

The Subfield Qualifying Exam in Political Methodology
Revised October, 2004

Students selecting political methodology as their second field of doctoral study are expected to (1) demonstrate a mastery of foundational material covered in PLSC 501, 502 and 503, and (2) to demonstrate proficiency in one or more areas of advanced research methodology related to three additional graduate-level methodology courses.

Passage of the methodology exam indicates the student has a degree of breadth of knowledge and training in political methodology in addition to technical competence in a single specialty. Passage also indicates that the student has the written and verbal communication skills necessary to present results of analysis or teach basic techniques to a variety of audiences (e.g., peers at a professional meeting, undergraduates in a statistics class, analysts at a government agency, etc.) High scores shall be reserved for students who demonstrate unusual achievement and clear promise and competence as an instructor of political methodology at the graduate level. To successfully pass the exam, answers must be highly technically accurate, use exact terminology and numeric values when appropriate, and convey mastery of both the intuition behind and details of relevant techniques and/or estimators.

The methods examination consists of a written and an oral section. Section I tests the student on the foundational materials covered in the political science methods sequence (PLSC 501, 502 and 503) in a written exam that lasts three hours. Section II tests the student on a select advanced topic in an oral exam that also lasts approximately one hour and fifteen minutes; this exam is open to other faculty within the department. Responsibility for evaluating the qualifying examination shall rest entirely with the designated subfield examiners. However, the highly specialized nature of some research techniques may make it desirable for the examiners to seek the advice and input of faculty who have used or taught courses related to the student's method of specialization. Therefore, additional faculty from Political Science or other departments may be invited to read and evaluate the written examination and participate in the oral portion of the examination. Below are details on each of the two sections.

Section I. Written Exam on Foundational Materials

The written exam shall cover questions that are drawn primarily from, but are not limited to, courses PLSC501-503. While the focus of the exam is on the linear model, students should be prepared to answer questions on probability, models for binary dependent variables, basic time series models, and models for pooled/panel data. Questions from the PLSC504-equivalent course may also be tested. Similar to qualifying examinations in other fields, the written exam will be closed book. The written exam will be three hours long and takes up about 60% of the full credit of the whole methods qualifying exam.

Students skipping 501 or 502 because they completed an equivalent course elsewhere should consult with subfield examiners to confirm the emphasis in content for their particular exam session.

A core material resource list is included in the appendix of this document. This list of resources includes material from these courses over the last few years, supplemented by references to readings with similar breadth and depth. The resource list is intended to give the student alternative presentations of the same fundamental material. Unlike other subfield reading lists, the student **should not** read all of these books and chapters in preparing for the written exam. However, these books and chapters do cover many of the key topics in political methodology. Hence, the student is expected to be familiar with the shared content of the material.

The examination committee recognizes that the content and emphasis of these courses varies from year to year. Therefore, the student can expect to be examined on topics that correspond to the emphasis and scope of the versions of the methods sequence that he or she experienced. Finally, the student should always expect one research design question to appear on the written exam.

Section II. Oral Exam on Advanced Topic

The oral component of the exam is built around one advanced methodological technique within an area of specialization (e.g., ordered logit/probit models in the class of models for qualitative outcomes, competing risk models in event history analysis, X spatial analysis, Y small-N case studies, problems with response bias in survey research, or evolutionary games in formal modeling). The focus of this portion of the examination shall be an advanced technique in one published or publishable research article or scholarly book that exemplifies high-quality work using this technique. The oral exam will be one hour and fifteen minutes long and takes up about 40% of the full credit of the whole methods qualifying exam.

The research article or book shall be selected by the student and formally proposed to, and approved by, the year's field examiners. For the purpose of the oral exam, the student is encouraged to present on an advanced technique employed in one of their own papers.

The student must submit a short written proposal that identifies the particular methodological topic to be presented, the significance of the methodological technique, the article(s) or book for the oral presentation on the technique, and the reasons why it is appropriate for examining the student's competence on that topic. Students must also generate a syllabus for a 15 week graduate seminar (even if such a seminar is not offered at Penn State) on the area of specialization, highlighting the advanced methodological technique that is the focus of the presentation. Examples of the area of specialization might include:

Event history analysis	Game theory
Questionnaire design	Sampling designs
Structural equations models	Elite interviewing
Case study design	Time series analysis
Interpretive field research	Computational modeling
Geographic Information Analysis	Spatial models

This list is intended to be illustrative only. In addition, the breadth/depth balance may differ from student to student (a single topic of “survey research” might encompass sampling and questionnaire design).

A sample syllabus of a related graduate course taken at Penn State or elsewhere must be attached to the proposal as an appendix (in the absence of such a syllabus, the student may create their own brief outline and reading list that approximates a 15 week graduate seminar at Penn State).

The student should turn into the Graduate Coordinator the selected article or book and the proposal for the oral exam about 10 weeks prior to the occurrence of the exam. The Graduate Coordinator then relays the materials to the faculty examiners. Both the selection of the advanced technique and the selection of the article or book must be approved by the field examiners before the student can take part in the oral exam. In the event that the selection and proposal are rejected by the field examiners, the student must submit another proposal for approval. It is the student’s responsibility to select an article or book and submit a proposal that can be approved by the field examiners. Thus the student is encouraged to consult the field examiners informally before filing a formal proposal.

Based on the selected materials and the approved proposal, the student will prepare an oral presentation of 30-45 minutes that:

- a. Describes the substantive problem that motivates the use of the technique;
- b. Explains the technique employed, with sufficient technical details;
- c. Illustrates the use of the technique with an applied example, including the research design, the execution and the interpretation of the analysis;
- d. Identifies the important limitations of the technique.

In other words, the student should prepare a mock class session that introduces the advanced technique and illustrates it with one published application.

The student shall assume that the “audience” is familiar with material covered in PLSC 501-502-503 but may not be familiar with the specialization chosen. The core of the oral presentation, lasting 15-25 minutes, is to explain the technique as in part “b” above. Secondly, 10-15 minutes may be devoted to part “c” – interpreting and explaining the empirical results and conclusions of the selected article or book. Parts “b” and “c”, therefore, should be both highly elaborate in terms of technical details and accessible. For example, the student should present the derivation of a statistical technique or an equilibrium solution of a game theoretic model if necessary.

After the presentation, the subfield examiners and other attending faculty members shall examine the student for about thirty minutes. Questions will be generally linked to the presentation and its substance, but may include those related to basic issues of research design, causal inference, validity, and case selection (i.e., topics covered in PLSC 501), questions about the foundations of the specialized technique (e.g., relevant issues based on probability and the general linear model from 502 and 503), as well as questions specifically focused on the topic of the oral presentation.

Appendix: Selected Reading List for Foundational Materials in Political Science Methods Sequence

Readings for PLSC 501: Methods of Political Analysis

Topics include: Introductions to research design, measurement, scientific inquiry, experiments and quasi-experiments, formal modeling, and cross-level inference.

General Overviews of Research Design:

Kerlinger, Fred N. 1986. *Foundations of Behavioral Research*. Holt, Rinehart & Winston. Third Edition.

Johnson, Janet Buttolph and Richard Joslyn. 2001. *Political Science Research Methods* (4th edition) Washington D.C.: Congressional Quarterly Press.

Frankfort-Nachmias, Chava and David Nachmias. 1999. *Research Methods in the Social Sciences* Sixth Edition. New York: St. Martin's.

Note: the Kerlinger volume is encyclopedic and written for graduate students, though many examples come from psychology; Johnson & Joslyn and Nachmias & Nachmias are lower-level overviews that may contain more examples from political science.

Philosophy of Science and the History of the Discipline:

Susser, Bernhard. 1992.. *Approaches to the Study of Politics*. Macmillan.

Thomas Kuhn. 1996. *The Structure of Scientific Revolutions* 3rd. Chicago: University of Chicago Press.

Experiments and Quasi-Experiments:

Campbell, Donald T. and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Alternative Perspectives on "Scientific" Inquiry

Lave, Charles, and James G. March. 1975. *An Introduction to Models in the Social Sciences*. New York: Harper and Row. (chapters 2-3)

Gary King, Robert Keohane, and Sidney Verba, 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research* Princeton: Princeton University Press.

Stinchcombe, Arthur L. 1968. *Constructing Social Theories*. Chicago: University of Chicago Press. Chapter 2.

King, Gary. 1989. *Unifying Political Methodology*. New York: Cambridge University Press. Chapter 1. "Introduction"

Baumgartner, Frank R., and Beth L. Leech. *Basic Interests*. Princeton: Princeton University Press., chapter 9 "Learning from Experience."

Formal Modeling:

Gates, Scott and Brian D. Humes. 1997. *Games, Information, and Politics: Applying Game Theoretic Models to Political Science*. Ann Arbor: University of Michigan Press.

Morton, Rebecca B. 1999. *Methods and Models*. New York: Cambridge University Press.

Cross-Level Inference:

King, Gary 1997. *A Solution to the Ecological Inference Problem*. Chapter 1. Princeton: Princeton University Press.

Achen, Christopher H. and W. Phillips Shively. 1995. *Cross-Level Inference*. Chicago: Chicago University Press. Chapter 1.

Area Studies:

Bates, Robert H. 1997. *Area Studies and the Discipline. A Useful Controversy?* PS 30: 166-169.

Readings for PLSC 502

Elementary Statistics Texts with no calculus or matrix algebra:

Note: These cover probability, random variables, descriptive statistics through multivariate OLS, and the essentials of statistical inference (sampling distributions, hypothesis tests, etc). These comprise a small subset of many equivalent texts.

Anderson & Finn. 1996. *The New Statistical Analysis of Data*, Springer.

Devore & Peck. 1993. *Statistics: The Exploration and Analysis of Data*. Duxbury.

Ott & Longnecker. 2001. *An Introduction to Statistical Methods and Data Analysis*. Duxbury.

Wonnecutt & Wonnecutt. 1990. *Introductory Statistics*. Wiley.

Hays. 1984. *Statistics*. Harcourt-Brace. (this is the most thorough and mathematically rigorous of these texts but the most recent, fifth, edition contains many typos).

Agresti, Alan and Barbara Finlay. 1997. *Statistical Methods for the Social Sciences*, 3rd Edition. Prentice Hall (more applied, with minimal statistical theory and probability).

The Logic of Causal Inference from Multivariate Analysis:

Note: These books discuss the logic of causal inference and how different methods of conceptualizing third variables (as antecedent, intervening, or as an interaction) lead to different ways of “unpacking” the observed association between an independent and dependent variable. The Davis volume is more cursory but has the advantage of being in print.

Morris Rosenberg. 1968. *The Logic of Survey Analysis*. Basic Books.

Davis 1980. *The Logic of Causal Order*. Sage.

Introductions to Applied Multiple Regression:

Note: These books provide (without calculus or matrix algebra) an introduction to OLS with two or more independent variables. Emphasis is on the logic of least squares, interpretation of coefficients and fit statistics, and an introduction to violations of regression assumptions. The Allison book is both more basic (starts in a more elementary way) and more thorough (provides more examples and places related methods (logit, event history, etc) in context of the simple linear model).

Allison, Paul. 1999. *Multiple Regression: A Primer*. Pine Forge Press.

Lewis-Beck, Michael. 1980. *Applied Regression*. Sage.

Mathematical statistics and probability theory:

Note: These books cover probability, estimation, hypothesis testing, linear regression, and analysis of variance with a strong technical and theoretical focus. The books below are in order of *decreasing* technical/mathematical level.

Degroot, Morris H. 1986. *Probability and Statistics*, 2nd Edition. Addison-Wesley.

Wackerly, Dennis D., William Mendenhall III and Richard L. Scheaffer. 2002. *Mathematical Statistics with Applications*, 6th ed. Duxbury Press.

Relevant sections from Hays, Wonnecott & Wonnecott, and Ott (listed above).

Other: Some students have found study guides to be helpful.

Murray R. Spiegel. *Shaum's outline series: Probability and Statistics.*

Readings for PLSC 503: Multivariate Analysis

The general linear model and extensions:

Note: The shared content in these readings is very high with moderate variation in presentation style, comprehension, and degree of technical difficulty. Greene is highly comprehensive, highly technical, and always up to date (includes datasets and program files). It is an excellent resource for tricky and specific econometric problems. Along with Griffith et al., this text uses matrix algebra. Griffiths is a comprehensive text offering good verbal explanation (like Kennedy) together with some math. Maddala is a very technical (without requiring matrix algebra) and classic introduction that is very clear if one can live with symbols and math. It is older, so has less detail on newer techniques or tests. Gujarati and Wooldridge can be read without any knowledge of matrix algebra and the “word to equation ratio” is much higher than in Maddala. Kennedy provides a very accessible treatment of most of the material covered in 503, relying more on verbal explanation and clarification than on equations/symbols. He also offers sensible suggestions for modeling issues, such as how to deal with assumption violations.

Greene, William H. 2003. *Econometric Analysis*, 5th Edition. Prentice Hall.

Maddala, G.S. 1992. *Introduction to Econometrics*, 2nd edition. New York: MacMillan Publishing Company.

Gujarati, Damodar N. 2002. *Basic Econometrics*, 4th ed. New York: McGraw-Hill, Inc.

Wooldridge, Jeffrey M. 2002. *Introductory Econometrics: A Modern Approach*, 2nd edition. South-Western College Publishing.

Griffiths, Hill, Judge. 1993. *Learning And Practicing Econometrics*. Wiley.

Kennedy, Peter. 1998. *A Guide to Econometrics*. MIT press.

Sage Monographs:

Note: Several Sage monographs cover this material in more detail and in a more applied fashion. Examples are more likely to be social science examples, even political science.

Fox, John. 1992. *Regression Diagnostics*.

Berry, William D. 1993. *Understanding Regression Assumptions*.

Achen, Christopher H. 1982. *Interpreting and Using Regression*.

Aldrich, John H. and Forrest D. Nelson. 1984. *Linear Probability, Logit and Probit Models*

Hardy, Melissa A. 1993. *Regression with Dummy Variables*.

Saysrs, Lois. 1989. *Pooled Time Series Analysis*.

Namboodiri, Krishnan. 1984. *Matrix Algebra: An Introduction*.

Ostrom, Charles. *Time Series Analysis*.

Aiken, L. S. and S. G. West. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage [From Sage but not part of the monograph series]

Sample questions

The following questions are designed to give examinees a sense of the scope and nature of the questions that will appear on the written portion of the exam. To successfully pass the exam, answers should be highly technically accurate, use exact terminology and numeric values when appropriate, and convey mastery of both the intuition behind and details of relevant techniques and/or estimators.

1. A colleague comes to you with the following research question and data. The specific question is: Does interest group density influence congressional attention to environmental issues? The data set includes the number of interest groups, the number of hearings Congress holds on the topic, and a measure of media attention to the issue, all measured annually for the last 50 years. What advice do you give him? Why? Your answer should discuss research design, estimation, and inference.
2. You are planning a dissertation that investigates the influence of campaigns on election outcomes. Your dissertation advisor suggests that you consider the various approaches below. Explain the virtues and disadvantages of two of these each research design as it applies in this example.
 - a. Run experiments of the kind conducted by Iyengar and colleagues to see if voters change opinions on the candidates after viewing particular types of campaign advertisements.
 - b. Analyze the large volume of tracking polls conducted over the 2000 election campaign to capture the influence of events and media coverage on the campaign trail.
3. Explain the logic underlying maximum likelihood estimation and provide a substantive example in which MLE might be used fruitfully. Be sure to explain why MLE is appropriate in the circumstance.
4. What issues should the analyst be particularly concerned about when applying a general linear model to time series data? To cross-national survey data? Why? How can the analyst sensitive to these problems proceed in this framework? Make reference to substantive examples in your answer.
5. You are asked to teach the equivalent of our course PLSC 503, what will you put in your syllabus and why? Provide a description of the main sections of the course and explain the material to be covered and why.
6. Consider the following model:
$$Y = B_0 + B_1X_1 + B_2X_2 + u$$
(Or denoted as: $Y = BX + u$)
 - (A) What is the Gauss-Markov theorem regarding this model? What are the assumptions underlying the Gauss-Markov theorem? Use notation for each assumption and explain each assumption with one or two sentences.

- (B) What is the OLS estimator for this model? (Notation of either matrix or algebraic form is fine.)
- (C) Show that the OLS estimator is unbiased. Be specific as to which assumptions you need to show unbiasedness.

7. Given the model in Question 6,

- (A) what is the estimator for the variance of B_1 ?
- (B) Discuss what the components of the variance of B_1 are. What is the implication of each of these components for testing hypothesis: $B_1=0$.
- (C) Define heteroskedasticity for this model. Explain why in the presence of heteroskedasticity, the OLS estimator of B_1 is still unbiased but the standard errors are biased. What is the implication for hypothesis testing?
- (D) Suggest a method for correcting heteroskedasticity

8. Consider the following OLS regression model:

$$Y_i = B_0 + B_1X + u_i$$

Now, let us assume that some third variable, Z , is correlated with X and also is known (from prior research) to be a cause of Y .

(A) According to standard econometric texts, what can we say about the estimator B_1 as a desirable estimator of the effect of X upon Y ?

Now consider the following data set of 150 countries:

i	X Colony in 1935	Z Market economy in 1975	Y Democracy in 2000
1	1	0	10
2	0	1	6
3	1	0	8
4	1	1	3
5	0	1	10
6	0	0	3
7	0	1	9
...			
149	0	1	6
150	1	1	6

The researcher estimates the model:

$$DEMOCRACY_i = B_0 + B_1COLONY + u_i \quad (\text{or, } Y_i = B_0 + B_1X + u_i)$$

and claims that B_1 is an unbiased estimator of the effect of colonization status in 1935 on democracy in 2000. Another scholar claims that this is clearly untrue because the presence

of a market economy is correlated with colonization status and with democracy and this, therefore produces a biased estimate of B_1 .

(B) Explain the merits of both positions and provide advice to researchers facing similar problems.