A Review of Textbooks for Teaching Graduate Research Methods

Wendy L. Martinek, *Binghamton University*

It is not exactly novel to say that selecting a textbook for a graduate student methods class is a challenging task. Nonetheless, it *is* a challenging task. There are a lot of variables to consider, chief among them are a program's curricular structure and the purposes that motivate that structure. Reflecting those varied curricular goals, there is a sizable amount of variation in the number of methods courses graduate programs require their students to take as well as in what the content of those courses looks like (Schwartz-Shea 2003; Zorn 2003). Alas, this means that there is no one-size-fits-all option for a methods text. The key questions for guiding the selection of a text include: is the class the only such course students will take or is it one of a set of methods classes required of students? If the latter, are the courses sequenced such that the instructor can count on students coming in with a common foundation of knowledge? Or, is the course in question actually the one intended to get all students on the same page before they tackle more advanced methods? Is the methodological training provided to students intended to make them skilled applied researchers or to position them to become political methodologists in their own right? Which text is the right text depends on the answers to these questions. In this review, I focus on five econometrics texts, four established texts with new (or reasonably recent) editions and one debut entry in the field (see table 1).

For classes that provide a comprehensive but less technical introduction to regression, Gujarati and Porter's *Basic Econometrics* and Wooldridge's *Introductory Econometrics: A Modern Approach* are staples on graduate methods syllabi. *Basic Econometrics* first appeared almost four decades ago and is now a classic, with the most recent edition dating to 2009. In comparison, the first edition of *Introductory Econometrics* appeared in 2000, with the most recent edition coming out just this past year. Also appearing very recently is (what I presume will be the first but not the last edition of) Bailey's *Real Stats: Using Econometrics for Political Science and Public Policy*. Each of these three textbooks is appropriate for a one-semester course in programs that require only a single semester of econometric training. But they are also suitable for a first- or second-semester course in a multi-course sequence that becomes progressively more technical. In addition, the topics included in each are plentiful enough (particularly in *Basic Econometrics* and *Introductory Econometrics*) that any one of the three could be used for a two-semester sequence where the purpose is to produce adept applied researchers rather than political methodologists per se. For classes aimed at

producing political methodologists (or applied researchers with a firm grasp on econometric theory), the more technically advanced choices are either Greene's *Economic Analysis* (the most recent edition of which dates to 2011) or Fox's *Applied Regression Analysis and Generalized Linear Models* (the most recent edition of which dates to 2015).[1] Each includes at least two semesters (or more) of material but, with careful curation of the assigned readings, each text could potentially also be used for a stand-alone one-semester course.

LESS TECHNICAL TREATMENTS

Gujarati has produced five editions of *Basic Econometrics*, with the last edition adding Porter as a coauthor.[2] With each edition, the focus has been on clear exposition that does not depend on students having a sophisticated mathematical or statistical background. The fact that Gujarati and Porter include a good deal of explanatory text in conjunction with the presentation of any equations makes a significant contribution to its accessibility. Further, the inclusion of numerous illustrative examples, a hallmark of every edition of this book, adds to its general accessibility. Regrettably, but not surprisingly given the disciplinary backgrounds of the authors, virtually all of the examples are drawn from economics, which makes them a bit less appealing for political science students. This is a minor concern, however, since the examples do serve their intended purpose (i.e., contributing to the clarity of the exposition of econometric concepts and applications).

In terms of substantive coverage, approximately one third of the Gujarati and Porter text is devoted to the development of the linear model, with another quarter focused on violations of the assumptions underlying that model. The remainder of the text deals with varied topics, including models for non-linear regression, qualitative response data, panel data, time series data, and simultaneous equations. *Basic Econometrics* is especially good when it comes to its treatment of violations of the assumptions underlying the linear model. Gujarati and Porter first deconstruct the various violations, making explicit the implication of each violation for the qualities of our estimators. They then provide a menu of diagnostics and means for addressing those violations. This feature makes the text useful as both an instructional text and a reference tool. It admittedly also runs the risk of encouraging students to see violations as merely technical issues to be addressed with canned off-the-shelf solution rather than developing students' ability to consider violations in meaningful substantive terms. Appropriate guidance from an instructor can, of course, help students avoid this problem.

As with *Basic Econometrics*, the appeal of Wooldridge's *Introductory Econometrics: A Modern Approach* is its accessibility to a non-technical audience. And, as with *Basic Econometrics*, the examples in *Introductory Econometrics* are primarily drawn from economics. Nonetheless, they do a perfectly serviceable job of contributing to the text's accessibility, even for

with panel data, while those relying on the former could do a reasonable job relying only on what appears in the text.

Like *Basic Econometrics* and *Introductory Econometrics*, Bailey's *Real Stats* is self-consciously and purposefully directed at students with minimal technical background. Written by a political scientist for political science, public policy, and law

The fact that Gujarati and Porter include a good deal of explanatory text in conjunction with the presentation of any equations makes a significant contribution to its accessibility.

a political science audience. The objective in each edition of *Basic Econometrics* has been to provide a bridge between how econometrics is conventionally taught in the classroom and how it is used by researchers in the conduct of actual empirical work. In other words, the text attempts to teach the material by presenting the opportunities and challenges of econometrics through the lens of applied research. This does not mean that Wooldridge eschews theoretical treatments of the material altogether. For example, in the second chapter, devoted to the simple regression model, the author introduces the methods of moments and illuminates the connection between the methods of moments and the general linear model. The result is a richer treatment that is still quite comprehensible to a less technical audience.

The first third of *Introductory Econometrics* is devoted to regression in the context of cross-sectional data. The remainder of the text covers time series and panel data methods, instrumental variables, simultaneous equations, and limited dependent variables. It concludes with a brief chapter on executing an empirical research project. Though that chapter is not a substitute for something like Powner (2015), it is a nice feature to have embedded in a textbook. Especially appealing is Wooldridge's treatment of consistency and efficiency, which are given their own (albeit brief) chapter in addition to being threaded throughout the rest of the book as appropriate. Also appealing is the two-chapter coverage of estimations with panel data. The treatment of panel data in the Wooldridge text is more substantial than that appearing in the Gujarati and Porter text. Students relying on the latter would most likely need a supplementary text to do something meaningful

audiences the text is intended as a sharp digression from the standard model of econometrics training. The mix of examples Bailey relies on draws heavily from political science and public policy, making them more relevant to a political science readership as compared to those in either the Gujarati and Porter text or the Wooldridge text.[3] More importantly, *Real Stats* employs the standard regression model as a unified framework. In other words, the standard regression model is the foundation on which each new concept (e.g., hypothesis testing, difference of means) rests. In his preface for instructors, Bailey makes two persuasive arguments for this approach. First, it means that, once students are familiar with the standard regression model, they "start from a comfortable place and can see the continuity that exists" as they tackle new material (Bailey 2015, xxvii). That is, students will be better positioned to integrate new material into a cohesive understanding of the econometric tools they are learning because there is a common frame of reference in which each new piece of material can be rooted. And, second, it affords the opportunity for instructors and students to revisit and review the standard regression model throughout the semester. Given what we know about the importance of repetition in the acquisition of any new technical skill, this is a very good thing.

In terms of the substantive content of the Bailey text, it opens (unconventionally but appealingly) with two brief chapters that introduce students to causal inference and good data practices, respectively. The first presents randomized experimental methods as the "gold standard" against which to consider the causal claims made by observational studies. The second is intended to foster the development of good habits when working with data (start by knowing the data!) and is explicit in inculcating students with strong replication norms (do not trust analyses that cannot be replicated!). The next five chapters are devoted to developing the standard regression model, with subsequent chapters focused on panel data (particularly fixed effects), instrumental data, experiments, regression discontinuity, and limited dependent variables.

In the penultimate chapters, gathered together in a section entitled "Advanced Material," the Bailey text covers time series followed by what is dubbed advanced OLS and then advanced panel data. The time series chapter is a mere 32 pages long but does an adequate job of introducing students (albeit briefly) to the key issues and common solutions

Table 1

Books Reviewed in this Article

Bailey, Michael A. 2015. *Real Stats: Using Econometrics for Political Science and Public Policy*. New York, NY: Oxford University Press.

Fox, John. 2015. *Applied Regression Analysis and Generalized Linear Models*, 3rd Edition. Thousand Oaks, CA: Sage.

Greene, William H. 2011. *Econometric Analysis*, 7th Edition. Upper Saddle River, NJ: Prentice Hall.

Gujarati, Damodar and Dawn Porter. 2009. *Basic Econometrics*, 5th Edition. New York, NY: McGraw-Hill Education.

Wooldridge, Jeffrey M. 2015. *Introductory Econometrics: A Modern Approach*, 6th Edition. Mason, OH: Cengage Learning.

for time-series-related challenges. The advanced OLS chapter includes more technical aspects of OLS that were set aside in the earlier chapters in favor of a focus on the practical issues associated with the standard regression model. The placement of this chapter near the end of the text makes sense given Bailey's asserted goal of focusing first and foremost on the substantive and the practical. An instructor could, however, profitably use that chapter earlier; for example, after covering multiple regression in chapter 5. Similarly, an instructor could use the chapter on advanced panel data (which includes dynamic models for the analysis of panel data and random effects models) after covering fixed effects in chapter 8.

None of these texts—*Basic Econometrics*, *Introductory Econometrics*, or *Real Stats*—assumes much prior knowledge on

with advanced technical backgrounds would appreciate and benefit from such a glossary.

MORE TECHNICAL TREATMENTS

The first edition of Greene's *Econometric* Analysis came out in 1990 and quickly found a place on a good many graduate methods syllabi. Greene's avowed objective is to serve as an introduction to applied econometrics *and* to provide a theoretical foundation that would allow students to "recognize new variants of the models learned about here as merely natural extensions that fit within a common body of principles" (Greene 2012, xxxiv). In other words, the text is directed at teaching students both the practice (how to undertake high quality econometric work) and the theory (what motivates econometric tools) of econometrics.

In other words, the text is directed at teaching students both the practice (how to undertake high quality econometric work) and the theory (what motivates econometric tools) of econometrics.

the part of students. Some basic college-level algebra and a reasonable foundation in elementary probability are the only prerequisites. The Gujarati and Porter text includes virtually no material about probability per se. The Bailey text includes a handful of pages on the topic while the Wooldridge text usefully includes an appendix covering the fundamentals of probability. In truth, the brevity of that appendix means that it works best as a refresher or ready reference for students rather than as an accompaniment to supplemental instruction about probability. In short, when adopting any of these three texts, instructors will need to assign an additional text for instruction on probability unless that has been covered in a prior course.

Additionally, none of these three texts require students to be familiar with matrix algebra. The Bailey text eschews matrix algebra all together. Both the Gujarati and Porter text and the Wooldridge text include appendices that provide a review of the fundamentals of matrix algebra as well as the linear regression model in matrix form. These are certainly nice to have but they are only useful if an instructor merely wishes to take a brief detour to ensure her students have seen (but not necessarily mastered) regression in matrix form. If an instructor intends to rely on matrix algebra in any substantial or sustained way over the course of the semester, then the materials included in these appendices will not be sufficient.[4]

Parenthetically, both the Wooldridge and the Bailey texts include glossaries, while the Gujarati and Porter text does not. While a glossary cannot substitute for clarity of exposition in the body of a text, and certainly does not make or break a text in terms of its utility as an instructional tool, a well-done glossary (like the glossaries included in *Introductory Econometrics* and *Real Stats*) is an important value-added element of a text. And, while the benefit is probably greatest for students with non-technical backgrounds, I would argue that even those

Every edition of Greene's text has been marked by breadth in its substantive coverage. After an introductory chapter, the *Econometric Analysis* devotes four chapters to the linear regression model, least squares, and hypothesis tests. It then proceeds to cover a host of more advanced and/or specialized topics. These include, for example, systems of equations, models for panel data, maximum likelihood estimation, Bayesian estimation, and discrete choice data. Though it is hard to imagine one second- or third-semester course including all of these specialized topics, their inclusion in the text gives instructors the opportunity to tailor their syllabi to account for their program's curricular structure and the needs of their particular students. It also means that the text will serve as a handy reference for students when they take advanced courses devoted to any of those specialized topics. For example, the chapters on serial correlation and nonstationary data would pair well with Box-Steffensmeier et al. (2014) in a class devoted to time series analysis. And, the material on event counts and duration data would nicely supplement Box-Steffensmeier and Jones (2004) in a class focused on those topics.

One downside of the breadth of Greene's text is that the volume of material has the potential to overwhelm students, including very strong students. This means that instructors would do well to select their reading assignments carefully as students will likely require more-than-usual specificity when it comes to the readings.[5] In addition, *Econometric Analysis* absolutely assumes students are not only familiar with but conversant with basic calculus and matrix algebra. Accordingly, to minimize student (and instructor) frustration, students should be comfortable with both before using *Econometric Analysis*. The text does include a terrific matrix algebra appendix that covers, among other things, the algebraic manipulation of matrices, the geometry of matrices, and the matrix approach to solving a system of linear equations.[6]

But that appendix is best used as a reference for students to refresh themselves as needed on particular elements of matrix algebra rather than as a substitute for some experience with matrix algebra.[7]

Two additional features of *Econometric Analysis* merit particular note. First, with each successive edition, Greene has incorporated additional substantive examples. They are generously sprinkled throughout each chapter and are invaluable for translating econometric theory into applied econometrics for students. The most recent edition includes more than 200 of them. Parenthetically, as with *Basic Econometrics* and *Introductory Econometrics*, the examples are drawn from economics. Second, virtually every chapter includes a set of useful exercises and applications instructors can assign for homework or which students can use to practice. There is usually a sufficient number and variety in each set that instructors can pick and choose among them to suit their needs. The exercises do, however, have the draw back of focusing almost exclusively on the technical aspects of the material without connecting back to a substantive context.

Like the Greene text, Fox's *Applied Regression Analysis and General Linear Models* (now in its 3rd edition) assumes students have a firm grip on probability and statistical inference. It also assumes that students are familiar with basic calculus and comfortable with matrix algebra. There are appendices devoted to calculus and matrix algebra. As with Greene's matrix algebra appendix, however, the Fox appendices are best used for students to refresh themselves on material they have seen before rather than for learning material they have not seen before.[8]

In terms of its substantive coverage, the Fox text opens with a chapter that juxtaposes statistical models with social reality, observation with experimentation, and populations with samples. It then moves to a trio of chapters gathered under the rubric of "Data Craft." These include a conceptual introduction to regression; examining univariate, bivariate, and multivariate data; and transforming data (e.g., transforming nonlinear data). The second part of the text takes up regression, both its estimation and inference, and includes dummy variable regression and analysis of variance. This part also has chapters devoted to the statistical theory of linear models and their vector geometry. The third part then tackles various model diagnostics. The fourth and fifth parts address, respectively, the generalized linear model and extensions of both the linear and generalized linear model. The fifth part is impressive in the number and variety of topics covered: time series, nonlinear regression, nonparametric regression, robust regression, missing data (including Bayesian multiple imputation), bootstrapping, and model selection. Finally, the last section of the text (new to the 3rd edition) considers mixed effects models for clustered data.

Considered as a whole, arguably the strongest part of the text is the attention it pays to model fit. The chapters devoted to influential data in the regression context and the detection and handling of multicollinearity, non-normality, non-constant error variance, and non-linearity are useful for helping students to conceptually grasp the issues as well as guide their search for solutions. Also appealing is the chapter on model selection. The treatment of model selection does lay out the technical details of model selection but Fox carefully distinguishes between model selection in the context of prediction and in the context of interpretation. This is crucially important for helping students to think in substantive and not just technical terms. Another plus (from the perspective of a political science audience) is the fact that the examples Fox (a sociologist) uses are drawn from a broader range of disciplines that is typical of standard econometrics texts.

The fact that the Fox text includes treatments of numerous advanced topics (e.g., nonparametric regression analysis, bootstrapping) should be appealing, but only as long as instructors (and students) bear in mind that those treatments are not intended to be exhaustive, merely introductory. As Fox notes in his preface with regard to these topics, his "aim … is to provide (1) enough information so that readers can begin to use these methods in their research and (2) sufficient background to support further work in these areas should readers choose to pursue them" (2015, xviii). In this regard, it is much like Greene in that the material it includes on any given advanced topic would most useful if paired with supplemental material on that topic.

Incidentally, while the Fox text is quite good on its own, its appeal is enhanced when it is coupled with Fox and Weisberg's *An R Companion to Applied Regression* (2011).[9] Though it is not a comprehensive users' manual for R, nor is it intended to be such, this text does provide easy-to-follow instructions for accessing and using R. It covers reading and manipulating data, transforming data, fitting linear and generalized linear models, executing diagnostics, and graphing. Moreover, it has a companion website that includes, among other things, a really useful appendix on each of the following: nonlinear regression, robust regression, nonparametric regression, time-series regression, Cox regression for survival data, multivariate linear models, mixed-effects models, structural-equation models, and bootstrapping.

It is nice to have choices and, in the case of econometrics texts for graduate-level classes, instructors are fortunate to have several very good choices available to them. Each of the five texts discussed here offers different combinations of attractive features. Which combination is the best choice depends on the goals of the particular class and the curricular structure of the particular program. ∎

NOTES

1. The first edition of Fox's text, which came out in 1997, was titled *Applied Regression Analysis, Linear Models, and Related Methods*. The revised title came with the 2nd edition and reflected an important change from the original edition. In particular, Fox provided an extended treatment of generalized linear models, a topic addressed in a much more abbreviated fashion in the original edition.

2. Porter is also now coauthor with Gujarati of *Essentials of Econometrics* (2010). *Essential of Econometrics* is pitched at a lower level of technical detail than *Basic Econometrics* and is intended to provide a practical exposition of key concepts. While *Basic Econometrics* is more appropriate for graduate-level training in econometrics, *Essential of Econometrics* would be more useful at the advanced undergraduate level.

3. Bailey is also the author of *Real Econometrics* (2016). His strong commitment to using examples that will be as appealing as possible for his intended audience is evidenced by the fact that *Real Stats* relies on political science examples while Bailey's *Real Econometrics* relies on economics examples.

4. The relevant chapters in Gill (2006) would be useful as a supplement in this context.

5. Amato (1995) also makes this point in his review of the second edition of *Econometric Analysis*.

6. In addition to the appendix on matrix algebra (as well as one on the data sets used in the illustrative applications and another with the standard statistical tables), the text includes the following appendices: B: Probability and Distribution Theory; C: Estimation and Inference; D: Large-Sample Distribution Theory; and, E: Computation and Optimization.

7. Even in this case, students might be better served by referencing, first, the relevant sections of Moore and Siegal (2013) or Gill (2006). Both have expositions that are very accessible and are accompanied by exercises to help students master the material.

8. The complete set of appendices appearing in the Fox text includes: A. Notation; B. Matrices, Linear Algebra, Vector Geometry; C. An Introduction to Calculus; and, D. Probability and Estimation. The appendix devoted to notation appears in the text itself as well as on line while the others appear on line exclusively.

9. The original edition of that text was authored by Fox and was entitled *An R Companion to S-PLUS Companion to Applied Regression*.

REFERENCES

Amato, Timothy. 1995. "Review of William Greene's *Econometric Analysis*." *Political Methodologist* 7 (1): 27–28.

Bailey, Michael. 2015. *Real Stats: Using Econometrics for Political Science and Public Policy*. New York, NY: Oxford University Press.

——. 2016. *Real Econometrics: The Right Tool to Answer Important Questions*. New York, NY: Oxford University Press.

Box-Steffensmeier, Janet M., John R. Freeman, Matthew P. Hitt, and Jon C. W. Pevehouse. 2014. *Time Series for the Social Sciences*. New York, NY: Cambridge University Press.

Box-Steffensmeier, Janet M. and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York, NY: Cambridge University Press.

Fox, John. 2015. *Applied Regression Analysis and Generalized Linear Models*, Third Edition. Thousand Oaks, CA: Sage.

——. 2011. *An R Companion to Applied Regression*, Second Edition. Thousand Oaks, CA: Sage.

——. 1997. *Applied Regression Analysis, Linear Models, and Related Methods*. Thousand Oaks, CA: Sage.

Fox, John and Sanford Weisberg. 2011. *An R Companion to Applied Regression*, Second Edition. Thousand Oaks, CA: Sage.

Gill, Jeff. 2006. *Essential Mathematics for Political and Social Research*. New York, NY: Cambridge University Press.

Greene, William H. 2011. *Econometric Analysis*, Seventh Edition. Upper Saddle River, NJ: Prentice Hall.

Gujarati, Damodar and Dawn Porter. 2009. *Basic Econometrics*, Fifth Edition. New York, NY: McGraw-Hill Education.

——. 2010. *Essentials of Econometrics*, Fourth Edition. New York, NY: McGraw-Hill Education.

Moore, Will H. and David A. Siegel. 2013. *A Mathematics Course for Political and Social Research*. Princeton: Princeton University Press.

Powner, Leanne C. 2015. *Empirical Research and Writing: A Political Science Student's Practical Guide*. Thousand Oaks, CA: CQ Press.

Schwartz-Shea, Peregrine. 2003. "Is This the Curriculum We Want? Doctoral Requirements and Offerings in Methods and Methodology." *PS: Political Science & Politics* 36 (3): 379–86.

Wooldridge, Jeffrey M. 2015. *Introductory Econometrics: A Modern Approach*, Sixth Edition. Mason, OH: Cengage Learning.

Zorn, Christopher. 2003. "Some Thoughts on Graduate Political Science Curricula in Quantitative Methods." *Political Methodologist* 11 (2): 16–20.