

The Political Methodologist

NEWSLETTER OF THE POLITICAL METHODOLOGY SECTION
AMERICAN POLITICAL SCIENCE ASSOCIATION
VOLUME 16, NUMBER 2, SPRING 2009

Editors:

PAUL M. KELLSTEDT, TEXAS A&M UNIVERSITY
kellstedt@polisci.tamu.edu

DAVID A.M. PETERSON, TEXAS A&M UNIVERSITY
dave@polisci.tamu.edu

GUY D. WHITTEN, TEXAS A&M UNIVERSITY
whitten@polisci.tamu.edu

Editorial Assistant:

MARK D. RAMIREZ, TEXAS A&M UNIVERSITY
mdramir@politics.tamu.edu

Contents

Notes from the Editors	1
Book Reviews	2
Nathaniel Beck: Review of <i>Microeconometrics: Methods and Applications</i> , by A. Colin Cameron and Pravin K. Trivedi	2
Shawn Treier: Review of <i>Semiparametric Regression for the Social Sciences</i> , by Luke Keele	5
Mark Pickup: Review of <i>An Introduction to State Space Time Series Analysis</i> , by Jacques J.F. Commandeur and Siem Jan Koopman	7
Computing and Software	10
Michael Malecki: Replicable Publication-Ready Model Output with R package <code>apsrtable</code>	10
Announcements	13
Janet Box-Steffensmeier and Erin McAdams: Visions in Methodology Workshop Supports Women in Political Methodology	13

Notes From the Editors

The proliferation of technology, particularly the world wide web, is generating an immense amount of accessible data for scholars. It is now easier than ever before for researchers to obtain data via the web from government organizations, individual candidates, media archives, blogs, and the web-pages of other researchers. With this growth in the availabil-

ity of data comes a concurrent growth in statistical methodology and statistical software. Fortunately, this growth in data, statistical methods, and software is being accompanied by books that explain how to solve contemporary data analytic problems with the latest statistical methods and software.

This issue of *The Political Methodologist* includes three book reviews that should be of interest to scholars wishing to add new techniques to their methodological toolbox or simply looking for that new book to use in a graduate-level methodology seminar. Nathaniel Beck provides a thorough review of *Microeconometrics: Methods and Applications* by Cameron and Trivedi. Yet, Professor Beck's article is not just a book review, but a comparison of textbooks commonly used for advanced graduate student econometric courses. For scholars wanting to know what book to adopt in a post-regression course or simply thinking about switching textbooks, Beck's side-by-side analysis is invaluable. The next two book reviews focus on statistical methods that are not frequently used in political science, although we think they should be. Shawn Treier reviews Luke Keele's book, *Semiparametric Regression for the Social Sciences*. The review provides a brief introduction to the importance of semiparametric regressions and why scholars might want to consider them when constructing models. In particular, the review and book illustrate the utility of semiparametric models in lieu of simple parametric transformations in regression models. Mark Pickup provides an overview of *An Introduction to State Space Time Series Analysis* by Commandeur and Koopman. As Pickup notes in his review, state space models provide a more general modeling strategy than conventional econometric and Box-Jenkins time-series approaches. The state space set-up encompasses these existing approaches, while allowing more

flexibility in estimating non-stationary time series, time-varying parameters, and missing data. Thus, we hope to see more use of these models in the future.

In the computing and software section, Michael Malecki provides detailed instructions on his R package `apsrtable`. The package allows users to generate L^AT_EX code in R to create tables suitable for publication in the *American Political Science Review* and related journals. The package provides more flexibility in creating tables than conventional means leaving us to believe that *TPM* readers will find this package extremely helpful. Finally, we end with an announcement from Janet Box-Steffensmeier and

Erin McAdams on the *Visions in Methodology Workshop* this past October at The Ohio State University. The last few years have seen an increase in specialized methodology conferences and we welcome readers wishing to share their experiences at these conferences with the rest of the methodology section to submit them to *TPM*.

We would like to thank each of this issue's contributors. We hope you enjoy the content of this issue as much as we do and, as always, welcome ideas for future editions of *TPM*.

The Editors

Book Reviews

Review of Cameron and Trivedi Microeconometrics: Methods and Applications

Nathaniel Beck

New York University
nathaniel.beck@nyu.edu

Microeconometrics: Methods and Applications. A. Colin Cameron and Pravin K. Trivedi. Cambridge University Press, 2005, 1056 pages. \$86.00, ISBN 978-0-521-84805-3 (hardcover).

Microeconometrics Using Stata. A. Colin Cameron and Pravin K. Trivedi. Stata Press, 2008, 692 pages. \$65.00, ISBN 1-59718-048-3 (paperback).

It is a good time to be teaching Ph.D. quantitative methods, at least in terms of the texts one can choose from. In 2002 Jeffrey Wooldridge's *Econometric Analysis of Cross Section and Panel Data* (MIT Press) ("W") joined William Greene's *Econometric Analysis* ("G") which is now in its sixth edition from Prentice-Hall. These two texts were joined in 2005 by Cameron and Trivedi's *Microeconometrics: Methods and Applications* ("C&T"). I have used all three texts for the course that follows the linear regression course (of the type often misleadingly called "maximum likelihood," but, more accurately, "purpose built models"), and have been quite happy with all three. They are different, and so I will take this opportunity to spell out some differences.

All three books are massive; W is the slimmest at 750 pages, with C&T coming in at about 1000 pages and G at almost 1200 pages. None of these books can be covered in a semester course. Thus in my case I had to pick and

choose amongst the topics. On this dimension C&T wins easily; the text is quite modular, and so one can pick and choose almost ignoring the chapter ordering given by the authors. G is closer to C&T on this score, but one has to frequently track back to previous chapters. Cameron and Trivedi themselves provide a guide for using the text in a one quarter course which takes the various chapters out of numerical order. While the tight integration of W is intellectually a joy, it surely makes it hard to pick and choose chapters for those (most instructors) who do not want to use the whole book.

The other obvious difference between the books is that G has a substantial time series section (over 150 pages) while the other two books are entirely cross sectional. This has obvious consequences for the choice of a text. I personally find that my course already has more than enough without doing time series, so I am happy with the purely cross sectional books, but your mileage may vary.

G shows its heritage. While the 6th edition is much different from the first, it lacks some important modern topics, in particular, matching, missing data and various "robust" methods. These topics all get their own chapters in C&T, and W has a long chapter on issues of selection bias and matching. Perhaps because we have seen a lot of matching work since W was published, I find its notation and treatment to be non-standard and to not cohere well

with standard discussions of matching (though on its own terms it shows Wooldridge's usual insightfulness). Cameron and Trivedi's chapter is consistent with the now relatively standard Rubin treatment of causality, and so makes for a very nice introduction to the subject, allowing students to then proceed with further readings in this area. (Matching is a fast moving area, so no textbook that is not revised constantly can keep up with all the new work in this area). While one might not have time to use the bootstrapping or missing data chapters, students are certainly better off being able to consult these chapters as needed.

Each of the books has some advantages over the others. C&T has, not surprisingly, excellent coverage of count data and event history, and W has the most comprehensive treatment of panels (and particularly issues of unmeasured heterogeneity). G is more comprehensive, with both C&T and W limited (intentionally) to issues related to microeconomics (largely human resource economics). W is most closely tied to microeconomic foundations (which may not be a plus for readers of this review). Fortunately many of the issues faced by political scientists using cross sectional methods are similar to human resource issues, so this is not a huge issue in deciding whether to use these texts in a political science course. Obviously G's time series chapters (and related time series cross section chapters) would appeal to those interested in comparative political economy. None of the books deal with spatial econometrics (which should be in the political methodologist's tool kit, and, for whatever reason, the analysis of networks seems not to be within the current state of microeconomics).

All three books have more than enough in terms of asymptotic proofs for my needs (and I think the needs of almost all readers of this review). Each of the books has a slightly different style of presenting proofs: G is the most straightforward; W is probably mathematically (as well as economically) deeper, though the proofs often refer back to assumptions made several pages before; C&T have a nice mix of providing the appropriate proofs and intuitions (often referring the reader to more technical treatments for dotting some i's). The C&T style works well for me, but, again, your mileage may vary. Almost all of our students will find the proofs in all of these books challenging (until they learn the trick that almost all asymptotic proofs work by approximating functions of the likelihood by the first two terms of its Taylor expansion, at which point the wizardry no longer seems quite so impressive). There are clearly more advanced texts (in terms of asymptotics) than any of these three, and there are clearly texts that demand less than these three, but these books seem at the right level for my classes.

Turning specifically to C&T, the overall structure of the book (other than its very helpful modularity) will not surprise anyone. There is a very nice consistency to the

chapters, each of which concludes with a helpful "Practical Considerations" section, a useful bibliography (neither too skimpy nor too encyclopedic) and a set of pencil and paper exercises which nicely complement the data exercises I rely on.

After a 50 page introduction, there are approximately 400 pages on general econometric methods. We have here good standard treatments of maximum likelihood, GMM, numerical methods, simulation and Bayesian methods. While a Bayesian would not find this an adequate compromise, any student could read the Bayesian chapter and pick up some familiarity with Markov chain Monte Carlo methods. In keeping with the modern microeconomic nature of the book, there are also good (though not exhaustive) treatments of quantile regression, semiparametric methods and bootstrapping. This section also contains an intelligent discussion of both hypothesis tests and model selection (intelligent in that the reader gets some understanding of the theoretical issues related to testing and model selection, not just a set of techniques).

The heart of the book for me is the next 230 pages, which cover the standard models often estimated by maximum likelihood: binary choice, multinomial choice, selection, survival and count data. As befits a text written by the authors of the standard book on count data, the count data chapter is both excellent and more thorough than in many other books. There are also three full chapters on survival models, including the important topics of repeated spells and a very thorough discussion of heterogeneity. The other chapters here are similarly thorough.

Each of the chapters has a good applied focus (without in any sense being a cookbook). A student going through these chapters will have a good understanding of the issues involved with using these various methods and models. The relevant derivations are all in the chapters, but they do not overwhelm the discussion of more practical applied issues. Each chapter starts with a specific example drawn from economics; fortunately these examples are not all that far removed from examples using political data. This section of the book has enough applied and theoretical material so as to make ancillary texts on these issues unnecessary (more so if one uses the Stata companion book discussed below).

The next section is about 100 pages on panel models. Given that the book is on microeconomics, this section is about large-N small-number-of-wave panels, which makes it less interesting for me (and comparative political economists in general, although a book relevant for comparative political economy could not be entitled *Microeconometrics*). The treatment here is quite reasonable, and the various panel models are nicely covered; good use is made of GMM and there is a nice discussion of non-linear panel models.

Finally, under the rubric of "Further Topics" there

is a useful set of chapters (about 100 pages total) on stratified/clustered samples, treatment evaluation (the counterfactual model), measurement error and missing data. It is nice to have these, and I use the latter three chapters in my course. It is rare to see a good missing data chapter in a text, and this chapter shows Cameron and Trivedi's interest in helping the applied researcher to do better econometric research.

Stata Press has just released a companion volume to C&T, *Microeconometrics Using Stata* ("MUS"). While I have obviously not yet used MUS, it looks like it will make a real contribution to my course. MUS is much more than a Stata manual for the relevant Stata routines for the methods discussed in C&T; there is a good discussion of a variety of applied issues relevant to those methods (graphics, diagnostics, all the important things that may get ignored in a theoretical course). MUS also makes up for what I see as shortcomings in Stata's documentation; most graduate students buy Stata without the full documentation, and it is hard to figure out which sophisticated things to do without that documentation which is not available online. (The Stata commands are well documented internally, but students need more than a full description of the various command options.) MUS also makes it unnecessary for students to have access to various Stata references, such as Stata texts on maximum likelihood or programming.

MUS also provides discussion and code for writing one's own maximum likelihood code (in both Stata's programming language and Mata), as well as writing code for simulations, matrix demonstrations for OLS and other numerical methods. It is, of course, critical that students understand how to actually code maximum likelihood routines, and to understand that the various routines that are hard coded in Stata are neither magic nor require an advanced degree in numerical methods. For a number of years I used R for this part of the course, but I always found it hard to both teach statistics and R at the same time (at NYU we start with Stata, so students come to my course with no knowledge of R). We can debate the merits of R and Stata until the cows come home, but, with Mata (and the somewhat newer optimizer), students can write the same type of code in Stata as they can in R. Given that they come in knowing Stata, and that it is a lot easier to deal with real data sets in Stata (using variable names, dealing with missing data, etc.), I am very happy to teach my course using Stata. The only problem was that I had to supply the documentation and code ("I" being a euphemism for "my excellent TAs"). MUS provides that documentation, with data and code avail-

able for easy installation into Stata. The combination of the good discussion of simulation in C&T and the code provided in MUS will make it much easier for me to have many more simulation exercises this year. Thus students will be in a much better position to see that what we teach is not magic but sound statistical modeling.

MUS also has excellent treatment of various applied topics, being particularly strong on instrumental variable estimation, panel data, discrete choice and event count data. It does not, however, provide complete practical advice in the Stata context for all the topic covered in C&T. In particular, it does not cover event history analysis so some additional applied readings (perhaps Box-Steffensmeier and Jones, *Event History Modeling*) would be helpful. MUS also does not discuss the topics in the last portion of C&T. However, Cameron and Trivedi have made available all the data (and most of the Stata routines) used in their book on their own website (<http://cameron.econ.ucdavis.edu/mmabook/mma.html>).

I used C&T last year and it worked quite well and I will use it again this year. To my mind it is a very nice mix of good econometric theory combined with topics and discussion relevant to the applied cross-sectional researcher. Since the various microeconomic methods discussed work quite well for cross sectional political data, political scientists should not be put off by the title. I am even more excited about teaching it in the coming semester, since MUS makes it so much easier to get students better able to write the own maximum likelihood routines, do simulations, and do better applied work. To my mind the combination of the two books is a winner, and a winner in a strong field.

References

- Box-Steffensmeier, Janet M. and Bradford S. Jones. 2004. *Event History Modeling*. New York: Cambridge University Press.
- Cameron, A. Colin and Pravin K. Trivedi. 2008. *Microeconometrics Using Stata*. College Station, TX: Stata Press.
- Greene, William H. 2008. *Econometric Analysis*, 6th ed. Upper Saddle River, NJ: Prentice Hall.
- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Boston, MA: The MIT Press.

Review of Keele Semiparametric Regression for the Social Sciences

Shawn Treier

University of Minnesota
satreier@umn.edu

Semiparametric Regression for the Social Sciences. Luke Keele. Wiley, 2008, 230 pages. \$80.00, ISBN 978-0-470-31991-8 (hardcover).

Linear and generalized linear models are the preeminent tools for quantitative research in the social sciences and dominate the coverage in our methods courses. The typical social science student will learn the consequences and corrections for most of the violations of the regression assumptions, such as heteroskedasticity, autocorrelation, and the various problems which result in $E[\mathbf{X}'\varepsilon] \neq 0$ (omitted variables, measurement error or endogeneity in the regressors). The most basic of these assumptions though, the linearity of the model, attracts relatively little attention. When linearity is in question, the corrections are limited to logarithmic transformations, quadratic models, and other power transformations estimated in a linear (in the parameters) model or nonlinear least squares models with *a priori* specified functional forms. With generalized linear models, students learn to apply the canonical parametric functions linking the linear predictor to the expected value of the dependent variable without any consideration of alternative forms.

A more flexible approach would be completely nonparametric, with the linear regression model

$$y = \beta_1 + x_i\beta_2 + \varepsilon_i \quad (1)$$

replaced by

$$y_i = m(x_i) + \varepsilon_i, \quad (2)$$

where $m(\cdot)$ is an unknown function to be *estimated*. The expected value of the dependent variable is no longer restricted to be a linear additive function of the independent variables; instead of $E[y_i|\mathbf{x}] = \mathbf{x}_i\boldsymbol{\beta}$, we have a completely nonlinear, fully interactive specification $E[y_i|\mathbf{x}] = m(\mathbf{x})$. Of course, in standard social science applications with many independent variables, removing all functional form assumptions is impractical; instead, *semiparametric* methods which relax linear additivity for only a few independent variables or retain additivity among all variables will prove quite useful.

Luke Keele's book *Semiparametric Regression for the Social Sciences* serves as an introduction to these models. This book distinguishes itself from other texts covering nonparametric and semiparametric regression in several ways. First, the asymptotic theory and computational details of these models are extremely complicated; many texts on

these methods focus extensively on these aspects, making them impenetrable to many political scientists. Second, the applications in these books, especially in the statistics literature, are far removed from the realm of social science. Not only are the applications quite foreign to the reader, but the focus is normally on prediction rather than the interpretation of marginal effects. The strength of Keele's book is that it offers clear, straightforward explanations of the models, illustrated with social science (primarily political science) applications. Applied social science researchers should be able to incorporate these methods in their own research relatively easily after reading this book.

Chapter 1 provides the basic motivation for the model and sets forth the primary issues addressed by the book. Chapters 2–4 introduce nonparametric regression methods. The reader should note the book covers only regression models; while the same approach is applicable for smoothing histograms of the dependent variable alone, all of the methods here relate \mathbf{x} to \mathbf{y} . Chapter 2 begins by illustrating with an extremely simple estimate based on moving averages. The author transitions to a description of *weighted* moving averages, with the weights defined by *kernel* functions — i.e., kernel smoothing estimates. Keele provides a nice exposition of the kernel function, defines the most popular choices, and explains clearly how the choice of the size of windows (spans or bin widths) and kernel function affects the estimates. Finally, he contrasts these unweighted and weighted moving average models with local polynomial regression models (loess and lowess).

Chapters 3 and 4 cover nonparametric spline regression models. Chapter 3 explains the basis functions and knots which define the spline regression model in a very clear manner, as well as the differences between several types of splines. Chapter 4 describes the automatic selection of spline functions, obviating the need for researcher selected tuning parameters for the degree of smoothness. In general, the models considered in chapters 2 and 3 (kernel smoothing, loess, lowess, splines) are terms recognized by many researchers who do not understand the differences between these models. Here, the differences in modelling choices are explicated in an understandable fashion. The explanations frequently connect back to the linear regression model, which is quite familiar to the target audience of the book. There is also a strong emphasis on conducting statistical inference, testing between canonical parametric models and

the local polynomial or spline models. The author also describes multivariate versions of these models, but observes a more restricted model is preferable in these situations. The methods are illustrated with a political science example, explaining challengers' vote share in the 1992 House elections as a function of Perot vote share in the district (Data from Jacobson and Dimock 1994).

The primary purposes of the first four chapters is to establish the foundation for understanding semiparametric regression methods, which utilize the methods from these chapters to estimate the nonparametric components. Chapters 5 and 6 describe additive semiparametric models, where the marginal effects for each continuous variable are nonlinear, and derived from an unspecified univariate function:

$$y_i = \alpha + f_2(x_2) + f_3(x_3) + f_4(x_4) + \cdots + f_k(x_k) + \varepsilon_i \quad (3)$$

Analogous to a generalized linear model, chapter 6 details a generalized additive model, where

$$E[y_i | \mathbf{x}_i] = g(\alpha + f_2(x_2) + f_3(x_3) + f_4(x_4) + \cdots + f_k(x_k)) \quad (4)$$

and $g(\cdot)$ is the known link function between the expected value and the additive component. The functions $f_j(x_j)$ are estimated using the univariate methods of chapters 2–4, with the functional form of interactions between continuous variables estimated with the multivariate nonparametric regression. Keele does a very nice job in describing the estimation algorithm (backfitting) for this model as well as presenting realistic social science examples. Chapter 5 extends the 1992 House elections example to include the full set of independent variables, and chapter 6 reexamines four published empirical applications.

Chapter 7 juxtaposes three important methods: generalized additive mixed models, Bayesian statistics and matching based on propensity scores. The mixed model framework, in addition to estimating hierarchical generalized additive models, can also be used to estimate (from a classical perspective) all of the models covered in this book. It is also the best modelling approach for a Bayesian spline model. The short introduction is suitable, since the interested reader can pursue Ruppert, Wand, and Carroll (2003) further for the mixed model. Similarly, those interested in more detail on the Bayesian implementation should read Congdon (2006, ch. 10). The propensity scores example contains two examples: one which has little improvement over the standard parametric approaches and another where the generalized additive model is superior.

One weakness of the book is that it focuses almost exclusively on the statistics literature regarding nonpara-

metric and semiparametric models, and could benefit from incorporating models from the econometric literature. One prominent omission is single index models, which relaxes the assumptions of the generalized linear model in a very different manner than generalized additive models (Horowitz 1998, chs. 2-3; Yatchew 2003, ch. 7). Here, the linear additive component is retained, but the link function is assumed to be unknown.¹ For example, in a binary model, one still estimates the linear coefficients β , but also estimates the probability function; one no longer assumes *a priori* the model is logistic or normal. This model could be used instead of alternative parametric (symmetric or asymmetric) links, such as the scobit from Nagler (1994). The flexibility of this approach could also address some of the concerns of Achen (2002).²

In addition to binary or ordinal models, single index (or multiple index) models can be applied to tobit and sample selection models, which are highly sensitive to the parametric choices of the error terms (which, in turn, determine the functional form of the mean function). Unfortunately, these models do not have available routines in R, although some single index models could potentially be estimated using projection pursuit regression techniques, such as `ppr()`. The models could serve as a nice introduction to optimizing in R though.

Chapter 8 covers bootstrapping, both in general and applied to semiparametric regression models. Keele provides a nice summary of the various implementations of the bootstrap and examples of this technique for semiparametric regression. The author probably should discuss though, in a very informal manner, the asymptotics of these models. The nonparametric and semiparametric regression models have slower rates of convergence than the familiar $N^{-\frac{1}{2}}$ rate. This is a particularly important issue, since many social science data sets may have much fewer observations than the typical application in statistics.

The issue of asymptotic refinement could also be addressed more directly. The bootstrap itself in many instances merely obtains the asymptotic distribution³, but under certain conditions obtain asymptotic refinements which are more accurate than the first order asymptotic approximations. An informal description on when this occurs, especially for semiparametric regression, would be useful. There are also a variety of issues which arise in applying the bootstrap to nonparametric and semiparametric models, such as correcting for bias through explicit bias correction or (more commonly) by undersmoothing (Horowitz 2001, section 4.2).

¹One could potentially estimate the link function on a generalized additive model as well (Ichimura and Todd 2007, p. 5416).

²One still might want to specify some theoretical restrictions on the functional form; one possible restriction is monotonicity with respect to the index.

³One of the examples in this chapter for semiparametric regression appears to reproduce the asymptotic distribution. Even in these situations, the bootstrap might be useful by avoiding the computational complexities of some variance calculations. For example, the asymptotic standard errors of some nonparametric densities require nonparametric estimates of derivatives of the estimated functions.

The book finishes with an appendix on statistical software. Relegating the implementation steps to an appendix rather than inline is a good choice by the author. The methods covered in the book will be unfamiliar to many of the readers, so copious amounts of R code interspersed with the model descriptions would be distracting. The appendix identifies relevant R packages and provides example code. All of the files for the analyses in the book are downloadable, and are extremely clean and easy to follow. Some of the variable names do not quite match the datasets, although that is simple to rectify.

One slight improvement the author might consider is to devote more space in the appendix to summarize the most important commands and options from each package. The author explains the `gam()` function very well, but other functions and packages could be outlined here as well. One instance where there is confusion is with the function `locfit`, which estimates a local polynomial regression. Its usage in the book seems slightly nonstandard; instead of a model formula of $y \sim \text{lp}(x)$, the author uses $y \sim x$, which will result in exactly the same fit — except when one changes the default values of the function.

In general, this is an excellent applied introduction to the topic. Keele presents a very basic, uncluttered description of a set of models which really should be used more often in the social sciences. This book will not only serve well as an assigned text in an advanced topics class, but is appropriate for self-study by most social science researchers who use quantitative methods.

References

Achen, Christopher H. 2002. "Toward a New Political

Methodology: Microfoundations and ART." *Annual Review of Political Science*, 5:423-450.

Congdon, Peter. 2006. *Bayesian Statistical Modelling*. 2nd ed. John Wiley and Sons.

Horowitz, Joel L. 1998. *Semiparametric Methods in Econometrics*. Springer-Verlag.

Horowitz, Joel L. 2001. "The Bootstrap." In *Handbook of Econometrics*. Vol. 5 Elsevier, Chapter 52.

Ichimura, Hidehiko and Petra E. Todd. 2007. "Implementing Nonparametric and Semiparametric Estimators." In *Handbook of Econometrics*. Vol 6B Elsevier, Chapter 74.

Jacobson, Gary and Michael Dimock. 1994. "Checking Out: The Effects of Bank Overdrafts on the 1992 House Elections." *American Journal of Political Science*, 38(3):601-624.

Nagler, Jonathan. 1994. "Scobit: An Alternative Estimator to Logit and Probit." *American Journal of Political Science*, 38(1):230-255.

Ruppert, David, M.P. Wand and R. J. Carroll. 2003. *Semiparametric Regression*. Cambridge University Press.

Yatchew, Adonis. 2003. *Semiparametric Regression for the Applied Econometrician*. Cambridge University Press.

Review of Commandeur and Koopman An Introduction to State Space Time Series Analysis

Mark Pickup

Simon Fraser University and University of Oxford
mark.pickup@gmail.com

An Introduction to State Space Time Series Analysis. Jacques J.F. Commandeur and Siem Jan Koopman. Oxford University Press, 2007, 240 pages. \$55.00, ISBN 978-0-19-922887-4 (hardcover).

An Introduction to State Space Time Series Analysis by Jacques J.F. Commandeur and Siem Jan Koopman (2007) is, in their own words, for "readers who are neither familiar with time series analysis nor with state space meth-

ods." Although maybe not immediately apparent, this book has the potential to substantially contribute to political science.

In a previous work, James Durbin and Koopman (2001) outlined the drawbacks of state space modeling: "In our opinion, the only disadvantages are the relative lack in the statistical and econometric communities of information, knowledge and software regarding these models."

This is also an accurate depiction of the disadvantages of state space modeling within the political science community. Consequently, Commandeur's decision to publish, with Koopman, his personal notes on state space methods in a manner that promises to help close the gaps in information and knowledge within the social sciences is most welcome.

Due to the gap in information, knowledge and software, state space time series models have not received the same attention within the social sciences as have autoregressive distributive lag, error correction, vector autoregression and other classic models. Still, they have been usefully applied in political science to such problems as the estimation of Supreme Court justice ideal points (Martin and Quinn, 2002), smoothing and forecasting public opinion (Pickup and Johnston, 2008; Jackman, 2005; Beck, 1990) and controlling for measurement error in economic popularity modeling (McAvoy, 1998). They are also used by political methodologists to develop dynamic event count models (Brandt and Williams, 2001) and estimate time series relationships with latent variables in dynamic measurement models (Armstrong, 2008; Kellstedt, McAvoy and Stimson, 1996). Moreover, anyone that has run an ARIMA has probably unknowingly benefited from the state space approach, as most popular statistical packages convert the ARIMA model into state space form for the purpose of estimation. The potential application of state space modeling within political science extends far beyond these few examples.

The state space approach to modeling time series is particularly useful for including time varying parameters, trends, cycling and structural breaks. It is a powerful method for filtering and smoothing time series, and forecasting. It is more flexible in solving problems of nonstationarity, missing data and measurement error than classical approaches. It is also more easily extended to the multivariate case. It is very useful for the estimation of non-linear dynamic models. Finally, it subsumes almost all classical linear time series models, in that most classical models can be expressed in state space form.

For those with a background in time series analysis, state space models (also termed dynamic models) relate observations on a variable to unobserved (read latent) states by an *observation* (or *measurement*) equation. For the Gaussian case:

$$y_t = z_t \alpha_t + \epsilon_t, \quad \epsilon_t \sim NID(0, \sigma_\epsilon^2)$$

The term z_t is an $m \times 1$ design matrix and α_t is an $m \times 1$ state vector, where m denotes the number of unobserved states. The unobserved states are modeled as dynamic processes by the *state* equation.

$$\alpha_{t+1} = T_t \alpha_t + \eta_t, \quad \eta_t \sim NID(0, Q_t)$$

In the state equation T_t is an $m \times m$ transition matrix, η_t it is a $m \times 1$ vector of state disturbances with 0

means and variances defined by the $m \times m$ diagonal matrix Q_t . The observation and state disturbances are assumed to be serially and mutually independent. It is straightforward to include covariates in the observation and/or state equations and the parameters on the covariates can themselves be modeled by a state equation (allowing them to vary over time).

An Introduction to State Space Time Series Analysis is not a lengthy read and it truly is an introduction. It begins with a brief overview of classical linear regression analysis—assuming prior knowledge—and although matrix algebra is used in a few sections, it is not integral to following the text. The book runs through the most commonly used state space models, explaining the terminology used, the process usually employed in model building and the necessary diagnostics (chapters 2-7). It repeatedly employs two simple examples throughout the text—traffic fatalities and inflation. It provides the intuition behind using the Kalman filter to estimate state space models in a maximum likelihood framework and the application of the Kalman approach to filtering, smoothing and forecasting time series (chapter 8). Of potentially great use to those interested in going further with state space methods is the chapter outlining a general unified notation for both univariate and multivariate state space modeling (chapters 8 and 9). This is the notation commonly found in more advanced texts, where it is not necessarily explained in as great detail as it is in *An Introduction to State Space Time Series Analysis*.

The text also includes a detailed comparison of state space modeling to Box-Jenkins ARIMA modeling (chapter 10). It focuses particularly on the fact that the time series of interest is not required to be stationary, and the ease with which state space models handle missing data, time varying parameters and multivariate extensions. It also highlights the fact that state space modeling is a structural approach in which unobserved components such as trends and cycles are explicitly modeled and diagnosed rather than just treated as a nuisance to be differenced away, as they are in the Box-Jenkins approach. By the time one has finished reading this chapter, it is difficult not to view state space modeling as superior to ARIMA modeling. Some but not all of the comparisons also apply to time series approaches other than ARIMA; however, this is not made explicit. This is not surprising, as this is an econometrics text intended for an audience that is used to employing ARIMA modeling. As this approach is not as popular in political science, the comparison is of less value and one is left wondering just how state space methods compare to a broader range of time series approaches.

As for the application of state space methods, the book includes a chapter on performing state space time series analysis using two programs: STAMP and SsfPack (chapter 11). The first is a component of Oxmetrics 5.0,

available from Timberlake Consultants. The second is available for free and is a library of C functions which can be linked to C programs or the programming language Ox. The lack of familiarity with Ox based programs in political science limits the accessibility of these programs. Therefore, the lack of software described by Durbin and Koopman in 2001 remains somewhat of a difficulty. It should be noted though that while it is not covered in this book, there is a state space package available in R called `sspir` (Dethlefsen and Lundbye-Christensen, 2006). The call to define the state space model is designed to be similar to the `glm` call. Further, as political methodologists become more familiar with `WinBUGS`, implementing state space methods within a Bayesian framework will become relatively straightforward. Overall, the limitations of software accessibility promise to decline.

As an introduction to a complex subject, this book succeeds spectacularly. It by no means answers every (or even many) of the more advanced questions one may have about state space methods. For example, only Gaussian state space models are covered, and one of the more interesting extensions of state space methods is to non-Gaussian models. For those interested in a more advanced treatment of state space methods, Durbin and Koopman (2001) have provided such a text, as has Andrew Harvey (1989), Mike West and Jeff Harrison (1997), and others. However, given the lack of familiarity with such approaches within political science, the basic introduction provided by Commandeur and Koopman is a useful asset even to those used to more advanced methods. Despite starting slowly, the text takes the reader right through to multivariate time series analysis and provides the necessary background to read the more advanced texts. Therefore, if one is interested in utilizing state space methods, this book provides an excellent entry point. For those not so much interested in using state space methods as they are simply curious exactly what they are all about this book also provides a surprisingly easy-to-read overview.

References

- Armstrong, David A. 2008. "Measuring the Democracy-Repression Nexus." Presented to the workshop on *Producing Better Measures by Combining Data Cross-Temporally* at the University of Oxford.
- Beck, Nathaniel. 1990. "Estimating Dynamic Models Using Kalman Filtering." *Political Analysis* 1(1): 121-56.
- Brandt, Patrick T. and John T. Williams. 2001. "A Linear Poisson Autoregressive Model: The Poisson AR(p) Model." *Political Analysis* 9(2): 164-184.
- Dethlefsen, Claus and Søren Lundbye-Christensen. 2006. "Formulating State Space Models in R with Focus on Longitudinal Regression Models." *Journal of the Statistical Software* 16, <http://www.jstatsoft.org>.
- Durbin, James and Siem Jan Koopman. 2001. *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
- Harvey, Andrew C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Jackman, Simon. 2005. "Pooling the Polls over an Election Campaign." *Australian Journal of Political Science* 40(4): 499-517.
- Kellstedt, Paul, Gregory E. McAvoy and James A. Stimson, 1996. "Dynamic Analysis with Latent Constructs." *Political Analysis* 5(1): 113-150.
- Martin, Andrew D. and Kevin M. Quinn, 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2): 134-153.
- McAvoy, Gregory E. 1998. "Measurement Models for Time Series Analysis: Estimating Dynamic Linear Errors-In-Variables Models." *Political Analysis* 7(1): 165-186.
- Pickup, Mark and Richard Johnston, 2008. "Campaign Trial Heats As Election Forecasts: Measurement Error and Bias in 2004 Presidential Campaign Polls." *International Journal of Forecasting* 24(2): 272-284.
- West, Mike and Jeff Harrison. 1997. *Bayesian Forecasting in Dynamic Models*. New York: Springer-Verlag.

Computing and Software

Replicable Publication-Ready Model Output with R package `apsrtable`

Michael Malecki

Washington University in St. Louis

malecki@wustl.edu

Overview

Formatting fitted model objects is not trivial and should not be done at the last minute before submitting a paper. It is too easy for modeling details to get lost, or simply for values to be omitted, transposed, or otherwise misplaced. Thus automation is a boon, if not a requirement, for replicability.

Fitted model objects in R contain a lot of information, but often we want to look at several of them side-by-side. It would be nice to name the models, and include some model-fit information, and to name the covariates. Nested and non-nested models and naming become a problem for `cbind` and `rbind`, and last but not least, we want this information in some legible format—either plain text or, preferably, \LaTeX . My R package `apsrtable`, installable from CRAN, seeks to solve all of these problems, creating tables ready to be published in the *American Political Science Review*. I might have named it after *Political Analysis*, but the latter often features other innovative presentations of data and results, such that “a *PA*-style table” is less identifiable.

Like many graduate students—including probably most who read *TPM*—I have typeset *a lot* of fitted models in the course of methods training, leaving aside the relatively few tables that make their way into final seminar, conference, or journal papers. We were “strongly encouraged” to prepare homework assignments using \LaTeX and create “professional-looking” tables of model output, with the expectation that we would get faster at it, while becoming attuned to the intricacies of \LaTeX ’s tabular and matrix environments. I found this exercise as tedious as formatting a bibliography by hand. So, I set out to automate R’s output of “professional-looking” tables. To complete the bibliography analogy, I also wanted to make the automated solution more user-friendly than \BibTeX BSTs—though “an order of magnitude easier than BST” is still an extremely low target.

Usage

To demonstrate `apsrtable`’s features, I’ll replicate some of the results from Persson and Tabellini’s (2003) *The*

*Economic Effects of Constitutions*¹, in particular models 1-3 of Table 6.1 (p. 159) (Figure 1).

Figure 1. Persson and Tabellini (2003), p.159, detail.

Table 6.1
Size of government and constitutions: Simple regression estimates

	(1)	(2)	(3)
Dependent variable	CGEXP	CGEXP	CGEXP
<i>PRES</i>	-6.08 (1.97)***	-5.29 (1.92)***	
<i>MAJ</i>	-3.29 (1.73)*	-5.74 (1.95)***	
<i>PROPRES</i>			-7.08 (2.70)**
<i>MAJPAR</i>			-7.30 (3.02)**
<i>MAJPRES</i>			-10.36 (2.70)***
Continents	No	Yes	Yes
Colonies	No	Yes	Yes
Sample	1990s, broad	1990s, broad	1990s, broad
Number of observations	80	80	80
Adjusted R^2	0.58	0.63	0.63

Note: Robust standard errors in parentheses. All regressions incl

```

model1 <- lm(cgexp ~
  pres + maj +
  lyp + gastil + age + trade +
  prop65 + prop1564 + federal +
  oecd, data=pt)
model2 <- lm(cgexp ~
  pres + maj +
  lyp + gastil + age + trade +
  prop65 + prop1564 + federal +
  oecd + col_uka + col_espa + col_otha +
  africa + asiae + laam, data=pt)
model3 <- lm(cgexp ~
  pro.pres + maj.parl + maj.pres +
  lyp + gastil + age + trade +
  prop65 + prop1564 + federal +
  oecd + col_uka + col_espa + col_otha +
  africa + asiae + laam, data=pt)

```

¹The data are available from the authors, though they may contain some coding errors in the institutional interaction variables. A corrected replication dataset is available here.

The result of simply calling `apsrtable(model1, model2, model3)` shouldn't be very far from usable output. We get decimal-aligned columns that are also lined up in the \LaTeX source, errors in parentheses, and a single star indicating $p < .05$. But, we want to include "robust" standard errors, omit a bunch of controls, give the variables less cryptic names, and, to complete the exercise, add more stars. All of these are extremely easy to do with `apsrtable`, but looking at the default output is a helpful baseline and a recommended first step every time.

Alternate Standard Errors However you want to go about estimating alternative standard errors, simply insert a vector or the full new variance-covariance matrix into the fitted model object and name it `se2`. To get the same standard errors that Persson and Tabellini report (also, I believe, the Stata default), use the `sandwich` package:

```
model1$se <- vcovHC(model1,type="HC1")
```

and so on for each model. By default, `apsrtable` will include only the robust standard errors when present, but the argument `se` can either ignore them and print the model's original standard errors ("vcov") or both ("both").

Order of Covariates Including nested and non-nested models side-by-side, lining up the covariates that are included and leaving blank those that are not, is the main feature of the package. For this replication (and most of the time) the default "lr" left-to-right incorporation of terms is correct. For special cases, two other options, "rl" and "longest" are provided and discussed briefly under implementation details. For typical use, it is best to nail down the desired order in which the models are passed to the function, and then the value of `order`.

Omitting controls Variables can be omitted from the display either by name or index. It is easiest in this case to supply the argument `omit=c(1,4:17)`. However, the order of coefficients can change, so a list of quoted character names of coefficients to omit may be safer, and certainly makes your code more transparent and thus easier to maintain and replicate.

Names of Models and Covariates Models by default are named "Model 1", "Model 2", etc., but numbering can be changed arbitrarily (via argument `model.counter`), or a vector of meaningful names can be supplied as `model.names` (if it comes up short, numbering accounts for the named ones).

Covariate name replacement takes place after the list of included coefficients from all the models has been

generated. Therefore, again it is a good idea to look at the default, settle on the values of `order` and `omitcoef`, and only then supply a vector of coefficient display names as the argument `coef.names`.

Stars The default behavior is to indicate coefficient significance at the level $p < .05$ with a single superscript asterisk. Two arguments allow you to increase the sparkle of almost any table. To set a different level for a single star, supply it to `lev`. To include a dagger for .10, a star for .05, two stars for .01, and three stars (!) for .001, supply the argument `stars="default"`. I admit the name is confusing, but "default" here indicates the R default rather than the *APSR* and `apsrtable` default.

Other arguments The other arguments are documented in the package, but two are worth mention here. The handling of notes about standard errors and significance indicators is discussed below. Also, `Sweave=TRUE` should be used whenever you include a call to `apsrtable` from within a `Sweave` document. Otherwise output will include the table environment code and empty caption and label, which should be included in the `Sweave` document itself if emacs `reftex` is to keep track of tables. An internal test for `Sweave` might appear in future versions, but the default output is meant for cutting and pasting from an `R` session into a \LaTeX document, which seems to be fairly common workflow.

Some Implementation Details

Three features of `apsrtable` bear mention for how I implemented them: covariate aggregation across models, table notes, and model summaries, which provides a flexible extension framework.

Covariate Order

Authors have immense control over the order in which covariate rows are presented in the final table. Through a combination of the order of terms in the model objects, and the choice of `order` in the call to `apsrtable`, one should be able to create tables that never need "rearrangement" of rows between output and publication. First, it builds a list of all variable names, and then notes the position of the variables in each model with respect to this final order. Next, the `omitcoef` list is marked, and finally, the names of remaining terms are replaced with the optional list of `coef.names`, which, of course, may contain \LaTeX markup. The column of variable names is set in text mode, so any math should be delimited by $\$$, and because it's R, backslashes have to be doubled. A label " β_0 Inter-

²Most model objects are simply lists rather than formal classes (that is, they use the S3 class system). Models fit with `lmer` are formally `mer` class objects but changes in how `lme4` methods `display` and `summary` work has prevented incorporation of that class of models so far.

cept” would be supplied as “ β_0 Intercept”.

The default left-to-right order starts with the first model; any terms included in the second not in the first are appended to the order, and so on. Right-to-left (`order="r1"`) takes the initial order from the rightmost model; and `order="longest"` starts with the order of terms in the model with the most terms, wherever it is in the list of models, then appends any others always left-to-right.

Table Notes

Some kinds of notes pertaining to the table depend on values used to generate it. In particular, informing readers of “robust” standard errors, and the indicators and “level” of statistical significance lies in the gap between content and presentation. Authors also commonly indicate the source of data in a note beneath a table (technically a multicolumn span within the tabular environment).

The `notes` argument allows you to specify a list of functions or character strings. R’s “lazy evaluation” means that these functions are evaluated only when `eval` is called explicitly inside `apsrtable`. R’s variable scoping meant that the variables in the `apsrtable()` call reside 3 levels up the call stack, so custom functions in the `notes` list can only access them by specifying the environment at `sys.frame(-3)`. Of course, simple character strings are valid in the `notes` list and will be typeset along with the results of the dynamic functions.

Extending `apsrtable` with `modelInfo`

Besides the coefficients, standard errors, and statistical significance star(s), model summaries and some indication of goodness-of-fit should be included with each model, which I call `modelInfo`. But, different researchers prefer to include different statistics describing models as a whole. Therefore, `apsrtable` comes with some reasonable defaults and a simple mechanism to change them. The same mechanism makes extending the package to other types of models relatively easy.

`ModelInfo` methods are called on the list of model summaries (that is, the result of `summary(model)`). At a minimum, one can select from the information contained therein to generate a list of named character strings that comprise the `modelInfo` that will be included in the final output. The formatter doesn’t care what you include here—it simply matches the names (always left-to-right) across models and prints it all. If the `summary` object does not contain some data that you need for a model summary statistic, modify the relevant summary method (e.g. `summary.lm()`).

So, what does a `modelInfo` method look like? It is a formal S4 method with two arguments: a model summary object, and `digits`, passed down explicitly from the `apsrtable` call. Therefore to change the list of model information, just change what the `modelInfo` function re-

turns for a given model class. Future versions will probably include a selection of presets, but changing by a quick call to `setMethod` is straightforward and my aim here is to make clear how to get the `modelInfo` you want. For `lm` objects, the default prints the N , R^2 , adjusted R^2 , and residual standard deviation. The example in `?modelInfo` shows you how to print only the N and residual sd; the easiest way to find the default is via `getMethod("modelInfo", "summary.lm")`. The defaults for `lm` and `glm` are also included as named private functions, to make reversion easier (see the `modelInfo` example).

To return to the Persson and Tabellini example, they summarize the controls used (but not displayed) in the `modelInfo` section. It is simple to write a `modelInfo` function for `lm` summaries that tests for the presence of a particular name among coefficients. Besides this, they report the N as “Number of Observations” (so we want to change the name of that element in the list), and the adjusted R^2 . So, we just create a custom `modelInfo` method that returns these items.

```
## Add robust se to the models
library(sandwich)
model1$se <- vcovHC(model1,type="HC1")
model2$se <- vcovHC(model2,type="HC1")
model3$se <- vcovHC(model3,type="HC1")

## Create and register custom modelInfo for lm

setMethod("modelInfo", "summary.lm", function(x) {
  env <- parent.frame()
  digits <- evalq(digits,env)
  model.info <- list(
    "Continents" = (
      ifelse(!is.na(charmatch("laam",rownames(coef(x)))),
        "\\mathrm{Yes}", "\\mathrm{No}")),
    "Colonies" = (
      ifelse(!is.na(charmatch("col",rownames(coef(x)))),
        "\\mathrm{Yes}", "\\mathrm{No}")),
    "Number of\\tabularnewline Observations" = (
      formatC(sum(x$df[1:2]),format="d")),
    "adj. $R^2$" = (
      formatC(x$adj.r.squared,format="f",digits=digits))
  )
  class(model.info) <- "model.info"
  return(model.info)
} )
```

With the new `modelInfo` method registered, the final call to `apsrtable()` is shown below, and the results typeset as Table 1.

```
apsrtable(model1,model2,model3,
omitcoef=c(
  "(Intercept)","lyp", "gastil", "age", "trade",
  "prop65", "prop1564", "federal" ,
```

```

"oecd" , "col_uka", "col_espa", "col_otha",
"africa", "asiae", "laam"),
coef.names=c("Presidential", "Majoritarian",
"Proportional Presidential",
"Majoritarian Parliamentary",
"Majoritarian Presidential"),
align="left", stars="default",
notes=list(se.note(), stars.note())
)

```

The example above shows how the `modelInfo` methods can be used to change the display for linear model summary objects. In theory, any model object for which `coef`, `vcov`, and `summary` methods should work with `apsrtable`. Extending means mainly creating `modelInfo` methods for other classes of model summaries.

Conclusion

I saw a striking need to produce publication-ready output as a part of the modeling process, especially as com-

plexity grew. Not surprisingly, I learned after publishing the package on CRAN and announcing it to PolMeth, that I was not alone. Martin Elff has a similar function (`mtable`) in his `memisc` package, which can also produce non- \LaTeX output. Some of the details of his implementation caused me to refactor parts of my own. My current workflow relies heavily on `Sweave` to ensure that the code I run is actually connected to both the theoretical arguments I make and the results I present. I rely on `apsrtable` to produce reliable, replicable, relatively painless tables of results. I hope that *TPM* readers can save themselves (and their graduate students!) some typesetting frustration by using `apsrtable`, and I look forward to extensions and improvements that users may provide.

References

- Persson, Torsten and Guido Tabellini. 2003. *The Economic Effects of Constitutions*. Boston, MA: The MIT Press.

Announcements

Visions in Methodology Workshop Supports Women in Political Methodology

Janet Box-Steffensmeier and Erin McAdams

The Ohio State University and The Ohio State University
jboxstef+@osu.edu and mcadams.19@osu.edu

Twenty methodologists attended the 2008 Visions in Methodology conference for Women in Political Methodology from October 2 to 4, 2008, in Columbus, Ohio. The conference was sponsored by the Methodology, Measurement, and Statistics Program and the Political Science Program at National Science Foundation as well as the Department of Political Science and The College of Social and Behavioral Sciences at Ohio State University. The conference was the first in a series of workshops designed to bring together female faculty whose research interests are related to political methodology. Janet Box-Steffensmeier and Corrine McConnaughy were the local hosts, while Erin McAdams was the conference coordinator.

The three-day workshop was highlighted by scholarly presentations and discussions on both methodological and substantive topics. Attendees also participated in discussion sessions based on a set of reading material on topics related to gender and careers in academia. Oral biographies were presented by featured senior scholars at the 2008 conference—Elisabeth Gerber (University of Michigan) and

Sara McLaughlin Mitchell (University of Iowa). Attendees were determined by a competitive proposal process and represented both junior and senior faculty across several subfields in political science. More information about the VIM conference, including the full program, can be found at: <http://polisci.osu.edu/faculty/jbox/vim.htm>

The conference was part of a broader goal of supporting activities for women in the field of political methodology that are funded by the National Science Foundation. These activities are intended to create networks and opportunities, as well as to plug the “leaky pipeline” for women in the field of political methodology in which women are underrepresented. This new initiative implements recommendations from the recent National Academy of Sciences reports, the APSA Workshop on the Advancement of Women in Academic Political Science, and the 2006 Political Methodology Long Range Strategic Planning Committee Report. The two key themes that emerged from these reports included the need for better networking and more systematic mentoring. The Women in Political Methodology conferences are

thus specifically designed to balance opportunities for scientific advancement, networking and professional mentoring on an intimate scale.

The next VIM will be sponsored and held at the

University of Iowa in 2009 and hosted by Sara McLaughlin Mitchell and Caroline Tolbert.



PHOTO CAPTION: Back Row (l to r): Janet Box-Steffensmeier (Ohio State University); Meg Shannon (University of Mississippi); Suzanna Linn (Penn State University); Jennifer Victor (University of Pittsburgh); Jennifer Jerit (Florida State University); Leslie Schwindt-Bayer (University of Missouri). Middle Row (l to r): Sara McLaughlin-Mitchell (University of Iowa); Sona Golder (Florida State University); Carole Wilson (University of Texas at Dallas); Amber Boydston (University of California, Davis); Elisabeth Gerber (University of Michigan); Dawn Brancati (Washington University in St. Louis). Front Row (l to r): Caroline Tolbert (University of Iowa); Miki Kittilson (Arizona State University); Wendy Tam Cho (University of Illinois at Urbana-Champaign); Laura Langbein (American University); Betsy Sinclair (University of Chicago); Gina Reingardt (Texas A&M University); Hyeran Jo (Texas A&M University). Not pictured: Melanie Springer (Washington University in St. Louis), Corrine McConaughy (Ohio State University), and Erin McAdams (Ohio State University and Photographer).

Department of Political Science
2010 Allen Building
Texas A&M University
Mail Stop 4348
College Station, TX 77843-4348

The Political Methodologist is the newsletter of the Political Methodology Section of the American Political Science Association. Copyright 2009, American Political Science Association. All rights reserved. The support of the Department of Political Science at Texas A&M in helping to defray the editorial and production costs of the newsletter is gratefully acknowledged.

Subscriptions to *TPM* are free to members of the APSA's Methodology Section. Please contact APSA (http://www.apsanet.org/section_71.cfm) to join the section. Dues are \$25.00 per year and include a free subscription to *Political Analysis*, the quarterly journal of the section.

Submissions to *TPM* are always welcome. Articles should be sent to the editors by e-mail (tpm@polisci.tamu.edu) if possible. Alternatively, submissions can be made on diskette as plain ascii files sent to Paul Kellstedt, Department of Political Science, 4348 TAMU, College Station, TX 77843-4348. See the *TPM* web-site, <http://polmeth.wustl.edu/tpm.html>, for the latest information and for downloadable versions of previous issues of *The Political Methodologist*.

TPM was produced using L^AT_EX on a Mac OS X running iTeXMac.



President: Philip A. Schrodt

University of Kansas
schrodt@ku.edu

Vice President: Jeff Gill

Washington University in St. Louis
jgill@wustl.edu

Treasurer: Jonathan Katz

California Institute of Technology
jkatz@hss.caltech.edu

Member-at-Large: Robert Franzese

University of Michigan
franzese@umich.edu

Political Analysis Editor: Christopher Zorn

Pennsylvania State University
politicalanalysis@la.psu.edu