations of regression models and provide solutions where applicable	(70-80%) • Mostly correct definition of terms • Mostly correct interpretation of regression output • Identify most potential limitations of regression models and provide solutions where applicable	(0-60%) • Incorrectly define terms • Incorrectly interpret regression output • Identify only a few potential limitations of regression models and fail to provide solutions where applicable
Brief Essay	(90-100%) • Identify key concepts required to address the question • Correctly define terms • Discuss the OLS assumptions involved in the question • Clearly describe the steps you would take to correctly specify the regression model	(70-80%) • Identify most key concepts required to address the question • Mostly correct definition of terms • Mostly complete discussion of the OLS assumptions involved in the question • Description of the steps you would take to correctly specify the regression model covers most required steps	(0-60%) • Missed several key concepts required to address the question • Incorrectly define terms • Incomplete or incorrect discussion the OLS assumptions involved in the question • Missed several key steps to correctly specify the regression model

Assignment Rubric: Empirical Paper (20%)

The paper for 881 must:
1) answer a specific interesting question,
2) include an abstract,
3) analyze data using an ordinary least squares model,
4) conduct certain specification tests (listed below), and
5) not exceed 10 pages of text (double-spaced, font 10-12, and 1-inch margins; cover page, abstract, references, and tables are not part of the 10-page limit).

Papers should have the following sections in this order:
1) A structured abstract, which should follow the format required for Health Services Research, and be of no more than 200 words. (Note: for those students not focusing on health policy, you can model your abstract on another journal's style. Please indicate which journal format you are using).
2) Introduction: explain the question to be answered and why it is important. For this assignment, voluminous citations of the literature are not only unnecessary, but may crowd out more important empirical sections. I strongly encourage you to limit your background research to a single motivating paragraph which explains the importance of the problem.
3) Conceptual model: the conceptual model is not the same as the statistical model, which is described in the Methods section below. The conceptual model usually takes a disciplinary perspective. This section should give a brief description of how you approached the problem (e.g., "This paper uses social control theory to examine the effect that living in a community has on health care services use;" "This paper examines the demand function for mental health services by examining the role of price, income, and price of competing services"). This section should not describe the empirical model, the data, nor the variables. For Masters students who may not have had disciplinary training, this section should be a brief paragraph explaining the reasoning behind your model set up (e.g. "region of the country has been shown to be an important predictor of health care use. Some regions have more medical centers or greater access to health insurance, which translate to greater health care expenditures.").
4) Methods, including results of
   a. an examination of multicollinearity and
   b. a test for either heteroskedasticity or autocorrelation.
   c. You should also discuss other tests as appropriate (e.g., tests of functional form). Note that these tests may cause your final model to be different from the initial model, which is why they belong in the Methods section, rather than the Results section. You'll want to consider the possibility of all violations of the basic assumptions whether you provide tests or not (these could appear in Methods or Limitations, generally depending on whether you can do something about them (Methods) or not (Limitations).
5) Data: discuss your data carefully, especially how the sample was chosen. Define the dependent and independent variables, why they were selected, why transformations were or were not used, and include a table of summary statistics. If you use MEPS, you can assume that the reader is familiar with the MEPS Full Year Characteristics File and keep your description of this data to a minimum. If you use other data sets (including other files in the MEPS) you will want to briefly describe the data. If your analysis file is only going to include a subset of the original data set (e.g., due to missing data for some variables or because your focus is on a subpopulation, such as the elderly), you should assess the representativeness of the subset used or describe any imputation done in this section.

6) Results:
   a. Generally, Results sections start with a description of the empirical characteristics of the variables you described earlier. This generally means you should have a Table 1 of variable means/percents, and standard deviations for continuous variables only. Keep this section brief; do not verbalize every line in the table. If your key variable is binary or categorical, you may want to present your Table 1 results separately by the key variable (e.g., first column is "overall mean," second column is "mean for those with private insurance," and the third column is "means for those without private insurance" if private insurance is your key independent variable). In this case, tests (such as t-tests or chi-squared tests) are often conducted. Feel free to do these, especially if you are interested in publishing your work, but these tests are not covered and thus are not required.
   b. Discuss the final regression model results, answering the question posed at the beginning. In particular, you must do four tests in the term paper:
      i. a formal test of main hypothesis;
      ii. one other interesting hypothesis test (e.g. F-test, Chow test, linear restriction);
      iii. a test for multicollinearity; and
      iv. a test for either heteroskedasticity or autocorrelation.

      Be clear about the test you are conducting and its null hypothesis and be precise about the results. You should also give the marginal effect for the key variable(s), noting that it may be a formula if you have used a non-linear functional form or interactions. Other regressions that are run for things like residual analysis, removal of outliers, etc., do not have to be reported in tables, but should be mentioned briefly in Methods and Results. However, if you want to draw the reader's attention to a finding of interest (for example, the results on the split samples from a Chow Test, if the differences were significant and interesting) you may include other tables.
   c. Do a prediction for some person/people of interest. If you logged your dependent variable, be sure to give the predicted amount in the original scale (e.g., dollars). If you have used a two-part model, do the prediction for both parts, separately.
7) Discussion/Conclusion: Discuss the implications of your findings on health care policy. Do not use this section to reword/summarize your Results section. Be sure to include a paragraph with a brief discussion of problems/limitations, both in the original data set and in your analysis.
8) Tables should be in the appendix, not in the text.
9) The appendix should contain the complete Stata code for generating your analyses starting from the raw data set (only include the do file, not the log file, and include only the analyses relevant to your final paper. This section should be single spaced, please).

A one-page summary is due on mid semester. The summary should indicate the following:
1) The data set you will use for your analysis (if not using MEPS, mention whether you have this data set in hand or not)
2) The continuous dependent variable you will analyze
3) The key independent variable (note, in rare cases there may be more than one key variable)
4) Whether your paper will be subject to IRB approval

Criteria	Fully Met	Partially Met	Not met
Style (10 points)	(9-10 points) • An abstract which is accurate and interesting • An introduction that motivates the topic • No spelling or grammatical errors, especially of the type discussed in the Style Guide • A conceptual model which motivates the topic	(7-8 points) • An abstract which is accurate but not focused • An introduction that provides background but too broad or narrow • One or two spelling or grammatical errors • A conceptual model that is internally consistent even if not well-matched to topic	(0-6 points) • An incomplete and/or disjointed abstract • An introduction that is unfocused and/or off-topic • Multiple spelling or grammatical errors • Lack of a conceptual model that is internally consistent and tied to topic
Data and Sample (15 points)	(13-15 points) • A concise section describing the data and sample well (i.e., reader could replicate) • A table of descriptive statistics which is easy to read • A brief discussion of the key variables of interest	(10-12 points) • Mostly complete description of the data and sample, potentially a little wordy • A table of descriptive statistics organized logically • Discussion of the key variables of interest is complete but not concise	(0-9 points) • An incomplete or overly long description of the data and sample • A table of descriptive statistics that is confusing or hard to read • An incomplete or overly long and unfocused discussion of the key variables of interest
Empirical Model (20 points)	(18-20 points) • A clear, concise section describing the regression model, using math as needed • Appropriate variables in the model • A clear discussion of the form of the dependent variable, whether it was transformed or not, and why • A discussion of the key explanatory variable, its hypothesized influence, and a discussion of the functional forms used and why they were chosen	(14-17 points) • Adequate section describing the regression model • Majority of appropriate variables in the model • A discussion of the form of the dependent variable that is mostly complete in its detail and motivation • A discussion of the key explanatory variable that is mostly complete in its detail and motivation	(0-13 points) • Section describing the regression model is incomplete or unclear • Model missing appropriate variables or including inappropriate variables • Incomplete discussion of the form of the dependent variable, whether it was transformed or not, and why • Incomplete discussion of the key explanatory variable

Criteria	Fully Met	Partially Met	Not met
Results (20 points)	(18-20 points) • A table of results that is easy to read and contains all relevant information • Correct interpretation of coefficients and significance tests • Correct indication of the marginal/differential effect of key variable(s) • A prediction calculated from the regression model, which is well described, accurate, and interesting	(14-17 points) • A table of results that is organized logically and mostly complete regarding relevant information • Mostly correct interpretation of coefficients and significance tests • Mostly correct indication of the marginal/differential effect of key variable(s) • A prediction calculated from the regression model, which is well described and accurate	(0-13 points) • A table of results that is hard to read and does not contain all relevant information • Incorrect interpretation of coefficients and significance tests • Incorrect indication of the marginal/differential effect of key variable(s) • A prediction calculated from the regression model, which is not well described or inaccurate
Specification Tests (20 points)	(18-20 points) • Four (or more) tests as required which test interesting and relevant hypotheses • A clear statement of each of the tests conducted, the null hypothesis of each test, and the result, conclusion, and remedy (if appropriate)	(14-17 points) • Three tests as required which test hypotheses • A mostly complete statement of each of the tests conducted, the null hypothesis of each test, and the result, conclusion, and remedy (if appropriate)	(0-13 points) • Two or fewer tests as required which test hypotheses • An incomplete statement of each of the tests conducted (i.e., missing the null hypothesis of each test, the result, conclusion, and remedy)
Discussion / Conclusion (15 points)	(13-15 points) • An interesting discussion of the key results from the model and their implications • A brief mention of any other violations of the core assumptions that may affect the interpretation of the results and other relevant limitations	(10-12 points) • A discussion of the key results from the model • A mostly complete discussion of any other violations of the core assumptions that may affect the interpretation of the results and other relevant limitations	(0-9 points) • Incomplete or unorganized discussion of the key results from the model • An incomplete discussion of any other violations of the core assumptions that may affect the interpretation of the results and other relevant limitations