Quantitative Empirical Methods (QEM) Exam Reading List (Last updated: June 2021)

Coursework

The primary QEM coursework is a two-course doctoral sequence:

- PLSC 500 (Foundations of Statistical Inference) and
- PLSC 503 (Theory and Practice of Quantitative Methods).

These courses, along with the exam and your future quantitative research, require mathematical maturity, and thus:

- PLSC 529 (Mathematics for Political Science)

is also a core course that serves as a corequisite for PLSC 500 and a prerequisite for PLSC 503. (PLSC 529 is also core for the Formal Theory subfield, for which the coursework is strongly recommended in conjunction with the QEM sequence.) Mastery of the materials covered in PLSC 500 and 503 are essential to passing the QEM exam. PLSC 508 (or equivalent advanced course on research design, e.g., PLSC 511 or S&DS 517) is strongly recommended for students intending to take the exam. Attendance at the MacMillan-CSAP Workshop on Quantitative Research Methods, in addition to at least one other subfield seminar series, is also strongly recommended.

Additional courses in mathematics, formal theory, political economy, probability, statistics, qualitative and archival methodology, and the empirical social sciences in the department and around the university can help you develop your skills. Coursework in the Qualitative and Archival Methodology subfield is recommended as a natural complement to the sequence that will help hone your empirical instincts. Comfort with political economy and formal theory will help you use economic reasoning when engaging with empirical research. More generally, the QEM field spans the breadth of political science, and you are expected to have some familiarity with major debates across the subfields. Your instructors and advisors will be able to help you find suitable courses depending on your interests.

Exam Format

A typical set of instructions for the exam are reproduced below. The instructions may change from exam to exam.

You have seven hours to complete the exam. This exam consists of three parts.

Back up your assertions with mathematics where appropriate and show your work. Good answers will provide a direct answer that illustrates an understanding of the question, and calculations or statistical arguments to validate the answer. Where applicable, exceptional answers will include all of these as well as proofs that are technically complete, including formally articulating sufficient assumptions and regularity conditions. Questions will not be weighted equally. A holistic score will be assigned to the exam. Therefore, it is important to demonstrate your understanding of the material to the best of your ability.
Part 1 (Short Answer Section) consists of seven short answer questions. Advice: Note there are multiple correct answers to some questions. We encourage you to give the most complete (but still succinct) solution possible. Do not leave sub-parts of questions unanswered.

Part 2 (Essay Section) contains a recent, well-regarded empirical article. We will ask you to offer an evaluation of its methodological approach and presentation of results. In particular, we will advise you to pay particular attention to the identification conditions (either explicit or implicit), the associated estimation strategy, and possible threats to inference. Your response may be anywhere from 500 to 1500 words.

The only aids permitted for Parts 1 and 2 are (i) one page of double-sided notes, (ii) a word processor on one of the Statlab computers to write up your answers (you may also write up your answers to Part 1 using pencil/pen and paper). After handing in your answers for Parts 1 and 2 of the exam, you may begin Part 3 (though feel free to look ahead). You may hand in Parts 1 and 2 whenever you wish, but we recommend spending no longer than five hours on Parts 1 and 2.

Part 3 (Computer Assisted Section) will involve using statistical software to answer one longer exercise with five associated questions. A complete answer to Part 3 will include code and output, as well as your written answers. Advice: We recommend that you explain what you are trying to do in comments in your code. Even if you are not able to execute your program correctly, you can receive partial credit for explaining clearly what you wanted to do and why.

For Part 3, you are permitted (i) unrestricted use of your own computer with access to the internet or (ii) use of a Statlab computer with access to the internet. The only restriction for Part 3 is that you may not interact with anyone, online or otherwise. For Part 3 (Computer Assisted Portion) of the exam, please turn in a hard copy of your code to Colleen, and also email a digital copy of the code to colleen.amaro@yale.edu.

Reading List

Each section is roughly ordered in increasing difficulty. Some entries are listed more than once across sections.

Probability:
- Peter M. Aronow and Benjamin T. Miller. 2020. Foundations of Agnostic Statistics. Cambridge University Press. Ch. 1-2. (Note: the “Further Readings” sections at the end of each chapter of Aronow and Miller may often be helpful.)

Data Science:
- Peng, Roger D., Caffo, Brian, and Leek, Jeff. R Programming. Coursera course at: https://www.coursera.org/learn/r-programming (Note: this is free, choose the “Audit” option.)

Core Statistics and Econometrics:
- Aronow and Miller. Ch. 3-5.
- Wackerly, Mendenhall, and Scheaffer.
- Wasserman. All of statistics. Chs. 6-15, 19-22
- Hansen. Econometrics.

Research Design and Causal Inference:
- Gerber, Alan S. and Green, Donald P., 2012. Field experiments: Design, analysis, and interpretation. WW Norton. Read with:
- Sekhon, Jasjeet S., and Rocío Titiunik. 2012. When natural experiments are neither natural nor experiments. American Political Science Review. 35-57.
- Aronow and Miller. Ch. 6-7.
- Hansen. Econometrics.