

**PPA 207 – QUANTITATIVE METHODS**

**MASTER’S PROGRAM IN PUBLIC POLICY AND ADMINISTRATION  
MASTER’S PROGRAM IN URBAN LAND DEVELOPMENT**

**CALIFORNIA STATE UNIVERSITY, SACRAMENTO**

**SPRING 2020**

Professor: Rob Wassmer, Ph.D.

E-Mail: [rwassme@csus.edu](mailto:rwassme@csus.edu)

Twitter: @rwassme

Home Page: <https://www.csus.edu/faculty/w/rwassme/>

Office: 230 Sac State Downtown Center (3<sup>rd</sup> and S Streets)

Office Phone: (916) 278-6304 (campus) or (916) 278-4556 (downtown)

Office Hours: Monday and Thursday, 4:00 - 5:30 p.m. and by appointment if necessary

Class Location: Monday, 6 - 8:50 p.m., 110 Sac State Downtown

Texts:

Please purchase all texts and complete the assigned reading for our first meeting. I expect you to do all the assigned reading before each meeting. Also, bring the appropriate text(s) to class on the dates covered and your laptop with EXCEL and STATA loaded on it.

(1 - Required) *Applied Statistics Using STATA* (2017), Mehmet Mehmetoglu and Georg Jakobsen, Sage Press; purchase at [Amazon](https://www.amazon.com) or other internet bookseller (used~\$30).

(2 - Required) *Naked Statistics* (2011), Charles Wheelan, Norton Press; purchase at [Amazon](https://www.amazon.com) or other internet bookseller (used ~ \$5).

(3 – Required) *Big Data: A Very Short Introduction* (2017), Dawn Holmes, Oxford Press, purchase at [Amazon](https://www.amazon.com) or other internet bookseller (Kindle ~ \$6).

(4 - Required) *STATA/IC 16 Grad Plan Statistical Package*, purchase at <http://www.stata.com/order/new/edu/gradplans/student-pricing>, based upon being a Sac State student. Prices are \$48/6 months and \$225/perpetual. Get it in time to load for first class. Unless you are certain you want a permanent copy, I suggest getting the 6-month version and updating to perpetual within the 6 months if you change your mind.

(5 – Should have from PPA 200 or 240A) *Writing Literature Reviews: A Guide for Students of the Social and Behavioral Sciences*, Jose Galvin, 4<sup>th</sup> Edition or later, Pycszak Publishing; purchase at [Amazon](https://www.amazon.com) or internet bookseller (used ~ \$10).

**You will also need a laptop computer upon which STATA and EXCEL loaded.**

Prerequisites:

If you have not taken a previous course in statistics, or it has been awhile since doing so, review this tutorial:

<http://www.cpp.edu/~dj Moriarty/b211/b211%20Basic%20Statistics%20Review%20-%20Part%20One.pdf> . I will also review this material through *Naked Statistics* and the first chapter of *Applied Statistics Using Stata*.

This class also requires the use of Excel spreadsheets. If you require some practice in this program, review this tutorial: <https://www.tutoriacomputer.com/microsoft-excel-2016-tutorial-free-and-online/> .

Data Sources:

Type	Date Set Name	Description	File Name & Location
Housing Sales	<i>Multiple Listing Service (MLS) Data</i>	Data on characteristics and selling price of all homes that sold in Sacramento County in Fall 2013	MLS_Data_PPA207_SP20 @ Canvas  (2019 data forthcoming)
CC Education	California Community College Chancellor’s Office <i>Data Mart</i>  and  <i>Scorecard</i>	These data sets used to answer questions posed by administrators, educators, parents, students, state leadership, and professional organizations by providing information on students, courses, student services, outcomes and faculty and staff	CC_Data_PPA207_SP20 @ Canvas  <a href="http://datamart.cccco.edu/datamart.aspx">http://datamart.cccco.edu/datamart.aspx</a>  <a href="https://scorecard.cccco.edu/scorecard.aspx">https://scorecard.cccco.edu/scorecard.aspx</a>
Health	<i>California Health Interview Survey (CHIS)</i>	Info on health risk behaviors, preventive health practices, and health care access primarily related to chronic disease and injury	<a href="http://healthpolicy.ucla.edu/chis/about/Pages/about.aspx">http://healthpolicy.ucla.edu/chis/about/Pages/about.aspx</a>  (create in class)
Various	<i>Inter-University Consortium for Political and Social Research (ICPSR)</i>	With over 700 member institutions and over a half million files of social science research, the ICPS is the world’s largest social science data archive	<a href="http://www.ssrn.org/data/icpsr_direct">http://www.ssrn.org/data/icpsr_direct</a>  (create in class)

The final assignment for this class is a regression-based research paper. Such a paper requires a data set with at least 300 or more observations on a dependent variable, and the various explanatory variables that you expect to cause variation in the dependent variable. The table above contains a description of data sets examined/used in class. If you have another data set

that you think feasible for your final paper, please discuss it with me in office hours during the first or second week of class as you will need to be putting it together ASAP in an Excel spreadsheet.

By no later than the eighth week of class (March 16), you will need to discuss with me the data set you intend use for your course paper. See the list above for potential databases. I suggest you begin exploring these databases immediately and think about a dependent variable of policy/administration relevance.

Canvas:

On Canvas, I will post a PowerPoint of material covered in each meeting and a description of homework (HW) that is due at the following meeting. The course notes in PowerPoint will be available the Friday evening before the class meets. The HW due the following Monday will be available the Wednesday evening before it is due. I will post all grades on Canvas.

Learning Objectives:

At the end of PPA 207, for a student attending all classes, and completing all assignments, my expectation is competency in these five learning goals:

(1) Understand and apply the analytic tool of statistical analysis, through regression analysis, to offer insights into a policy or administration concern.
(2) Learn the statistical knowledge and skills necessary to produce a credible regression analysis and pass judgement on others as credible.
(3) Learn how to locate, access, and transfer to a statistical analysis package, the data necessary to complete an informative to policy/administration regression analysis.
(4) Effectively review empirically based literature to assist and inform in the creation of a regression analysis.
(5) Practice writing a regression-based research study in a manner that is theoretically sound and understandable to a non-statistical audience.

The purpose of this course is to expose the MPPA or MSULD student to some of the empirical methods used in the analysis and formulation of government policies. These include: (1) types and sources of data, (2) descriptive statistics, (3) regression analysis and interpretation, (4) how to review the relevant empirical literature before beginning a study of this type, (5) some of the basic issues/problems that can arise in regression analysis, and (6) how to write up empirical results.

My goal is not to turn you into an expert on statistics and regression analysis, but I do wish to provide you with a working knowledge of the most basic applied techniques in these areas. Even if you never expect to apply these techniques directly in your anticipated career path, it is very

likely you will interpret and comment on reports that contain policy analyses based on these techniques.

### Method:

Pedagogy includes in-class lectures, discussions, and STATA-based exercises. In addition, I ask you to complete written answers to assigned (homework) HW exercises. There is a HW due the first night we meet. Check Canvas a week before we first meet to access it. Submit the week's homework assignment through Canvas only if you will be in class. Your top 11 HW grades count. Therefore, do not ask to submit HW early or late if you will not be in class. You may miss three HWs and your HW grade will not decline.

Each class meeting will begin with a discussion of HW in student groups and coming together as a class to share answers. After 90 minutes of class time, we will take a 15-minute break. I will devote class time to covering the use of the Excel and STATA computer packages. **Purchase and install STATA (and Excel if you do not have it) on your laptop before our first meeting and bring your laptop to class the first night (and every night).** There are limited outlets in the classroom, so keep it charged, or bring a power strip to share an outlet.

### Academic Honesty

When you do any writing for this class, or any class at Sacramento State, it is important that you are aware of what plagiarism is, and how its practice can become grounds for dismissal from the university. Details are available at <https://csus.libguides.com/plagiarism> . The following is also helpful:

*Plagiarism is a form of cheating. At Sacramento State plagiarism is the use of distinctive ideas or works belonging to another person without providing adequate acknowledgement of that person's contribution. Regardless of the means of appropriation, incorporation of another's work into one's own requires adequate identification and acknowledgement. Plagiarism is doubly unethical because it deprives the author of rightful credit and gives credit to someone who has not earned it. Acknowledgement is not necessary when the material used is common knowledge. Plagiarism at Sacramento State includes but not limited to:*

- 1. The act of incorporating into one's own work the ideas, words, sentences, paragraphs, or parts thereof, or the specific substance of another's work without giving appropriate credit thereby representing the product as entirely one's own. Examples include not only word-for-word copying, but also the "mosaic" (i.e., interspersing a few of one's own words while copying another's work), the paraphrase (i.e., rewriting another's work while still using the other's fundamental idea or theory); fabrication (i.e., inventing or counterfeiting sources), ghost-writing (i.e., submitting another's work as one's own) and failure to include quotation marks on material that is otherwise acknowledged; and*
- 2. Representing as one's own another's artistic or scholarly works such as musical compositions, computer programs, photographs, paintings, drawing, sculptures, or similar works.*

I will also use the *Turn-It-In* plagiarism check on Canvas for your final paper. So please run your paper through it before submitting final draft. Many times, plagiarism is inadvertent and this check flags potential sentences that when written, you make have not considered as potential plagiarism. Remember the job of an analyst is often to read the findings of others and put it in their own, more simplified, words for a client or the public. I desire that you to practice that here through your HW assignments and final paper.

Paper:

To receive a passing grade in this course, each student must complete a regression-based paper on the topic of his/her choosing (subject to my approval). Details on what to include in the paper is below. I will post a few examples of well-crafted previous student papers on Canvas.

Grades:

You should complete the homework assigned before the night it is due and participate in the group/class discussion on it. Failure to do this will result in a zero on that HW averaged in with the grades received on other HW. I will assign a grade to each of your written answers to HW and derive an overall average grade on your HW from your top 11 grades. **Label all HW answers with your name, date turned in, question asked, and type the answer – except for requested graphs or diagrams that can be hand drawn and turned in the night of class.**

Your final grade is based upon the following proportions:

Participation in class and visits to my office hours/class breaks	20%
Midterm Literature Review Assignment	20%
Final Regression Paper Assignment	30%
Average of 11 Weekly HW Grades	30%

Homework Scoring:

<b>Percent Correct</b>	<b>Letter Grade</b>	<b>Number Grade</b>
100-97	A+	4.3
96-93	A	4.0
92-89	A-	3.7
88-85	B+	3.3
84-81	B	3.0
80-77	B-	2.7
76-73	C+	2.3
72-69	C	2.0
68-65	C-	1.7
64-61	D	1.0
<61	F	0.0

## Schedule:

I strongly suggest that you attend all Monday meetings. If you do not attend, you will not be able to turn in the HW due that week (**no exceptions**). I will drop your three lowest HW grades to provide some flexibility. By **March 16** (preferably earlier), each of you will also need to visit my office or call me for a 10-minute discussion of the plan for your paper and the progress you have made on it.

There will be no midterm exam. Instead, I ask that you turn in by April 6 the six-page literature review that is part of your paper. Some details on this are below, more will be forthcoming in class. I will also use the homework exercises that are due at the start of each class meeting to judge your progress throughout the semester. The final is a paper that is due no later than May 11.

The readings out of Mehmetoglu and Jakobsen (M&J), Holmes (H), and Wheelan (W) are below. I also offer other supplemental sources of background information that should assist in your understanding of the material. See the companion website <https://study.sagepub.com/mehmetogluandjakobsen> for M&J that also suggests instructional videos for every chapter in the book. I reserve the right to add additional readings/videos accessed by the Internet or posted on Canvas.

*In the PowerPoint notes I will indicate the names of students responsible for being the “go to” for whom I will rely upon to help me stimulate class discussion. Thus, if so designated, make sure you are well versed in this material.*

### Meeting 1 – Monday, January 27 – Introduction to Statistics for Policy Analysis

Syllabus Review

(Nova Video) Prediction by Numbers

<https://www.pbs.org/wgbh/nova/video/prediction-by-the-numbers>

(Planet \$ Podcast) Wheelan on What Causes What?

<https://www.npr.org/transcripts/178635250>

(W) Introduction: Hate Calculus but Love Stats, and Chapter 1: What’s the Point

“Does the Likely Demographics of Affordable Housing Justify NIMBYism?” with Imaz Wahid,  
*Housing Policy Debate*, 2019, 29(2), 343-58.  
@ Canvas

### Meeting 2 – Monday, February 3 – Descriptive/Deceptive Statistics and STATA Interface

(W) Chapter 2: Descriptive Stats, and Chapter 3: Deceptive Description

(M&G) Chapter 1: Research and Statistics

(STATA You Tube)

STATA Interface <https://www.youtube.com/watch?v=EhqzAzr5ThI&feature=youtu.be>

(Cheat Sheet) Data Analysis with Stata

<https://www.stata.com/bookstore/statacheatsheets.pdf>

**Meeting 3 – Monday, February 10 – Correlation/Probability, and STATA Basics**

(W) Chapter 4: Correlation, and Chapter 5: Basic Probability

(M&G) Chapter 2, pp. 15-30: What is STATA?, Entering and Importing Data, and Data Management

(STATA You Tube)

Import Excel Data [https://www.youtube.com/watch?v=N5ZFgzN2\\_7c](https://www.youtube.com/watch?v=N5ZFgzN2_7c)

Analyze Data <https://www.youtube.com/watch?v=YAVq99iUTTI&feature=youtu.be>

Descriptive Stats <https://www.youtube.com/watch?v=kKFbnEWwa2s>

**Meeting 4 – Monday, February 17 – Probability/Data and Bivariate Inference in STATA**

(W) Chapter 6: Problems with Probability, and Chapter 7: Importance of Data

(M&G) Chapter 2, pp. 30-42: Descriptive Statistics and Graphs, and Bivariate Inferential Statistics

(STATA You Tube)

Scatterplots <https://www.youtube.com/watch?v=GhVGpe3lb3E>

Correlation <https://www.youtube.com/watch?v=o7ko844ff-g>

Tables/Crosstabs <https://www.youtube.com/watch?v=3WpMRtTNZsw>

One-Way ANOVA: <https://www.youtube.com/watch?v=XEFGGkFRdD4>

**Meeting 5 – Monday, February 24 – Central Limit Theorem/Inference and Lit Review Start**

(W) Chapter 8: Central Limit Theorem, and Chapter 9: Inference

(G) Chapter 1: Writing Reviews, Chapter 2: Considerations in Writing Reviews, Chapter 3: Selecting Topic, Chapter 4: General Guidelines, Chapter 5: Quant Literature, and Chapter 7: Building Tables, Chapter 8: Synthesizing Lit, and Chapter 9: Guidelines

**Meeting 6 – Monday, March 2 – Bivariate Linear Regression**

(M&G) Chapter 3: Simple (Bivariate) Regression

(STATA You Tube)

Simple Linear Regression <https://www.youtube.com/watch?v=HafqFSB9x70>

Intro to Margins (Continuous) <https://www.youtube.com/watch?v=L9-PWY79aVA>  
Intro to Margins (Categorical) <https://www.youtube.com/watch?v=XAG4CbIbH0k>

**Meeting 7 – Monday, March 9 -- Lit Review Finish and Other Data**

(G) Chapter 10: Coherent Essay, Chapter 11: Style, Chapter 12: Feedback, Chapter 13: Reference List, and Chapter 14: Self Editing

“Differences in California Community College Success Rate Across Campuses and Time: Do Policy Variables Matter?” with Meredith Galloway, Working Draft.  
@ Canvas

California Community College Chancellor’s Office Data Mart and Scorecard  
<http://datamart.cccco.edu/datamart.aspx>  
<https://scorecard.cccco.edu/scorecard.aspx>

Inter-University Consortium for Political and Social Research (ICPSR)  
[http://www.ssrlic.org/data/icpsr\\_direct](http://www.ssrlic.org/data/icpsr_direct)

**Meeting 8 – Monday, March 16 – Multiple Regression**

*Deadline for deciding upon data set/topic for final paper*

(W) Chapter 11: Regression Analysis

(M&G) Chapter 4: Multiple Regression

(Nature) Scientist Rise Up Against Statistical Significance  
<https://www.nature.com/articles/d41586-019-00857-9>

**Meeting 9 – Monday, March 23 – Regression Mistakes and Corrections**

(W) Chapter 12: Common Regression Mistakes, and Conclusion: Five Questions That Statistics Can Help Answer

(M&G) Chapter 7: Linear Regression Assumptions and Diagnostics

(STATA You Tube)  
Multicollinearity Diagnostics <https://www.youtube.com/watch?v=jo-es2s6Xpw>  
Heteroskedasticity Tests <https://www.youtube.com/watch?v=PS2X8kPCRkE>

**Meeting 10 – Monday, April 6 – Big Data and Dummy Variable Regression**

*Literature Review Due*

(Ted Talk) Big Data is Better Data <https://www.youtube.com/watch?v=8pHzROP1D-w>



(H) Chapter 1: Data Explosion, Chapter 2: Why Special, Chapter 5: Medicine,  
Chapter 6: Business, Chapter 7: Security, and Chapter 8: Society  
@Canvas

(M&G) Chapter 5: Dummy-Variable Regression

**Meeting 11 – Monday, April 13 – Big Data Concerns and Regression Interaction Effects**

(Econ Talk Podcast) Weapons of Math Destruction  
<https://www.econtalk.org/cathy-oneil-on-weapons-of-math-destruction/>

(M&G) Chapter 6: Interaction-Moderation Effects Using Regression

(STATA You Tube)  
Intro to Margins (Interactions) [https://www.youtube.com/watch?v=43uX4D\\_7uaI](https://www.youtube.com/watch?v=43uX4D_7uaI)

**Meeting 12 – Monday, April 20 – Logistic Regression**

(M&G) Chapter 8: Logistic Regression

(STATA You Tube)  
Logistic Binary Predictors <https://www.youtube.com/watch?v=rSU1L3-xRk0>  
Logistic Continuous Predictors [https://www.youtube.com/watch?v=vmZ\\_uaFI mzQ](https://www.youtube.com/watch?v=vmZ_uaFI mzQ)

**Meeting 13 – Monday, April 27 – Panel Data Regression**

(M&G) Chapter 10: Panel-Data Analysis

(STATA You Tube)  
Panel Regression Analysis <https://www.youtube.com/watch?v=8PUVipH7HFE>

**Meeting 14 – Monday, May 4 – Demonstrating What You Learned**

(Consultant Paper) “Cost of State Regulations on California Small Business Study”  
@Canvas

Remainder of Class Devoted to Review/Questions Necessary to Complete Your Paper

**Final Paper Due – Monday, May 11 at 6 pm**

PPA 207  
Spring 2020  
Midterm  
**Literature Review Assignment**  
**(Subject to Revision)**

Your assignment is to write a six-page, double-spaced, typed literature review that is in 11 font and with one-inch margins all around. It is due in electronic form to me by no later than 6 p.m. on Monday, April 6. Each portion of a day late will result in a one-grade deduction. Please submit as a WORD attachment to an e-mail ( [rwassme@csus.edu](mailto:rwassme@csus.edu) ). I describe below the requirements for the paper. I also list the points earned by satisfying each of the requirements. I will use this grading rubric for the assignment. All references to Galvin below refer to the fourth edition.

<b>Required Element</b>	<b>Points Available</b>	<b>Points Earned</b>
(1) Use a <b>minimum of five regression-based articles</b> drawn from academic and preferably refereed journals and closely aligned with your research topic. You may also use more articles that are not regression based.	<b>10</b>	
(2) Include a <b>reference list</b> at the end of the literature review (that does not count toward your page limit of six) that is in <b>APA style</b> . References made throughout the review should also follow APA style.	<b>10</b>	
(3) Read your articles in the manner described in Chapter 4 in Galvin for “General Guidelines for Analyzing Literature” and organize your literature review around <b>three different themes designated as separate sections in your review</b> .	<b>10</b>	
(4) Review Chapter 5 in Galvin on “Analyzing Quantitative Research Literature.” Note particularly <b>Guidelines 4 (cause and effect issues covered), 9 (differences in variable measurement), 10 (sampling issues), 12 (magnitude and statistical significance of regression coefficients), and 13 (flaws in studies observed) and incorporate these suggestions into your analysis and write up</b> .	<b>10</b>	
(5) Review Chapter 7 in Galvin “Building Tables to Summarize Literature.” You are to <b>include a well-crafted table of the type described here</b> in your review. If you put it in an appendix, it need not count against your six-page limit.	<b>20</b>	
(6) As discussed in Chapter 8 of Galvin “Synthesizing Literature Prior to Writing a Review,” your <b>literature review’s “voice” is suitable for academic writing and differences among studies are noted (Guideline 5), obvious gaps discussed (Guideline 1), relevant theories discussed and how studies advance them (Guidelines 7 and 8), summaries are offered after each section and at end (Guideline 9), conclusions/implications, and suggestions for future research [your own PPA 207 paper] are included (Guidelines 10 and 11)</b> .	<b>20</b>	
(7) You have a coherent essay according to Chapter 10 in Galvin. This means an <b>overview at start (Guidelines 1 and 2), annotations avoided (Guideline 4), subheadings used (Guideline 5), conclusion at end (Guideline 8), and argument flows well (Guideline 9)</b> .	<b>10</b>	
(8) Style and mechanics follow Galvin’s suggestions in Chapter 11. In particular <b>Guideline 3 (no overuse of direct quotations), Guideline 4 (correct APA use of citations), Guideline 6 (spell out acronyms), Guideline 9 (avoid slang), Guideline 11 (check your draft using Microsoft Grammar Editor, avoid passive voice), Guideline 12 (concise and descriptive title), and Guideline 14 (absolutely no plagiarism)</b> .	<b>10</b>	
<b>TOTAL</b>	<b>100</b>	

## **PPA 207 CHECKLIST FOR FINAL PAPER**

**Professor Rob Wassmer, Public Policy and Administration, Sacramento State  
Spring 2020  
(Subject to Revision)**

Your grade on the final paper comes from how well you satisfy the items on this list. You must turn in an electronic copy (to [rwassme@csus.edu](mailto:rwassme@csus.edu)) by 6 pm on May 11, 2020. Each partial day after this time and date, it is late and will result in a one lower grade deduction (that is, at 6:05 pm on May 15, it is one day late).

*A one after a criterion below means that you satisfied it 100%. A decimal value means that you satisfied it at that level. A zero indicates not done and an "X" indicates not relevant. These indicators are the basis of my assignment of your overall grade. Also, see my comments on the electronic version of your paper.*

### **OVERALL**

- There is a cover page with title, your name, and date handed in. \_\_\_\_
- There are seven major sections in your paper, marked by roman numerals and section titles. \_\_\_\_
- First paragraph of a section is an introductory paragraph that briefly describes what is in it. \_\_\_\_
- Sub-sections used within your seven sections and they contain headings. \_\_\_\_
- There are no spelling errors. \_\_\_\_
- There are no grammatical errors (use Microsoft Word Editor or Grammarly with passive voice detection on). \_\_\_\_
- Smooth transitions between paragraphs. \_\_\_\_
- One-inch margins, 11 Times Roman font, and double-spaced. \_\_\_\_
- Paper includes a list of references at end of paper in APA style. \_\_\_\_
- Follow the APA style given in Hacker's *A Pocket Style Manual* (or described at <https://owl.english.purdue.edu/owl/resource/560/01>). \_\_\_\_
- Paper written such that educated layperson working in public policy can follow. \_\_\_\_
- Paper written in first-person "active" voice. \_\_\_\_

### **I. EXECUTIVE SUMMARY (1 – 2 pages)**

- Follows the suggestions offered in "Executive Summaries Complete the Report," – <http://www.csun.edu/~vcecn006/summary.html> \_\_\_\_

### **II. INTRODUCTION (2 - 3 pages)**

- The first paragraph clearly contains your research question. What are you trying to discover through regression analysis? What is the dependent variable? What is (are) the key explanatory variable(s)? \_\_\_\_

- The remainder of your introduction motivates the reader to continue by placing your question in the context of current events and public policy. \_\_\_\_
- Cite at least two newspaper/magazine, think-tank, advocacy, etc. articles that point out the populist importance of determining the impact of your key explanatory variable(s) on the dependent variable. Use search engines like <http://www.sacbee.com/> , <http://www.latimes.com/> , and/or <http://www.sfgate.com/> . \_\_\_\_
- Include at least one figure/diagram (not a numeric table) that helps the reader understand patterns in your dependent variable and/or relationship(s) with your key explanatory variable(s). \_\_\_\_
- The last paragraph contains a description of what contained in the remaining six sections of your paper. A one-sentence description for each section is appropriate. \_\_\_\_

### **III. LITERATURE REVIEW (5 - 6 pages)**

- Contains a description of at least five regression-based research articles in the area of your policy topic. You can find this research by searching the Sacramento State Library's Web Page of literature databases. I would suggest using ECONLIT and EBSCOhost as two literature sources that will have regression studies in them. Google Scholar is also an excellent resource. Search using keywords that include "regression" and your topic. \_\_\_\_
- Divide your literature review into at least three labeled themes (or subsections). \_\_\_\_
- **Attach a copy of your graded midterm to the back of paper with my comments included and your electronic note below each of them as to how handled.** \_\_\_\_
- Address all the comments I offered on your midterm in a new draft of the lit review that you contain in the paper. (If you wish to ignore something, write a note below the comment as to why.) \_\_\_\_

### **IV. MODEL (2 - 3 pages)**

- Offer a motivation for your choice of a dependent variable. How does it relate to your research question? \_\_\_\_
- Specifically describe where your dependent and explanatory variables come from (sources), unit of observation, dates, and any concerns that arose in using this data. \_\_\_\_
- Include a description of the factors expected to cause variation in your dependent variable. The factors should first be listed as broad causes (say causes A, B, C, etc.) and the specific variables which represent broad causes {  $A = f(x_1, x_2, x_3)$ ,  $B = f(x_4, x_5)$ ,  $C = f(x_6, x_7, x_8)$ , etc. } \_\_\_\_
- Describe the variables you use to specifically proxy for each of the broad causes ( $x_1$  through  $x_8$ )? Justify your choices. \_\_\_\_
- Do not use acronyms anywhere in your paper to describe  $x_1, x_2$ , etc., instead write out a short 3 to 5-word description. \_\_\_\_
- What is the expected direction of effect for each of the specific causes (positive, negative, uncertain) you have before regression run? Justify with a verbal cause and effect table and description. \_\_\_\_

## **V. DATA (2 - 3 pages)**

- Describe in paragraph form what is in Tables 1 – 3. \_\_\_\_
- Create a Table 1 that provides description and source for each variable used. (No direct STATA results allowed for any tables. Create tables in your own form and be consistent throughout. Place title on all tables.) \_\_\_\_
- Create a Table 2 that provides descriptive statistics for all variables used (name, mean, standard deviation, maximum, and minimum). \_\_\_\_
- Create a horizontal Table 3 that provides simple correlation coefficients between all explanatory variables. If large, place this in an appendix. \_\_\_\_

## **VI. REGRESSION ANALYSIS (3 - 4 pages)**

- List your regression results in table form in the standard manner taught in class. (No direct STATA results allowed.) \_\_\_\_
- First, give your starting OLS results with no corrections. This should begin with the lin-lin, then quadratic, log-lin, and then log-semilog (if possible) forms. \_\_\_\_
- Pick the “best” functional form based upon the number of statistically significant regression coefficients and use it in remaining corrections. \_\_\_\_
- Discuss how you checked for multicollinearity. Was it an issue, and if it was, how you corrected for it? Be sure to include VIF values and refer to partial correlation coefficients. \_\_\_\_
- Include relevant location or other relevant dummy variables. Discuss your findings. \_\_\_\_
- Include an interaction terms using your key explanatory variable(s). Discuss your findings. \_\_\_\_
- Check for heteroskedasticity in your regression by presenting and describing the Breusch-Pagan Test. If heteroskedasticity is present in your regression analysis, provide the appropriately corrected results. \_\_\_\_
- If your dependent variable is dichotomous (0 or 1), report both OLS (lin-lin and quadratic which represent linear probability models) and Logistic regression results. Describe what both mean and which of the two is more appropriate. \_\_\_\_
- It is possible to use panel data techniques for your analysis. If not, justify why you believe so. If it can, use it throughout. \_\_\_\_ (*Extra Credit 10 points*)

## **VII. CONCLUSION (2-3 pages)**

- Considering your final regression result (with all the appropriate corrections), turn statistically significant regression coefficients into 90% confidence intervals. Report them in a table that lists the explanatory variables in an order from the largest positive influence, to the largest negative influence. (Alternatively, chose the appropriate measure if using logistic.) \_\_\_\_
- For your statistically significant coefficients, how do they compare to the expected signs you described in model section? If findings are different, give a reason it may be the case. \_\_\_\_
- For your statistically significant coefficients, describe the relevance of variable based upon its magnitude. \_\_\_\_

- Interpret the R-Squared (OLS) or hit ratios (Logit). \_\_\_\_
- What does your regression results indicate as an answer to your research question? \_\_\_\_
- What are the specific policy lessons learned from your results? Offer responses to the policy questions you raised in your introduction. \_\_\_\_
- Suggest improvements that you would undertake if you had more time. \_\_\_\_