Economics 671 Reading Guide
Fall 2015, version 1.0
Gray Calhoun

Introduction

- In this document, I'll often write things like, "you should know these definitions." I do not mean that you need to memorize and reproduce those definitions word for word. Instead, you should be able to describe what the definition means reasonably accurately and (more important) should be able to look at especially representative objects and determine whether they fit that definition or not.

  As an example: if you were asked to know the definition of "convex functions" I'd want you to be able to decide whether or not the function
  $$g(x) = a + bx$$
  is convex. (Here a and b are arbitrary constants.)

- Even if I tell you to skip a result or a section, it is still interesting and important and I'd encourage you to read it on your own. You'll be asked to read some of these sections after the RAT in preparation for specific activities and others will help you prepare for next semester.

- The examples in these books are pretty good and, at the very least, don't skip them when you do the reading. (Believe me, I understand the temptation.)

- Unless they are very short or I state otherwise, you can skip the proofs when you first read this material.

- In the past, this class has used Casella and Berger (2002) and Greene (2012) as its required textbooks. They both cover more material than the textbooks that we're using now but convey the intuition less well and are poorly organized. This is a serious problem now that we are using TBL to organize the class but you may still find those books useful and you can almost certainly borrow copies from other graduate students. I have also assigned Gallant (1997) as a recommended textbook in the past, and you may want to consult it for the probability and asymptotics material.
- Regarding the "recommended" textbooks for this class:
  - Freedman (2009) excels in building intuition for when regression is appropriate — when the statistical assumptions hold or break down and, even more important, when the assumptions that let you interpret a regression model as a model of the real world hold or break down. Freedman's perspective can probably be summarized as "regression is rarely appropriate," and his explanations are very educational. If you are interested in empirical research or econometrics, you should definitely read this book even if you don't have time until after the semester is over.
  - Thomson (2011) on the other hand, explains how to be a successful graduate student. You should read it now and reread it later on as necessary. The first chapter covers graduate school explicitly, but you'll also need to write a paper (2nd chapter), give a presentation (3rd), and probably write a referee report (4th) before you graduate. The book has very little to do with this class specifically, but you should read it during your first semester of graduate school at the latest, so it's on our syllabus. The first chapter is available for free from MIT press's homepage.[1]
- Dr. Bartalotti will use Wooldridge (2010) in his microeconometrics elective (Econ 673) and Dr. Bunzel and I will use Hamilton (1994) in our macroeconometrics elective (Econ 674). Both of these books are typically recommended for Econ 672 as well. If you plan to take these electives, you may want to purchase the textbooks now to use as additional references.
- Other interesting resources (probably for reading after the semester is over) include: Davidson (1994), Billingsley (1995), and Pollard (2002) (for probability and asymptotics); Lehmann and Casella (1998), van der Vaart (1998) and Lehmann and Romano (2006) (statistics); Jaynes (2003) and Geweke (2005) (Bayesian statistics, which we don't cover this semester); Pólya (1945) (mathematical problem solving); White (2000), Gelman and Hill (2007), and

[1] http://mitpress.mit.edu/books/guide-young-economist-0.

James et al. (2013) (aspects of regression); Fisher (1926), Rosenbaum (2009), and Imbens and Rubin (2015) (causal inference); Tufte (1990, 1997, 2006b,a), Cleveland (1993, 1994), and Wainer (2000, 2007) (statistical graphics, which we also don't cover this semester); Hunt and Thomas (2000), Friedl (2006), Sargent and Stachurski (2014), and Wickham (2014) (aspects of computation); Strunk and White (1999), University of Chicago Press Staff (2010), Miller (2013), and Williams and Bizup (2014) (writing); Kelsky (2015) (professional development); and Manski (1999, 2003, 2008), Geweke (2010) and the NBER Summer Institute Econometrics Lectures[2] (other aspects of applied econometrics). Each of these is listed in the bibliography at the end of the reading guide. This list is a bit of a brain dump but my personal favorites are Pólya (1945), Cleveland (1993, 1994), and Rosenbaum (2009).

[2] Videos and slides are available for free online at http://www.nber.org/SI_econometrics_lectures.html.

- When the semester is over, there are two "best" things to do to prepare for next semester's econometrics class over break. First, reread Ramanathan (1993) carefully, focusing on material that we skipped in Econ 671 and solving the exercises in each chapter. Second, read Freedman (2009) and solve as many exercises as you can. Next semester will cover Chapters 3–9 of Hayashi (2000), along with a small number of additional topics, so you can read ahead if you want. But I would suggest spending time on the other two books instead.
- Finally, if you are considering econometrics as a potential research area, you should do more than the reading outlined here: work through all of the examples and proofs independently to make sure you understand the concepts; work through as many of the practice and end-of-chapter exercises as you have time for; read all of the unassigned sections; do the empirical and computational textbook questions; etc. I'd also strongly encourage you to form a study group with other students interested in econometrics to discuss all of this material in more detail. (If I've done my job properly, you should be in different teams.)

1 Probability theory

The item numbers in this section refer to sections in Ramanathan (1993) listed as [Ram93] on the syllabus.

- 2.1, Sample space, sample points, and events. Understand definitions and do the practice problems.
- 2.2, Some results from set theory. You should do the practice problems (try problems 2.6 and 2.7 but you don't need to solve

them) and learn/remember the basic set theory results in 2.2. Know the definitions of a σ-field and the Borel σ-field.

- 2.3 Definitions and concepts of probability. Try to understand the axiomatic definition of a probability space and definitely understand the informal definitions of objective probability. You should understand the intuition behind the theorems but don't need to be able to reproduce their proofs and should do the practice problems. The material on conditional probabilities and Bayes Theorem should be review from undergrad and you should know them, and you should understand the definition of independence between events. Read and understand the intuition in the subsection on sequences of events.

- Intro to Chapter 3. Random variables are mathematical functions from a sample space to the real line. Accept this, even if you don't understand it. :-) In general, almost all of this chapter is (or should be) review from your undergraduate statistics and econometrics classes. You've seen the normal distribution; you've seen expectations and variances; you've seen densities; etc. They all work exactly the same as before, but there's a slightly different concept of probability underlying them here than there may have been when you first saw them.

- 3.1, Distribution function. Distribution functions should be review. Look at the theorems and try to do the practice problems.

- 3.2, Discrete distributions, and 3.3, continuous distributions. These sections should be review. You should know the specific distributions mentioned and I'm going to assume that you can reproduce them from memory. The bold terms at the top of page 32 are actually important and will show up repeatedly.

- 3.4, Transformations of random variables. You should read and try to understand the proofs of Theorems 3.4 and 3.5 — they basically give all of the intuition for the actual results and are easier to remember.

- 3.5, Characteristics of distributions. Most of this should be review. The subsection on the *Stieltjes Integral* defines some notation. Specifically, you should be able to derive the mean and variances in the practice problems, and you should be somewhat familiar with Chebychev's inequality.

- 3.6, Generating functions. You can skip this section. (We will get to it in class, though.)

2 Sampling and asymptotics

Again, the item numbers in this section of the reading guide refer to sections in Ramanathan (1993). Refer to his Appendix A if you need to review results from matrix algebra. You also should own a copy of Simon and Blume (1994) from math camp and for Econ 600 and you may want to refer to it as well.

This chapter covers "bivariate" distributions first and then covers generic "multivariate" distributions. For some parts, seeing the bivariate case first really helps a lot. For others, not so much.

- 5.1, Bivariate distributions. This material should be largely review and I expect you to understand all of the results.
- 5.2, Conditional expectation, and 5.3, conditional variance. Also review; you should be able to prove Theorems 5.5–5.7 and do the practice problems.
- 5.4, Bivariate normal distribution. Skip — this section is encompassed by 5.10, which covers the "multivariate normal." I find the matrix notation used in 5.10 much easier to understand. You can refer back to this section if it helps to understand parts of 5.10.
- 5.5, Bivariate transformations. You can skip this for the RAT, but we'll discuss it in class.
- 5.6, Convolution formula. You can skip this for the RAT, but we'll discuss it in class.
- 5.7, Mixture distributions. You can skip this for the RAT.
- 5.8, Bivariate characteristic function. You can skip this for the RAT, but we'll discuss it in class.
- 5.9, Multivariate density functions. Read the first two paragraphs carefully, skip the rest for now.
- 5.10, Multivariate Normal distribution. Read the basic properties of the mean vector and covariance matrix; you need to know the density of the multivariate normal and be roughly familiar with the properties listed under "marginal and conditional distributions..." (But you don't need to remember the exact form of the conditional mean and variance.)
- 5.11, Chi-square distribution, and 5.12, distribution of quadratic forms. This treatment of the chi-square is a bit overkill. Theorems 5.21 and 5.22 cover the essential points, and

you can use Theorem 5.21 for an easy proof of the mean and variance of the chi-square.

- 5.13, Multinomial distributions. Skip. You'll see this next semester.
- 6.1, Random Samples, and 6.2, Sample Statistic. These sections have simple definitions that you should know.
- 6.3, Sampling distributions. You should understand why Theorems 6.3 and 6.5 hold, but don't need to worry about the t and F densities yet. You should also understand what the term "sampling distribution" means.
- 6.4, Monte Carlo simulations of data. Skip for now (there is much more to simulating than this, anyway).
- 7.1, Types of convergence. Try to understand each of the different modes of convergence. (Figure 7.1 is useful but Figure 7.2 is terrible.) We'll discuss this in class.
- 7.2, Relationships among modes of convergence. Understand the relationships represented in Figure 7.4. You can skip the rest for now, but it's worth working through this material if you have time. We'll discuss this in class.
- 7.3, Weak Law of Large Numbers. You need to know what the WLLN means and you should understand the proofs of Theorems 7.10 and 7.11. The other proofs are more advanced.
- 7.4, Strong Law of Large Numbers. Again, understand what the SLLN means and how it is different than the WLLN.
- 7.5, Central Limit Theorem, and 7.6, multivariate CLT. Most important is to understand what the CLT means. Also understand Theorem 7.17 (which is illustrated by example 7.1) and Theorem 7.18, but don't worry about proofs.

3 Estimation and inference

The item numbers in this section of the reading guide refer to sections in Ramanathan (1993). We're not going to cover these chapters as thoroughly as we covered the previous chapters. There's a lot of theory here that we're going to skip over. Our emphasis is going to be on constructing and using these statistics instead.

- 8.1, Small sample criteria for estimators, and 8.2, large sample properties. You should know all of the definitions in this section except for "completeness."

- 8.3, Likelihood function. Read the definition of the likelihood function (the first two paragraphs) and skip the rest.
- 8.4, Principle of Maximum Likelihood. Very important! I expect to be able to give you a simple density function and have you calculate the MLE, even on the RAT.
- 8.5, Lower bounds for variances of estimators. Learn what the "score" is and understand it's relationship to the Cramér-Rao lower bound. I don't expect you to remember the regularity conditions listed in this section, but you should try to understand why they can be loosely characterized as "moment" and "smoothness" conditions. (Hint: it's because they deal with differentiability and existence of expectations.)
- 8.6, Exponential family of distributions. Skip.
- 8.7, Small sample properties of MLE. Skip; we'll calculate the sampling distributions numerically.
- 8.8, asymptotic properties of MLE. Skip. The "asymptotic properties" are really in Sections 8.11 and 8.12 anyway; this section is just a placeholder.
- 8.9, Joint estimation of several parameters, and 8.10, information matrix. Read carefully; this section gives you the multivariate equivalent to a lot of the material you've already read. Example 8.12 is a bit of foreshadowing and you may want to spend some time understanding it now.
- 8.11, Consistency of MLE, and 8.12, asymptotic normality of MLE. Skip or skim for future reference. We're not going to rely on these results (i.e. consistency and asymptotic normality of the MLE in general) in this class; you're going to cover it detail in the first half of 672. But these proofs are especially well organized, so these sections might be good references when you do study this material later.
- 8.13, Numerical procedures. Skip, but likely to be relevant next semester.
- 9.1, Basic concepts in hypothesis testing. You need to know all of these terms and definitions.
- 9.2, Neyman-Pearson Lemma. Understand p-values and skim the rest, but we'll try to cover this in class.
- 9.3, Monotone Likelihood Ratio. Skip.

- 9.4, Applications to the Normal distribution. Skip, but we'll obviously get to this in class.
- 9.5, Unbiased tests. Understand one-sided vs. two-sided testing and skim the rest.
- 9.6, UMPU tests for multiparameter exponential families. Skip.
- 9.7, Generalized Likelihood Ratio tests, and 9.8, examples. The Likelihood Ratio test is important! Just like with the MLE, I expect to be able to give you a (simple) family of densities and a null hypothesis and have you calculate the Likelihood Ratio test. (You may need to use the computer to calculate its critical values numerically, though.) Read the examples in 9.8 carefully for examples of how to do that. (9.7 is more than a little abstract.)
- 9.9 and 9.10 (both have more examples). Skip.
- 9.11, Wald, Likelihood Ratio, and LM tests. Skip. You may want to flag this section for reference in the first half of 672, though.
- 9.12, Test of Goodness of Fit. Skip.
- 9.13, Confidence Intervals. Read carefully and especially try to understand the connection between confidence intervals and hypothesis tests. You can take the claim about asymptotic normality of the MLE on faith for now.

4 Regression in finite samples

A key issue in applied economics is trying to understand causality, and we're going to look at that in detail in Part 6 of the class. Before that, we will focus on algebraic and mathematical properties of the linear regression model, so that we understand how the estimator works and what it returns when we use it, but we'll be focusing on relatively small and specialized problems. The material in this part should be a review of material that you covered as an undergraduate, but the notation may be slightly different since we will be using matrix notation.

This part of the course covers the math behind the OLS estimator — algebraic properties and statistical properties. "Algebraic" properties refers to properties of the OLS estimator that are always true, simply because of algebraic identities: what happens when another variable is added to the model, for example. "Statistical" properties are properties of the estimator that depend on the underlying data generating process. Right now we're only concerned with the math behind the OLS estimator, we're not worried about interpretation yet.

The item numbers in this section refer to subsections in Hayashi (2000) Chapter 1.

- 1.7, Empirical application. Reading this section (an empirical example on electricity supply) will help you understand the rest of the chapter's material and I strongly recommend reading it first. You should probably read it twice: first before reading the rest of the chapter at all, then again once you've finished the chapter. I will only ask very basic questions about the analysis on the RAT.
- 1.1, Classical linear regression model. The assumptions should be review; the use of matrix notation may be new, so spend some time trying to understand it since we're going to be using it heavily from now on.
- 1.2, Algebra of least squares. Read this section carefully; it's all important (even the "optional" section on "Influential Analysis.")
- 1.3, Finite-sample properties of OLS. The finite-sample properties of OLS should be review and you need to know them. The proofs of Proposition 1.1 (a) and (b) are mandatory.
- 1.4, Hypothesis testing under normality. Read but you can skip the proofs of results. You should also try to relate this section back to Chapter 5 of Ramanathan (1993).
- 1.5, Relation to maximum likelihood. This section of the book connects linear regression back to the previous unit on estimation, so make sure that you understand it.
- 1.6, Generalized Least Squares. Make sure you understand the derivation and motivation for GLS. It should be easier to rederive Proposition 1.7 on the fly rather than to memorize and reproduce it.

5 Regression asymptotics

Item numbers in this section refer to subsections of Hayashi (2000), Chapter 2.

- 2.11, Empirical application. Read the empirical exercise in 2.11 before reading the rest of the chapter and then read it again afterward. It will help you remember and understand the theoretical results if you can put them in context. I will only ask very basic questions about the analysis on the RAT.

- 2.1, Review of limit theorems. This should be review from earlier in the semester, but I may ask basic questions about it on the RAT. Skip it if you already know the material, but reread it otherwise.
- 2.2, Fundamental concepts in time-series analysis. This section extends the previous section (and the limit theory we've already seen) to cover time-series. The main takeaways are: the concepts of stationarity and ergodicity; the definition of "martingale difference sequences"; and the "CLT for Ergodic Stationary Martingale Difference Sequences."
- 2.3, Large-sample distribution of the OLS estimator. Your main focus should be understanding Propositions 2.1 and 2.2 and their proofs; you can skim/skip the rest. (We're going to go through the rest of the material in class activities, along with the material in Sections 2.4 through 2.8.) The assumptions will make more sense once you understand how they're used.
- 2.4, Hypothesis testing. Skip; we will cover this material in class.
- 2.5, Estimating $\mathrm{E}(\varepsilon_i^2 x_i x_i')$ consistently. Skip; we will cover this material in class.
- 2.6, Implications of conditional homoskedasticity. Skip; we will cover this material in class.
- 2.7, Testing conditional homoskedasticity. Skip; we will cover this material in class.
- 2.8, Estimation with parametrized conditional heteroskedasticity. Skip; we will cover this material in class.
- 2.9, Least squares projection. Read this section; it is important. Your goal in reading it should be to understand the last subsection, "OLS Consistently Estimates the Projection Coefficients." If you understand what that means and why it's true, then you have a good understanding of the rest of the section.
- 2.10, Testing serial correlation. You can skip this for the RAT.
- 2.12, Time regressions. You can skip this for the RAT.

6 Causal inference and modeling

Item numbers starting with "R" in this section refer to sections of Rubin (2008) and those starting with "IW" refer to sections of Imbens and Wooldridge (2009). Read Rubin first — it's shorter, less technical, and explains the material better.

The big distinction between this unit and the previous two units is that here:

- we explicitly want to estimate the effect of a future policy action;
- we do not want to assume that the effect will be the same for everyone treated or that the effect has a particular (linear) functional form; and
- we want to impose as little economic theory as we can to estimate the effect.

The theory in this unit (and in the Program Evaluation literature in general) tries to explain why randomized experiments are effective and how to mimic their important features in observational studies. If you've come across the phrase "natural experiments" before, you're somewhat familiar with this approach.

Before getting into specifics, your focus in this reading should be on the following:

- Understand the definitions of the terms *potential outcomes*, *propensity score*, and *average treatment effects* and understand the meaning of what's called the *unconfoundedness assumption*.
- Understand how the unconfoundedness assumption weakens the exogeneity assumptions we made in the previous two sections. It may be obvious that it is weaker than strict exogeneity (or it may not), but it is also weaker than sequential exogeneity. Understanding this relies on IW.5.1 and IW.5.3, so you should probably do all the reading first, and then try to understand how this assumption works.
- Understand why linear regression can consistently estimate the average treatment effect under the unconfoundedness assumption. What's really interesting is that linear regression does not necessarily estimate the other parameters consistently under this assumption, (you need the stronger exogeneity assumptions from Hayashi for that) but it still can estimate the causal effect of interest. THIS INSIGHT IS THE SINGLE MOST IMPORTANT PART OF THIS SECTION OF THE CLASS.

Obviously, the reading covers more than those bullets, but try to nail them down before worrying about understanding the rest in detail.

As an aside, I should also point out that the approach we're studying here is inappropriate or infeasible for many research questions in economics. Many people would argue that estimating a structural model of economic decision-making is the main point of econometrics, which is the opposite approach of the one we're studying now — the empirical study in Chapter 2 of Hayashi is one example, and it is hard to think of quasi-experimental studies that could look conclusively at the long-term deviations from the efficient markets hypothesis that are important in behavioral finance. Heckman is one of the best known proponents of this argument, and has written an accessible review article that spells out the general ideas (Heckman, 2008). But regardless of the sort of empirical research you do in the future, you should understand this approach and it is a very useful way to frame many estimation problems.

- R.1, Randomized experiments versus observational studies. Read for an overview of the material.
- R.2, Brief review of parts of the Rubin Causal Model. Read this section carefully. It introduces a bit of notation and terminology, but the notation is extremely important, much more so than usual. You should especially understand the terms "potential outcome," "assignment mechanism," and "propensity score."
- R.3, Design observational studies to approximate randomized trials — general advice. Simple guidance on how empirical research with observational data can mimic experimental studies. The major recommendation is somewhat buried in the second paragraph of R.3.3:

  > When designing a randomized experiment, we cannot look at any outcome measurements before doing the design, and this crucial feature of randomized experiments can be, and I believe must be, implemented when designing observational studies.

- R.4, Examples using propensity scores and subclassification. Make sure that you understand the role of the propensity score as described in R.4.2.
- R.5, A principal stratification example. This example is interesting but uses tools analogous to what you'll see next semester. You can skip it.
- R.6, Discussion. A two paragraph summary of the recommendations. You should probably read this section of the paper first, because it will give you a framework for the paper's specific details..
- IW.1, Introduction. Skim for context and an overview of the paper.
- IW.2, The Rubin Causal Model. Another review of the Rubin Causal Model. Skim it here if you understood Rubin's presentation.

IW.2.3 explains some of the issues in this literature that are unique to economics.

- IW.3, Estimands and hypotheses [on treatment effects]. Understand the different types of “average treatment effect” described in IW.3.1 conceptually; I won’t quiz you on the terminology, but you should try to understand why it might be useful to condition on different covariates or treatment statuses. You can skip IW.3.2–IW.3.4.
- IW.4, Randomized experiments. You can skip or skim this section.
- IW.5, Estimation and inference under unconfoundedness. This section describes how and when regression can be used to estimate average treatment effects. The introduction to this section gives a useful overview. You should understand the assumptions listed in IW.5.1, especially how they can be satisfied by linear regression models, and the first half of IW.5.3 (through the first paragraph of page 30.) The rest of the section is important and we’ll get to some of it in class activities, but you can skip it on first reading. There is some overlap with the Rubin paper.
- IW.6, Selection on unobservables. This section covers many of the estimation strategies developed recently and shows also how instrumental variables fits into this framework. You will see some of this material next semester and can skip it for now.
- IW.7, Multivalued and continuous treatments. Skip, but we may discuss aspects of this in class.
- IW.8, Conclusion. It’s one paragraph, read it.

Reading over Thanksgiving break

This is intentionally light and relatively nontechnical reading. The paper is meant to give you a different perspective than you might have from the textbooks, where everything is implied to work well and be accurate and reliable. You are not going to be directly tested on any of this reading, so in a sense it is optional.

- Freedman (1991). A skeptical discussion of linear regression and probably closest to my own views. A representative quotation from this paper is at the beginning of Section 3:

  > If regression is a successful methodology, the routine paper in a good journal should be a modest success story. However, the situation is quite otherwise.

- Rosenbaum (2009), Chapter 19 and Summary. Rosenbaum (2009) is an an excellent textbook about observational studies. Rosenbaum focuses on matching estimators, but we know (from reading Imbens and Wooldridge, 2009) that regression estimators behave similarly. I'm assigning Chapter 19 and a two-page summary, which you can download from the course webpage. Chapter 19 discusses "plans" for observational research and emphasizes multi-step plans that allow more than one research objective to be explored. Since almost every empirical paper has more than one outcome, it's worth thinking about how to do it. The summary briefly covers key aspects of designing and planning observational studies.
- Finally, it goes beyond what we study this semester, but this would be a good time to read the rest of Imbens and Wooldridge (2009) and Freedman (2009).

References

Billingsley, P. (1995). *Probability and Measure.* John Wiley & Sons, 3rd edition.

Casella, G. and Berger, R. L. (2002). *Statistical Inference.* Duxbury Press, 2nd edition.

Cleveland, W. S. (1993). *Visualizing Data.* Hobart Press.

Cleveland, W. S. (1994). *The Elements of Graphing Data.* Hobart Press, revised edition.

Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians.* Oxford University Press.

Fisher, R. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain*, 33:503–513.

Freedman, D. A. (1991). Statistical models and shoe leather. *Sociological Methodology*, 21:291–313.

Freedman, D. A. (2009). *Statistical Models: Theory and Practice.* Cambridge University Press, revised edition. Recommended textbook.

Friedl, J. (2006). *Mastering Regular Expressions.* O'Reilly.

Gallant, A. R. (1997). *An Introduction to Econometric Theory.* Princeton University Press.

Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge University Press.

Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*. John Wiley & Sons.

Geweke, J. (2010). *Complete and Incomplete Econometric Models*. The Econometric and Tinbergen Institutes Lectures. Princeton University Press.

Greene, W. H. (2012). *Econometric Analysis*. Prentice Hall, 7th edition.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hayashi, F. (2000). *Econometrics*. Princeton University Press. Required textbook.

Heckman, J. J. (2008). Econometric causality. *International Statistical Review*, 76(1):1–27.

Hunt, A. and Thomas, D. (2000). *The Pragmatic Programmer: From Journeyman to Master*. Addison-Wesley Professional.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Kelsky, K. (2015). *The Professor is In. The Essential Guide to Turning Your Ph.D. into a Job*. Three Rivers Press.

Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*. Springer, 2nd edition.

Lehmann, E. L. and Romano, J. P. (2006). *Testing Statistical Hypotheses*. Springer, 3rd edition.

Manski, C. F. (1999). *Identification Problems in the Social Sciences*. Harvard University Press.

Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer.

Manski, C. F. (2008). *Identification for Prediction and Decision*. Harvard University Press.

Miller, J. E. (2013). *The Chicago Guide to Writing About Multivariate Analysis*. University of Chicago Press, 2nd edition.

Pollard, D. (2002). *A User's Guide to Measure Theoretic Probability*. Cambridge University Press.

Pólya, G. (1945). *How to Solve it: A New Aspect of Mathematical Method*. Princeton University Press.

Ramanathan, R. (1993). *Statistical Methods in Econometrics*. Academic Press. Required textbook.

Rosenbaum, P. R. (2009). *Design of Observational Studies*. Springer.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *Annals of Applied Statistics*, 2(3):808–840.

Sargent, T. J. and Stachurski, J. (2014). *Quantitative Economics*. http://quant-econ.net.

Simon, C. P. and Blume, L. (1994). *Mathematics for Economists*. W.W. Norton and Company.

Strunk, W. J. and White, E. B. (1999). *The Elements of Style*. Longman, 4th edition.

Thomson, W. (2011). *A Guide for the Young Economist*. MIT Press, 2nd edition. Recommended textbook.

Tufte, E. R. (1990). *Envisioning Information*. Graphics Press.

Tufte, E. R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press.

Tufte, E. R. (2006a). *Beautiful Evidence*. Graphics Press.

Tufte, E. R. (2006b). *The Visual Display of Quantitative Information*. Graphics Press, 2nd edition.

University of Chicago Press Staff (2010). *The Chicago Manual of Style*. University of Chicago Press, 16th edition.

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Wainer, H. (2000). *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. Psychology Press.

Wainer, H. (2007). *Graphic Discovery: A Trout in the Milk and other Visual Adventures*. Princeton University Press, 2nd edition.

White, H. (2000). *Asymptotic Theory for Econometricians.* Academic Press, revised edition.

Wickham, H. (2014). *Advanced R.* Chapman & Hall.

Williams, J. M. and Bizup, J. (2014). *Style: The Basics of Clarity and Grace.* Longman, 5th edition.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data.* MIT Press, 2nd edition.

License and copying